Received: 10 January 2022 | Revised: 5 February 2022 | Accepted: 7 February 2022

DOI: 10.1002/jcla.24303

RESEARCH ARTICLE

WILEY

Role of m^5C RNA methylation regulators in colorectal cancer prognosis and immune microenvironment

Xiaojie Fang[1] | Chenyun Miao[2] | Tianni Zeng[3] | Weijian Chu[1] | Yi Zheng[1] | Xi Sun[1] | Xin Yin[1] | Yanyan Li[1]

[1]Department of Anorectal Surgery, Hangzhou TCM Hospital Affiliated to Zhejiang Chinese Medical University, Hangzhou, China

[2]Department of TCM Gynecology, Hangzhou TCM Hospital Affiliated to Zhejiang Chinese Medical University, Hangzhou, China

[3]Department of Medical Oncology, Hangzhou TCM Hospital Affiliated to Zhejiang Chinese Medical University, Hangzhou, China

Correspondence
Yanyan Li, Departments of Anorectal Surgery, Hangzhou TCM Hospital Affiliated to Zhejiang Chinese Medical University, 453 Tiyuchang Road, Xihu District, Hangzhou, Zhejiang 310007, China.
Email: liyanyanhz@hotmail.com

Funding information
YL received funding from Hangzhou Science and Technology Bureau (Grant number:20201203B202; URL: http://kj.hangzhou.gov.cn/) and she conceived and designed the study

Abstract

Background: RNA modification has become one of the hot topics of research as it can be used for tumor prognosis. However, its role in various biological processes is still poorly understood. The aim of this study was to investigate the role of m^5C and m^1A regulators on colorectal cancer prognosis using bioinformatics tools. The association between these regulators and differences in patient survival as well as the clinicopathological characteristics and tumor immune microenvironment in colorectal cancer tissues were assessed.

Methods: We selected publicly available colorectal cancer data sets from The Cancer Genome Atlas and used the "limma" package in R to identify differentially expressed genes. The least absolute shrinkage and selection operator regression model was used to calculate the prognostic risk, and a risk prediction model was constructed, to help assess the prognostic values of the differentially expressed genes. Finally, using TISCH and TIMER, we assessed the extent of cellular infiltration in colorectal cancer.

Results: We explored NSUN6 and DNMT3A expression using UALCAN and HPA and found that their expression is significantly increased in colorectal cancer tissues and correlated with sex and TP53 mutation status. Moreover, we found NSUN6 and DNMT3A were related to the infiltration of six major immune cells, with DNMT3A being closely related to dendritic cells, $CD4^+$ T cells, and B cells, whereas NSUN6 to B cells and $CD8^+$ T cells.

Conclusion: Our findings suggest that m^5C regulators can predict the clinical prognostic risk and regulate the tumor immune microenvironment in colorectal cancer.

KEYWORDS
5-methylcytosine, colorectal cancer, prognostic model, RNA methylation, tumor immune microenvironment

Xiaojie Fang, Chenyun Miao and Tianni Zeng contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2022 The Authors. *Journal of Clinical Laboratory Analysis* published by Wiley Periodicals LLC.

1 | INTRODUCTION

Colorectal cancer (CRC) is one of the most commonly diagnosed gastrointestinal malignancies in the world. Its incidence and mortality rates have been continuously increasing, and it is now considered the second leading cause of death with an oncological origin.[1] According to the data reported by GLOBOCAN, approximately 1.8 million new CRC cases are diagnosed per year, 50% of which are fatal.[2] CRC has a high degree of malignancy, causing distant visceral metastasis through the blood and lymphatic system, resulting in poor prognosis.[3] Chemotherapy and surgery are the major treatment strategies for CRC; however, owing to the significant heterogeneity, the clinical effects of current chemotherapeutic drugs are still far from satisfactory.[4] Therefore, the development of more efficient methods is urgently needed.

Tumor immune microenvironment (TIM) is a crucial factor contributing to the occurrence, development, and prognosis of tumors.[5] The TIM contains various cell types (infiltrating immune cells, vascular cells, and mesenchymal stem cells) and associated cytokines/chemokines[6] and is a complex and dynamic system. The immune inflammatory response varies among patients,[7] and high levels of Th1 cells and dendritic cells (DCs) in tumor tissues are related with a good prognosis of CRC.[8] In addition, M1 macrophages secreting pro-inflammatory cytokines (TNF-α, IL-1-β, and IL-12) can suppress colon cancer cell growth and promote apoptosis, whereas M2 macrophages enhance tumor metastasis via production of anti-inflammatory cytokines, such as TGF-β and IL-10.[9]

RNA methylation, the most important RNA epigenetic modification in non-coding RNA (ncRNA) and messenger RNA (mRNA) of eukaryotic species, which modulates RNA splicing, translation, and other biological processes, accounts for 60% of RNA modifications.[10] To date, more than five types of RNA methylation—N1-methyladenosine (m^1A), N6-methyladenosine (m^6A), eukaryotic 5-methylcytosine (m^5C), 7-methylguanosine (m^7G), and RNA 2′-O-methylation (Nm)—have been identified, of which m^5C RNA methylation is the second most common type, after m6A methylation.[11,12] Although m^5C modification was first discovered in the 1970s,[13] little is known about its role in various biological processes. With the development of gene sequencing technology, m^5C modification has recently gained increased attention. m^5C is ubiquitous in eukaryotic tRNAs and rRNAs and participates in RNA export and ribosome translation.[14] The expression of m^5C regulators is also linked to a variety of human cancers,[15,16] and m^1A is considered important in the regulation of tumor development.[17-19]

This study was designed to investigate the role of m^5C and m^1A regulators in CRC prognosis (using bioinformatics methods) and to analyze the association between these regulators and differences in survival as well as the clinicopathological characteristics and TIM in CRC tissues. Furthermore, this study helps to explore novel biomarkers that predict the therapeutic efficacy of current treatments and benefit therapeutic modulation.

2 | MATERIALS AND METHODS

2.1 | Data sources

First, CRC transcriptomic and relevant clinical data were downloaded from TCGA (https://tcga-data.nci.nih.gov/tcga/) and organized separately as the training cohort. A total of 602 research samples (48 healthy individuals and 554 CRC patients) were included. The independent gene expression data set GSE39582 was obtained from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) and used as the validation cohort.

2.2 | Differentially expressed genes (DEGs) between CRC and normal tissues

We reviewed the relevant literature to determine m^5C and m^1A regulators; 24 regulators were obtained, of which 9 were m^1A regulators and 15 were m^5C regulators (Table 1). The expression matrix and clinical information of m^1A and m^5C regulators in 554 CRC cases and

TABLE 1 RNA-modified m^1A and m^5C regulators

Regulator	Type
m^1A	
TRMT6	Writer
TRMT61A	Writer
RRP8	Writer
ALKBH1	Reader
ALKBH3	Reader
YTHDF1	Eraser
YTHDF2	Eraser
YTHDF3	Eraser
YTHDC1	Eraser
m^5C	
TRDMT1	Writer
NSUN1	Writer
NSUN2	Writer
NSUN3	Writer
NSUN4	Writer
NSUN5	Writer
NSUN6	Writer
NSUN7	Writer
DNMT1	Writer
DNMT2	Writer
DNMT3A	Writer
DNMT3B	Writer
ALYREF	Reader
YBX1	Eraser
TET2	Eraser

48 normal cases obtained from TCGA were used for further analysis. The "limma" package in R software (4.0.5) was used to screen for differentially expressed m^1A and m^5C regulators between the normal and tumor tissue groups. For DEGs, p-values < 0.05, and $|\log_2(FC)|>1$ were used as the cut-off values. Heatmap and violin plots were generated for visualization.

2.3 | GEO database validation of differential expression

Differential gene analysis between tumor and normal tissues was performed using the "limma" package. Subsequently, visualization of the differences in expression between the two groups was performed using heatmaps.

2.4 | Construction of protein–protein interaction (PPI) network

To select key modules and hub genes, the PPI network was constructed using the search tool for the retrieval of interacting genes (STRING) platform (https://cn.string-db.org/)(confidence score 0.4).

2.5 | Establishment of the prognostic risk model

First, univariate Cox regression analysis and the least absolute shrinkage and selection operator (LASSO) algorithm were used to assess the associations between m^5C regulators and clinical prognosis of CRC. Next, the following formula was used to calculate the prognostic risk scores for each patient:

$$\text{Risk score} = \text{coefficient}\,1 * \text{value}1 + \text{coefficient}\,2 * \text{value}\,2$$

The obtained value was the relative expression level of each gene calculated using the comparative CT method ($2^{-\Delta\Delta Ct}$). We further categorized patients with CRC from TCGA database into low- and high-risk groups based on their median risk scores. Finally, Kaplan–Meier survival curves were used to assess the difference in survival between the two groups.

