Am J Transl Res 2020;12(3):912-922
www.ajtr.org /ISSN:1943-8141/AJTR0102296

Original Article

Role of m^{5}C-related regulatory genes in the diagnosis and prognosis of hepatocellular carcinoma

Yuting He[1,2,3,4*], Xiao Yu[1,2,3,4*], Jie Li[1,2,3,4*], Qiyao Zhang[1,2,3,4], Qingyuan Zheng[1,2,3,4], Wenzhi Guo[1,2,3,4]

[1]Department of Hepatobiliary and Pancreatic Surgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, Henan, P. R. China; [2]Key Laboratory of Hepatobiliary and Pancreatic Surgery and Digestive Organ Transplantation of Henan Province, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, Henan, P. R. China; [3]Open and Key Laboratory of Hepatobiliary & Pancreatic Surgery and Digestive Organ Transplantation at Henan Universities, Zhengzhou 450052, Henan, P. R. China; [4]Henan Key Laboratory of Digestive Organ Transplantation, Zhengzhou 450052, Henan, P. R. China. [*]Equal contributors.

Received September 13, 2019; Accepted February 19, 2020; Epub March 15, 2020; Published March 30, 2020

Abstract: Hepatocellular carcinoma (HCC) is one of the most common malignancies globally, and is frequently associated with a poor prognosis. 5-methylcytosine (m^{5}C) is a common epigenetic modification with many critical roles in eukaryotes. However, the expression and functional roles of m^{5}C regulators are largely unknown. In this study, we utilized The Cancer Genome Atlas (TCGA) to determine the expression, gene signatures, and prognostic values of m^{5}C-related genes. We confirmed that the frequency of mutation events of m^{5}C regulatory genes was high in HCC (35/363). Dysregulation of m^{5}C-related genes was also associated with a higher HCC stage. Moreover, a strong relationship was found between the expression of m^{5}C regulatory genes and HCC patient survival. High expression of *NSUN4* and *ALYREF* correlated significantly with survival outcome. We developed a two-gene signature of m^{5}C regulators with HCC prognostic value based on the least absolute shrinkage and selection operator (LASSO) and multivariate Cox regression models. Gene set enrichment analysis (GSEA) results indicated that high expression of *NSUN4* was associated with methylation and demethylation processes. Meanwhile, high expression of *ALYREF* was clearly related to cell cycle regulation and mitosis. In conclusion, our results revealed that m^{5}C-related genes play an essential role in tumor progression in HCC. Further detection of m^{5}C methylation could provide a novel method for HCC targeted therapy.

Keywords: Hepatocellular carcinoma, m^{5}C, TCGA

Introduction

According to a recent report, hepatocellular carcinoma (HCC) is the third most common cause of cancer-related death globally (8.2% of all cases), and thus is a severe public health problem [1, 2]. Despite considerable progress in the development of therapeutic strategies for HCC, the five-year survival rate of advanced HCC remains poor due to late-stage diagnosis, metastasis, and postsurgical recurrence [3-5]. Therefore, investigation into the molecular mechanism of tumorigenesis and tumor progression is critical for identifying novel biomarkers for early detection and effective therapeutic targets for treating patients with HCC.

Growing evidence suggests that post-transcriptional modification of RNA plays essential roles in different cancers [6, 7]. Genetic and epigenetic alterations of RNA and histones have been widely studied in tumor progression, leading to the development of many therapeutic modalities, including histone deacetylase inhibitors and drugs targeting hypoxia-related pathways [8]. Although numerous RNA modifications have been documented, our understanding about their regulation and function remains limited, especially for mRNAs, due to technical limitations in accurately locating modifications across the genome. Common RNA methylation sites include 5-methylcytosine (m^{5}C), 7-methylguanosine (m^{7}G), m^{1}G, m^{2}G, m^{6}G, N1-methyladenosine (m^{1}A), and m^{6}A [9-16]. m^{5}C modification promotes splicing and translation [17]. In comparison, m^{1}G, m^{2}G, and m^{1}A modifications, which occur at the first or second codon, repress protein synthesis, while tRNA m^{7}G methylation is required for mRNA translation into proteins.

Reversible m5C modification, which is one of the longest-known post-transcriptional modifications of RNA, plays essential roles in regulating mRNA alternative splicing, export, stabilization localization, and translation [18, 19]. The regulatory role of m5C in mRNAs is beginning to be revealed, as an increasing number of studies has demonstrated a broad effect of m5C on tRNAs, rRNAs, and mRNAs. m5C methylation involves a series of regulators, including m5C methyltransferases, demethylases, and "readers". The methyltransferase "writer" complex increases methylation at the C5 position of RNAs, then different "reader" proteins recognize and bind the methylated mRNAs, while the "eraser" protein reverses m5C modification by degrading the written methylation [20]. The m5C modification involves adenosine methyltransferases, demethylases, and RNA-binding proteins called m5C "writers", including NSUN1-7, DNMT1, DNMT2, DNMT3A, and DNMT3B, m5C "erasers" such as TET2, and m5C "readers" such as ALYREF [21]. Accumulating lines of evidence have demonstrated that m5C modification plays a pivotal regulatory role in various important biological processes, including the structure, stability, and translation of mRNAs [22, 23]. However, the specific gene signatures and prognostic values of m5C-related regulators in cancers, including HCC, remain largely unknown.

In this study, we analyzed a TCGA dataset for m5C-related regulators involved in HCC. Our results revealed a close correlation between genetic alterations of changes in the spectrum of 13 m5C regulatory genes and clinical pathological features, including HCC survival. Moreover, the changes observed in the m5C regulatory gene were significantly associated with a higher tumor stage. Based on the least absolute shrinkage and selection operator (LASSO) and multivariate Cox regression models, we built a two-gene signature of m5C regulators that aligns with an HCC prognostic value to effectively predict the prognosis of patients with HCC. In addition, high *NSUN4* or *ALYREF* expression was associated with critical biological processes in HCC. In conclusion, we identified changes in m5C-related genes, which could affect key regulatory molecules that contribute to HCC progression.

