Roles of m5C RNA Modification Patterns in Biochemical Recurrence and Tumor Microenvironment Characterization of Prostate Adenocarcinoma

Zhipeng Xu1,2†, Shuqiu Chen1,2†, Yuxi Zhang3*, Ruiji Liu1,2,4* and Ming Chen1,2,4*

1 Department of Urology, Affiliated Zhongda Hospital of Southeast University, Nanjing, China, 2 Surgical Research Center, Institute of Urology, Southeast University Medical School, Nanjing, China, 3 Department of Radiation Oncology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China, 4 Department of Urology, Zhongda Hospital Lishui Branch, Nanjing, China

Background: Prostate cancer is the second most common cancer with a high risk of biochemical recurrence (BCR) among men. Recently, 5-methylcytosine (m5C) modification has attracted more attention as a new layer of RNA post-transcriptional regulation. Hence, we aimed at investigating the potential roles of m5C modification regulators in the BCR of prostate adenocarcinoma (PRAD).

Methods: CNV data, mutation annotation data, mRNA expression profiles, and clinical data were downloaded from TCGA and GEO databases. Kaplan-Meier curves analysis, log-rank test, univariate and multivariate Cox regression, and time-dependent ROC curves analysis were performed to evaluate the prognostic factors. Principal components analysis (PCA) was applied to validate the distinction between subgroups. Gene set variation analysis (GSVA) was used to investigate the underlying pathways associated with m5C modification patterns. Single sample gene set enrichment analysis (ssGSEA) was utilized to assess the infiltration of distinct immune cells. Tumor Immune Dysfunction and Exclusion (TIDE) prediction was carried out to assess the potential response to immune checkpoint blockade (ICB) therapy. The m5C modification signature was constructed via LASSO Cox’s proportional hazards regression method.

Results: After comprehensively analyzing various types of data from TCGA dataset, and exploring the differential expression and prognostic value of each m5C regulator, we identified m5C modification patterns based on 17 m5C regulators. Two patterns presented a significant difference in the risk of BCR, the tumor microenvironment (TME), and immunotherapy response in PRAD. We found that TET2, which was highly expressed in adjacent normal tissues compared to tumor tissues, was closely associated with many infiltrating immune cells. The m5C modification signature was constructed for the clinical application. Risk score calculated by m5C signature was associated with T stage, N stage, Gleason score, and the possibility of BCR (HR, 4.197; 95% CI, 3.016-
BACKGROUND

Since the 1960s, RNA modifications have been well documented as the key factor in regulating the RNA metabolism under various biological processes (1, 2). Many researchers have recently focused on revealing the specific molecular functions of RNA modifications such as N6-methyladenosine (m6A) (3). With the in-depth research on RNA modifications, 5-methylcytosine (m5C), which was rarely studied, has attracted more attention. m5C has been demonstrated as a widespread modification throughout the coding and non-coding RNA (4, 5). In one report, the m5C modification in the 3'UTR of heparin binding growth factor (HDGF) mRNA can promote the stability of HDGF mRNA, which contributes to bladder cancer progression (6). In another report, Aly/REF export factor (ALYREF) has been identified as a m5C modification binding protein (7). RNA m5C modification is dynamically modulated by relevant proteins in human cells (8). m5C writers are defined as the enzymes contributing to the formation of cytosine-5-methylation, which includes the NSUNs family of proteins and the DNA methyltransferase (DNMTs) family members (8, 9). RNA m5C modification demethylases are named as m5C erasers, including the ten-eleven translocator (TETs) family members and Alpha-Ketoglutarate-Dependent Dioxygenase AlkB Homolog 1 (ALKBH1) (10). Moreover, m5C readers, including ALYREF and Y-box binding protein 1 (YBX1), can recognize the mRNA with m5C modification and take part in the biological function of RNA m5C modification (8, 11). In 2017, Xin et al. validated that the m5C reader, ALYREF, regulated mRNA export via a m5C-dependent manner (7). Recently, some researchers have also found that ALYREF could stabilize pyruvate kinase muscle isozyme M2 (PKM2) mRNA via a m5C-dependent manner in bladder cancer (12). A previous study found that YBX1 stabilized its target mRNAs via an m5C-dependent manner (13). Similarly, it has also been reported that YBX1 could stabilize its target mRNAs by recognizing the m5C methylation site in the HDGF mRNA and by recruiting ELAV like RNA binding protein 1 (ELAVL1) in bladder cancer (6). Previous studies show the need for exploring the relationship between the m5C modification and various biological processes.