2.6 | Clinical profile and correlation between the clinicopathological characteristics and gene expression

The UALCAN web portal (http://ualcan.path.uab.edu/) is an open-access and online platform used to visualize gene expression alterations occurring between cancer and paired normal tissues with respect to clinicopathological characteristics based on TCGA database. Using the UALCAN database, we analyzed the data according to clinical pathology parameters such as sex, sample type, and TP53 mutation status.

2.7 | Immunohistochemical analysis and gene set enrichment analysis (GSEA)

We obtained protein expression data for DNMT3A and NSUN6 from the Human Protein Atlas (HPA) database (https://www.proteinatlas.org). Staining was evaluated qualitatively based on the proportion of stained cells (<25, 25%–75%, or >75%), and the staining intensity (negative, weak, medium, and strong) were categorized based on the following grades: no staining, weak staining, and strong staining. Mutations in the prognostic-related m^5C regulators were also analyzed using the cBioPortal database (http://www.cbioportal.org/). We subsequently divided CRC samples in the CRC cohort into DNMT3A high-expression group (237 samples) and NSUN6 high-expression group (224 samples), and GSEA was performed. Terms enriched in hub genes were considered statistically significant at $p < 0.01$ and FDR < 0.05.

2.8 | Association between m^5C regulators and tumor microenvironment-related cells (TISCH)

Tumor immune single-cell hub (TISCH) (http://tisch.comp-genomics.org) is a large-scale single-cell RNA-sequencing database that characterizes tumor microenvironment (TME) at a single-cell resolution. This database was used to investigate TME heterogeneity in various data sets and cells.

2.9 | Association between prognostic-related m^5C regulators and tumor-infiltrated lymphocytes

TIMER (https://cistrome.shinyapps.io/timer/) is an online platform for the analysis of immune cells in filtrates of multiple tumors. The immune penetration algorithm can be used to calculate the infiltration abundance of six immune cells ($CD4^+$ T cells, B cells, $CD8^+$ T cells, macrophages, neutrophils, and DCs) in TCGA. Using the "Immune" module, the relationship between multiple factors and immune cell infiltration could be comprehensively analyzed.

2.10 | Statistical analysis

All statistical analyses were performed using R software (version 4.1.2). The Kaplan–Meier survival curve was used to analyze overall survival (OS), and the chi-square test was used to analyze the correlation between the risk signature and clinical characteristics. Univariate and multivariate Cox regression analyses were used to determine the prognostic value of the risk signature. The area under

the receiver operator characteristic curve (AUC) analyses were used to assess the accuracy of the prognostic signature. Statistical significance was set at $p < 0.05$.

3 | RESULTS

3.1 | Differential expression of m^1A and m^5C between CRC and normal tissues

First, we analyzed the expression of m^1A and m^5C in CRC samples and normal tissue samples. The results showed that most m^1A and m^5C regulators were differentially expressed between the two groups. Among them, seven m^1A regulators (*TRMT6*, *TRMT61A*, *RRP8*, *ALKBH3*, *YTHDF1*, *YTHDF2*, and *YTHDC1*) were highly expressed in CRC tissues (Figure 1A). Meanwhile, 11 m^5C regulators (*NSUN2*, *NSUN5*, *NSUN6*, *NSUN3*, *DNMT3B*, *NSUN7*, *DNMT1*, *NSUN4*, *DNMT3A*, *ALYREF*, and *YBX1*) were highly expressed in CRC tissues, while TET2 expression was downregulated in CRC tissues. ($p < 0.001$) (Figure 1B).

3.2 | Relationship between m^1A and m^5C regulators and OS in patients with CRC

The relationship between m^5C and m^1A regulators and OS was investigated using a univariate Cox regression analysis. The results showed that the two m^5C regulators *NSUN6* (hazard ratio [HR] = 1.109, 95% confidence interval [CI] = 1.046–1.176, $p < 0.001$) and *DNMT3A* (HR = 1.046, 95% CI = 1.002–1.092, $p = 0.041$) were at high risk (Figure 2A). Subsequently, we constructed a prognostic risk model using these two genes (Figure 2B, C), and the coefficients were obtained using the LASSO algorithm (Table 2). The integrated risk score for each patient was calculated as follows:

$$\text{Risk score} = 0.093 \times \text{NSUN6} + 0.020 \times \text{DNMT3A}$$

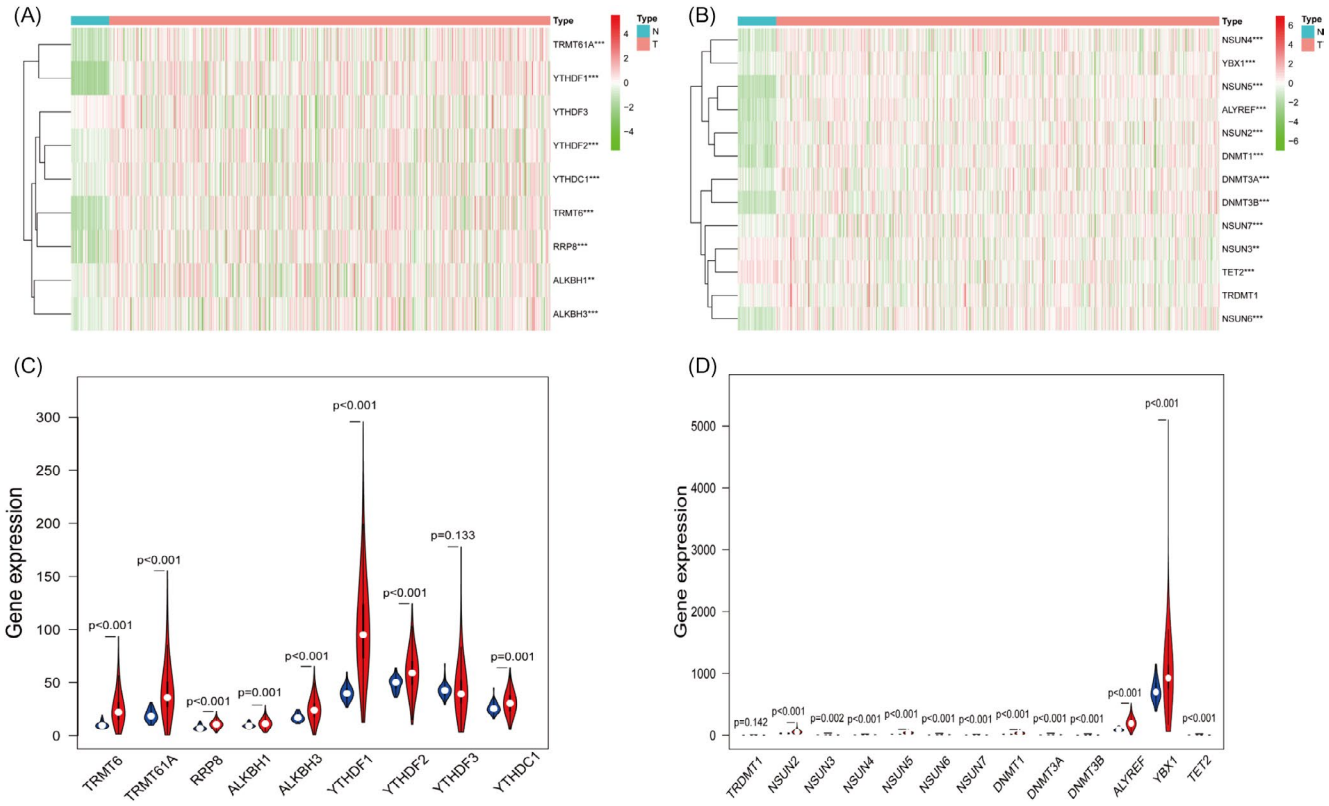
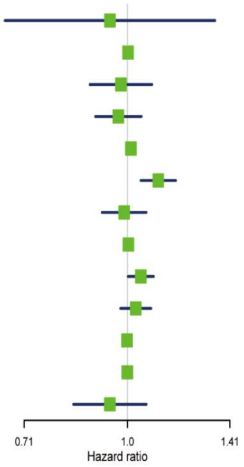


FIGURE 1 Expression of m^1A and m^5C regulators in colorectal cancer patients and normal individuals. Heatmap and violin plot depicting the differences in (A and C) m^1A RNA methylation and (B and D) m^5C RNA methylation regulator expression between the two groups. N stands for normal samples; T stands for tumor samples; blue violin stands for normal samples; and red violin stands for tumor samples. $^*p < 0.001$, $^{**}p < 0.01$, $^*p < 0.05$