Materials and methods

Liver hepatocellular carcinoma (LIHC) dataset acquisition and process

All HCC copy number variants (CNV), single nucleotide variants (SNV), mRNA, and all corresponding clinical data used in our study were downloaded from the TCGA data portal (http://gdc-portal.nci.nih.gov/) as described previously [24]. We obtained 423 and 364 samples for transcriptome and SNV data, respectively. For CNV data, 375 samples were analyzed by the R package RTCGA. There were 377 clinical samples from which clinical information data was acquired. After integrating the data and excluding samples with a survival time of less than 90 days, a total of 319 HCC samples were available for further analysis.

Gene set enrichment analysis (GSEA)

GSEA, which is available in Java, was used to determine which gene sets were associated with m5C gene expression in the TCGA dataset, as described previously [25].

The least absolute shrinkage and selection operator (LASSO)

LASSO is an accepted method for regression analysis of high-dimensional data. In our study, LASSO was adopted to choose optimal predictive biomarkers for m5C regulatory genes in TCGA-LIHC. The selected factors in the LASSO regression model were analyzed by multivariate analysis. Two m5C regulatory genes (*NSUN4* and *ALYREF*) were filtrated as the target genes. Kaplan-Meier survival analysis, risk prediction and receiver operating characteristic (ROC) curves were applied to evaluate the prognostic value of *NSUN* and *ALYREF*.

Statistical analysis

SPSS 23.0 software (IBM Corp, Armonk, NY, USA) and R language were used for the statistical analyses. Kaplan-Meier and log-rank tests were adopted to analyze the survival time of patients. The chi-squared test was used to evaluate the association of CNV and SNV with m5C regulatory genes and clinicopathological characteristics. *P*-values less than 0.05 were considered statistically significant.

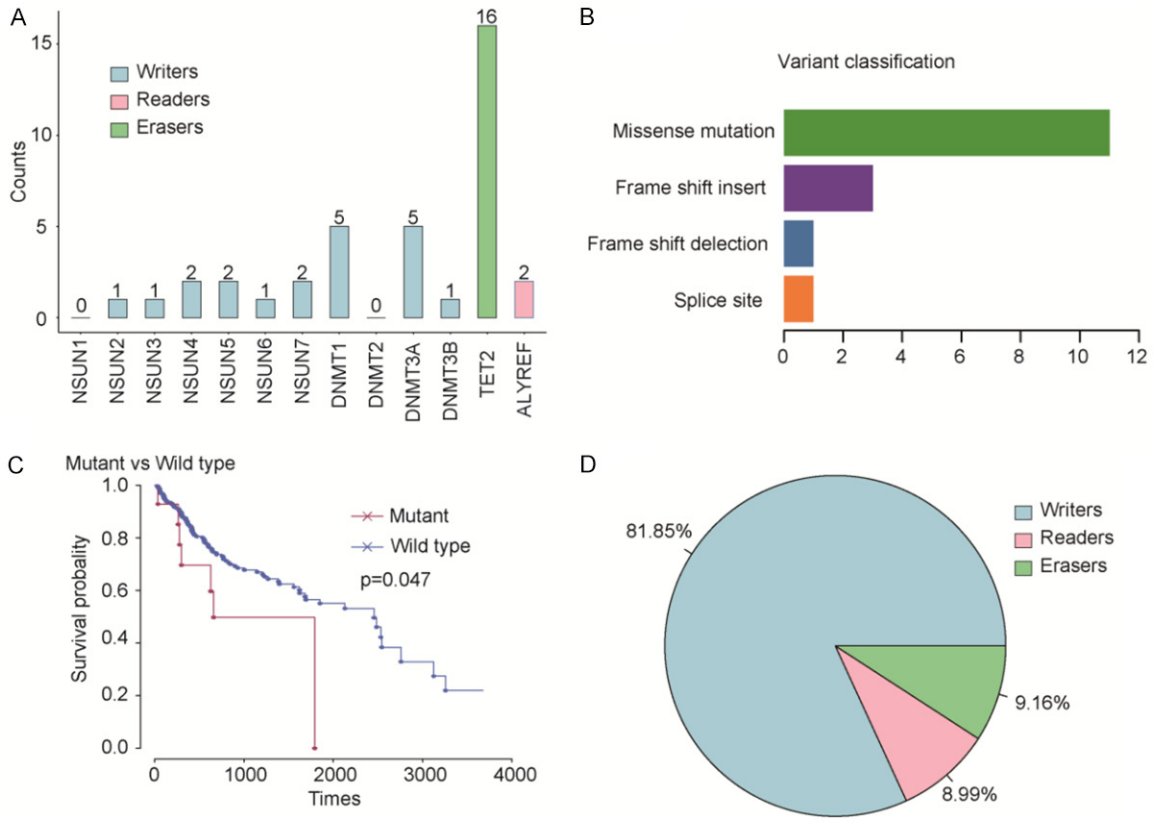

Figure 1. Relationship between mutations of m^{5}C regulatory genes and hepatocellular carcinoma (HCC). A. Frequency of mutations of different m^{5}C regulatory genes in HCC; *TET2* had the highest frequency of mutation events (16/363). B. Classification of mutations of different m^{5}C regulatory genes in HCC; missense mutations were the most frequent mutation event (68.75%). C. Kaplan-Meier analysis of the association between mutations of m^{5}C regulatory genes and survival in patients with HCC; patients with genetic mutations had poorer prognosis. D. Frequency of CNV of m^{5}C regulatory genes in HCC; "writer" genes had the highest frequency of CNV (81.85%).

Results

Mutations and CNVs of m^{5}C-related genes in HCC patients

Among the 363 HCC cases with available sequencing data, mutations in m^{5}C-related genes were observed in 35 independent samples. More specifically, the m^{5}C "eraser" *TET2* had the highest frequency of mutation events (16/363) followed by the m^{5}C "writer" genes *DNMT2* and *DNMT3A* (both 5/363). The frequency of mutation of the m^{5}C "eraser" gene was over eight times that of the average number of "writer" and "reader" genes. Moreover, we found that m^{5}C "writer" genes displayed a broad range of genetic mutations compared with the m^{5}C "eraser" and "reader" genes (**Figure 1A**). For the SNV, we found that 7 of 13 m^{5}C regulator genes possessed functional changes, with missense mutations being the most frequent mutation event (68.75%) (**Figure 1B** and **Table 1**). Additionally, we observed that the seven patients exhibiting genetic mutations had the worst survival in patients with HCC (**Figure 1C**).