It is estimated that prostate cancer is the second most frequent cancer and in 2020 was the fifth leading global cause of cancer death in men (14). Among several histological subtypes of prostate cancer, prostate adenocarcinoma (PRAD) is the most common one (15). With the development of prostate cancer detection and management, patients with prostate cancer at an early stage exhibit a 99% ten-year overall survival (OS) rate (16), whereas more than half of patients undergo biochemical recurrence (BCR) characterized by elevated prostate-specific antigen (PSA) levels after radical therapy (17, 18). It is reported that approximately 35% of patients will experience BCR within 10 years after radical prostatectomy (19, 20). Recently, more and more studies have revealed that the tumor microenvironment (TME), including various immune and non-immune factors, plays an essential role in prostate cancer tumorigenesis and progression (21–24). Exploring the related mechanisms of prostate cancer BCR has become clinically significant. Bora et al. have reported a positive correlation between chronic inflammation and prostate cancer grade (25). Tumor-associated macrophage infiltration can act in prostate cancer as a predictive factor for BCR (26). Some studies have found that regulatory T-cells infiltration is associated with a worse BCR-free survival rate (27, 28). Similarly, it has been demonstrated that higher B-cells infiltration in prostate cancer, and positively correlates with tumor grade and BCR (29). Focusing on the changes in TME may thus conduce to identifying biomarkers for the diagnosis and treatment of prostate cancer.

In this study, we have aimed at clarifying the potential roles of m5C modification regulators inPRAD. By comprehensively analyzing the copy number alteration data, mutation annotation data, and gene expression profiles from The Cancer Genome Atlas (TCGA) database, we could identify the potential

Keywords: m5C RNA modification, biochemical recurrence, tumor microenvironment, immunotherapy, prostate adenocarcinoma (PRAD)
pathways related to m5C modification. Based on this, we further explored the association between m5C modification and TME. Subsequently, we constructed an m5C-related signature for assessing the risk of BCR in PRAD and revealed its underlying connection with immunotherapy. This signature was also validated in the external testing dataset (GSE70770). Our results may provide new sights for further research on m5C modification and personalized management of PRAD.

METHODS

Data Source
In this study, the datasets were obtained from the TCGA and Gene Expression Omnibus (GEO) databases. We downloaded the copy number variation (CNV) data, mRNA expression data, and clinical data from the TCGA website (https://portal.gdc.cancer.gov/). For the mRNA expression profiles, we converted the FPKM value into the TPM value for further analysis, as it is identical to the microarray values (30, 31). The mRNA expression profiles consisted of 540 samples, including 489 tumor samples and 51 normal samples of which the clinical information was extracted, and BCR was taken as the positive event. After having excluded two samples with uncertain prognostic information, three samples with M1 stage, and four repeated samples, we ultimately included 480 tumor samples and 51 normal samples in the present study. To construct an external testing dataset, we searched the datasets related to PRAD BCR in the GEO database. The GSE70770 dataset, including 200 tumor samples, was eventually selected for its perceived accuracy. The expression matrix and clinical information of GSE70770 dataset were then downloaded and applied to validate the prognostic value of m5C modification signature.

Analysis of the CNV Data and Mutation Annotation Data
We determined the CNV by using the GISTIC 2.0 software developed by Craig et al. (32). Then, the CNVs of each m5C regulator, including amplification and deletion, were visualized with the R package 'RCircos'. Mutation annotation format (MAF) of somatic mutation data of TCGA PRAD cohort was applied for further analysis with the R package 'maftools' (33).

Gene Expression Analysis of m5C Regulators and Survival Analysis
According to the relevant literatures, we extracted 17 m5C modification regulators, including 11 writers (NOP2, NSUN2, NSUN3, NSUN4, NSUN5, NSUN6, NSUN7, DNMT1, DNMT3A, DNMT3B, and TRDMT1), four erasers (TET1, TET2, TET3, and ALKBH1) and two readers (ALYREF and YBX1) (8–11). The ‘limma’ package was used to compare the m5C regulators’ expression among different groups. To evaluate the prognostic value, we used the related methods, including Kaplan-Meier curves analysis, log-rank test, univariate and multivariate Cox regression, and time-dependent ROC curves analysis, by the R packages ‘survival’, ‘survminer’, and ‘survivalROC’.

Identification of m5C Modification Patterns and Molecular Characteristics
The ‘ConsensusClusterPlus’ package was applied to identify the distinct m5C modification patterns in the TCGA-PRAD cohort. According to the relative change in area under the CDF curve and the internal consistency of distinct k values, we determined the number of patterns. In order to assure the stability of the cluster, the step was repeated 1000 times (34). The principal components analysis (PCA) was applied to validate the distinction between different m5C modification patterns with the R package ‘ pca3d’ and ‘rgl’ (35, 36).