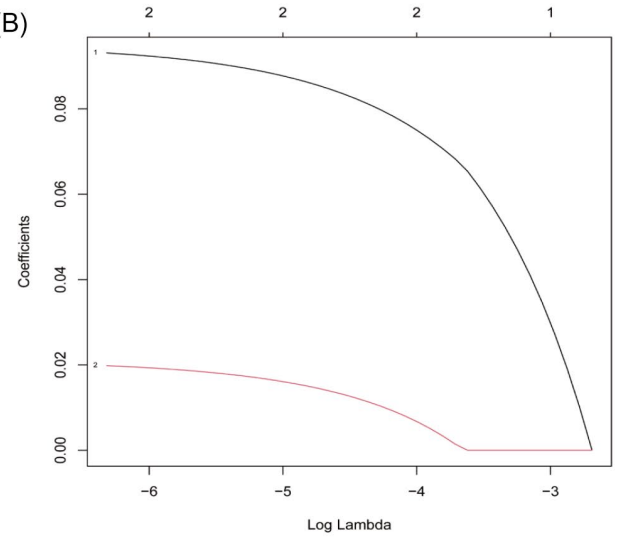
FIGURE 2 Prognostic risk score model constructed based on *NSUN6* and *DNMT3A*. (A) Univariate Cox regression analysis results reveal that the expression of *NSUN6* and *DNMT3A* are independent risk factors of CRC. (B) Least absolute shrinkage and selection operator (LASSO) model to estimate the coefficients of each variable. (C) In the LASSO model, 10-fold cross-validation is used to tune parameter choices. (D) Kaplan–Meier survival curve showing the differences in overall survival between high- and low-risk groups. (E) Receiver operator characteristic curve of the prediction model

(A)

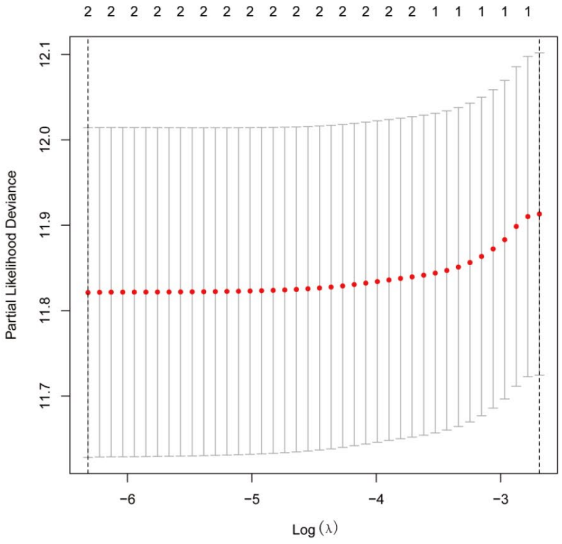
	pvalue	Hazard ratio
TRDMT1	0.745	0.943(0.663−1.341)
NSUN2	0.733	1.002(0.991−1.013)
NSUN3	0.679	0.978(0.882−1.085)
NSUN4	0.438	0.970(0.898−1.048)
NSUN5	0.155	1.012(0.996−1.028)
NSUN6	<0.001	1.109(1.046−1.176)
NSUN7	0.772	0.989(0.919−1.065)
DNMT1	0.659	1.003(0.990−1.016)
DNMT3A	0.041	1.046(1.002−1.092)
DNMT3B	0.285	1.028(0.977−1.082)
ALYREF	0.375	0.999(0.996−1.002)
YBX1	0.417	1.000(0.999−1.000)
TET2	0.345	0.943(0.834−1.065)



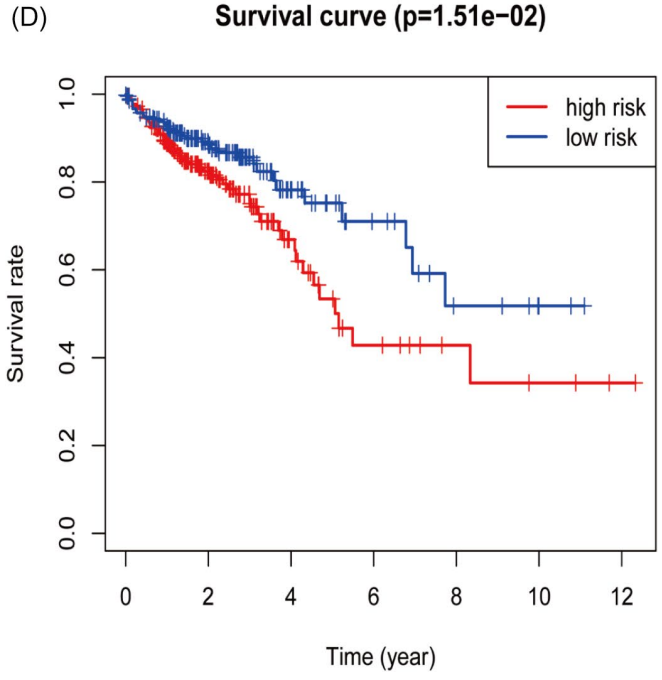
(B)



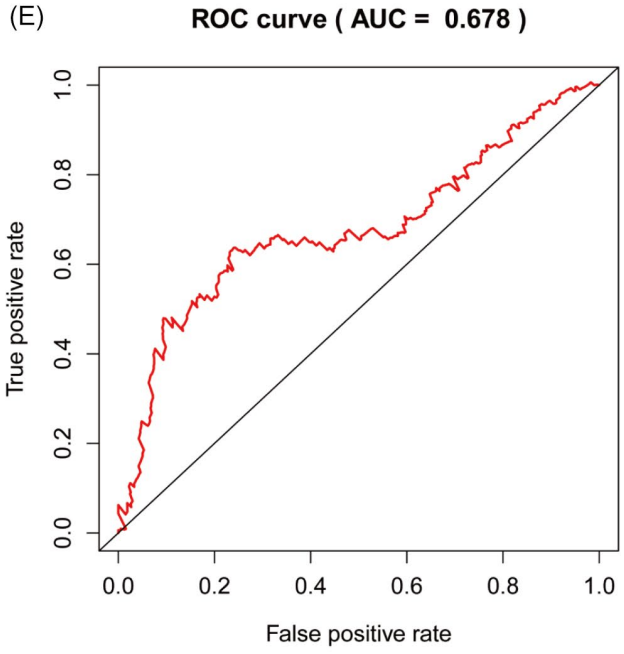
(C)



(D)



(E)



According to the results of the survival analysis, the high-risk group patients had a significantly lower OS rate than the low-risk group patients ($p < 0.05$; Figure 2D). The area under the receiver operator characteristic curve was 0.678 (Figure 2E), indicating a good predictive performance. However, there was no significant correlation observed between the expression of m^1A regulators and OS.

TABLE 2 Genes selected to build risk signature and their corresponding coefficients

Genes	Coefficients
NSUN6	0.093091494593057
DNMT3A	0.0198262658117633

3.3 | Validation of differential expressed m^5C regulators using GEO

The GSE39582 data set was used to further validate the difference in the expression of m^5C regulators in CRC and normal tissues. The expression levels of *NSUN2* ($p < 0.001$), *DNMT1* ($p < 0.001$), *ALYREF* ($p < 0.05$), *NSUN4* ($p < 0.001$), *YBX1* (<0.001), *DNMT3A* ($p < 0.001$), *NSUN5* ($p < 0.001$), and *DNMT3B* ($p < 0.001$) were significantly higher in tumor tissues than in the normal tissues. Conversely, the expression of *NSUN3* ($p < 0.001$), *TET2* ($p < 0.001$), and *NSUN6* ($p < 0.05$) was lower in cancer tissues than in normal tissues (Figure 3A). However, there was no significant difference in *TRDMT1* and *NSUN7* expression between the two groups ($p>0.05$).

