

## Research Article

# m5C-Related Signatures for Predicting Prognosis in Cutaneous Melanoma with Machine Learning

Maoxin Huang,<sup>1</sup> Yi Zhang,<sup>1</sup> Xiaohong Ou,<sup>2</sup> Caiyun Wang,<sup>1</sup> Xueqing Wang,<sup>1</sup> Bibo Qin,<sup>1</sup> Qiong Zhang,<sup>1</sup> Jie Yu,<sup>1</sup> Jianxiang Zhang<sup>ID</sup>,<sup>3</sup> and Jianbin Yu<sup>ID</sup><sup>1</sup>

<sup>1</sup>Department of Dermatology, The first Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450052, China

<sup>2</sup>Department of Gastroenterology, The first Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450052, China

<sup>3</sup>Department of Oncology, The first Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450052, China

Correspondence should be addressed to Jianxiang Zhang; ywczgk3302@163.com and Jianbin Yu; yjbdoctor@sina.com

Received 27 May 2021; Accepted 28 July 2021; Published 5 August 2021

Academic Editor: Quan Cheng

Copyright © 2021 Maoxin Huang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Background.** Cutaneous melanoma (CM) is one of the most life-threatening primary skin cancers and is prone to distant metastases. A widespread presence of posttranscriptional modification of RNA, 5-methylcytosine (m5C), has been observed in human cancers. However, the potential mechanism of the tumorigenesis and prognosis in CM by dysregulated m5C-related regulators is obscure. **Methods.** We use comprehensive bioinformatics analyses to explore the expression of m5C regulators in CM, the prognostic implications of the m5C regulators, the frequency of the copy number variant (CNV), and somatic mutations in m5C regulators. Additionally, the CM patients were divided into three clusters for better predicting clinical features and outcomes via consensus clustering of m5C regulators. Then, the risk score was established via Lasso Cox regression analysis. Next, the prognosis value and clinical characteristics of m5C-related signatures were further explored. Then, machine learning was used to recognize the outstanding m5C regulators to risk score. Finally, the expression level and clinical value of USUN6 were analyzed via the tissue microarray (TMA) cohort. **Results.** We found that m5C regulators were dysregulated in CM, with a high frequency of somatic mutations and CNV alterations of the m5C regulatory gene in CM. Furthermore, 16 m5C-related proteins interacted with each other frequently, and we divided CM patients into three clusters to better predicting clinical features and outcomes. Then, five m5C regulators were selected as a risk score based on the LASSO model. The XGBoost algorithm recognized that NOP2 and NSUN6 were the most significant risk score contributors. Immunohistochemistry has verified that low expression of USUN6 was closely correlated with CM progression. **Conclusion.** The m5C-related signatures can be used as new prognostic biomarkers and therapeutic targets for CM, and NSUN6 might play a vital role in tumorigenesis and malignant progression.

## 1. Introduction

Cutaneous melanoma (CM) is one of the most common, aggressive, and life-threatening types of malignant primary skin cancers and is prone to distant metastases associated with a high mortality rate [1, 2]. Surgical management, chemotherapy agents, and immunotherapy have been considered the fundamental treatment of CM in recent years [3, 4]. Although great achievements have been made in CM treatment, the 5-year overall survival with metastatic melanoma patients remains poor, attributed to late diagnosis [5, 6], rapid metastasis, and poor response to treatment

[7, 8]. Therefore, exploring powerful prognostic predictors and a novel therapeutic target is urgent and crucial to improving CM diagnosis and treatment.

Additionally, 5-methylcytosine (m5C) is a widespread in posttranscriptional modification of RNA which has been observed in substantial RNA species, including rRNAs and tRNAs, mRNAs, eRNAs, and noncoding RNAs in the cytoplasm and mitochondria [9–18]. The methylation of m5C involves a series of regulators, including m5C methyltransferases, demethylases, and “readers.” The methyltransferase “writer,” including NSUN1–7, DNMT1, DNMT2, DNMT3A, and DNMT3B, increases methylation

at the C5 position of RNAs. Then, different “reader” proteins, such as TET1-3, recognize and bind the methylated mRNAs, while the “eraser” protein, such as ALYREF and YBX1, reverses m5C modification by degrading the written methylation [19, 20]. Some studies have indicated that m5C modification plays an important role in many biological functions, including ribosome assembly, tRNA stability, mRNA export, protein transcription, and stem cell regulation [21, 22]. It has been reported that the m5C gene mutations are closely associated with various human diseases such as nervous system disorders, metabolic disease, and virus infections [23, 24]. In addition, dysregulated m5C regulators have been observed in human cancers such as breast, gallbladder, and bladder cancer [25–28]. However, the potential tumorigenesis mechanism and prognosis in CM by dysregulated m5C-related regulators is obscure.

In this study, we explored the mRNA level of expression of m5C regulators by using The Cancer Genome Atlas (TCGA) database and Gene Expression Omnibus (GEO) database. Then, we analyzed the frequency of m5C regulators’ copy number variant (CNV) and somatic mutations. The CM patients were divided into three clusters to better predict clinical features and outcomes via consensus clustering of m5C regulators. The established prognostic gene signature showed a five-gene prognostic signature, which had an effective prediction ability of the progression and prognosis of CM patients, comprising of NOP2, NSUN3, NSUN6, DNMT2, and YBX1 via Lasso Cox regression analysis. Moreover, the protein expression levels of NSUN6 were significantly upregulated in the CM, associated with the advanced TNM stage. Further survival analysis indicated that NSUN6 could be an independent risk factor for the prognosis of CM patients. NSUN6 expression was significantly related to the pathologic state and TNM staging of CM. These results showed that m5C-regulator-based prognostic signature can be used as new prognostic biomarkers and therapeutic targets for CM. Additionally, NSUN6 might play vital roles in tumorigenesis and malignant progression.

## 2. Materials and Methods

**2.1. Datasets.** The RNA-seq transcriptome data of CM were downloaded from The Cancer Genome Atlas (TCGA, <https://tcga-data.nci.nih.gov/tcga>) and the Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/GSE3189>, GSE7553, GSE31909, GSE46517, GSE98394, and GSE114445). The characteristics of these CM patients and normal people are presented in Tables S1 and S2.

**2.2. Tissue Samples.** The 40 tumors and 8 normal tissue microarrays were purchased from Shaanxi Avila Biotechnology Co., Ltd. (Xi’an, China). Also, our studies were approved by the Ethics Committee of the First Affiliated Hospital of Zhengzhou University.

**2.3. Immunohistochemistry (IHC).** Immunohistochemistry was performed as previously reported [29, 30]. Semi-quantification analysis of immunohistochemical staining was

performed using ImageJ (NIH, Bethesda, MD). Two experienced pathologists who were blinded to the clinicopathological data evaluated the immunostaining samples separately, and they scored the samples according to the proportion of positive cells as follows: no staining, 1+; weak staining, 2+; moderate staining, 3+; and strong staining, 4+. Table S3 lists the information of antibodies used in this study.

**2.4. Consensus Clustering.** Consensus clustering was based on the m5C regulators’ profiles in CM using the R package “Consensus Cluster Plus,” and to guarantee the stability of clustering, 1000 times repetitions were performed. The appropriate number of clusters was calculated via the cumulative distribution function (CDF) and consensus matrices.

**2.5. Pathway Enrichment Analysis.** Spearman’s analysis was conducted to evaluate the correlation between 16 m5C regulators and their function, and interaction analysis was predicted by the String website (<https://string-db.org/>). A comprehensive gene function analysis was performed with Metascape (<http://metascape.org/>). To further explore the potential mechanism of the m5C regulators in CM, we performed the gene set variation analysis (GSVA), the Kyoto Encyclopedia of Genes and Genomes (KEGG), and Gene Ontology (GO) analysis [31].

**2.6. Construction and Validation of Risk Score.** To construct the powerful prognostic factors, the five m5C regulators were selected by the Least Absolute Shrinkage and Selection Operator (LASSO) Cox regression model [32]. The Cox regression analysis was conducted for 16 genes, and 0.2 was set as the cutoff *P* value to prevent omissions. The risk score for each patient in both the training (TCGA) and validation (GSE65904) datasets was calculated using the following formula:

$$\text{risk score} = \sum_{i=1}^n \text{Coef} * xi. \quad (1)$$

Coefi means the coefficient, and xi is the expression value of the z-score transformation of each selected m5C regulator.

**2.7. Machine Learning Algorithm.** We used the machine learning model, extreme gradient, (XGBoost) with the Shapley additive explanation method to explore the importance of m5C regulators to the risk score.

**2.8. Statistical Analysis.** All statistical analyses were performed in the R statistical computing language (R version 4.0.4) unless otherwise indicated. The differences between the two independent groups were determined using Student’s *t*-test (unpaired, two tailed). The chi-square test and rank-sum test were used for qualitative variables. The Kaplan–Meier overall survival (OS) and progression-free survival (PFS) analyses were performed using the log-rank test based on the median. Univariate and multivariate Cox regression analyses were performed to assess independent

prognostic factors.  $P < 0.05$  was considered statistically significant (\* $P < 0.05$ , \*\* $P < 0.01$ , and \*\*\* $P < 0.001$ ).