The gene set variation analysis (GSVA) was performed to investigate the underlying pathways involved in m5C modification using the R package ‘GSEABase’, ‘GSVA’, and ‘qheatmap’ (37). In this study, we downloaded the ‘c2.cp.kegg.v7.2.symbols.gmt’ gene set as the background from the MSigDB database. Differentially enriched pathways were screened out by using the R package ‘limma’ with a threshold of adjusted p-value < 0.05 and $|\log2(Fold \ change)| > 1$.

Evaluation of the Abundance of Various TME Cells
To evaluate the tumor component in each sample, we applied the ESTIMATE algorithm to calculate the score of tumor purity, stromal, immune, and ESTIMATE score in each sample, and the score between different subgroups was further analyzed with R package ‘limma’. The single sample gene set enrichment analysis (ssGSEA) was used to assess the infiltration of distinct immune cells. According to the research of Charoentong et al., the gene set signature, including 18 types of adaptive immune cells and 10 types of innate immune cells, was applied to the ssGSEA analysis (38). This step was conducted and visualized with the R package ‘GSEABase’, ‘GSVA’, ‘ggpubr’, ‘reshape2’, and ‘limma’. The estimated score that was generated in the ssGSEA analysis represented the infiltration of each type of immune cell in each sample.

Assessing the Differential Expression of Immune Checkpoints, the Response for Immunotherapy, and Drug Sensitivity Analysis
In order to analyze the expression of immune checkpoints in distinct cohorts, we extracted 15 immune checkpoint molecules (GZMA, LAG3, IDO1, CXCL10, CD274, HAVCR2, IFNG, CD8A, TNF, CTLA4, PRF1, PDCD1, GZMB, TBX2, and CXCL9) according to the study of Peng et al. (39). Jiang et al. developed the Tumor Immune Dysfunction and Exclusion (TIDE), a computational method that modeled the tumor immune evasion mechanism, to assess the potential response to immune checkpoint blockade (ICB) therapy (40). TIDE prediction was performed on the website: http://tide.dfci.harvard.edu/. Different expression of immune checkpoint molecules, different microsatellite instability (MSI) score, and
different TIDE score were determined with the R package 'limma'. Moreover, the R package 'pRRophetic' was used to analyze the expression profile of TCGA-PRAD cohort in the drug sensitivity analysis (41).

Construction and Evaluation of the m5C Modification Signature
To construct an m5C modification signature with better prognostic values, we focused on differentially expressed genes (DEGs) between two m5C modification patterns. The DEGs were filtered with a threshold of adjusted p-value < 0.001 and $|\log2(\text{Fold change})| > 1$. Univariate and multivariate COX analysis was used to further select the candidate genes with an independent prognostic value from the DEGs. We then employed the least absolute shrinkage and selection operator (LASSO) Cox’s proportional hazards regression method to screen out the candidate genes with the R package ‘glmnet’ and ‘survival’ (42). The proper genes for the m5C modification signature were selected with the lowest Akaike’s information criterion (AIC) (43). Eventually, the risk score base on the m5C distribution data, respectively. In three or more group analyze the expression profile of TCGA-PRAD cohort in the drug sensitivity analysis (41).

RESULTS
Genetic Variation of m5C Regulators in PRAD
In this study, we evaluated 17 m5C regulators including 11 writers, four erasers, and two readers. The landscape of CNV and somatic mutations and m5C regulators was analyzed in PRAD. CNV was determined as amplification (segment mean > 0.2), diploid (-0.2 < segment mean < 0.2), and deletion (segment mean < -0.2). As shown in Figure 1A, ALYREF, NSUN5, DNMT3A, and NSUN2 had a higher frequency of copy number deletion, while a higher frequency of copy number amplification was identified in the other m5C regulators. Figure 1B shows the location of CNV of 17 m5C regulators on chromosomes in PRAD. Then, we explored the relationship between CNV incidence and mRNA expression of m5C regulators, which verified that CNV played an essential role in regulating the expression of m5C regulators. As shown in Figure 1C and Supplementary Figure S1A, CNV amplification was significantly associated with a higher gene expression of ALYREF, DNMT3A, DNMT3B, NSUN2, NSUN5, and TET1. In turn, CNV deletion was significantly associated with a lower gene expression among most m5C regulators. Additionally, the mutations of m5C regulators were identified in only 13 samples among 484 samples in the TCGA-PRAD cohort (Supplementary Figure S1B). Considering the frequency of CNV and somatic mutations of m5C regulators was overall low, this indicated that m5C regulators could be conserved in PRAD. The analysis of the gene expression in PRAD tissue and normal prostate tissue validated that most m5C writers and readers were highly expressed in PRAD, while the expression of erasers except TET1 and TET3 was significantly lower in PRAD (Figure 2A). Besides, the PCA analysis was performed and visualized by the prCOMP and pcd3D algorithm. As shown in Supplementary Figure S1C, the 17 m5C regulators could make a clear distinction between the normal and tumor samples in the TCGA-PRAD cohort. The results above indicated that the different expression profiles and genetic variation landscape of m5C regulators could play an important role in PRAD tumorigenesis and progression.