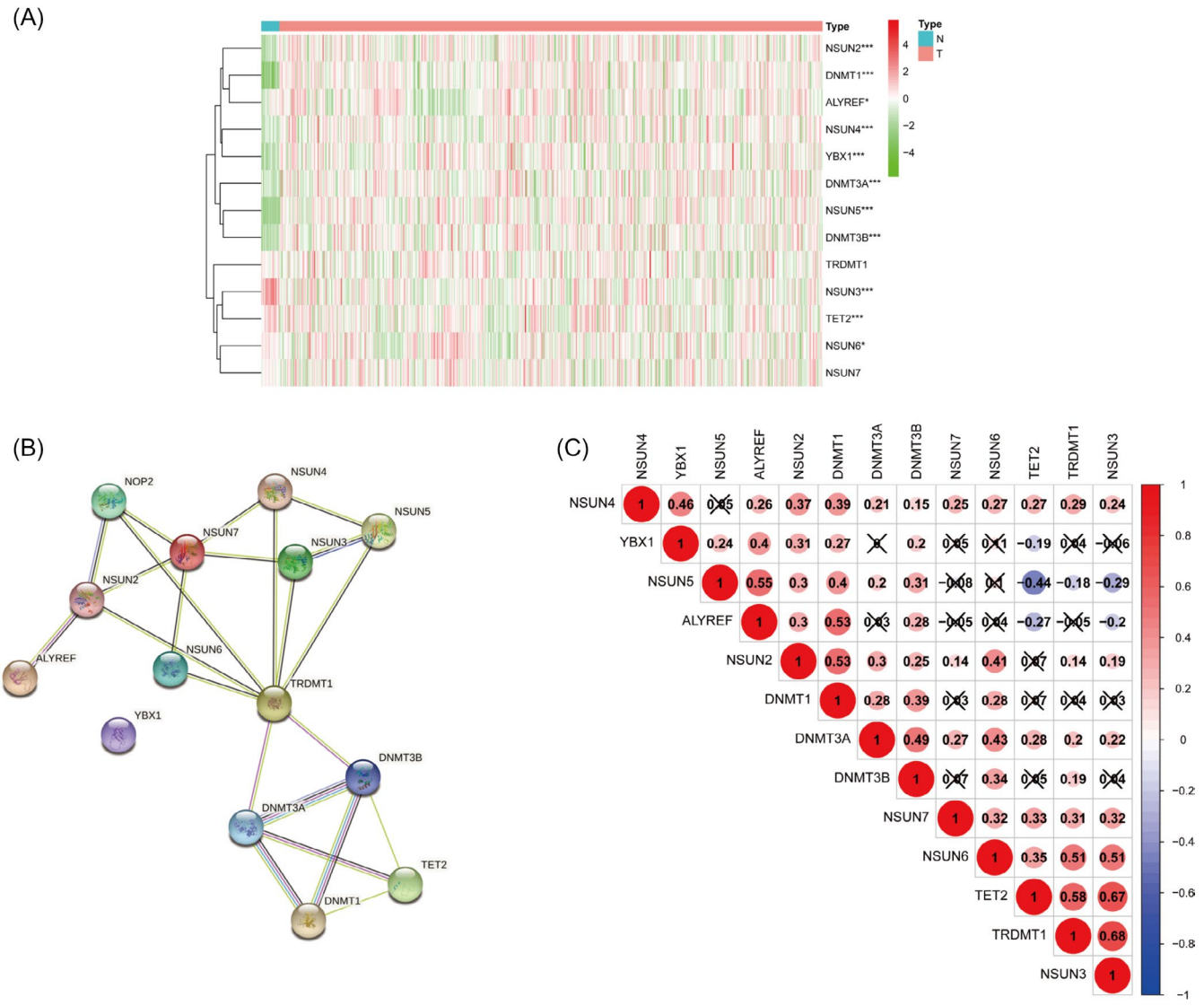


FIGURE 3 Differentially expressed genes (DEGs) in m^5C regulators validated using the GEO database and exploring the interactions and correlations among m^5C regulators were explored. (A) Expression of m^5C regulators in the data set GSE39582 compared between the tumor and normal groups. (B) Protein interactions among the m^5C regulators predicted using STRING. (C) Association using Pearson correlation analysis. ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$

3.4 | Interaction and correlation between m^5C regulators

As shown in Figure 3B, *TRDMT1* is the hub gene of this network and is closely related to other genes. Moreover, *TRDMT1* expression was significantly correlated with *NSUN3*, *NSUN4*, *NSUN5*, *NSUN2*, *DNMT3A*, *DNMT3B*, *NSUN7*, *NSUN6*, and *TET2* expression. The majority of genes showed strong correlations with each other, and the strongest correlation was found between *TRDMT1* and *NSUN3* (Figure 3C). These results suggest a correlation among m^5C regulators.

3.5 | Prognosis-related risk score is an independent risk factor for prognosis

We further investigated the association between the risk scores and clinicopathological characteristics. CRC samples with high-risk scores generally had elevated expression of *NSUN6* and *DNMT3A* (Figure 4A). In addition, significant differences between the high- and low-risk groups observed in terms of survival status and N staging ($p < 0.05$). Univariate Cox regression analysis demonstrated that age (HR = 1.038, 95% CI = 1.016–1.061, $p < 0.001$), pathological stage (HR = 2.546, 95% CI = 1.951–3.323, $p < 0.001$), T staging (HR = 3.240, 95% CI = 2.049–5.125, $p < 0.001$), N staging (HR = 2.219, 95% CI = 1.693–2.908, $p < 0.001$), M staging (HR = 5.290, 95% CI = 3.314–8.444, $p < 0.001$), and risk score (HR = 1.236, 95% CI = 1.074–1.422, $p = 0.003$) remained significantly associated with OS (Figure 4B), suggesting that they all could serve as independent risk factors for CRC. Conversely, no significant correlations ($p > 0.05$) were observed between sex and OS. In the multivariate Cox regression analysis, only age (HR = 1.057, 95% CI = 1.033–1.083, $p < 0.001$) and risk score (HR = 1.238, 95% CI = 1.069–1.433, $p = 0.004$) were found to be independent prognostic factors for CRC (Figure 4C).

3.6 | Relationship between *NSUN6* and *DNMT3A* expression and clinicopathological characteristics of patients with CRC

To further investigate the expression of *NSUN6* and *DNMT3A* in CRC and normal tissues, we examined their expression using the UALCAN database. The expression levels of both genes were significantly elevated in colon adenocarcinoma tissues ($p < 0.001$) (Figure 5). Although there was no significant difference in their expression between the two groups in terms of sex ($p>0.05$), *NSUN6* expression was significantly elevated in the TP53-mutant group ($p < 0.001$).

3.7 | Differences in DNMT3A and NSUN6 protein expression, gene mutation types, and GSEA

We used the HPA database to detect the expression of *NSUN6* and *DNMT3A* in CRC tissues and normal tissues. IHC results showed that *NSUN6* was highly expressed in both colon adenocarcinoma cells and normal colon gland cells. *DNMT3A* was expressed lowly in colon adenocarcinoma cells and highly in colon gland cells (Figure 6A). Using the cBioPortal database, we found that alterations in *NSUN6* and *DNMT3A* in 1510 samples from TCGA harbored missense mutations and deep deletions. The mutation frequencies were 0.9% for *NSUN6* and 1.8% for *DNMT3A*. *DNMT3A* alterations in TCGA, TCGA pan-cancer, and TCGA firehose legacy data were all mutations, while *NSUN6* alterations in TCGA, TCGA pan-cancer, and TCGA firehose legacy data included both mutations and deep deletions (Figure 6B, C). We subsequently performed GSEA to investigate the signaling pathways associated with the differential expression of *NSUN6* and *DNMT3A* in CRC. Single-gene GSEA showed that high expression of *DNMT3A* is associated with ascorbate and aldehyde metabolism, pentose and glucose interconversion, cell adhesion, and systemic lupus erythematosus. Pathways enriched in the *NSUN6* upregulation group were involved in maturity-onset diabetes of the young, pentose, and gluconate interconversions and spliceosome (Figure 6D).

3.8 | Correlation between TME and m^5C regulators in CRC

We analyzed the degree of invasion of the risk-related genes *NSUN6* and *DNMT3A* in TME-associated cells using the TISCH database. *NSUN6* showed higher infiltration in exhausted CD8 T cells, proliferating T cells, and myofibroblasts, and *DNMT3A* showed the highest degree of infiltration in myofibroblasts (Figure 7A, C). In the TISCH database, GSE139555 was divided into 18 cell clusters and 12 cell types, allowing us to visualize the distribution and number of various TME-related cells (Figure 7B). The pie chart shows that B lymphocytes are the most abundant in GSE139555, followed by CD4Tconv cells.

3.9 | Correlation between m^5C regulators and immune cells

Using the TIMER database, we investigated the relationship between *NSUN6* and *DNMT3A* and the degree of infiltration of six immune cells. The analysis showed that *NSUN6* and *DNMT3A* were positively correlated with the degree of infiltration of all six immune cells (Figure 8A). B cells ($p < 0.001$) and $CD8^+$ T cell ($p < 0.05$) infiltration levels were significantly reduced in the Arm-level Deletion group compared to normal *NSUN6* somatic cells. Similarly, B cells ($p < 0.05$), $CD4^+$ T cells ($p < 0.01$), neutrophils ($p < 0.01$), and DCs ($p < 0.001$) were significantly reduced in the Arm-level Deletion group compared to those in the normal *DNMT3A* somatic cells (Figure 8B). Combining the above analysis results, we can conclude that *DNMT3A* is closely related to B cells, $CD4^+$ T cells, and DCs in CRC, while NSUN6 is closely related to B cells and $CD8^+$ T cells.