### 3. Results

**3.1. Variation and Prognosis Value of m5C Regulators in CM.** To probe the significant biological function of m5C regulators in CM, we explore the mRNA expression levels of m5C regulators in CM and normal tissues using public databases. We found that the mRNA expression levels of NOP2, YBX1, DNMT3A, NSUN5, ALYREF, and DNMT1 were significantly upregulated and NSUN6, NSUN7, TET2, and TET3 were significantly decreased in CM tissues (Figure 1(a)). Furthermore, to obtain insight into the causes of dysregulation of m5C regulators, we investigated alterations to the somatic mutations and the CNV alteration frequency of m5C regulators. Among all the 467 patients, 130 experienced m5C regulators mutations with a frequency of 27.84%, indicating a high frequency of somatic mutations of the m5C regulatory gene in CM (Figure 1(b)). We found that the copy number of CNV in NSUN5, DNMT3B, and ALYREF was amplified while NSUN6, TET1, TET2, and DNMT2 were deleted. NSUN5 and ALYREF consistently represented high mRNA expression levels, while NSUN6 and TET2 indicated low mRNA expression levels (Figure 1(c)). Survival analysis showed that NSUN3, DNMT2, and NSUN6 were potential protective factors and that YBX1 and NOP2 were potential risk factors for OS (Figures 1(d) and S1). Moreover, NOP2 was a potential risk factor for PFS in CM (Figures 1(e) and S2). Overall, as these results presented significant genomic variations of CM m5C regulators, m5C regulators had a potential prognosis value in CM.

**3.2. Unsupervised Consensus Clustering Analysis of m5C RNA Methylation Modulators.** To integrally investigate the role of m5C modification in CM, we analyzed the interaction among the 16 m5C-related regulators. Results indicated that the protein-protein interaction networks and m5C-related proteins interacted with each other frequently (Figure 2(a)). Additionally, we explore the correlation among the expression profile of 16 m5C regulators in a TCGA dataset by conducting a Pearson correlation analysis. These results indicated that most m5C regulators were positive correlations, among which the expression of ALYREF, DNMT2, DNMT3B, NSUN5, NSUN6, and TET1 showed a stronger correlation in CM (Figure 2(b)). Thus, there is a close relationship between the biological functions of 16 m5C regulators and CM. Based on the expression levels of the m5C modulators of CM patients, we observed the clustering stability of the TCGA dataset from  $k = 2$  to 6 (Figure S3) via unsupervised consensus.

**3.3. Clustering analysis.** Using  $k = 3$ , we identify these three clusters, designated as m5C cluster 1, m5C cluster 2, and m5C cluster 3 (Figure 2(c) and Table S4). Subsequently, we found that cluster 1 presented moderate expression in most of the m5C regulators and cluster 2 showed the highest

expression of USUN6, USUN7, TET1, TET3, DNMT3A, and DNMT3B; cluster 3 exhibited the lowest expression of NSUN3, USUN6, TET1, TET2, TET3, DNMT2, DNMT3A, and DNMT3B (Figure 2(d)). Then, we compared the prognosis between cluster 2 and cluster 3, finding a significant distinction between m5C cluster 2 and cluster 3. According to the prognostic analysis, cluster 2 had a more significant potential survival advantage than cluster 3 (Figure 2(e)). Moreover, a t-SNE dimensional reduction showed that it was feasible to segregate the m5C clusters into three discrete clusters (Figure 2(f)). These results indicate that m5C clusters might be more related to CM prognosis. Next, we compared the clinical features and clinicopathological factors between these three clusters. Chi-square analysis indicated that cluster 3 was significantly associated with higher pathological grade ( $P < 0.05$ ) and deeper Breslow depth value ( $P < 0.001$ ) and cluster 2 was significantly associated with orthotopic tumors, a lower pathological grade, and shallower Breslow depth value (Table S4).

**3.4. Biological Functional Annotation of m5C Clusters.** To further explore the biological behaviors of m5C clusters, we selected the top 500 upregulated expressed genes in cluster 3 based on the  $P$  value to annotate the biological function. Results indicated that the “Melanin biosynthetic process” and “Keratinization” were significantly enriched in m5C cluster 3. We displayed the top 20 significantly enriched biological processes in Figure 3(a). Simultaneously, the GSEA analysis indicated that certain tumor progression pathways (“DNA repair,” “the reactive oxygen species pathway,” “the p53 signaling pathway,” and “angiogenesis”) were significantly activated in m5C cluster 3 patients (Figure 3(b)). Moreover, GSEA analysis showed that “cellular response to reactive oxygen species,” “the melanin biosynthetic process,” “cell cycle DNA replication,” “angiogenesis,” “the Wnt signaling pathway,” and “the p53 signaling pathway” were significantly enriched in m5C cluster 3, exhibiting worse clinical features and prognosis (Figures 3(c)–3(h)). In summary, there was a significant association between clusters of m5C RNA methylation modulators and the progression of malignant CM.

**3.5. Prognostic Value and Construction of the Risk Score Signature of m5C Regulators.** We performed a prognostic risk score signature containing DNMT2, NSUN3, NSUN6, YBX1, and NOP2 using a LASSO Cox regression model according to the minimum criterion to evaluate the prognosis value of m5C methylation modification on individual CM patients accurately (Figure 4(a)). The coefficients of DNMT2, NSUN3, NSUN6, YBX1, and NOP2 were  $-0.01842$ ,  $-0.08767$ ,  $-0.07701$ ,  $0.00057$ , and  $0.18925$ , respectively. Thus, we calculated each for CM patients with the formula risk score =  $(-0.01842 \times \text{expression of DNMT2}) + (-0.08767 \times \text{expression of NSUN3}) + (-0.07701 \times \text{expression of NSUN6}) + (0.00057 \times \text{expression of YBX1}) + (0.18925 \times \text{expression of NOP2})$  in the training (TCGA) and validation (GSE65904) dataset.

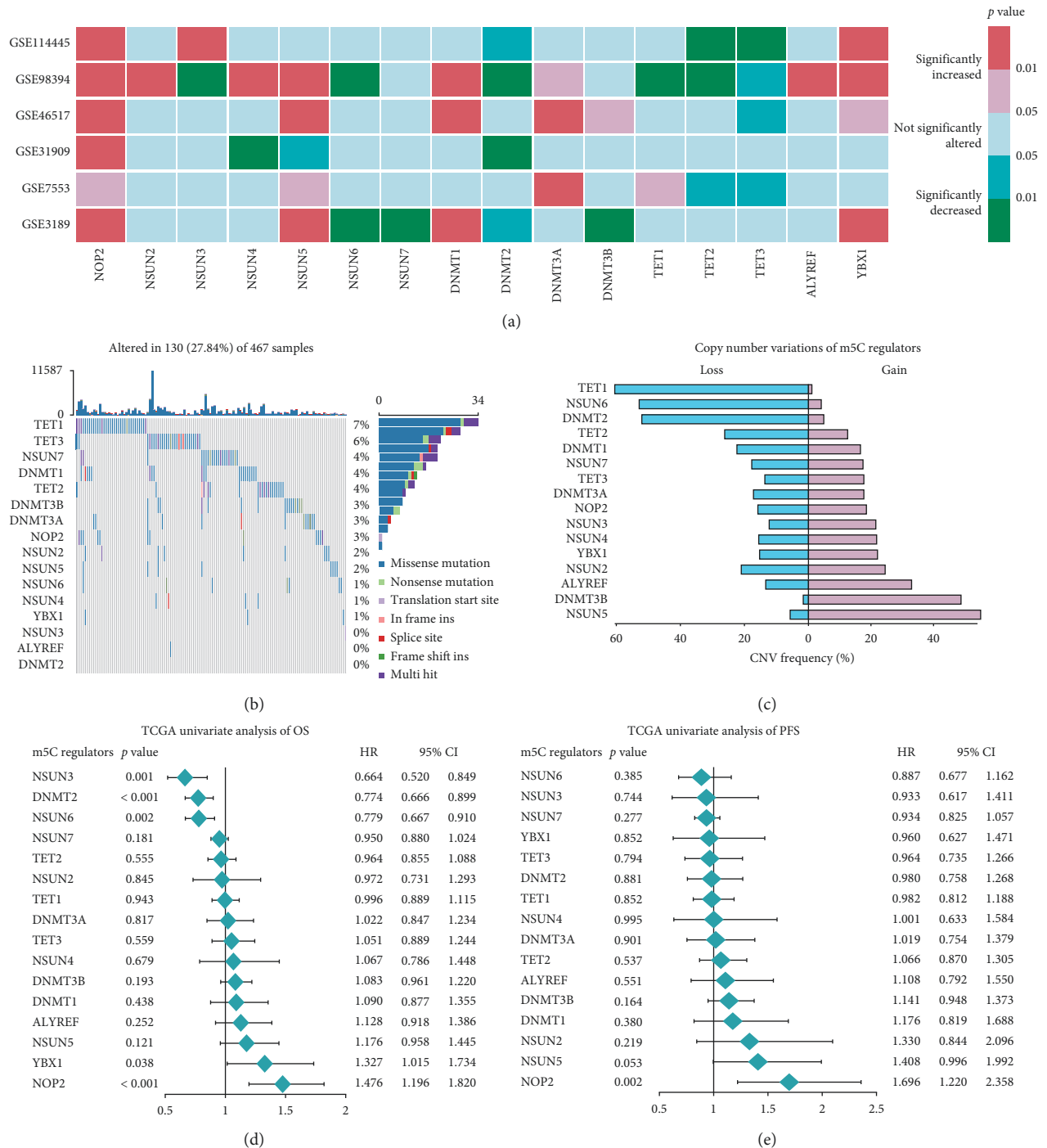
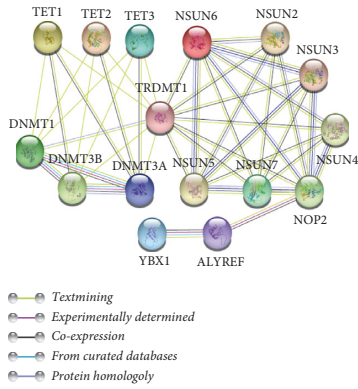