Furthermore, CNV was found to play an essential role in the changes of m^{5}C regulator genes. We found that "writers" exhibited the highest frequency of mutation in the 375 samples with CNV data (**Figure 1D**). In particular, *TET2* displayed the highest frequency of mutation (40.42%), followed by *ALYREF* (35.29%). However, the *NSUN1* and *DNMT2* genes were not detected (**Table 2**). Taken together, these results suggest that mutation of the m^{5}C genes may be frequent in HCC.

Effect of m^{5}C-related genes SUV and CNV in HCC prognosis and key regulatory molecules

We performed a Cox analysis to explore the effect of m^{5}C-related gene SNV and CNV on the

Table 1. Functional changes of m5C regulatory genes in hepatocellular carcinoma

Gene	Frame Shift Deletion	Frame Shift Insert	Missense Mutation	Splice Site	Total
DNMT3A	1	0	3	0	4
TET2	0	2	2	0	4
DNMT1	0	0	2	1	3
NSUN4	0	1	1	0	2
DNMT3B	0	0	1	0	1
NSUN6	0	0	1	0	1
NSUN7	0	0	1	0	1

Table 2. Copy number variants (CNV) of m5C regulatory genes in hepatocellular carcinoma

Function	Genes	Diploid	Deletion	Amplification	CNV sum	Amplification %	Deletion %	Percentage
Writer	NSUN1	-	-	-	-	-	-	-
	NSUN2	250	10	118	128	92.19%	7.81%	33.86%
	NSUN3	329	12	36	48	75.00%	25.00%	12.73%
	NSUN4	274	76	29	105	27.62%	72.38%	27.70%
	NSUN5	263	15	95	110	86.36%	13.64%	29.49%
	NSUN6	307	19	51	70	72.86%	27.14%	18.57%
	NSUN7	291	48	41	89	46.07%	53.93%	23.42%
	DNMT1	289	57	36	93	38.71%	61.29%	24.35%
	DNMT2	-	-	-	-	-	-	-
	DNMT3A	321	12	47	59	79.66%	20.34%	15.53%
	DNMT3B	291	2	85	87	97.70%	2.30%	23.02%
Eraser	TET2	227	146	8	154	5.19%	94.81%	40.42%
Reader	ALYREF	242	7	125	132	94.70%	5.30%	35.29%

prognosis of patients with HCC. As shown in **Table 3**, a higher cancer stage is associated with a worse prognosis; however, neither SNV or CNV were associated with the HCC prognosis. Studies have shown that alterations in *TP53*, *CYP1A1*, *NQO1*, *ALDH2*, and *EPHX1* play essential roles in the underlying mechanism regulating the pathogenesis of HCC [26-30]. Therefore, we investigated the effect of m5C-related gene SNV and CNV on these key regulatory molecules and their alteration. Our data demonstrate that *TP53* alteration was markedly associated with SNV and CNV of m5C-related genes (**Table 4**).

Next, we elevated the effect of SNV and CNV on the expression of m5C regulatory genes. We detected 11 of 13 m5C genes in the 423 samples analyzed. The level of mRNA expression of all 11 m5C regulatory genes was associated significantly with CNV. Furthermore, CNV-mediated amplification of nine genes (*NSUN2*, *NSUN3*, *NSUN4*, *NSUN5*, *NSUN6*, *NSUN7*, *DNMT3A*, *TET2,* and *ALYREF*) was linked to increased mRNA expression (**Figure 2A-I**). However, no differences were observed for two genes (*DNMT1* and *DNMT3B*) (**Figure 2J** and **2K**). These findings indicate that SNV and CNV of m5C-related genes could affect the expression of key regulatory molecules and contribute to HCC progression. Moreover, the expressions of all "eraser" and "reader" genes were associated significantly with CNV, suggesting that these genes play important roles in m5C regulation.

Relationship between the expression of m5C-related genes and clinical grade in HCC

To further investigate the expression of m5C-related genes in different clinical grades, we defined TNM I and II stages as low TNM stage (low stage), and TNM III and IV stages as high TNM stage (high stage). Then, we analyzed the TCGA database to further validate the expression of m5C-related regulators in different TNM stages. We found that the expression of *DNMT3A*, *NSUN4*, *NSUN5*, *DNMT1*, *TET2,* and *ALYREF* were substantially overexpressed in

Table 3. Results of Cox regression analysis to explore the effect of m5C-related gene single nucleotide variants (SNV) and copy number variants (CNV) in the prognosis of patients with hepatocellular carcinoma (HCC)

	beta	HR	(95% CI for HR)	Wald test	P
Stage	-1	0.37	(0.25-0.54)	26	3.20E-07
Gender	-0.3	0.74	(0.51-1.1)	2.2	0.13
SNV	0.19	1.2	(0.66-2.2)	0.39	0.53
CNV	0.17	1.2	(0.66-2.1)	0.33	0.57
Without SNV or CNV	0.079	1.1	(0.61-1.9)	0.07	0.79

Abbreviations: HR = hazard ratio.

Table 4. Association between key regulatory molecules and m5C-related gene single nucleotide variants (SNV) and copy number variants (CNV) in hepatocellular carcinoma

			Without SNV or CNV	With SNV or CNV	X^2	P
TP53		Wild type	38	183	7.22	0.007
	n=327	Altered	6	100		
CYP1A1		Wild type	325	0	-	-
	n=327	Altered	0	2		
NQO1		Wild type	326	0	-	-
	n=327	Altered	0	1		
ALDH2		Wild type	326	0	-	-
	n=327	Altered	0	1		
EPHX1		Wild type	43	278	0.138	0.711
	n=327	Altered	1	5		