Evaluation of m5C Modification Patterns Based on 17 Regulators
Kaplan-Meier survival analysis verified that the PRAD patients with a higher expression of NSUN2, NSUN5, NSUN6, NOP2, DNMT1, DNMT3A, DNMT3B, ALYREF, YBX1, TET2, or TET3 could show higher BCR incidence, while higher expression of TRD1T1, TET1 and ALKBH1 could be a good prognostic factor

Clinical Samples and Immunohistochemical (IHC) Staining
The four pairs of PRAD specimens and matched adjacent normal specimens used in our study were collected from The Affiliated Zhongda Hospital of Southeast University, China. Prior informed consent was obtained from all of the patients. This study was approved by the Ethical Committee of the Affiliated Zhongda Hospital of Southeast University for tissue specimens used for research purposes. The 4 μm paraffin-embedded tissue sections were used in IHC staining. The sections were incubated with primary antibodies at 4°C overnight and secondary antibodies conjugated with horseradish peroxidase at room temperature for half an hour. Next, the sections were stained with DAB solution for 10 min and counterstained with hematoxylin for 1 min. Images of the stained sections were obtained with a microscope (Nikon, Tokyo, Japan). The clinical index of four PRAD patients was displayed in Supplemental Table S1.

Statistical Analysis
Rstudio 1.4 and R for windows 4.1.0 software were used to perform the statistical analysis. We applied the Student’s t-test or Wilcoxon test in a two-group comparison of normally or skewed distribution data, respectively. In three or more group comparisons, we used the Kruskal–Wallis test or one-way ANNOVA for parametric or non-parametric comparison purposes. Fisher’s exact test was applied to determine the component differences in subgroups. All statistical tests were two-sided, and p-value < 0.05 was considered statistically significant. ****, p-value < 0.0001; ***, p-value < 0.001; **, p-value < 0.01; *, p-value < 0.05; ns, no significant difference.
for BCR in PRAD (Supplementary Figures S2, S3). A univariate Cox regression model was also utilized to evaluate the predictive values of m5C regulators for BCR in TCGA-PRAD cohort (Supplementary Table S2). The gene network of m5C regulators depicted their interaction, correlation of mRNA expression, and prognostic values for PRAD patients (Figure 2B). These data revealed that most m5C-related genes were the risk factors for BCR in TCGA-PRAD cohort. Unsupervised clustering was performed with ‘ConsensusClusterPlus’ R package to classify patients based on the expression of 17 m5C regulators. Two m5C modification patterns were finally identified, with 234 and 246 patients in patterns A and B, respectively (Figures 2C, D). The PCA analysis was utilized to validate the discrimination between two m5C modification patterns (Figure 2E). As shown in Supplementary Figure S4A, m5C modification pattern A presented higher expression of m5C writers and readers. The survival analysis on BCR indicated that pattern B showed an advantage in non-BCR survival time (Figure 2F).

Then, GSVA enrichment analysis was used to reveal the distinct biological process and potential pathways between two m5C modification patterns (Figure 2G; Supplementary Table S3). Compared to m5C modification pattern B, pattern A was significantly enriched in some pathways associated with the metabolism and function of RNA, such as ‘kegg.RNA.degradation’, ‘kegg.RNA.polymerase’, and ‘kegg.aminoacyl.tRNA.biosynthesis’, which further validated the influence of distinct m5C modification on RNA. We also found that pattern A presented the enrichment of some carcinogenic pathways including p53 signaling pathway, cell cycle, TGF-β signaling pathway, NOTCH signaling pathway, WNT signaling pathway, mTOR signaling pathway, VEGF signaling pathway, MAPK signaling pathway, ERBB signaling pathway.
pathway, and JAK-STAT signaling pathway. The activation of these pathways could be the reason for the shorter non-BCR survival time in m5C modification pattern A. Besides, it is interesting that pattern A was also markedly associated with immune-related pathways including antigen processing and presentation, chemokine signaling pathway, Toll-like receptor signaling pathway, T cell receptor signaling pathway, B cell receptor signaling pathway, and natural killer cell-mediated cytotoxicity.