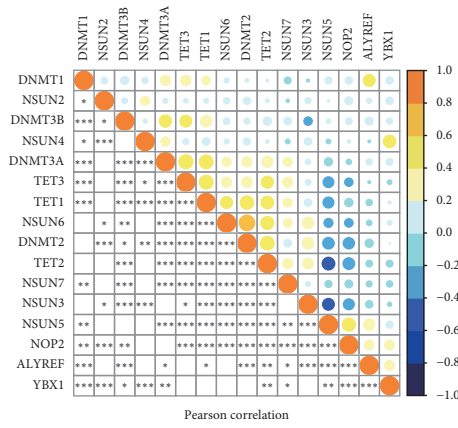
FIGURE 1: The variation and prognosis value of m5C regulators in CM. (a) The mRNA expression status of m5C regulators in CM. (b) The CNV frequency of m5C regulators in the TCGA-CM cohort. (c) The copy number of 16 m5C-related genes in CM. (d) Univariate Cox analysis of OS in CM. (e) Univariate Cox analysis of PFS in CM.

To further investigate the prognostic value of the risk score, CM patients in the TCGA were classified into high- or low-risk groups according to the median. The results demonstrate that patients with high-risk scores were correlated with a higher death rate (Figure 4(b)). Additionally, we found that the difference between high- and low-risk groups in OS and PFS was significant (Figures 4(c) and 4(d)). Moreover, to

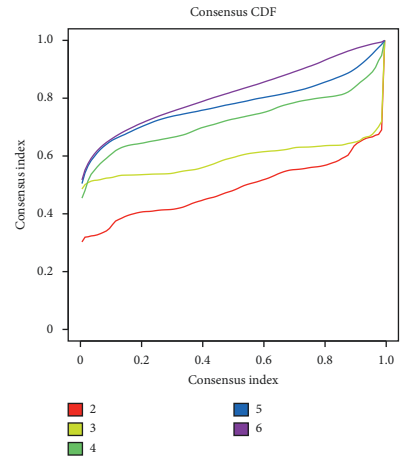
validate the predictive efficiency of the prognostic risk score, we used the same analysis in the GSE65904 dataset; the results were in line with the aforementioned findings (Figures 4(e)–4(g)). Our findings suggested that the signature based on five m5C regulators could function as a potent prognostic signature and effectively stratify CM patients based on risk scores providing new insight into targeted therapy.



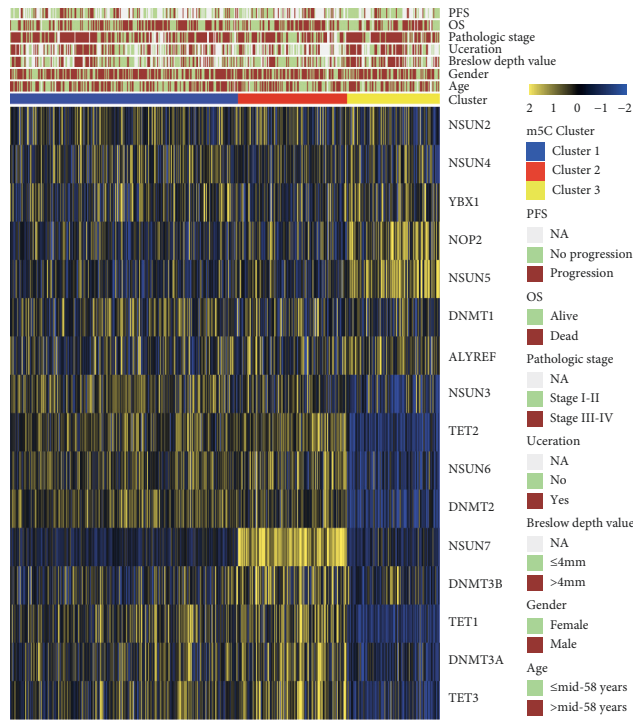
(a)



(b)



(c)



(d)

FIGURE 2: Continued.

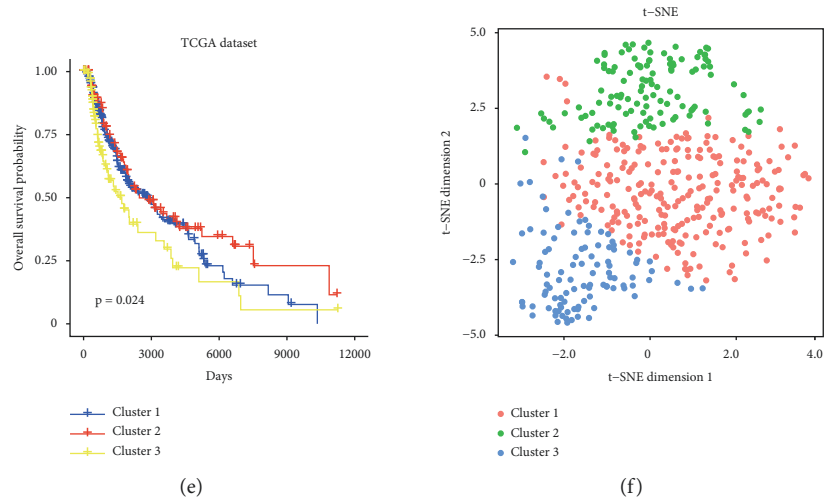
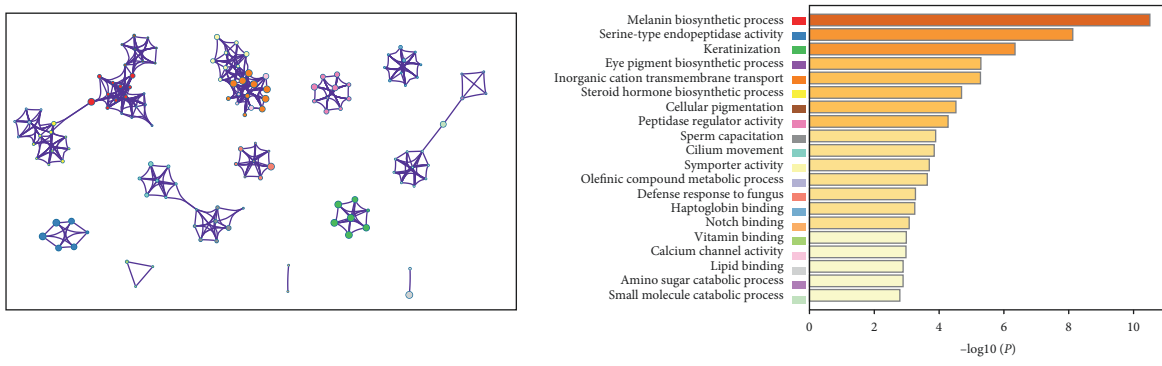
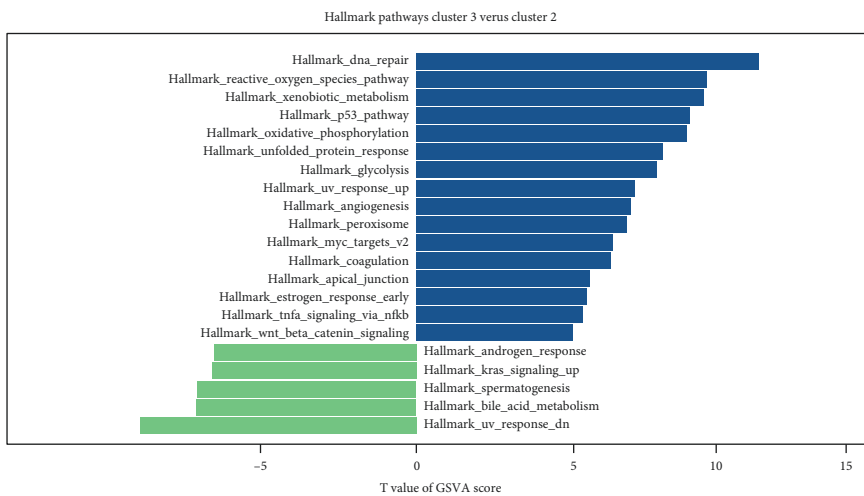


FIGURE 2: Unsupervised consensus clustering analysis of m5C RNA methylation modulators. (a) The protein-protein interactions of sixteen m5C regulators. (b) Pearson’s correlation analysis of the expression of 16 m5C regulators in CM. (c) Consensus clustering matrix for the most suitable groups,  $k = 3$ . (d) The relationship between the expression levels of m5C regulators and the clinical features in the three clusters of CM. (e) Overall survival analysis for CM patients stratified by the three clusters. (f) t-SNE plots of TCGA-CM RNA-sequence profiles for the three m5C clusters.