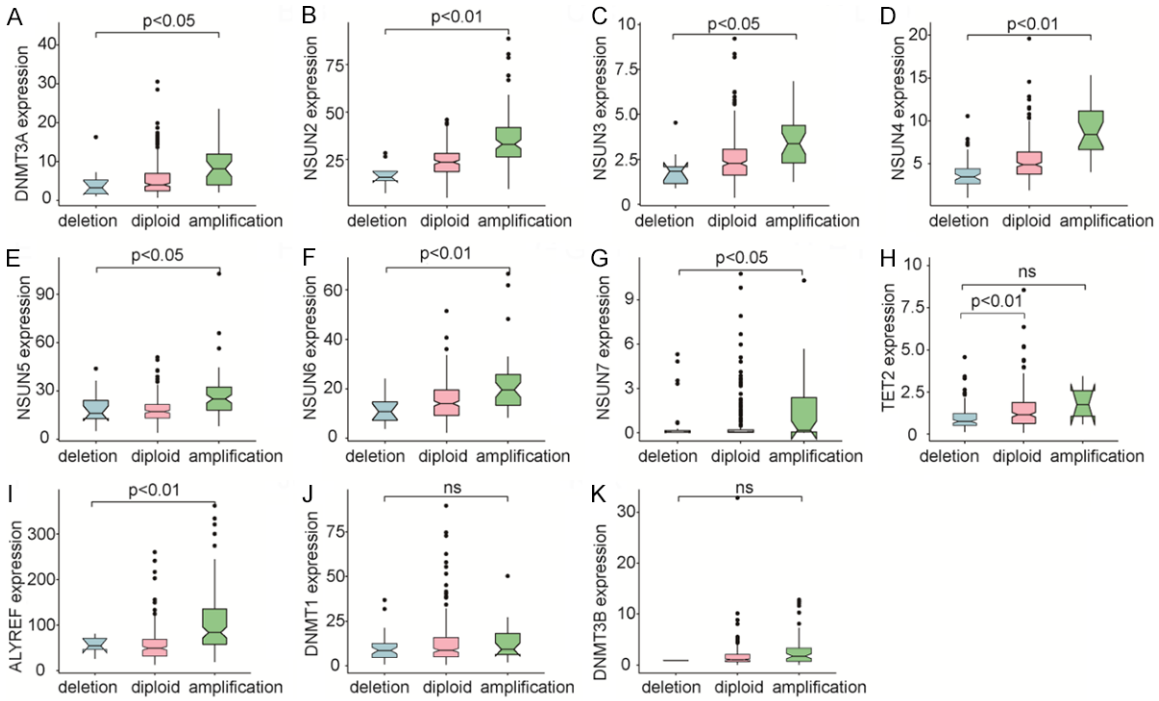

Figure 2. Analysis of m5C regulatory gene expression with copy number variants (CNV) and single nucleotide variants (SNV). A-I. Comparison of expression levels of different m5C regulatory genes in different SNV and CNV; results show that different SNV and CNV were significantly associated with nine genes (*NSUN2*, *NSUN3*, *NSUN4*, *NSUN5*, *NSUN6*, *NSUN7*, *DNMT3A*, *TET2,* and *ALYREF*). J, K. SNV and CNV were not significantly associated with two genes (*DNMT1* and *DNMT3B*).

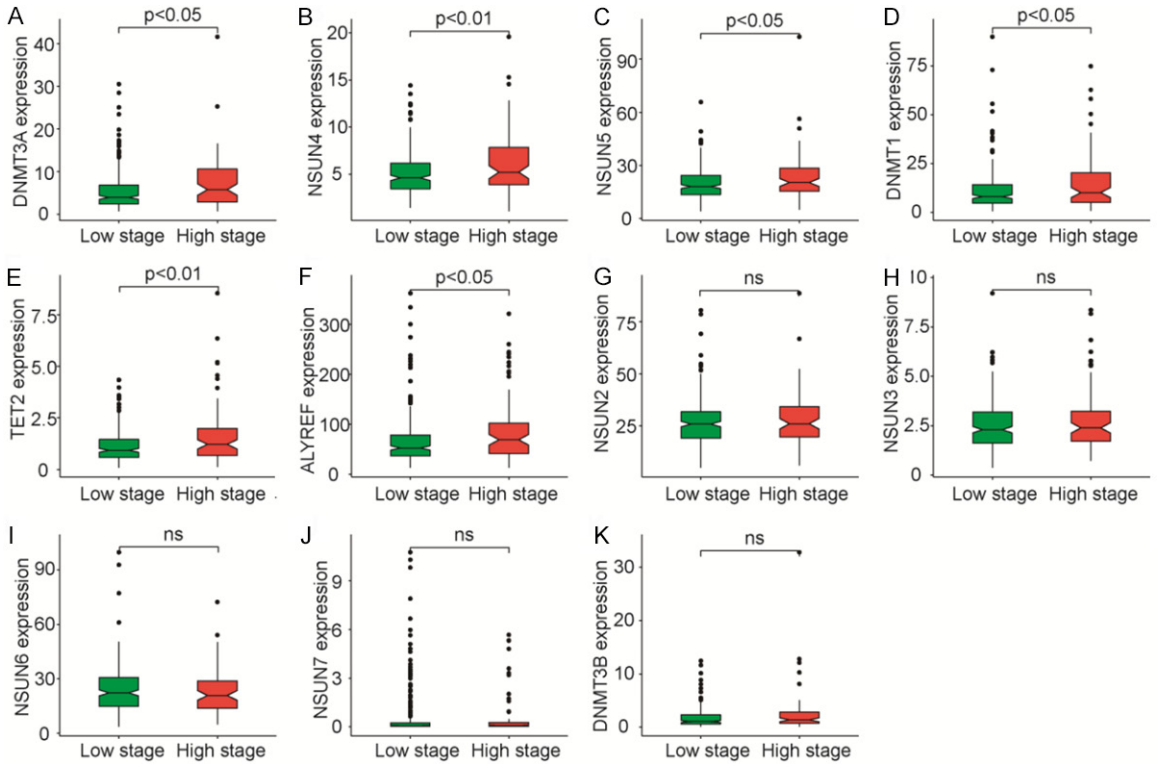

Figure 3. Analysis of m^{5}C regulatory gene expression by hepatocellular (HCC) TNM stages. A-F. Comparison of expression levels of m^{5}C regulatory genes in patients with different TNM stages of HCC. The expression levels of six genes (*DNMT3A*, *NSUN4*, *NSUN5*, *DNMT1*, *TET2,* and *ALYREF*) were overexpressed significantly among patients with high TNM stages. G-K. The expression levels of five genes (*NSUN2*, *NSUN3*, *NSUN6*, *NSUN7,* and *NSUN3B*) were not different across TNM stages.

high stage HCC tissues (**Figure 3A-F**). In contrast, no difference was observed in five "writer" genes (*NSUN2*, *NSUN3*, *NSUN6*, *NSUN7,* and *NSUN3B*) (**Figure 3F-J**). Taken together, most m^{5}C-related gene expression was up-regulated, suggesting the possibility of an important relationship between the expression of m^{5}C-related genes and HCC stage.