**TME Cell Infiltration Characteristics of m5C Modification Patterns**

The difference between two m5C modification patterns in immune-related pathways prompted the analysis of the TME cell infiltration. The evaluation based on the ESTIMATE algorithm revealed lower tumor purity, higher immune scores, and higher stromal scores of pattern A (Figure 3A). We used the ssGSEA algorithm to evaluate the abundance of diverse immune cells in the samples, and the results indicated that pattern A presented a higher abundance of almost all types of immune cells (Figure 3B). To explore the potential reason for the contradiction of higher BCR probability and more immune cells infiltration in pattern A, we focused on immune checkpoint pathways which were the inhibitory pathways hardwired into the immune system (44). Now, it has been widely recognized that immune checkpoints were the major mechanism of immune resistance, especially against T cells specific for tumor antigens (45). Hence, we estimated the expression of immune checkpoints in two m5C modification patterns. As shown in Figure 4A, immune checkpoints including GZMA, IDO1, CXCL10, HAVCR2, CD8A, CTLA4, PRF1, PDCD1, GZMB, and CXCL9 were higher expressed in pattern A than in pattern B, which indicated that the patients in pattern A could benefit from ICB therapy. Subsequently, we evaluated the immune checkpoints inhibitors (ICI) response by the tumor immune dysfunction and exclusion (TIDE) algorithm. As speculated, immunotherapy was more likely to be effective in patients in pattern A than pattern B according to the TIDE score (Figure 4B). There was no significant difference, however, in the MSI score between the two m5C modification patterns (Figure 4C). The analysis firmly supported that distinct m5C modification patterns were closely associated with TME and patients’ benefits from ICB therapy. In addition, the pPRerophetic algorithm was applied to evaluate the drug sensitivity. Among the drugs identified by R package ‘pPRerophetic’, docetaxel was the only drug recommended for clinical treatment of prostate cancer (46). Docetaxel is a kind of tubulin-binding taxanes (47). In 2004, two randomized controlled trials validated that docetaxel was the first chemotherapeutic agent with survival benefits in patients with metastatic castration-resistant prostate cancer (48, 49). As shown in Supplementary Figure S4B, the IC50 value of docetaxel was higher in pattern A than in pattern B, which prompted a higher drug sensitivity of docetaxel in pattern B.

**The m5C Eraser TET2 Was Closely Associated With Infiltrating Immune Cells**

We performed Spearman’s correlation analysis via TIMER algorithm (http://timer.cistrome.org/) to reveal the relationship between TME and each m5C regulator (Supplementary Figures S5, S6) (50). As shown in Figure 5A, the m5C eraser TET2 showed a significant correlation with many immune cells. IHC staining was applied to examine the expression of TET2 in four pairs of PRAD samples and adjacent normal samples. The expression of TET2 was higher in normal prostate epithelial cells than in PRAD cells, which was consistent with our results based on TCGA-PRAD cohort (Figure 5B).

**Construction of m5C Modification Signature**

According to the population-based classification in the TCGA-PRAD cohort, our studies have already validated the prognostic value of m5C modification and the relationship between m5C modification and TME in PRAD. In order to promote the clinical application of m5C modification patterns, we focused on constructing an m5C modification signature for individual tumor cases. We screened out 4877 DEGs (log2(Fold change) > 1 and FDR < 0.001) between two m5C modification patterns (Supplementary Table S4). Univariate COX analysis identified 1230 prognosis-related genes out of 4877 genes (Supplementary Table S5). Multivariate COX analysis further identified 248 genes with the independent prognostic value as the candidate genes for constructing the m5C modification signature (Supplementary Table S6). The LASSO Cox regression algorithm was applied to these candidate genes in the TCGA-PRAD cohort. Eventually, 33 genes were identified based on the minimum criteria to construct the m5C modification signature (Figure 6A and Supplementary Table S7). The function for calculating the risk score was built according to the coefficients of 33 genes generated by LASSO. As shown in Supplementary Table S8, the risk score of each sample in the TCGA-PRAD cohort was calculated, and patients in the TCGA-PRAD cohort were divided into two risk groups according to the median risk score.

Subsequently, the Kaplan-Meier curve analysis and the log-rank test assessed the difference in non-BCR survival time between the two groups (Figure 6B). Univariate and multivariate COX analysis indicated the m5C modification signature as the independent prognostic factor for BCR in the TCGA-PRAD cohort (Figures 6C, D). The results of the time-dependent ROC curves analysis showed that the AUC for m5C modification signature was 0.910, 0.916, and 0.886 for one-, three-, and five-year non-BCR survival time in the TCGA-PRAD cohort (Figure 6E). These results verified that the m5C modification signature was a robust and independent prognostic factor, and this signature presented higher predictive accuracy for BCR than age, pathologic T stage, pathologic N stage, and Gleason score in PRAD. We further validated the association between m5C modification patterns and the signature. As shown in Figure 6F, pattern A exhibited a significantly higher risk score than pattern B. Considering that pattern A was associated
FIGURE 2 | Evaluating the m5C modification patterns based on 17 m5C regulators. (A) Gene expression of m5C regulators in PRAD tissue and normal prostate tissue. **p-value < 0.01; *p-value < 0.05; ns, no significant difference. (B) Gene network of m5C regulators. (C) M5C modification patterns identified with K-means clustering based on m5C regulators. (D) The CDF curve of the clustering result. (E) PCA analysis on the distinction between two m5C modification patterns. (F) Kaplan-Meier survival analysis for two m5C modification patterns based on the BCR data of TCGA-PRAD cohort. (G) GSVA enrichment analysis determines the distinct biological process and potential pathways between two m5C modification patterns. Red, activated pathways; blue, inhibited pathways.
with shorter non-BCR survival time, the results again verified the prognostic value of the m5C modification signature.