(a)



(b)

FIGURE 3: Continued.

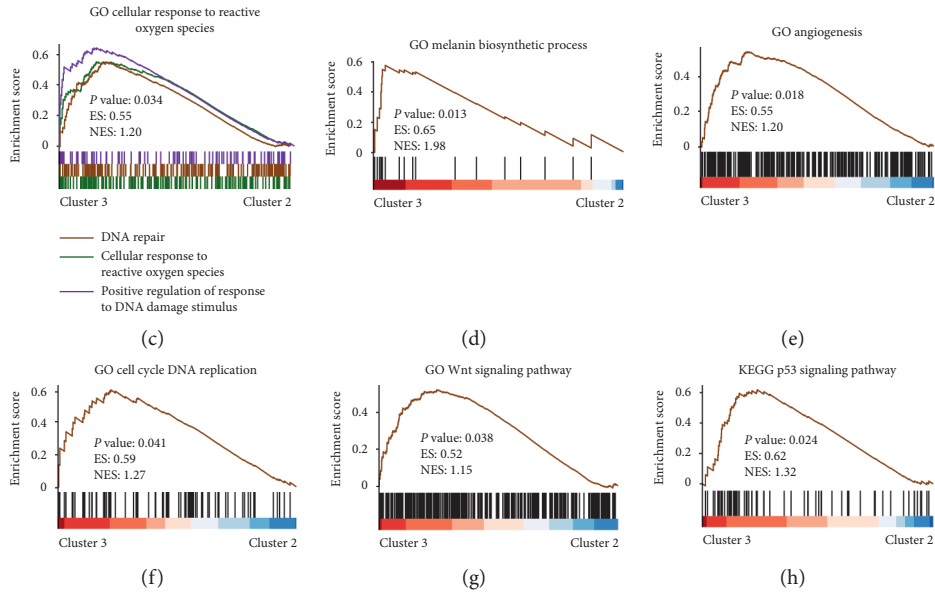


FIGURE 3: Functional annotation of m5C clusters. (a) The network and the top 20 significantly enriched biological pathways in the m5C clusters. (b) The GSEA between m5C clusters 2 and m5C clusters 3 using hallmark gene sets. (c–h) GSEA showed that the “cellular response to reactive oxygen species,” “melanin biosynthetic process,” “cell cycle DNA replication,” “angiogenesis,” “Wnt signaling pathway,” and “p53 signaling pathway” were significantly enriched in m5C cluster 3.

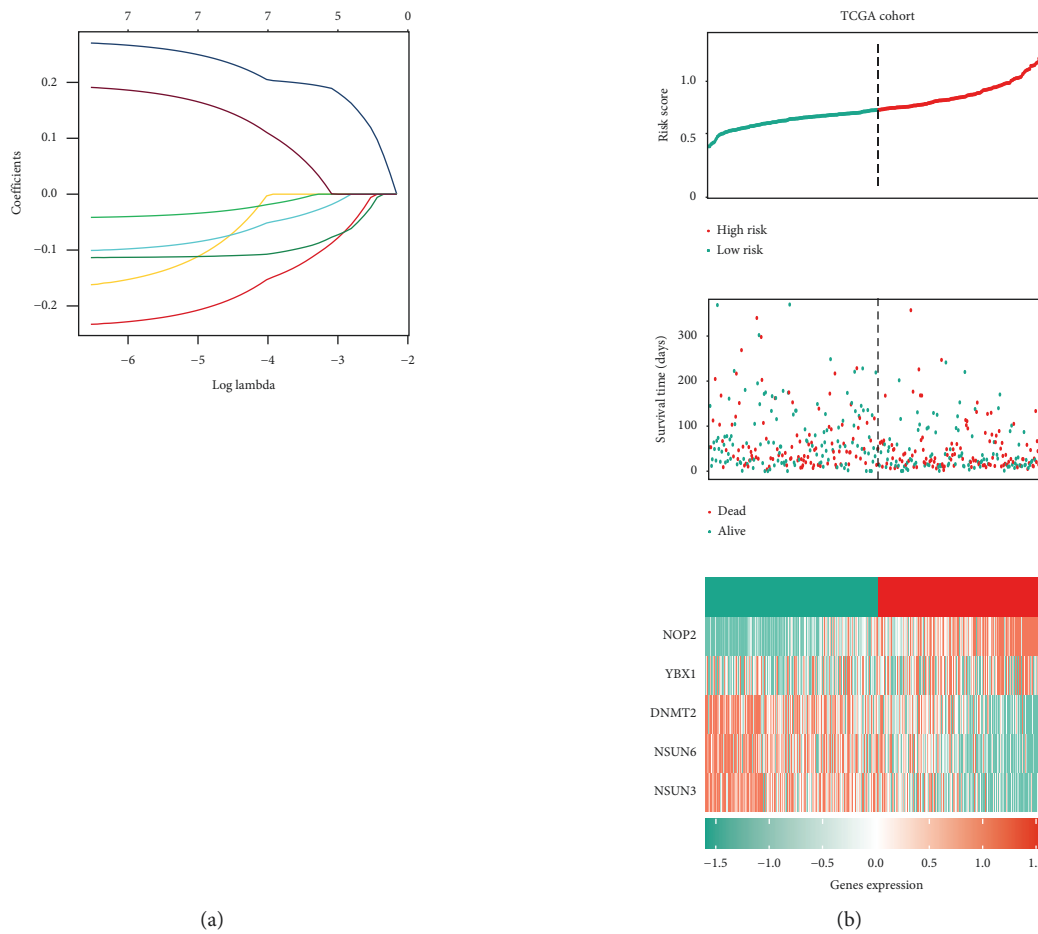


FIGURE 4: Continued.