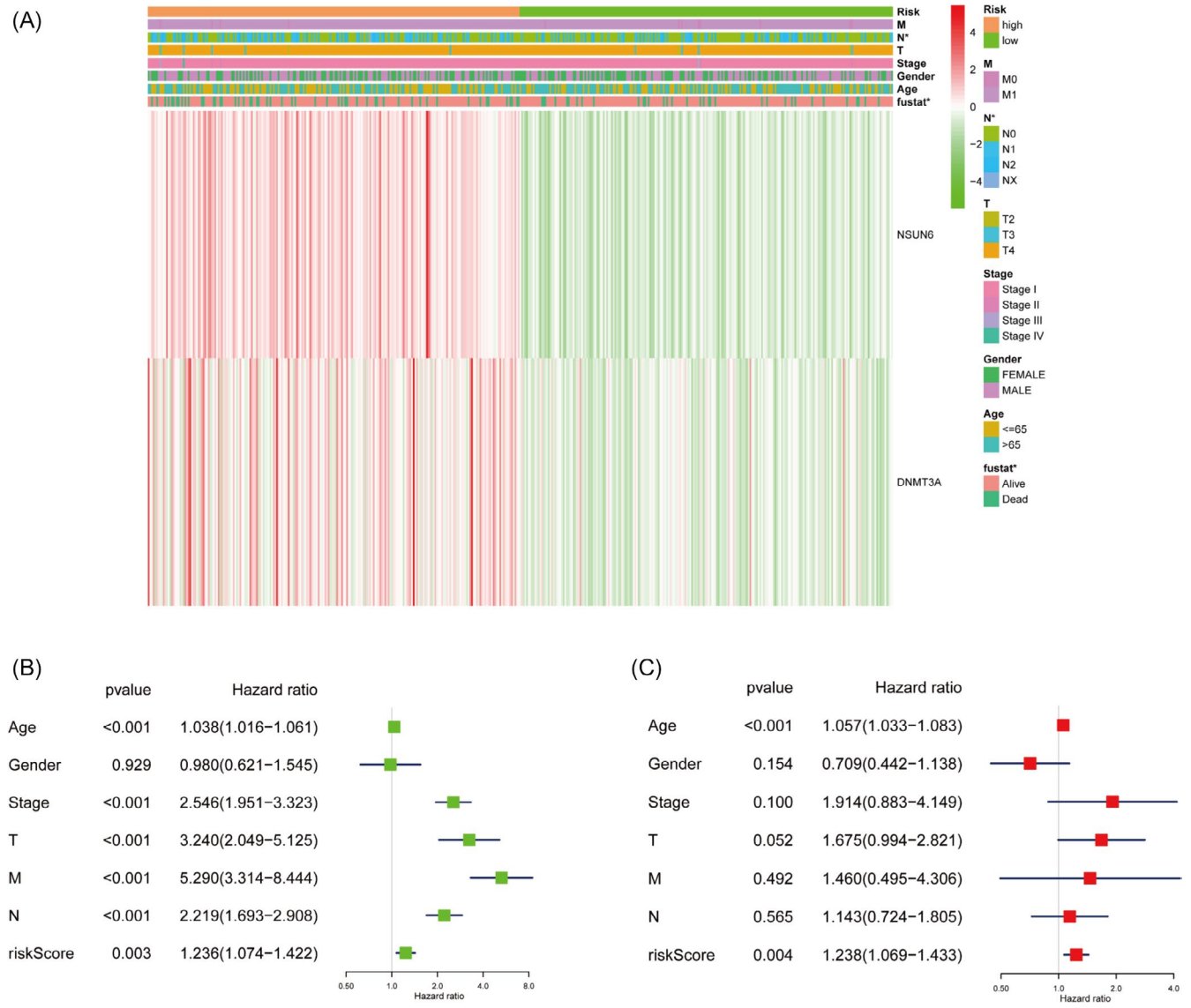


FIGURE 4 Prognostic value of risk score and its relationship with clinicopathological characteristics of CRC. (A) Differences in clinicopathological characteristics and risk scores between the high- and low-risk groups. (B) Risk score and clinicopathological characteristics analyzed using a univariate Cox regression model. (C) Risk scores and clinicopathological characteristics analyzed using a multivariate Cox regression model. $^{*}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

4 | DISCUSSION

To date, more than 170 chemically distinct types of RNA modifications have been identified, with m^6A, m^5C, and m^1A being the most prominent ones. RNA modifications mainly interact with three classes of regulators, writers, readers, and erasers.[20] m^1A methylation regulators include three writers (TRMT6, TRMT61A, and SRRP8), two readers (ALKBH1 and ALKBH3), and four erasers (YTHDF1, YTHDF2, YTHDF3, and YTHDC1), whereas m^5C methylation is controlled by 12 writers (TRDMT1, NSUN1, NSUN6, NSUN4, NSUN5, DNMT1, NSUN7, NSUN2, NSUN3, DNMT2, DNMT3A, and DNMT3B), one reader (ALYREF), and two erasers (YBX1 and TET2). In this study, m^1A and m^5C regulators were found to be differentially expressed in CRC tissues. Among them, the m^5C regulators, *NSUN6* and *DNMT3A*, were considered prognostic signatures based on the Cox and LASSO analyses. Thus, *NSUN6* and *DNMT3A* were used to develop a reliable prognostic risk-score model for patients with CRC. Moreover, we thoroughly investigated the association between these m^5C regulators and the TIM.

Epigenetic modifications of ncRNAs are important factors contributing to the development of CRC, with methylation being the most important post-transcriptional modification of ncRNAs.[21] As a methyl group at the first position of adenosine, m^1A modification and relevant long non-coding RNAs play key roles in CRC.[22] As shown in our results, seven of nine m^1A regulators—*TRMT6*, *TRMT61A*, *RRP8*, *ALKBH3*, *YTHDF1*, *YTHDF2*, and *YTHDC1*—were differentially expressed between CRC and normal tissues. Among them, *YTHDF1* and *YTHDC1* were studied intensively. A previous study demonstrated that the knockdown of *YTHDF1* significantly inhibits the tumorigenicity of CRC cells and the growth of murine xenograft tumors based

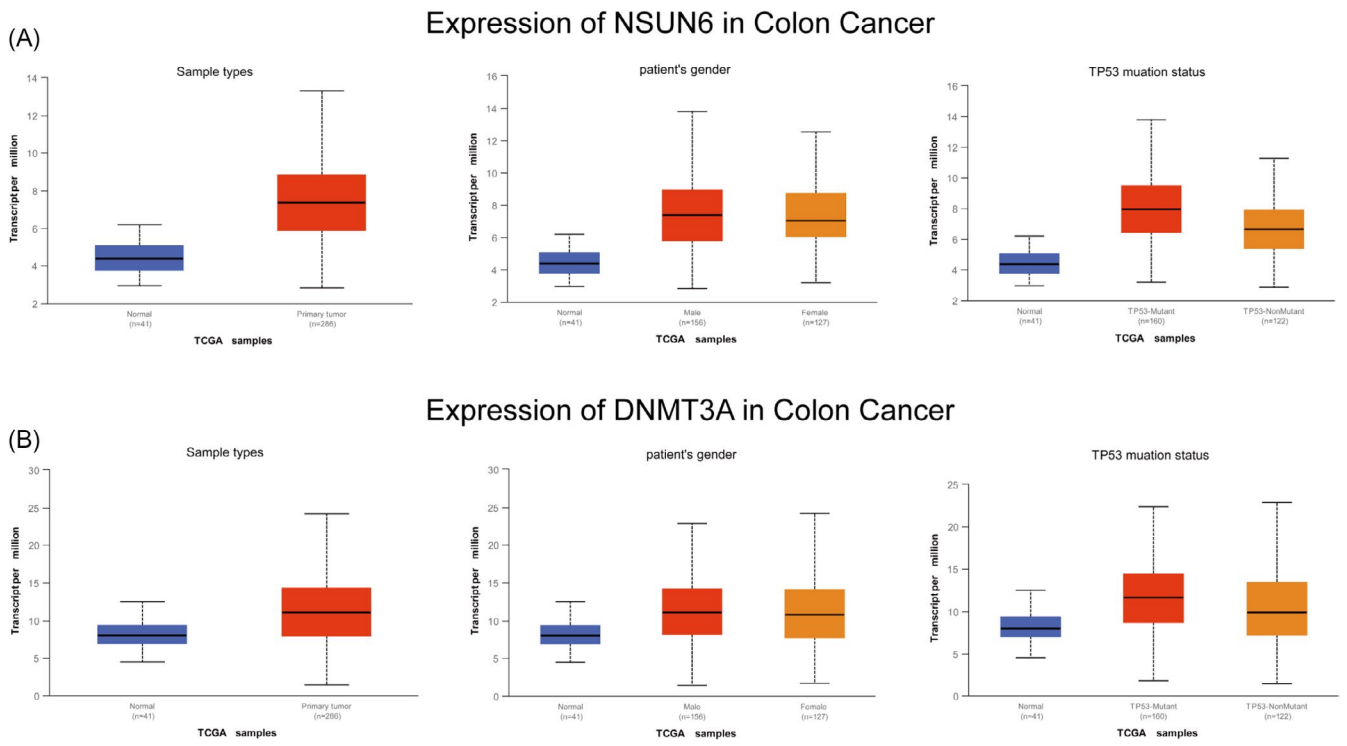


FIGURE 5 Expression of *NSUN6* and *DNMT3A* and clinicopathological characteristics in patients with colorectal cancer. (A) *NSUN6* and (B) *DNMT3A* expression in different sample types based on patient sex and TP53 mutation status

on in vitro and in vivo experiments.[23] YTHDF1 could also affect the GLS1-glutamine metabolic axis to reduce cisplatin sensitivity in CRC cells.[24] YTHDC1 binds to SLC7A5, thereby promoting the proliferation and migration of CRC cells.[25] However, no m^1A regulator was screened as a prognostic signature in the univariate Cox regression analysis in this study, which might be due to data selection bias in TCGA database.