We assessed the association between the m5C modification signature and the clinical characteristics in the TCGA-PRAD cohorts. The results demonstrated that elderly patients, higher Gleason score, advanced T stage, and advanced N stage were significantly associated with a higher risk score (Figures 6G, H, 7A, B). We performed the Kaplan-Meier curve analysis in the subgroups with different clinical characteristics to further evaluate the prognostic value of the m5C modification signature in the distinct population. As shown in Figures 7C–F, 8A–D, the m5C modification signature exhibited a prognostic power in various subgroups except for the patients with N1 stage. In the TCGA-PRAD cohort, there were only 76 patients with N1 stage, which could be the reason for the failure of the m5C modification signature in the N1 stage subgroup. These results demonstrated that the m5C modification signature had the potential to act as a biomarker for assessing the clinical characteristics and predicting the BCR in PRAD patients.

**Correlation of m5C Modification Signature With TME and Immunotherapy**

In order to investigate the TME characteristics of the m5C modification signature, we determined the association between the m5C modification signature and the infiltration of diverse immune cells (Figure 8E). The risk score based on the m5C modification signature was positively associated with the abundance of most immune cells. Subsequently, we analyzed the expression profiles of immune checkpoints in the high- and low-risk score groups. As
shown in Figure 9A, the expression of all immune checkpoints was markedly higher in the high-risk group. However, we found that there was no significant difference in the TIDE score and MSI score between the high- and low-risk score groups (Figures 9B, C).

**Validation of m5C Modification Signature in the Testing Dataset**

To further assess the stability of the m5C modification signature, we applied the constructed signature to the other independent PRAD cohorts. The relevant datasets were searched in the GEO database. Eventually, the GSE70770 dataset was chosen as the testing dataset due to its accurate clinical information. According to the function for calculating the risk score, we acquired the risk score for each patient in the testing dataset (Supplementary Table S9). By taking the median risk score as the cutoff point, the patients in the testing dataset were divided into two groups. As shown in Figure 9D, the patients with higher risk scores were significantly associated with shorter non-BCR survival time. The results of univariate and multivariate COX analysis indicated that the m5C modification signature also had the independent prognostic value for BCR in the testing dataset (Figures 9E, F). Furthermore, we performed the analysis of the correlation between the m5C modification signature and clinical characteristics, and Figures 10A, B indicated that higher risk score based on the m5C modification signature was closely associated with higher Gleason score and advanced T stage. Moreover, we further evaluated the predictive power of the m5C modification signature in patients with specific clinical characteristics. As shown in Figures 10C–F a higher risk score was significantly associated with shorter non-BCR survival time in patients with Gleason score ≤ 7, and the consistent result was validated in patients with T3 or T4 stage.

Considering the correlation between m5C modification and tumor immune in the TCGA-PRAD cohort, we also analyzed this correlation in the testing dataset. The ssGSEA algorithm was applied to assess immune cells infiltration, and the results validated that the risk score was significantly associated with the abundance of several types of immune cells including natural killer cells, macrophages, monocytes, and so on (Figure 11A). Among immune checkpoints, HAVCR2, LAG3, PRF1, and
TBX2 presented significantly higher expression in the high-risk group than in the low-risk group (Figure 11B). These data indicated that the correlation between m5C modification and tumor immune was not less significant in the testing dataset than in the TCGA-cohort.

**DISCUSSION**

Owing to the advance in detection technologies, accumulating research has verified that m5C modification played a key role in the post-transcriptional modification of gene expression, which was closely related to tumor formation, maintenance, and progression (10, 51, 52). As reported, the m5C writer NSUN2 was proven to stabilize the GRB2 mRNA via m5C dependent manner in esophageal cancer (53). Some studies have validated that m5C modification played an important role in bladder cancer progression via modulating the stabilization of mRNA (6, 12). Moreover, increasing evidence has recently proven that RNA modification regulators had the potential to act as the biomarkers for diagnosis and prognostic monitoring in cancer (54–56). The m5C reader ALYREF expression was dysregulated in several cancers, and higher ALYREF expression was associated with poor prognosis in hepatocellular carcinoma (57). Some studies indicated that m5C writers NSUN3 and NSUN4 were associated with the infiltration of some immune cells, such as CD8+ T cells and neutrophils (58). However, there remained a lack of comprehensive analysis of the prognostic value and functional annotation of m5C regulators in prostate cancer.