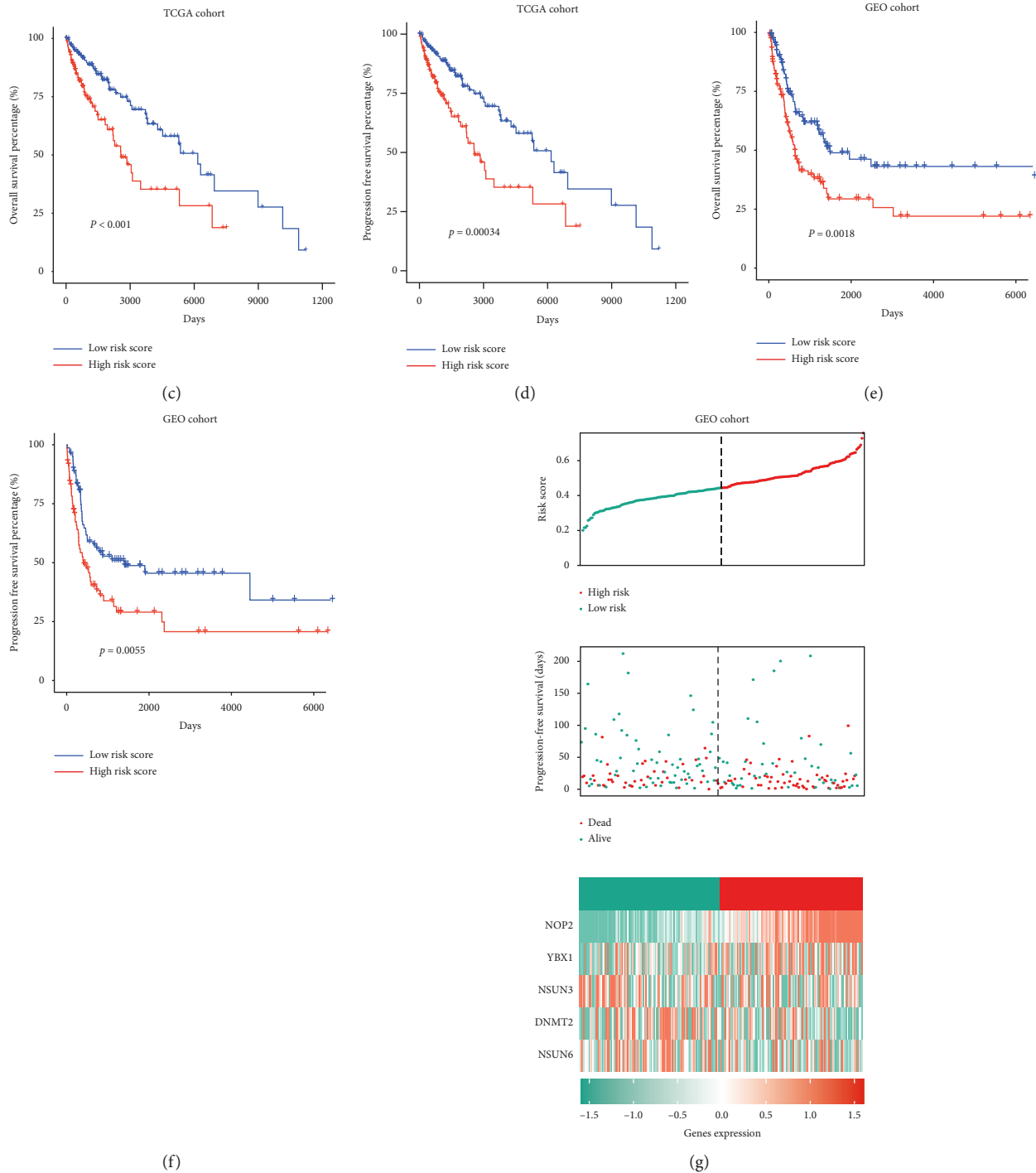


FIGURE 4: Prognostic value and construction of the risk score signature of m5C regulators. (a) LASSO Cox regression was used to calculate the coefficients for five m5C-related genes. (b) Patients with high-risk scores were correlated with higher death rates among CM patients. (c, d) Kaplan–Meier analysis of OS and PFS in patients with high- and low-risk score groups in the TCGA dataset. (e, f, g) Validation of the risk score signature in the GSE65904 dataset using the same analysis.

**3.6. Relationship between the Risk Score and Clinical Characteristics in CM Patients.** Next, we tested the correlation of m5C-related genes and risk scores with clinical characteristics. Heatmaps represented the expression of m5C

regulators and clinical features in low- and high-risk patients in the TCGA dataset (Figure 5(a)). Results illustrated that high expression of NOP2 and USUN5 and low expression of USUN6, DNMT2, TET2, and USUN3 were



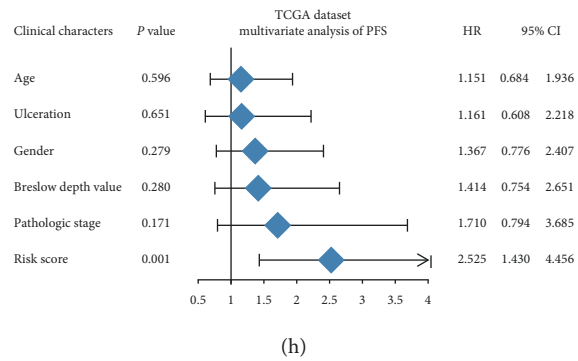
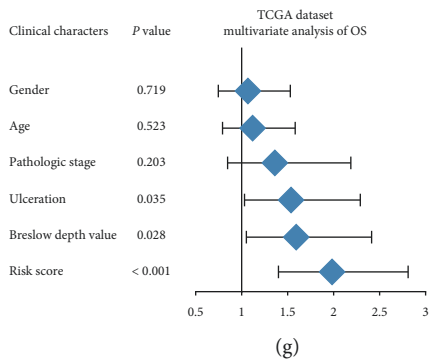
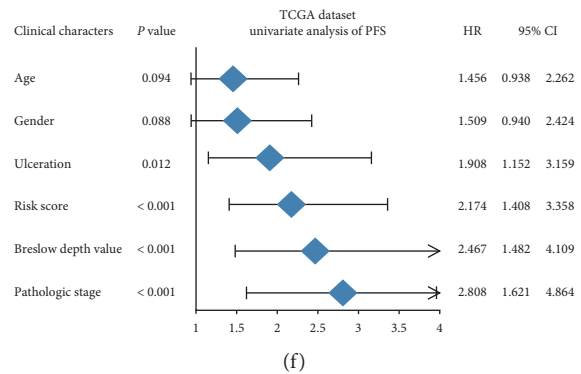
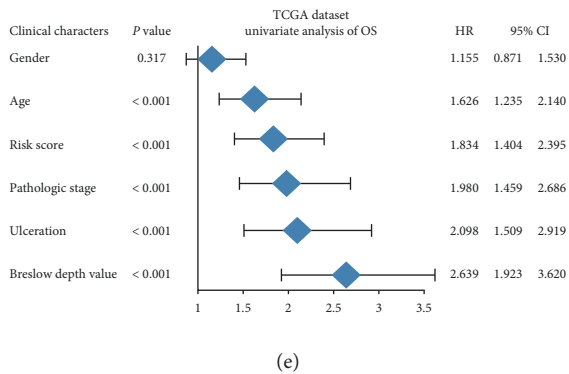
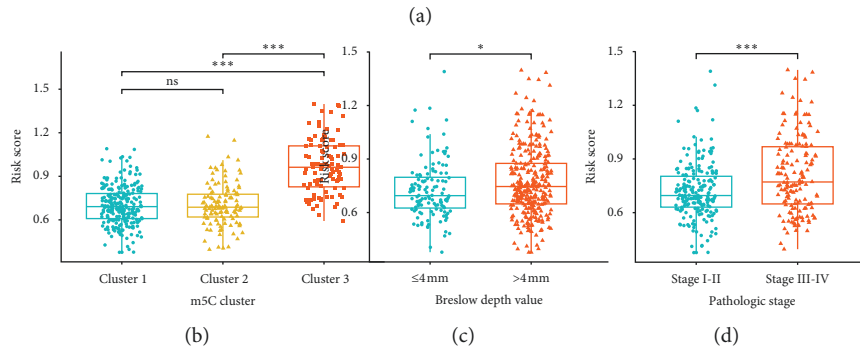
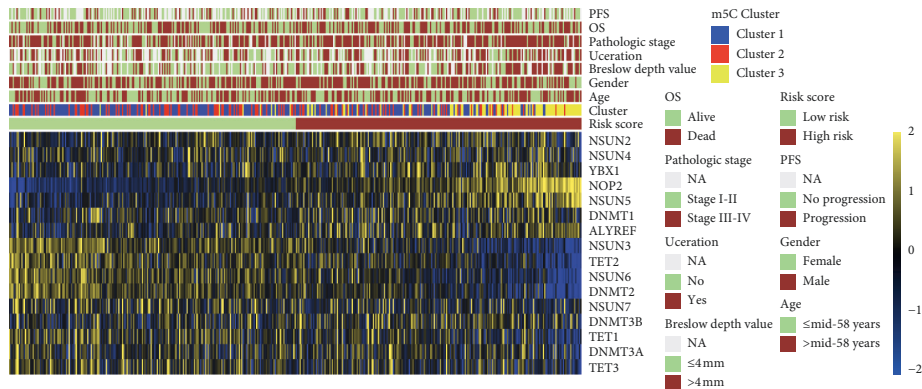


FIGURE 5: The relationship between the risk score and clinical characteristics. (a) The relationship between sixteen m5C expression profiles stratified by risk score and clinical characteristics in CM patients. (b–d) The relationships between the risk scores and clinical characteristics including m5C clusters, Breslow depth value, and pathologic stage. (e, g) The risk score is an independent prognostic factor of OS in the TCGA dataset. (f, h) The risk score is an independent prognostic factor of PFS in the TCGA dataset.

associated with the high-risk score. In addition, we analyzed the association between the risk scores and each clinicopathological characteristic. As shown in Figure 5, the risk scores of cluster 2 and cluster 3 were significantly

different in the Breslow depth value and pathologic stage (Figures 5(b)–5(d) and Table S5). Moreover, univariate and multivariate Cox regression analyses of OS and PFS in CM were conducted, indicating that the risk score was an

independent predictor of prognosis and progression (Figures 5(e)–5(h) and Table S6). Results revealed that the risk scores were significantly correlated with the malignancy of CM.

**3.7. Importance of m5C Regulators to the Risk Score Using Machine Learning.** To build a classifier that could recognize the outstanding contributors of risk score, we applied the XGBoost algorithm to build the model. NOP2, NSUN6, NSUN3, and DNMT2 were the top four important regulators evaluated by SHapley additive exPlanation value (SHAP), as shown in Figures 6(b)–6(f). Each point is the feature value of the specific gene. The features are sorted by the sum of the values of all samples. We found that NOP2 and NSUN6 were the most important regulators in our risk score.