Relationship between expression of m^{5}C-related genes and HCC prognosis

Our analysis also revealed a significant association between HCC TNM stage and patient prognosis (**Figure 4A** and **4B**). Thus, we used CNV and SNV as the research object to analyze the relationship between m^{5}C regulatory genes and patient survival in HCC. No obvious differences were observed between CNV or SNV of the m^{5}C regulatory genes and patient survival (**Figure 4C** and **4D**). Considering the positive correlation between the expression of some m^{5}C regulatory genes and CNV, we performed a univariate Cox regression analysis to explore the relationship between different m^{5}C regulatory gene expression levels and patient prognosis. Our results show that the expression level of eight genes (*NSUN4*, *ALYREF*, *DNMT3A*, *DNMT1*, *DNMT3B*, *NSUN5*, *NSUN2*, and *NSUN3*) correlated significantly with patient prognosis (P<0.05) (**Table 5**). Moreover, the expression level of six of the eight genes (*NSUN4*, *ALYREF*, *DNMT3A*, *NSUN5*, *NSUN2*, and *NSUN3*) correlated significantly with their CNV changes. We also found that gene expression information from 13 m^{5}C regulatory genes can be used to assess patient risk (**Figure 4E**) and that the area under the curve (AUC) for 1, 3, and 5 years is greater than 0.7 (**Figure 4F**) with multi-factor Cox regression analysis. Furthermore, we adopted LASSO analysis to predict the potential survival predictor. As showed in **Table 6**, only *NSUN4* and *ALYREF* were valid survival predictors in the regression analysis. These results indicate that the expression of m^{5}C regulatory genes can serve as a prognostic marker for HCC.

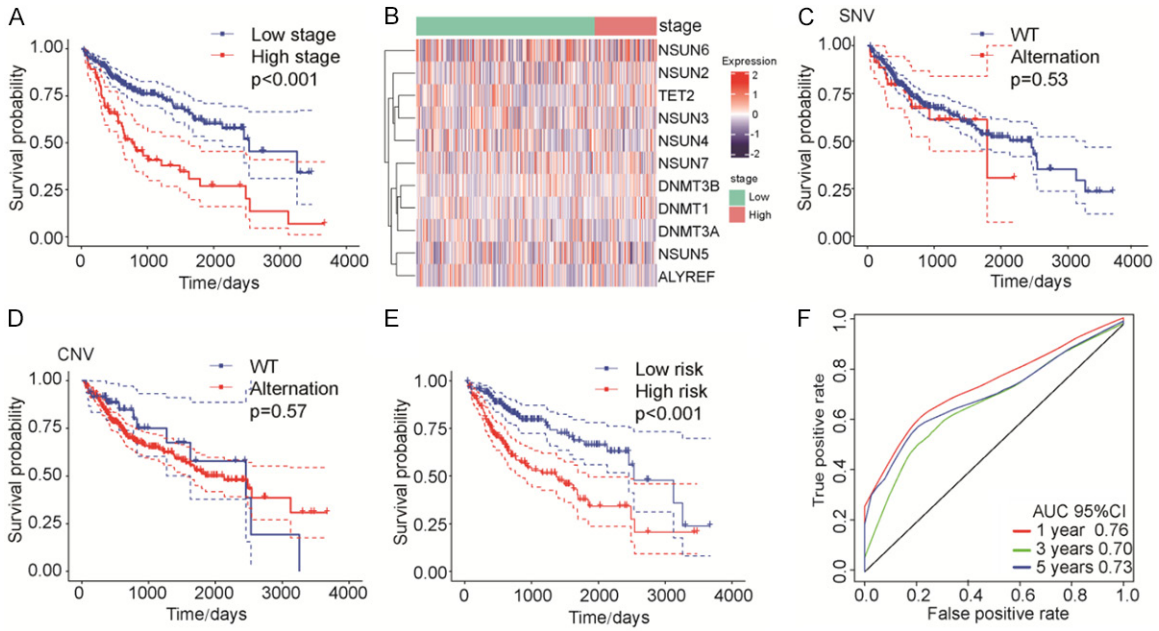

Figure 4. Relationship between m5C regulatory genes and prognosis prediction in hepatocellular carcinoma (HCC). A. Kaplan-Meier analysis shows that survival time was longer in patients with low TNM stages compared to patients with high TNM stages. B. The heat map shows different expression levels of m5C regulatory genes at different TNM stages. C, D. Kaplan-Meier curves of SNV or CNV and prognosis in patients with HCC. E. Kaplan-Meier curves of risk score and prognosis of patients with HCC; there were no difference in patients with SNV or CNV. F. The results of the receiver operating characteristic (ROC) curve of prognosis based on m5C regulatory genes at 1-year, 3-year, and 5-year survival were over 0.7.

Table 5. Results of Cox regression analysis to explore the effect of m5C-related gene expression in the prognosis of patients with hepatocellular carcinoma (HCC)

	beta	HR	(95% CI for HR)	Wald test	P	CNV sig
NSUN4	0.17	1.2	(1.1-1.3)	25	4.90E-07	yes
ALYREF	0.006	1	(1-1)	19	1.00E-05	yes
DNMT3A	0.069	1.1	(1-1.1)	18	2.70E-05	yes
DNMT1	0.024	1	(1-1)	16	5.70E-05	no
DNMT3B	0.077	1.1	(1-1.1)	11	0.0011	no
NSUN5	0.018	1	(1-1)	9.6	0.002	yes
NSUN2	0.022	1	(1-1)	7.7	0.0056	yes
NSUN3	0.18	1.2	(1-1.4)	6.5	0.011	yes
TET2	0.098	1.1	(0.93-1.3)	1.3	0.25	yes
NSUN7	0.045	1	(0.93-1.2)	0.57	0.45	yes
NSUN6	-0.0037	1	(0.97-1)	0.11	0.74	yes

Abbreviations: HR = hazard ratio; CNV = copy number variant.