In this study, the differential expression of 17 m5C regulators between the normal tissue and tumor tissue was determined. The m5C writers and readers dominantly presented higher expression, while most m5C erasers presented lower expression. Survival analysis verified that m5C writers and readers were
FIGURE 6 | Establishment of the m5C modification signature. (A) LASSO coefficient profiles and cross-validation for tuning the parameter selection in the LASSO analysis. (B) Kaplan-Meier survival analysis on the association between the BCR and different m5C modification patterns in TCGA-PRAD cohort. (C) Univariate COX analysis on the prognostic value of risk score. (D) Multivariate COX analysis on the prognostic value of risk score. (E) Time-dependent ROC curves analysis for evaluating the prognostic value of risk score in TCGA-PRAD cohort. (F) Comparison of the risk score between the m5C modification patterns in TCGA-PRAD cohort. (G) The analysis on the association between the risk score and age in TCGA-PRAD cohort. (H) The analysis on the association between the risk score and pathologic T stage in TCGA-PRAD cohort.
predominantly associated with the early BCR in PRAD. These data prompted that the activation of m5C modification may promote the tumorigenesis and progression of PRAD. Two m5C modification patterns based on the 17 m5C regulators were further identified. Interestingly, the two patterns exhibited distinct TME characteristics. Pattern A with higher BCR risk was associated with a higher abundance of immune cells, higher ESTIMATE score, and lower tumor purity. Considering the contradiction between poor prognosis and higher immune cells infiltration, we assumed that the immune checkpoints may contribute to blocking the anti-tumor immune response in pattern A. As assumed, the immune checkpoints showed higher expression in pattern A. Lower TIDE score also suggested the potential response to ICB therapy in pattern A. Besides, pattern A presented higher expression of m5C writers and readers. Hence, although pattern A with higher m5C modification had a worse prognosis, the anti-tumor immunity in the cases in pattern A or with higher m5C modification may be restored after ICB therapy.

FIGURE 7 | Evaluation of the m5C modification signature in TCGA-PRAD cohort. (A) The analysis on the association between the risk score and pathologic N stage in TCGA-PRAD cohort. (B) The analysis on the association between the risk score and Gleason score in TCGA-PRAD cohort. (C-F) Kaplan-Meier curve analysis in the subgroups with different clinical characteristics. (C) patients with age <=60; (D) patients with age >60; (E) patients with T2 stage; (F) patients with T3 or T4 stage.
FIGURE 8 | Evaluation of the m5C modification signature in TCGA-PRAD cohort. (A–D) Kaplan-Meier curve analysis in the subgroups with different clinical characteristics. (A) patients with N0 stage; (B) patients with N1 stage; (C) patients with Gleason score ≤7; (D) patients with Gleason score >7. (E) Correlations between risk score and immune cell infiltration in TCGA-PRAD cohort via Spearman analysis. A negative correlation was marked with blue and positive correlation with red. *p-value < 0.05; ns, no significant difference.
FIGURE 9 | Assessing the immunotherapy features in different risk score groups based on the m5C modification signature. (A) The expression of immune checkpoints in high and low-risk groups. (B, C) The predicted immunotherapy response rate of high and low-risk groups is based on TIDE analysis. (B) TIDE score; (C) MSI score. (D–F) Survival analysis was performed to evaluate the prognostic value of risk score in the testing cohort. (D) Kaplan-Meier survival analysis; (E) Univariate COX analysis; (F) Multivariate COX analysis. ***p-value < 0.001; **p-value < 0.01; *p-value < 0.05; ns, no significant difference.
The data above suggest a potential correlation between m5C modification and immunotherapy, which was clinically significant in the era of immunotherapy and remained to be further validated through experimental research.

Herein, we tried to promote the clinical application of m5C modification. The m5C modification signature was established as a risk score model. Survival analysis verified the robust and independent prognostic value of the m5C signature in both the training dataset (TCGA-PRAD cohort) and testing dataset (GSE70770 cohort). Our study also revealed that the risk score based on m5C signature was closely associated with main clinical characteristics, such as pathologic T stage, pathologic N stage, and Gleason score. Besides, our study verified that a higher risk score calculated by the m5C signature also suggested a higher abundance of TME cells and higher immune checkpoints expression. However, no significant correlation was determined between TIDE score and risk score. Despite this, these results still hinted at the clinical value of the m5C signature in evaluating individual prognosis and TME cells characteristics of patients with PRAD.