**3.8. Close Correlation of CM Progression and Low Expression of USUN6.** Subsequently, immunohistochemical staining revealed the decreased protein expression level of NSUN6 in CM tissues which was consistent with our findings (Figure 7(a)). According to the staining intensity, NSUN6 staining was scored from 1+ to 4+ (Figure 7(b)). Furthermore, NSUN6 expression was significantly related to the pathologic state and TNM staging of CM in TCGA and tissue microarray cohorts (Figures 7(c)–7(f)). These findings strongly suggested that USUN6 might serve as a tumor suppressor molecule in CM.

## 4. Discussion

CM is the primary skin cancer mortality cause worldwide owing to its high metastasis, invasiveness, and annually increasing incidence [1, 5]. Although present studies have shown that many complex multistep processes contribute to the initiation, progression, and metastasis of CM, its pathogenesis has not been well determined and effective prognostic biomarkers in CM are still lacking [4]. Therefore, it is valuable to understand the underlying molecular mechanisms, which contribute to the prognostic prediction and therapeutic target of CM. In recent years, m5C modification has been found to play a significant role in the biological function of many cancers. Additionally, the dysregulation of m5C regulators has been observed in numerous cancers including breast, gallbladder, and bladder cancer [25, 26, 28, 33]. Xue et al. [27] recently clarified the m5C-related regulators are more abundant in head and neck squamous cell carcinoma and play vital roles in tumor progression. However, the potential tumorigenesis mechanism and prognosis of CM by dysregulated m5C-related regulators is still elusive.

Our study systematically analyzed the mRNA expressions of m5C regulators and characterized the mutations and copy number variations (CNVs) of m5C regulators in CM for the first time, utilizing data extracted from the public databases. We found that the mRNA expression levels of NOP2, YBX1, DNMT3A, and DNMT1 were significantly increased and NSUN6, NSUN7, TET2, and TET3 were

significantly decreased in CM tissues. Somatic mutations and CNV alterations indicated a high mutation frequency of the m5C regulatory gene in CM. Previous studies have demonstrated that some m5C regulators participate in the malignant progression of cancer such as NOP2 and YBX1. The literature suggests that NOP2 is highly expressed in prostate cancer, gallbladder cancer, and lung adenocarcinoma, while upregulated YBX1 plays a significant protumorigenic role in breast, renal, and gastric cancer [34–38]. Conversely, NSUN6 confers cellular fitness advantages and functions as a tumor suppressor in pancreatic cancer [39]. Furthermore, survival analysis of OS and PFS in the TCGA revealed that the m5C regulators had a potential value in evaluating prognosis in CM. In summary, our findings showed that dysregulated m5C regulators may play a key role in the initiation and progression of CM.

Accumulating evidence indicated that m5C RNA methyltransferase-related regulators were implicated in tumorigenesis and tumor development in ovarian cancer, cervical cancer, prostate cancer, glioma, and other several tumors [12, 40]. However, the current study did not investigate the relationship between m5C modifications and CM development. We investigated the association between m5C regulators, clinical features, and prognosis in CM for a comprehensive analysis, finding that m5C regulators had frequent crosstalk. Additionally, we analyzed the expression profiles of m5C regulators, and three clusters with different clinical features and prognoses were identified. Furthermore, we chose hallmark pathways to conduct an in-depth research. Consistent with our results, UV radiation is the primary environmental driver including oxidative stress and DNA damage. Following UV exposure, keratinocytes in a p53-dependent manner produced  $\alpha$ -melanocyte-stimulating hormone and stimulated the melanocortin 1 receptor to produce melanin. Meanwhile, it has been demonstrated that Wnt/ $\beta$ -catenin signaling and angiogenesis played a crucial role in the pathogenesis and progression of CM [41–44]. Therefore, further studies should be conducted on m5C regulators and cancer-promoting signaling pathways that form a potential comprehensive network, significantly influencing CM progression.

We constructed a risk score, including the expression of DNMT2, NSUN3, NSUN6, YBX1, and NOP2, using the scoring algorithm in the TCGA (training set) and the GSE65904 (verification set) to accurately evaluate the patient's prognosis. Based on the median, CM patients were divided into high- and low-risk groups. Further analysis revealed that risk score had suitable stratification abilities for predicting prognosis and tumor progression in CM patients. Moreover, the risk score could independently predict prognosis and progression in CM patients. Moreover, the high-risk group is highly correlated with the advanced pathologic stage and deeper Breslow depth. These results demonstrate that the m5C regulators might be regarded as a potential tool to predict progression and prognosis in CM patients.

To gain more information on the core signature among m5C regulators in CM, we performed machine learning via the XGBoost algorithm [45–48]. An interesting finding was

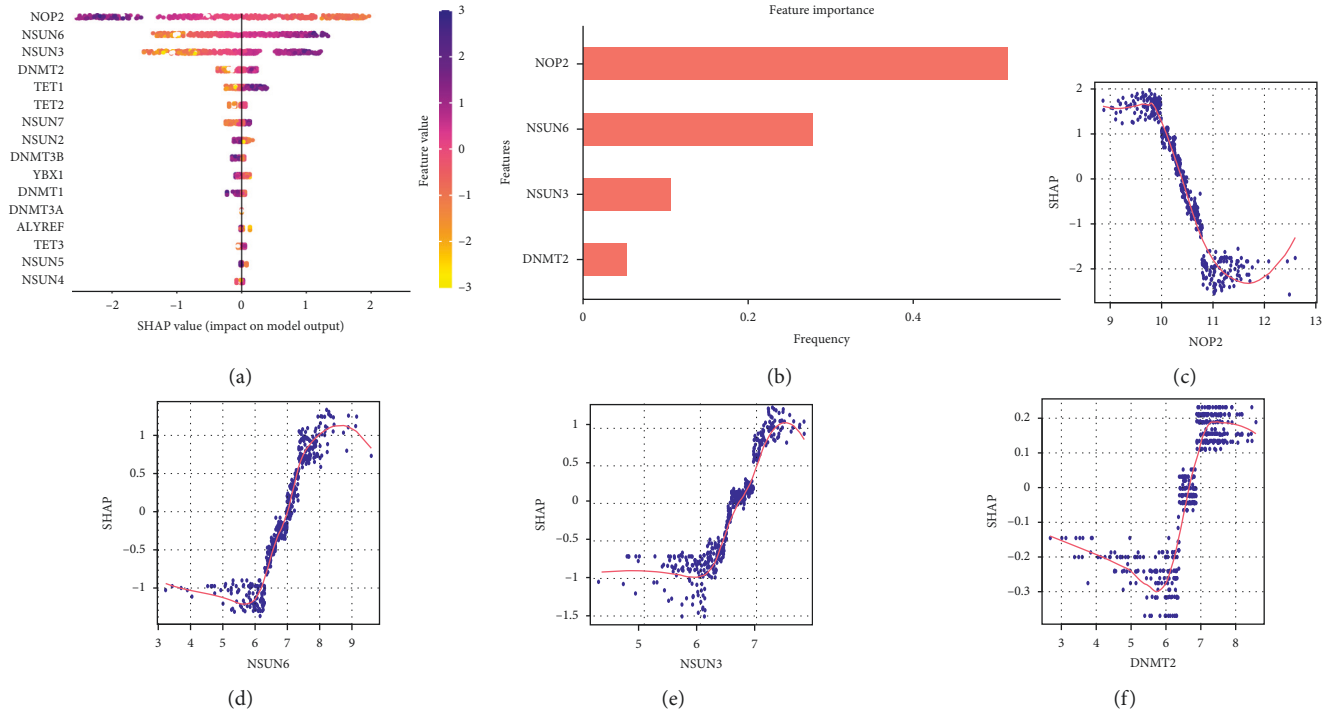


FIGURE 6: Importance of m5C regulators to the risk score by machine learning. (a) The importance of m5C regulators to the risk score by XGBoost algorithm. (b) The top 4 (NOP2, NSUN6, NSUN3, and DNMT2) important regulators evaluated by the SHAP value. (c-f) The individual SHAP value of NOP2, NSUN6, NSUN3, and DNMT2.