Dysregulation of m^5C modification is a crucial mechanism underlying tumorigenesis, and m^5C levels have been increasingly recognized as cancer markers.[26] Among the many regulators, the relationship between *NSUN2* and tumors is the most well-known. As previous articles have clarified, *NSUN2*, encoding an m^5C writer, is a downstream target gene of the oncogene *MYC*[27]; its expression level is correlated with the cell cycle in many cancer types, including breast cancer, skin cancer, and CRC.[28-30] Recently, circNSUN2, derived from the NSUN2-coding sequence, was identified as frequently upregulated in patients with CRC and can stabilize *HMGA2* mRNA to promote CRC liver metastasis.[31] Our results suggested that *NSUN6* and *DNMT3A* were risk factors significantly associated with prognosis, and a prognostic risk model was built using these two genes. The higher the expression levels of both genes, the lower the survival rate. Although there are few studies on the relationship between *NSUN6* and CRC, *DNMT3A* has been explored as a regulatory mechanism in CRC that functions via multiple targets and multiple pathways. *DNMT3A*, encoding a de novo DNA methyltransferase that methylates CpG dinucleotides,[32] is generally highly expressed in CRC.[33] Because it can be used to identify distal colon end-stage and microsatellite instability-positive tumors, this regulator has been considered a good diagnostic marker for patients with CRC.[34] It was found that DNMT3A could attenuate the proliferation of CRC cells by effectively downregulating the DAB2IP-activated MEK/ERK signaling pathway.[35] Li et al.[36] revealed the potential mechanism of DNMT3A effects in CRC as involving methylation of the *AGR2* promoter, thereby inhibiting the oncogenic activity of *AGR2* in CRC tumorigenesis and progression.

Our GSEA results suggested that the upregulation of *DNMT3A* expression is closely related to ascorbate and aldarate metabolism. As is well known, glucuronidation is a primary metabolism pathway that affects the xenobiotic metabolism of hormones and drugs,[37] including many anticancer agents. Recent research has demonstrated that glucuronidation represents an important mechanism of intrinsic drug resistance in CRC.[38] UDP-glucuronosyltransferase (UGT) polymorphisms that might affect the drug response and cancer susceptibility are associated with an increased risk in developing cancers.[39] Among numerous UGTs, UGT1A6 polymorphisms specifically increase CRC risk.[40] In addition, pathways enriched in the *NSUN6* upregulation group include maturity-onset diabetes of the young, pentose, and gluconate interconversions and the spliceosome. RNA splicing is essential for gene regulation. Selective splicing provides a way for cells to diversify their proteome; interestingly, spliceosome protein mutations can also promote cellular carcinogenesis.[41] Lv et al.[42] found that the spliceosome protein Eftud2 can mediate the effects of the NF-κB pathway in macrophages to promote tumorigenesis in colon tissues.

The TIM, including immune, stromal, and inflammatory cells, is related to tumorigenesis, progression, metastasis, recurrence,

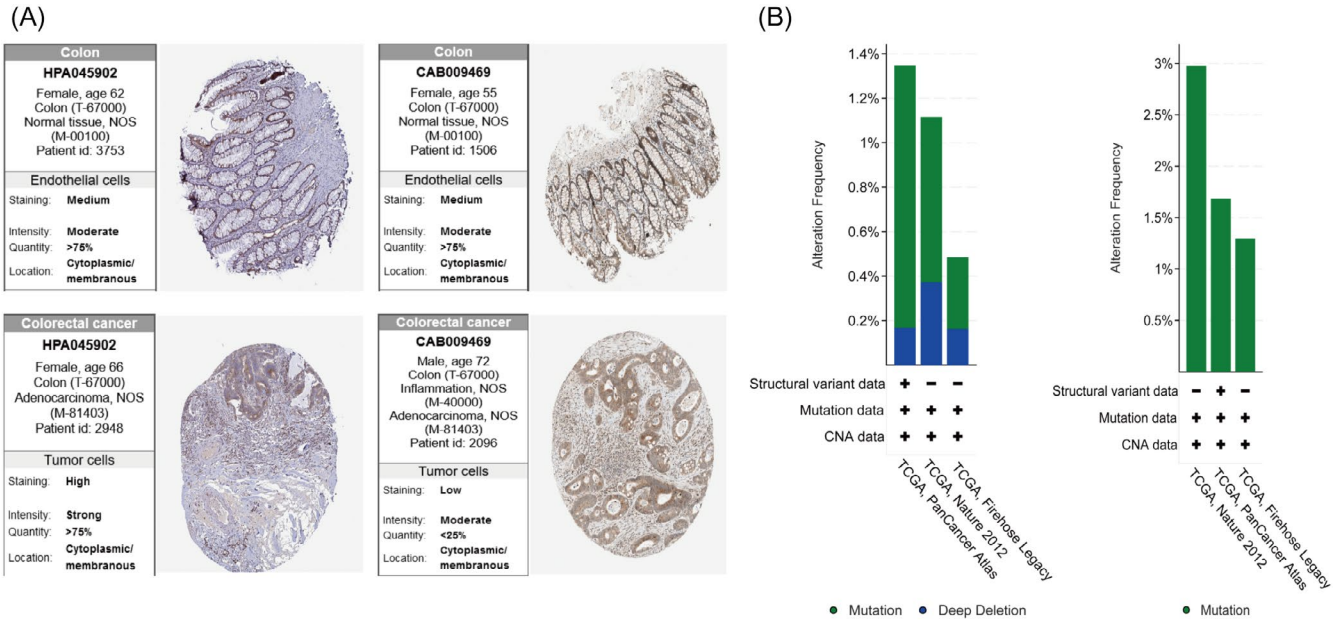


FIGURE 6 Immunohistochemical analysis, alteration frequency analysis, and gene enrichment analysis results of *NSUN6* and *DNMT3A* expression in CRC. (A) Expression of NSUN6 and DNMT3A in CRC tissues and normal tissues assessed in HPA database. (B and C) Frequency of *NSUN6* and *DNMT3A* gene alterations. (D) Pathway enrichment analysis in the DNMT3A high-expression and NSUN6 high-expression groups

and drug resistance, and influences outcomes in various tumors. As shown in our results, CD4Tconv, CD8T, CD8Tex, B, monocyte/macrophage, NK, plasma, and Treg cells are mainly involved in the TIM in CRC tissues. According to the single-cell RNA-sequencing results of CRC tissues, the proportions and functions of immune cells are altered in cancers compared to those in normal tissues.[43]