Survival and functional enrichment analyses of NSUN4 and ALYREF in HCC

Considering the results above, we investigated the correlation of *NSUN4* and *ALYREF* expression with HCC patient prognosis. As shown in **Figure 5A** and **5B**, the up-regulation of *NSUN4* and *ALYREF* expression was linked with worse overall survival (OS). Univariate Cox regression analysis revealed that the AUC (1, 3, and 5 years) of the two m5C regulatory genes was greater than 0.6 (P<0.001) (**Figure 5C**). Additionally, a cluster analysis based on *NSUN4* and *ALYREF* expression and risk values of patients suggested a predisposition of the two regulators to different cohorts of at-risk patients (**Figure 5D**). The expression of *NSUN4* and *ALYREF* were also associated with the TNM stage (**Figure 5E**). We used GSEA to predict the functional role of *NSUN4* and *ALYREF* in HCC progression. As shown in **Figure 5F** and **5G**, the high expression of *NSUN4* was remarkably associated with methylation and methylation processes. Meanwhile, the high expression of *ALYREF* was found to be related to cell cycle regulation and mitosis

Table 6. Least absolute shrinkage and selection operator (LASSO) analysis of m5C-related genes in hepatocellular carcinoma

duplicates	genes	CNV expression sig	functions	Survival sig	Stage sig
959	NSUN4	***	Writers	yes	**
955	ALYREF	***	reader	yes	**
653	NSUN5	*	Writers	yes	**
353	DNMT3A	***	Writers	yes	*
298	DNMT1	No	Writers	yes	*
121	TET2	***	eraser	no	**
119	NSUN7	No	Writers	no	No
109	DNMT3B	No	Writers	yes	No
86	NSUN6	***	Writers	no	No
57	NSUN2	***	Writers	yes	No
27	NSUN3	**	Writers	no	No

Abbreviations: CNV = copy number variant; *P<0.05; **P<0.001; ***P<0.001.

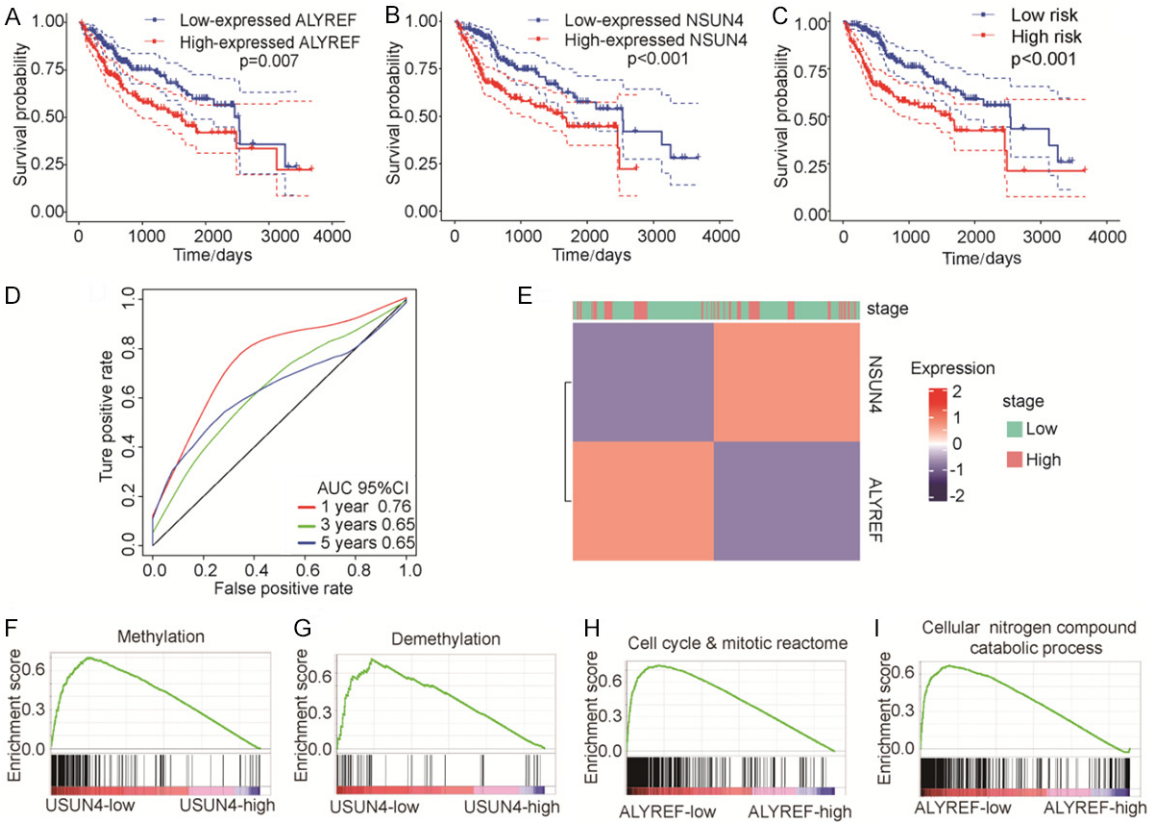

Figure 5. Role of ALYREF and NSUN4 expression in hepatocellular carcinoma (HCC) prognosis. A. Kaplan-Meier curves of differential ALYREF expression and prognosis in patients with HCC; patients with high *ALYREF* expression had poorer overall survival (OS). B. Kaplan-Meier curves of differential *NSUN4* expression and prognosis in patients with HCC; patients with high *NSUN4* expression had poorer OS. C. Kaplan-Meier curves of risk score and prognosis of patients with HCC; high-risk patients had poorer OS. D. Receiver operating characteristic (ROC) curves of a prognostic signature based on *ALYREF* and *NSUN4* expression for 1 year, 3 year, and 5 year survival. E. *ALYREF* and *NSUN4* expression at different TNM stages. F, G. High *NSUN4* expression was associated with methylation and demethylation. H, I. High *ALYREF* expression was associated with cell cycle, mitotic reactome and cellular nitrogen compound catabolic processes.

(**Figure 5H** and **5I**). These findings suggest that m5C-related genes play essential roles in critical biological processes of HCC.