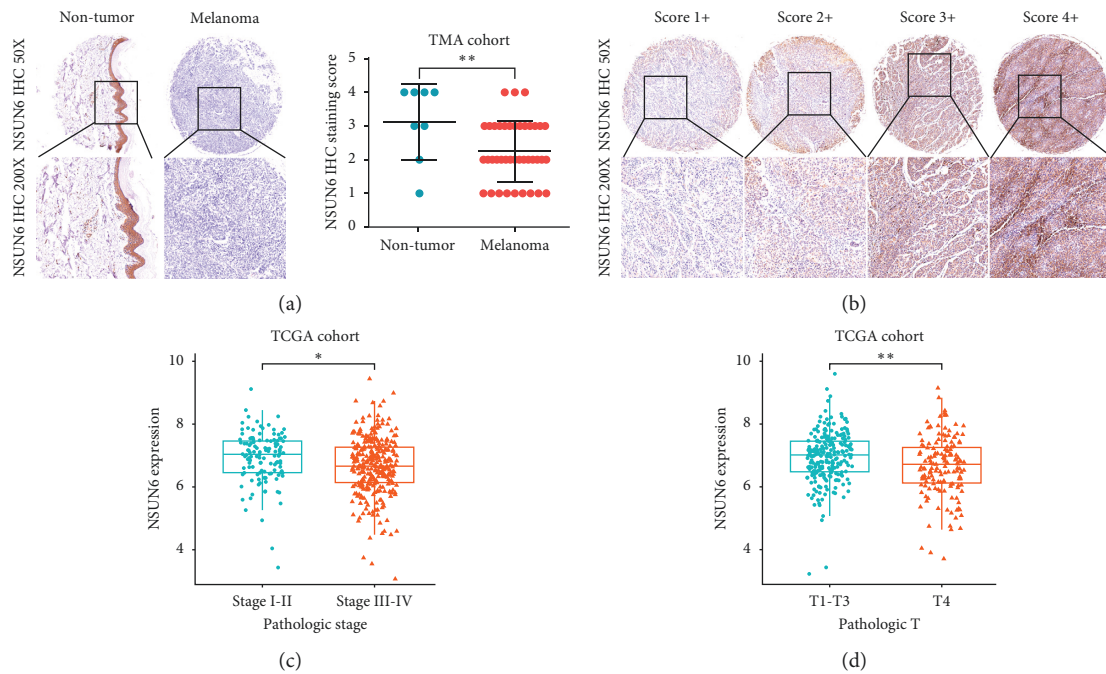


FIGURE 7: Continued.

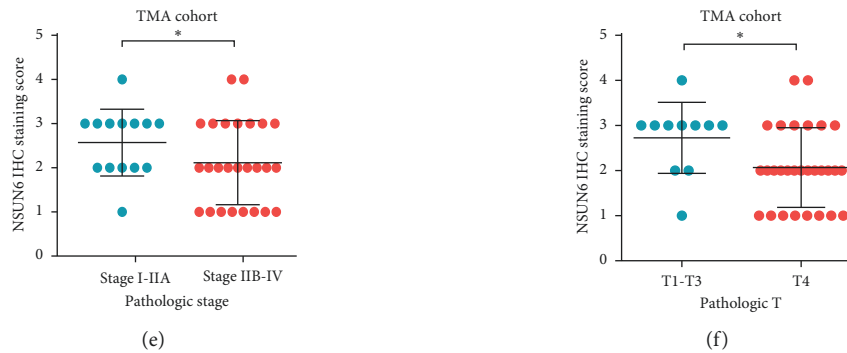


FIGURE 7: Low expression of USUN6 was closely correlated with CM progression. (a) Decreased expression of USUN6 in CM tissues. (b) Representative images of USUN6 staining in CM tissues. (c–f) The correlation between USUN6 expression levels and the pathologic state and TNM staging of CM.

that, of m5C regulators, NOP2 contributed the most to CM, while NSUN6 ranked second. NOP2 (also named NSUN1, p120) has been improved primarily for its protumorigenic roles in many cancers such as prostate cancer, gallbladder cancer, lung adenocarcinoma, and several other cancers, consistent with our study [34, 49, 50]. NSUN6 targets mRNA to transfer methyl and is higher expressed in healthy tissues than in tumors. Previous studies also found that NSUN6 could inhibit pancreatic cancer development [39, 51]. However, the role of NSUN6 in CM has not been documented thus far. To our best knowledge, our study would be a pioneering contribution to the suppressive role of NSUN6 in CM. Finally, using immunohistochemistry analysis, we found that the expression level of NSUN6 was lower in CM tissues than in the corresponding peri-CM tissues. In summary, our studies revealed that NSUN6 was a significant contribution suppressor molecule in CM.

## 5. Conclusions

We explored the relationship between m5C regulators and the progression and prognosis of CM. A risk score model was also established and validated to predict the progression and prognosis of CM patients. Machine learning algorithms indicated that NSUN6 was a significant contribution suppressor molecule in CM. Our results provide a unique approach to the application of novel diagnostic biomarkers and targeted therapy for CM.

## Data Availability

The data used to support the findings of this study are included within the article.

## Disclosure

Maixin Huang, Yi Zhang, and Xiaohong Ou share first authorship.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Authors' Contributions

Maixin Huang, Yi Zhang, and Xiaohong Ou contributed equally to this work.

## Supplementary Materials

Table S1: the characteristics of public databases used in this study. Table S2: demographics of the cohort. Table S3: information on antibodies used in this study. Table S4: clinical characteristics of patients with m5C clusters in the TCGA cohort. Table S5: clinical characteristics of patients with m5C risk score in the TCGA cohort. Table S6: the univariate and multivariate Cox regression analysis of m5C regulators, risk score, and clinical features in the TCGA cohort. Figure S1: Kaplan–Meier analysis of OS of sixteen m5C regulators. Figure S2: Kaplan–Meier analysis of PFS of sixteen m5C regulators. Figure S3: unsupervised consensus analysis of sixteen m5C regulators. (a–e) Consensus clustering matrix for  $k=2$ ,  $k=3$ ,  $k=4$ ,  $k=5$ , and  $k=6$ . (f) Consensus clustering cumulative distribution function for  $k=2$  to 6. (Supplementary Materials). (Supplementary Materials)

## References

- [1] J. Homsí, M. Kashani-Sabet, J. L. Messina, and A. Daud, “Cutaneous melanoma: prognostic factors,” *Cancer Control*, vol. 12, no. 4, pp. 223–229, 2005.
- [2] J. Dowling, S. P. McGregor, and P. Williford, “Update on current treatment recommendations for primary cutaneous melanoma,” *Dermatologic Clinics*, vol. 37, no. 4, pp. 397–407, 2019.
- [3] E. Maverakis, L. Cornelius, G. Bowen et al., “Metastatic melanoma—a review of current and future treatment options,” *Acta Dermato Venereologica*, vol. 95, no. 5, pp. 516–524, 2015.
- [4] L. E. Davis, S. C. Shalin, and A. J. Tackett, “Current state of melanoma diagnosis and treatment,” *Cancer Biology & Therapy*, vol. 20, no. 11, pp. 1366–1379, 2019.
- [5] M. M. F. Aubuchon, L. J. J. Bolt, M. L. G. Janssen-Heijnen, S. T. H. P. Verleisdonk-Bolhaar, A. Van Marion, and C. L. H. Van Berlo, “Epidemiology, management and survival outcomes of primary cutaneous melanoma: a ten-year overview,” *Acta Chirurgica Belgica*, vol. 117, no. 1, pp. 29–35, 2017.