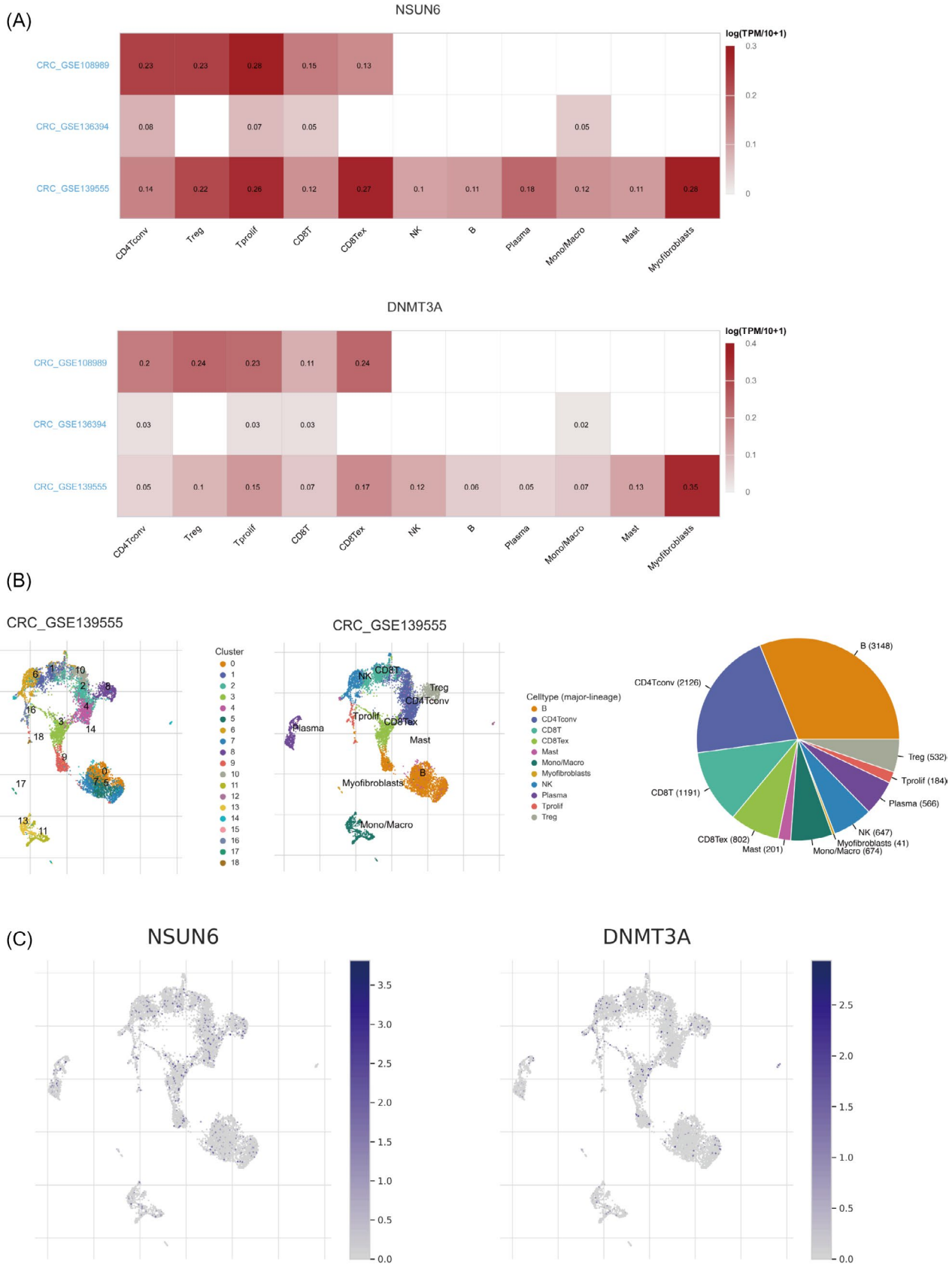


FIGURE 7 In the TISCH database, m^5C regulators were expressed in a variety of tumor microenvironment-associated cells. (A) Expression levels of NSUN6 and DNMT3A in CRC microenvironment-associated cells in the GEO data set. (B) Annotation of the cell types contained in the GSE139555 data set and the percentage of each cell. (C) Proportions of NSUN6 and DNMT3A in different cell types in GSE139555

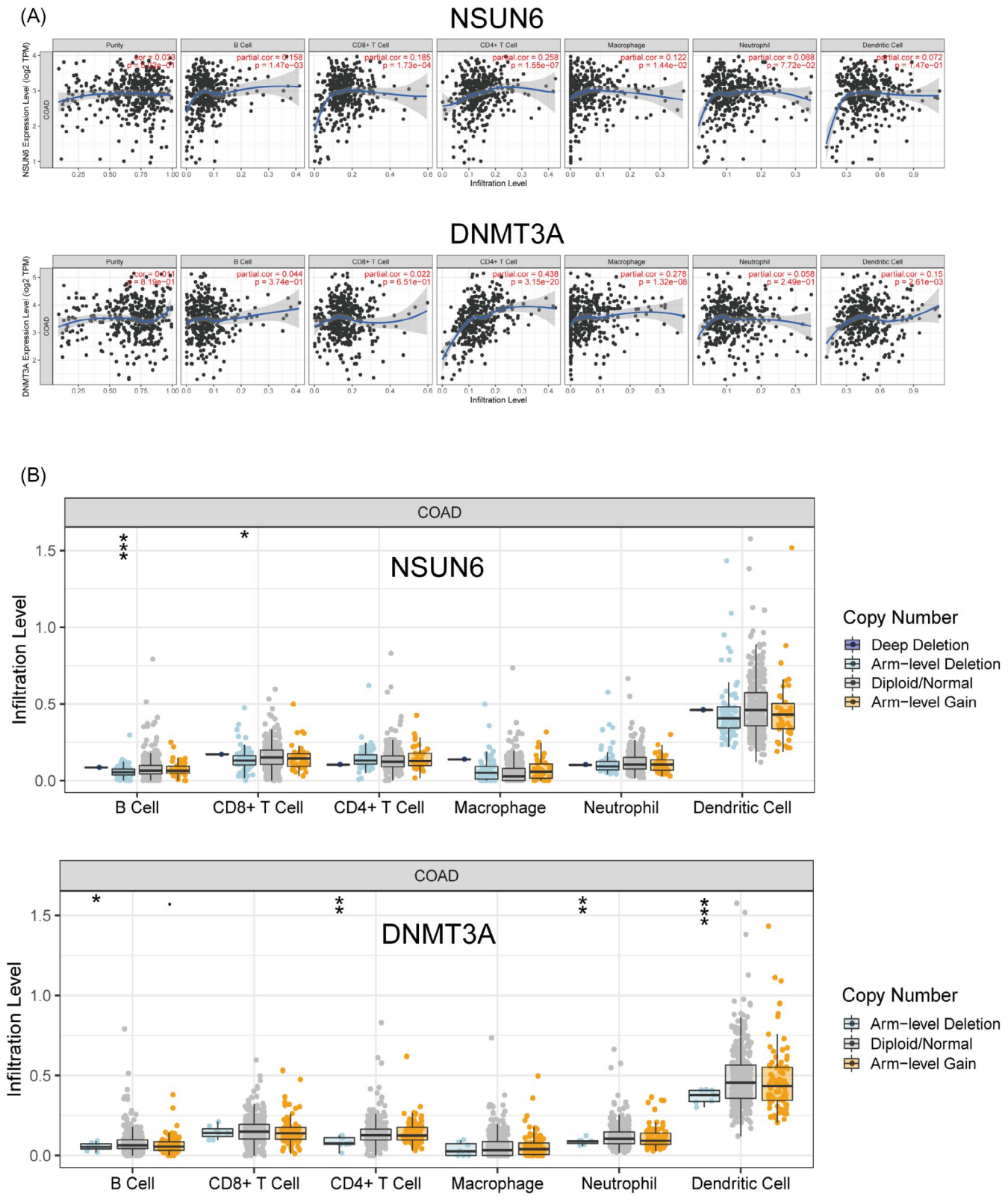


FIGURE 8 Correlation analysis of m^5C regulators with six major immune cell infiltration levels (TIMER). (A) Correlation analysis of NSUN6 and DNMT3A with six major immune cell infiltration levels after purity adjustment.[2] Correlation analysis of the changes in somatic cell copy number of NSUN6 and DNMT3A with the level of immune cell infiltration

Among them, the relationship between $CD4^+$ T cells and immunotherapy has received considerable attention. $CD4^+$ T cell levels are significantly higher in the peripheral blood of patients with CRC who respond well to immunotherapy.[44] Some scholars have even suggested that $CD4^+$ T cells might serve as a marker to predict the response of patients with CRC.[45]

NSUN6 and DNMT3A are expressed in major immune cells to different degrees, and the expression intensity of NSUN6 was higher than that of DNMT3A. More importantly, their expression was also tightly correlated with immune cells. However, there is scant evidence of the association between DNMT3A and B cells. B cell activation and plasma cell differentiation are both regulated by DNMT3A.[46] Conversely, little is known about the role of NSUN6 in immune cell fate determination, and therefore, this is a direction for further studies.

To our knowledge, this is the first study to explore the relationship between m^5C and m^1A regulators and CRC prognosis. Multiple online databases were used to analyze their differential expression in CRC tissues and to construct a prognostic risk model. However, this study has several limitations. First, there have been relatively fewer studies on colon cancer and rectal cancer. Second, the prognostic model did not distinguish among pathological types, making the model less feasible for clinical use. In addition, the small sample size of the GEO data sets used for validation is an evident limitation. Third, although TIMER (2.0) allows for correlation analysis of differentially expressed genes and immune cells in tumor tissues, no in vitro or in vivo experimental validation was performed. Therefore, further in-depth studies are required to address these issues.

5 | CONCLUSIONS

We discovered that m^5C regulators have the potential to effectively predict the survival of patients with CRC. In addition, NSUN6 and DNMT3A can regulate the TIM of CRC and have potential as therapeutic targets.

ACKNOWLEDGEMENTS

We would like to thank contributions from TCGA and GEO project team for the availability of the data. We also appreciate the Editage (www.editage.jp) for English language editing.