Discussion

In this study, we analyzed the roles of m5C-related regulators in HCC and explored the relationship between the gene expression of m5C regulators and HCC prognosis. Analyses of the TCGA databases revealed that mutation of m5C regulatory genes is linked to patient prognosis in HCC. These findings provide clues for the epigenetic understanding of m5C in HCC.

A growing body of evidence shows that m5C methylation involves a variety of cellular processes. m5C regulator genes play key roles in cell proliferation, developmental defects, and cell death [31, 32]. Although there have been many reports describing the importance of m5C methylation in many pathological processes, little is known about the relationship between m5C-related genes and HCC. We observed that *NSUN4* and *ALYREF* may be involved in regulating the methylation process and DNA recombination in HCC. Our results are consistent with Shinoda *et al.*, who reported that m5C modifications play critical roles in mitochondrial protein synthesis and produce respiratory chain complexes in mitochondrial tRNAs. Studies have also demonstrated that NSUN2 could introduce m5C into cytoplasmic tRNAs and mRNAs [33]. Additionally, Van Haute *et al.* found that NSUN2 was essential for m5C generation at positions 48, 49, and 50 of several mammalian mitochondrial tRNAs [34]. These published reports support our findings that NSUN4 and ALYREF may play essential roles in critical biological processes of HCC.

Our results also show that high expression of *DNMT3A*, *NSUN4*, *NSUN5*, *DNMT1*, *TET2*, and *ALYREF* are related to higher HCC stages, suggesting a robust relationship between the expression of m5C-related genes and clinical TNM stages in HCC. Furthermore, a recent study demonstrated the involvement of m5C in the development of human urothelial carcinoma of the bladder (UCB). Chen *et al.* found that NSUN2 and YBX1 could target the m5C methylation site in the *HDGF* 3'-untranslated region and drive UCB pathogenesis. Moreover, up-regulated *NUSN2*, *YBX1*, and *HDGF* expression predicted poorer survival rate in patients with UCB. In the present study, we found that high levels of *NSUN4* and *ALYREF* expression were associated with poorer outcomes in patients with HCC. In conclusion, these findings suggest that m5C regulator genes may be novel biomarkers for diagnosis and predictors of prognosis.

To further define the molecular mechanisms of HCC progression and develop novel strategies for prognosis prediction, we performed a Cox regression analysis to identify the independent prognostic factors in HCC. Our results show that the expression of eight genes correlated significantly with patient prognosis, with the expression of six genes linked to significant changes in CNV. Univariate Cox regression and LASSO analyses showed that the expression of *NSUN4* and *ALYREF* were independent prognostic factors in HCC. Taken together, the expression of m5C regulatory genes can serve as prognostic markers for HCC, and SNV and CNV of m5C-related genes may affect key regulatory molecules to mediate HCC progression.

Conclusion

Our results demonstrate that the role of m5C-related regulators in HCC are dysregulated and associated with patient survival. *NSUN4* and *ALYREF* may be tumor oncogenes in HCC and may be useful as novel prototype therapeutic agents and potential biomarkers in patients with HCC.

Acknowledgements

We thank the patients and investigators who participated in TCGA for providing the data analyzed in this study. This study was supported by funds from the National Natural Science Foundation of China (81671958 and 81902832); Science and Technology Innovation Talents in Henan Universities (19HASTIT003); Henan Medical Science and Technology Research Project in 2018 (SBGJ2018002); the Science and Technology research project of Henan province (192102310117); and the Tianqing Liver Diseases Research Fund (TQGB20200073).

Disclosure of conflict of interest

None.

Address correspondence to: Drs. Yuting He and Wenzhi Guo, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, Henan, P. R. China. Tel: +86-371-67967126; E-mail: fccheyt1@zzu.edu.cn (YTH); Tel: +86-371-67967128; E-mail: guowz66@163.com (WZG)

References

[1] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA and Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018; 68: 394-424.

[2] Couri T and Pillai A. Goals and targets for personalized therapy for HCC. Hepatol Int 2019; 13: 125-137.

[3] Lee S, Kang TW, Cha DI, Song KD, Lee MW, Rhim H, Lim HK, Sinn DH, Kim JM and Kim K. Radiofrequency ablation vs. surgery for perivascular hepatocellular carcinoma: propensity score analyses of long-term outcomes. J Hepatol 2018; 69: 70-78.

[4] Hiraoka A, Kawamura T, Aibiki T, Okudaira T, Toshimori A, Yamago H, Nakahara H, Suga Y, Azemoto N, Miyata H, Miyamoto Y, Ninomiya T, Murakami T, Ishimaru Y, Kawasaki H, Hirooka M, Abe M, Matsuura B, Hiasa Y and Michitaka K. Prognosis and therapy for ruptured hepatocellular carcinoma: problems with staging and treatment strategy. Eur J Radiol 2015; 84: 366-371.

[5] Hua X, Sanjiv K, Gad H, Pham T, Gokturk C, Rasti A, Zhao Z, He K, Feng M, Zang Y, Zhang J, Xia Q, Helleday T and Warpman Berglund U. Karonudib is a promising anticancer therapy in hepatocellular carcinoma. Ther Adv Med Oncol 2019; 11: 1758835919866960.

[6] Nachtergaele S and He C. The emerging biology of RNA post-transcriptional modifications. RNA Biol 2017; 14: 156-163.

[7] Kiss T. Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. EMBO J 2001; 20: 3617-3622.

[8] Zhang LS, Liu C, Ma H, Dai Q, Sun HL, Luo G, Zhang Z, Zhang L, Hu L, Dong X and He C. Transcriptome-wide Mapping of Internal N(7)-methylguanosine Methylome in mammalian mRNA. Mol Cell 2019; 74: 1304-1316.e1308.

[9] Song J and Yi C. Chemical modifications to RNA: a new layer of gene expression regulation. ACS Chem Biol 2017; 12: 316-325.

[10] Liu Y and Santi DV. m5C RNA and m5C DNA methyl transferases use different cysteine residues as catalysts. Proc Natl Acad Sci U S A 2000; 97: 8263-8265.

[11] Andreev DE, Dmitriev SE, Terenin IM, Prassolov VS, Merrick WC and Shatsky IN. Differential contribution of the m7G-cap to the 5' end-dependent translation initiation of mammalian mRNAs. Nucleic Acids Res 2009; 37: 6135-6147.