- [6] J. Rockberg, J. M. Amelio, A. Taylor, L. Jørgensen, P. Ragnhammar, and J. Hansson, "Epidemiology of cutaneous melanoma in Sweden-stage-specific survival and rate of recurrence," *International Journal of Cancer*, vol. 139, no. 12, pp. 2722–2729, 2016.
- [7] L. Nájera, M. Alonso-Juarranz, M. Garrido et al., "Prognostic implications of markers of the metabolic phenotype in human cutaneous melanoma," *British Journal of Dermatology*, vol. 181, no. 1, pp. 114–127, 2019.
- [8] D. Grossman, N. Okwundu, E. K. Bartlett et al., "Prognostic gene expression profiling in cutaneous melanoma: identifying the knowledge gaps and assessing the clinical benefit," *JAMA Dermatology*, vol. 156, no. 9, pp. 1004–1011, 2020.
- [9] K. E. Bohnsack, C. Höbartner, and M. T. Bohnsack, "Eukaryotic 5-methylcytosine (m<sup>5</sup>C) RNA methyltransferases: mechanisms, cellular functions, and links to disease," *Genes (Basel)*, vol. 10, no. 2, 2019.
- [10] B. S. Zhao, I. A. Roundtree, and C. He, "Post-transcriptional gene regulation by mRNA modifications," *Nature Reviews Molecular Cell Biology*, vol. 18, no. 1, pp. 31–42, 2017.
- [11] H. Gan, "Progress in epigenetic modification of mRNA and the function of m<sup>6</sup>A modification," *Sheng Wu Gong Cheng Xue Bao*, vol. 35, no. 5, pp. 775–783, 2019.
- [12] P. Nombela, B. Miguel-López, and S. Blanco, "The role of m<sup>6</sup>A, m<sup>5</sup>C and Ψ RNA modifications in cancer: novel therapeutic opportunities," *Molecular Cancer*, vol. 20, no. 1, p. 18, 2021.
- [13] J. M. Bujnicki, "Sequence-structure-function studies of tRNA: m<sup>5</sup>C methyltransferase Trm4p and its relationship to DNA: m<sup>5</sup>C and RNA: m<sup>5</sup>U methyltransferases," *Nucleic Acids Research*, vol. 32, no. 8, pp. 2453–2463, 2004.
- [14] J. Li, "m<sup>6</sup>A regulator-associated modification patterns and immune infiltration of the tumor microenvironment in hepatocarcinoma," *Frontiers in Cell and Developmental Biology*, vol. 9, no. 1679, 2021.
- [15] J. Li, B. Rao, J. Yang et al., "Dysregulated m<sup>6</sup>A-related regulators are associated with tumor metastasis and poor prognosis in osteosarcoma," *Frontiers in Oncology*, vol. 10, p. 769, 2020.
- [16] K. Guan, X. Liu, J. Li et al., "Expression status and prognostic value of m<sup>6</sup>A-associated genes in gastric cancer," *Journal of Cancer*, vol. 11, no. 10, pp. 3027–3040, 2020.
- [17] Y. Zhang, X. Liu, L. Liu, J. Li, Q. Hu, and R. Sun, "Expression and prognostic significance of m<sup>6</sup>A-related genes in lung adenocarcinoma," *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, vol. 26, Article ID e919644, 2020.
- [18] X. Cui, Z. Wang, J. Li et al., "Cross talk between RNA N<sup>6</sup>-methyladenosine methyltransferase-like 3 and miR-186 regulates hepatoblastoma progression through Wnt/β-catenin signalling pathway," *Cell Proliferation*, vol. 53, no. 3, Article ID e12768, 2020.
- [19] Y. S. Chen, "Dynamic transcriptomic m<sup>(5)</sup> C and its regulatory role in RNA processing," *Wiley Interdisciplinary Reviews: RNA*, vol. 12, Article ID e1639, 2021.
- [20] L. Liu, X. Liu, Z. Dong et al., "N<sup>6</sup>-methyladenosine-related genomic targets are altered in breast cancer tissue and associated with poor survival," *Journal of Cancer*, vol. 10, no. 22, pp. 5447–5459, 2019.
- [21] M. D. Metodiev, H. Spähr, P. Loguercio Polosa et al., "NSUN4 is a dual function mitochondrial protein required for both methylation of 12S rRNA and coordination of mitoribosomal assembly," *PLoS Genetics*, vol. 10, no. 2, Article ID e1004110, 2014.
- [22] L. Van Haute, "NSUN2 introduces 5-methylcytosines in mammalian mitochondrial tRNAs," *Nucleic Acids Research*, vol. 47, no. 16, pp. 8720–8733, 2019.
- [23] A. M. Barciszewska, "Global DNA demethylation as an epigenetic marker of human brain metastases," *Bioscience Reports*, vol. 38, no. 5, 2018.
- [24] M. Wnuk, P. Slipek, M. Dziedzic, and A. Lewinska, "The roles of host 5-methylcytosine RNA methyltransferases during viral infections," *International Journal of Molecular Sciences*, vol. 21, no. 21, 2020.
- [25] A. Chellamuthu and S. G. Gray, "The RNA methyltransferase NSUN2 and its potential roles in cancer," *Cells*, vol. 9, no. 8, 2020.
- [26] P. Haruehanroengra, Y. Y. Zheng, Y. Zhou, Y. Huang, and J. Sheng, "RNA modifications and cancer," *RNA Biology*, vol. 17, no. 11, pp. 1560–1575, 2020.
- [27] M. Xue, Q. Shi, L. Zheng, Q. Li, L. Yang, and Y. Zhang, "Gene signatures of m<sup>5</sup>C regulators may predict prognoses of patients with head and neck squamous cell carcinoma," *American Journal of Translational Research*, vol. 12, no. 10, pp. 6841–6852, 2020.
- [28] Z. Dong and H. Cui, "The emerging roles of RNA modifications in glioblastoma," *Cancers (Basel)*, vol. 12, no. 3, 2020.
- [29] J. Duerr, "Immunohistochemistry," *WormBook*, pp. 1–61, 2006.
- [30] J. A. Ramos-Vara, "Technical aspects of immunohistochemistry," *Veterinary Pathology*, vol. 42, no. 4, pp. 405–426, 2005.
- [31] M. Kanehisa and S. Goto, "KEGG: kyoto encyclopedia of genes and genomes," *Nucleic Acids Research*, vol. 28, no. 1, pp. 27–30, 2000.
- [32] R. Alhamzawi and H. T. M. Ali, "The Bayesian adaptive lasso regression," *Mathematical Biosciences*, vol. 303, pp. 75–82, 2018.
- [33] S. P. Kar, J. Beesley, A. Amin Al Olama et al., "Genome-wide meta-analyses of breast, ovarian, and prostate cancer association studies identify multiple new susceptibility loci shared by at least two cancer types," *Cancer Discovery*, vol. 6, no. 9, pp. 1052–1067, 2016.
- [34] F. Wang, J.-H. Yuan, S.-B. Wang et al., "Oncofetal long noncoding RNA PVT1 promotes proliferation and stem cell-like property of hepatocellular carcinoma cells by stabilizing NOP<sub>2</sub>," *Hepatology*, vol. 60, no. 4, pp. 1278–1290, 2014.
- [35] L. Sun, W.-K. Liu, X.-W. Du et al., "Large-scale transcriptome analysis identified RNA methylation regulators as novel prognostic signatures for lung adenocarcinoma," *Annals of Translational Medicine*, vol. 8, no. 12, p. 751, 2020.
- [36] A. K. Jayavelu, T. M. Schnöder, F. Perner et al., "Splicing factor YBX1 mediates persistence of JAK2-mutated neoplasms," *Nature*, vol. 588, no. 7836, pp. 157–163, 2020.
- [37] Y. Ban, P. Tan, J. Cai et al., "LNCAROD is stabilized by m<sup>6</sup>A methylation and promotes cancer progression via forming a ternary complex with HSPA1A and YBX1 in head and neck squamous cell carcinoma," *Molecular Oncology*, vol. 14, no. 6, pp. 1282–1296, 2020.
- [38] E. Zhang, X. He, C. Zhang et al., "A novel long noncoding RNA HOXC-AS3 mediates tumorigenesis of gastric cancer by binding to YBX1," *Genome Biology*, vol. 19, no. 1, p. 154, 2018.
- [39] R. Yang, X. Liang, H. Wang et al., "The RNA methyltransferase NSUN6 suppresses pancreatic cancer development by regulating cell proliferation," *EBioMedicine*, vol. 63, Article ID 103195, 2021.
- [40] P. Wang, M. Wu, Z. Tu et al., "Identification of RNA: 5-methylcytosine methyltransferases-related signature for

- predicting prognosis in glioma,” *Frontiers in Oncology*, vol. 10, p. 1119, 2020.
- [41] J. Deshmukh, R. Pofahl, and I. Haase, “Epidermal Rac1 regulates the DNA damage response and protects from UV-light-induced keratinocyte apoptosis and skin carcinogenesis,” *Cell Death & Disease*, vol. 8, no. 3, Article ID e2664, 2017.
- [42] G. H. Mahabeleshwar and T. V. Byzova, “Angiogenesis in melanoma,” *Seminars in Oncology*, vol. 34, no. 6, pp. 555–565, 2007.
- [43] D. J. M. Smith, “The Melanocortin 1 receptor and its influence on naevi and melanoma in dark-skinned phenotypes,” *Australasian Journal of Dermatology*, vol. 60, no. 3, pp. 192–199, 2019.
- [44] C.-Y. Wei, M.-X. Zhu, Y.-W. Yang et al., “Downregulation of RNF128 activates Wnt/ $\beta$ -catenin signaling to induce cellular EMT and stemness via CD44 and CTTN ubiquitination in melanoma,” *Journal of Hematology & Oncology*, vol. 12, no. 1, p. 21, 2019.
- [45] W. Li, Y. Yin, X. Quan, and H. Zhang, “Gene expression value prediction based on XGBoost algorithm,” *Frontiers in Genetics*, vol. 10, p. 1077, 2019.
- [46] Q. Cheng, J. Li, F. Fan et al., “Identification and analysis of glioblastoma biomarkers based on single cell sequencing,” *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 167, 2020.
- [47] H. Zhang, F. Fan, Y. Yu et al., “Clinical characterization, genetic profiling, and immune infiltration of TOX in diffuse gliomas,” *Journal of Translational Medicine*, vol. 18, no. 1, p. 305, 2020.
- [48] Q. Cheng, C. Huang, H. Cao et al., “A novel prognostic signature of transcription factors for the prediction in patients with GBM,” *Frontiers in Genetics*, vol. 10, p. 906, 2019.
- [49] H. Busch, R. K. Busch, J. W. Freeman, and L. Perlaky, “Nucleolar protein P120 and its targeting for cancer chemotherapy,” *Bollettino della Societa italiana di biologia sperimentale*, vol. 67, no. 8, pp. 739–750, 1991.
- [50] Y. Gao, Z. Wang, Y. Zhu et al., “NOP 2/Sun RNA methyltransferase 2 promotes tumor progression via its interacting partner RPL 6 in gallbladder carcinoma,” *Cancer Science*, vol. 110, no. 11, pp. 3510–3519, 2019.
- [51] Z. Huang, “Prognostic significance and tumor immune microenvironment heterogeneity of m5C RNA methylation regulators in triple-negative breast cancer,” *Frontiers in Cell and Developmental Biology*, vol. 9, Article ID 657547, 2021.