CONFLICT OF INTEREST

The authors declare that the study was conducted without any business or financial relationship that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

YL conceived and designed the study. XF, CM, TZ and YZ collected data, analyzed data, and made the figures. XF and CM wrote the manuscript, XS and WC was responsible for modification. All authors read and approved the version of the manuscript submitted.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

ORCID

Yanyan Li https://orcid.org/0000-0003-3571-0689

REFERENCES

1. Iftekhar A, Berger H, Bouznad N, et al. Genomic aberrations after short-term exposure to colibactin-producing E. coli transform primary colon epithelial cells. *Nat Commun*. 2021;12(1):1003.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, 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(6):394-424.
3. Iqbal A, George TJ. Randomized clinical trials in colon and rectal cancer. *Surg Oncol Clin N Am*. 2017;26(4):689-704.
4. Skelton WP, Franke AJ, Iqbal A, George TJ. Comprehensive literature review of randomized clinical trials examining novel treatment advances in patients with colon cancer. *J Gastrointest Oncol*. 2020;11(4):790-802.
5. Zhan L, Feng HF, Liu HQ, et al. Immune checkpoint inhibitors-related thyroid dysfunction: epidemiology, clinical presentation, possible pathogenesis, and management. *Front Endocrinol (Lausanne)*. 2021;12:649863.
6. Pham TND, Spaulding C, Munshi HG. Controlling TIME: how MNK kinases function to shape tumor immunity. *Cancers (Basel)*. 2020;12(8):2096.
7. Binnewies M, Roberts EW, Kersten K, et al. Understanding the Tumor Immune Microenvironment (TIME) for effective therapy. *Nat Med*. 2018;24(5):541-550.
8. Tosolini M, Kirilovsky A, Mlecnik B, et al. Clinical impact of different classes of infiltrating T cytotoxic and helper cells (Th1, th2, treg, th17) in patients with colorectal cancer. *Cancer Res*. 2011;71(4):1263-1271.
9. Mantovani A, Marchesi F, Malesci A, Laghi L, Allavena P. Tumour-associated macrophages as treatment targets in oncology. *Nat Rev Clin Oncol*. 2017;14(7):399-416.
10. Jonkhout N, Tran J, Smith MA, Schonrock N, Mattick JS, Novoa EM. The RNA modification landscape in human disease. *RNA (New York, NY)*. 2017;23(12):1754-1769.
11. Motorin Y, Helm M. RNA nucleotide methylation *Wiley Interdiscip Rev RNA*. 2011;2(5):611-631.
12. Roundtree IA, Evans ME, Pan T, He C. Dynamic RNA modifications in gene expression regulation. *Cell*. 2017;169(7):1187-1200.
13. Dubin DT, Stollar V. Methylation of Sindbis virus "26S" messenger RNA. *Biochem Biophys Res Comm*. 1975;66(4):1373-1379.
14. Nombela P, Miguel-López B, Blanco S. The role of m(6)A, m(5)C and Ψ RNA modifications in cancer: Novel therapeutic opportunities. *Mol Cancer*. 2021;20(1):18.
15. Chen X, Li A, Sun BF, et al. 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. *Nat Cell Biol*. 2019;21(8):978-990.
16. Li Y, Li J, Luo M, et al. Novel long noncoding RNA NMR promotes tumor progression via NSUN2 and BPTF in esophageal squamous cell carcinoma. *Cancer Lett*. 2018;430:57-66.
17. Shi Q, Xue C, Yuan X, He Y, Yu Z. Gene signatures and prognostic values of m1A-related regulatory genes in hepatocellular carcinoma. *Sci Rep*. 2020;10(1):15083.
18. Zheng Q, Yu X, Zhang Q, He Y, Guo W. Genetic characteristics and prognostic implications of m1A regulators in pancreatic cancer. *Biosci Rep*. 2021;41(4):BSR20210337.
19. Han X, Wang M, Zhao YL, Yang Y, Yang YG. RNA methylations in human cancers. *Semin Cancer Biol*. 2021;75:97-115.
20. Wiener D, Schwartz S. The epitranscriptome beyond m(6)A. *Nat Rev Genet*. 2021;22(2):119-131.

21. Lu S, Ding X, Wang Y, et al. The relationship between the network of non-coding RNAs-molecular targets and n6-methyladenosine modification in colorectal cancer. *Front Cell Dev Biol*. 2021;9:772542.
22. Shi L, Chen W, Zhang Z, Chen J, Xue M. N1-methyladenosine profiling of long non-coding RNA in colorectal cancer. *IUBMB Life*. 2021;73(10):1235-1243.
23. Bai Y, Yang C, Wu R, et al. YTHDF1 regulates tumorigenicity and cancer stem cell-like activity in human colorectal carcinoma. *Front Oncol*. 2019;9:332.
24. Chen P, Liu XQ, Lin X, Gao LY, Zhang S, Huang X. Targeting YTHDF1 effectively re-sensitizes cisplatin-resistant colon cancer cells by modulating GLS-mediated glutamine metabolism. *Mol Ther Oncolytics*. 2021;20:228-239.
25. Tang S, Liu Q, Xu M. LINC00857 promotes cell proliferation and migration in colorectal cancer by interacting with YTHDC1 and stabilizing SLC7A5. *Oncol Lett*. 2021;22(2):578.
26. Haruehanroengra P, Zheng YY, Zhou Y, Huang Y, Sheng J. RNA modifications and cancer. *RNA Biol*. 2020;17(11):1560-1575.
27. Frye M, Watt FM. The RNA methyltransferase Misu (NSun2) mediates Myc-induced proliferation and is upregulated in tumors. *Curr Biol*. 2006;16(10):971-981.
28. Yi J, Gao R, Chen Y, et al. Overexpression of NSUN2 by DNA hypomethylation is associated with metastatic progression in human breast cancer. *Oncotarget*. 2017;8(13):20751-20765.
29. Blanco S, Bandiera R, Popis M, et al. Stem cell function and stress response are controlled by protein synthesis. *Nature*. 2016;534(7607):335-340.
30. Chellamuthu A, Gray SG. The RNA methyltransferase NSUN2 and Its potential roles in cancer. *Cells*. 2020;9(8):1758.
31. Chen RX, Chen X, Xia LP, et al. N(6)-methyladenosine modification of circNSUN2 facilitates cytoplasmic export and stabilizes HMGA2 to promote colorectal liver metastasis. *Nat Commun*. 2019;10(1):4695.
32. Zhang X, Wang X, Wang XQD, et al. Dnmt3a loss and Idh2 neomorphic mutations mutually potentiate malignant hematopoiesis. *Blood*. 2020;135(11):845-856.
33. Fouad MA, Salem SE, Hussein MM, et al. Impact of global DNA methylation in treatment outcome of colorectal cancer patients. *Front Pharmacol*. 2018;9:1173.
34. Khaligh A, Fazeli MS, Mahmoodzadeh H, et al. Improved microsatellite instability detection in colorectal cancer patients by a combination of fourteen markers especially DNMT3a, DCD, and MT1X. *Cancer Biomark*. 2021;31(4):385-397.
35. Zhou Y, Yang Z, Zhang H, et al. DNMT3A facilitates colorectal cancer progression via regulating DAB2IP mediated MEK/ERK activation. *Biochim Biophys Acta*. 2022;1868(4):166353.
36. Li J, Hu J, Luo Z, et al. AGR2 is controlled by DNMT3a-centered signaling module and mediates tumor resistance to 5-Aza in colorectal cancer. *Exp Cell Res*. 2019;385(1):111644.
37. Xie Z, Li T, Gan B, et al. Investigation of miR-136-5p key target genes and pathways in lung squamous cell cancer based on TCGA database and bioinformatics analysis. *Pathol Res Pract*. 2018;214(5):644-654.
38. Cummings J, Ethell BT, Jardine L, Burchell B. Glucuronidation of SN-38 and NU/ICRF 505 in human colon cancer and adjacent normal colon. *Anticancer Res*. 2006;26(3b):2189-2196.
39. Hu DG, Mackenzie PI, McKinnon RA, Meech R. Genetic polymorphisms of human UDP-Glucuronosyltransferase (UGT) genes and cancer risk. *Drug Metab Rev*. 2016;48(1):47-69.
40. Osawa K, Nakarai C, Akiyama M, et al. Association between polymorphisms in UDP-glucuronosyltransferase 1A6 and 1A7 and colorectal cancer risk. *Asian Pac J Cancer Prev*. 2012;13(5):2311-2314.
41. Lee SC, Abdel-Wahab O. Therapeutic targeting of splicing in cancer. *Nat Med*. 2016;22(9):976-986.
42. Lv Z, Wang Z, Luo L, et al. Spliceosome protein Eftud2 promotes colitis-associated tumorigenesis by modulating inflammatory response of macrophage. *Mucosal Immunol*. 2019;12(5):1164-1173.
43. Wang W, Zhong Y, Zhuang Z, et al. Multiregion single-cell sequencing reveals the transcriptional landscape of the immune microenvironment of colorectal cancer. *Clin Transl Med*. 2021;11(1):e253.
44. Spitzer MH, Carmi Y, Reticker-Flynn NE, et al. Systemic immunity is required for effective cancer immunotherapy. *Cell*. 2017;168(3):487-502.e15.
45. Li Z, Li S, Liang Y, et al. Predictive value of postoperative peripheral CD4+ T cells percentage in stage I-III colorectal cancer: a retrospective multicenter cohort study of 1028 subjects. *Cancer Manag Res*. 2020;12:5505-5513.
46. Barwick BG, Scharer CD, Martinez RJ, et al. B cell activation and plasma cell differentiation are inhibited by de novo DNA methylation. *Nat Commun*. 2018;9(1):1900.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Fang X, Miao C, Zeng T, et al. Role of m^5C RNA methylation regulators in colorectal cancer prognosis and immune microenvironment. *J Clin Lab Anal*. 2022;36:e24303. doi:10.1002/jcla.24303