[12] Knutson CG, Akingbade D, Crews BC, Voehler M, Stec DF and Marnett LJ. Metabolism in vitro and in vivo of the DNA base adduct, M1G. Chem Res Toxicol 2007; 20: 550-557.

[13] Keith G and Heyman T. Heterogeneities in vertebrate tRNAs(Trp) avian retroviruses package only as a primer the tRNA(Trp) lacking modified m2G in position 7. Nucleic Acids Res 1990; 18: 703-710.

[14] Trecant C, Dlubala A, George P, Pichat P, Ripoche I and Troin Y. Synthesis and biological evaluation of analogues of M6G. Eur J Med Chem 2011; 46: 4035-4041.

[15] Li X, Xiong X, Zhang M, Wang K, Chen Y, Zhou J, Mao Y, Lv J, Yi D, Chen XW, Wang C, Qian SB and Yi C. Base-resolution mapping reveals distinct m(1)A methylome in nuclear- and mitochondrial-encoded transcripts. Mol Cell 2017; 68: 993-1005, e1009.

[16] Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, Cesarkas K, Jacob-Hirsch J, Amariglio N, Kupiec M, Sorek R and Rechavi G. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature 2012; 485: 201-206.

[17] Estibariz I, Overmann A, Ailloud F, Krebes J, Josenhans C and Suerbaum S. The core genome m5C methyltransferase JHP1050 (M.Hpy99III) plays an important role in orchestrating gene expression in Helicobacter pylori. Nucleic Acids Res 2019; 47: 2336-2348.

[18] Liu RJ, Long T, Li J, Li H and Wang ED. Structural basis for substrate binding and catalytic mechanism of a human RNA: m5C methyltransferase NSun6. Nucleic Acids Res 2017; 45: 6684-6697.

[19] Jacob R, Zander S and Gutschner T. The dark side of the epitranscriptome: chemical modifications in long non-coding RNAs. Int J Mol Sci 2017; 18.

[20] Motorin Y, Lyko F and Helm M. 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. Nucleic Acids Res 2010; 38: 1415-1430.

[21] Yang X, Yang Y, Sun BF, Chen YS, Xu JW, Lai WY, Li A, Wang X, Bhattarai DP, Xiao W, Sun HY, Zhu Q, Ma HL, Adhikari S, Sun M, Hao YJ, Zhang B, Huang CM, Huang N, Jiang GB, Zhao YL, Wang HL, Sun YP and Yang YG. 5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m(5)C reader. Cell Res 2017; 27: 606-625.

[22] Sibbritt T, Shafik A, Clark SJ and Preiss T. Nucleotide-level profiling of m(5)C RNA methylation. Methods Mol Biol 2016; 1358: 269-284.

[23] Cui X, Liang Z, Shen L, Zhang Q, Bao S, Geng Y, Zhang B, Leo V, Vardy LA, Lu T, Gu X and Yu H. 5-methylcytosine RNA methylation in arabidopsis thaliana. Mol Plant 2017; 10: 1387-1399.
[24] He Y, Xue C, Yu Y, Chen J, Chen X, Ren F, Ren Z, Cui G and Sun R. CD44 is overexpressed and correlated with tumor progression in gallbladder cancer. Cancer Manag Res 2018; 10: 3857-3865.
[25] Xue C, He Y, Zhu W, Chen X, Yu Y, Hu Q, Chen J, Liu L, Ren F, Ren Z, Cui G and Sun R. Low expression of LACTB promotes tumor progression and predicts poor prognosis in hepatocellular carcinoma. Am J Transl Res 2018; 10: 4152-4162.
[26] Gomes S, Bosco B, Loureiro JB, Ramos H, Raimundo L, Soares J, Nazareth N, Barcherini V, Domingues L, Oliveira C, Bisio A, Piazza S, Bauer MR, Bras JP, Almeida MI, Gomes C, Reis F, Fersht AR, Inga A, Santos MMM and Saraiva L. SLMP53-2 restores wild-type-like function to mutant p53 through Hsp70: promising activity in hepatocellular carcinoma. Cancers (Basel) 2019; 11.
[27] Yu BW, Zhang LQ, Teng XL, Zhang Y, Zou LB and Ying HY. Association between the CYP1A1 polymorphisms and hepatocellular carcinoma: a meta-analysis. Genet Mol Res 2015; 14: 1076-1084.
[28] Lin L, Sun J, Tan Y, Li Z, Kong F, Shen Y, Liu C and Chen L. Prognostic implication of NQO1 overexpression in hepatocellular carcinoma. Hum Pathol 2017; 69: 31-37.
[29] Seo W, Gao Y, He Y, Sun J, Xu H, Feng D, Park SH, Cho YE, Guillot A, Ren T, Wu R, Wang J, Kim SJ, Hwang S, Liangpunsakul S, Yang Y, Niu J and Gao B. ALDH2 deficiency promotes alcohol-associated liver cancer by activating oncogenic pathways via oxidized DNA-enriched extracellular vesicles. J Hepatol 2019; 71: 1000-1011.
[30] Duan CY, Liu MY, Li SB, Ma KS and Bie P. Lack of association of EPHX1 gene polymorphisms with risk of hepatocellular carcinoma: a meta-analysis. Tumour Biol 2014; 35: 659-666.
[31] Song J, Zhai J, Bian E, Song Y, Yu J and Ma C. Transcriptome-wide annotation of m(5)C RNA modifications using machine learning. Front Plant Sci 2018; 9: 519.
[32] Chhabra R. miRNA and methylation: a multifaceted liaison. Chembiochem 2015; 16: 195-203.
[33] Shinoda S, Kitagawa S, Nakagawa S, Wei FY, Tomizawa K, Araki K, Araki M, Suzuki T and Suzuki T. Mammalian NSUN2 introduces 5-methylcytidines into mitochondrial tRNAs. Nucleic Acids Res 2019; 47: 8734-8745.
[34] Van Haute L, Lee SY, McCann BJ, Powell CA, Bansal D, Vasiliauskaite L, Garone C, Shin S, Kim JS, Frye M, Gleeson JG, Miska EA, Rhee HW and Minczuk M. NSUN2 introduces 5-methylcytosines in mammalian mitochondrial tRNAs. Nucleic Acids Res 2019; 47: 8720-8733.