However, some limitations are also worth noting. Firstly, the m5C signature did not exhibit a valid predictive value in TIDE score, while a significant difference was verified between two m5C modification patterns. The unexpected results may be attributed to the LASSO Cox regression which mainly preserved prognostic features for constructing the m5C signature. And the immune-related DEGs may eventually not be included in the construction of m5C signature via LASSO Cox’s regression method. This disadvantage limited the application of m5C signature in predicting the individual response to immunotherapy. Secondly, the clinicopathological
variables, especially immunotherapy-related data, obtained from the public database were not as comprehensive as those involved in clinical practice. Other potential values of m5C modification remain to be further studied, which can further reduce bias in the performance of m5C score signature. Thirdly, the findings in this study were based on bioinformatics research and have not been validated in relevant experiments. In the future, we will conduct in vitro and in vivo experimental verification of these findings.

FIGURE 11 | The immune features of m5C modification signature in the testing dataset. (A) Correlations between risk score and immune cell infiltration in the testing dataset via Spearman analysis. A negative correlation was marked with blue and a positive correlation with red. (B) The expression of immune checkpoints in high and low-risk groups. **p-value < 0.01; *p-value < 0.05; ns, no significant difference.
CONCLUSIONS

Herein, we revealed the roles of m5C modification patterns in the PRAD BCR and TME diversity. The m5C signature based on distinct modification patterns had the robust and independent prognostic power for PRAD BCR, and could be used to predict the individual response to immune therapy. Our comprehensive analysis of m5C modification can provide new insight into the field of PRAD research, and contribute to the understanding of TME and immunotherapy in the future.

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 authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The Ethical committee of the Affiliated Zhongda Hospital of Southeast University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Conceptualization, ZX and RL; methodology, ZX; software, YZ; validation, ZX, RL, and YZ; formal analysis, ZX and SC; resources, MC; data curation, ZX, RL, and SC; writing—original draft preparation, ZX; writing—review and editing, SC; visualization, ZX, RL, and MC; supervision, MC; funding acquisition, MC. All authors have read and agreed to the published version of the manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (81872089, 81871157, 82070773, 82102831), The National Key Research and Development Program of China (SQ2017YFSF090096), Jiangsu Provincial Key Research and Development Program (BE2019751), Innovative Team of Jiangsu Provincial (2017XKJQW07) and Natural Science Foundation of Jiangsu Province (BK20201271).

ACKNOWLEDGMENTS

The authors would like to acknowledge to the other colleagues for reviewing the manuscript. The authors also acknowledge to the researchers and staff of TCGA and GEO public databases for providing data on prostate cancer.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2022.869759/full#supplementary-material

REFERENCES

1. Cohn WE. Pseudouridine, a Carbon-Carbon Linked Ribonucleoside in Ribonucleic Acids: Isolation, Structure, and Chemical Characteristics. J Biol Chem (1960) 235:1488–98. doi: 10.1016/S0021-9258(18)94323-3
2. Roundtree IA, Evans ME, Pan T, He C. Dynamic RNA Modifications in Gene Expression Regulation. Cell (2017) 169:1187–200. doi: 10.1016/j.cell.2017.05.045
3. Huang W, Chen TQ, Fang K, Zeng ZC, Ye H, Chen YQ. N6-Methyladenosine Methyltransferases: Functions, Regulation, and Clinical Potential. J Hematol Oncol (2021) 14:117. doi: 10.1186/s13045-021-01129-8
4. Khoddami V, Cairns BR. Identification of Direct Targets and Modified Bases of RNA Cytosine Methyltransferases. Nat Biotechnol (2013) 31:458–64. doi: 10.1038/nbt.2566
5. Squires JE, Patel HR, Nouch M, Sibbritt T, Humphreys DT, Parker BJ, et al. Widespread Occurrence of 5-Methylcytosine in Human Coding and Non-Coding RNA. Nucleic Acids Res (2012) 40:5023–33. doi: 10.1093/nar/gks144
6. Chen X, Li A, Sun BF, Yang Y, Han YN, Yuan X, et al. 5-Methylcytosine Promotes Pathogenesis of Bladder Cancer Through Stabilizing mRNAs. Nat Cell Biol (2019) 21:978–90. doi: 10.1038/s41556-019-0361-y
7. Yang X, Yang Y, Sun BF, Chen YS, Xu JW, Lai WY, et al. 5-Methylcytosine Promotes mRNA Export - NSUN2 as the Methyltransferase and ALYREF as an M(5)C Reader. Cell Res (2017) 27:666–25. doi: 10.1038/cr.2017.55
26. Nonomura N, Takayama H, Nakayama M, Nakai Y, Kawashima A, Mukai M, et al. m5C RNA Modifiers in Metastatic Castration-Resistant Prostate Cancer Patients: Immune Correlates. Cancer Immunol Immunother (2014) 63:407–18. doi: 10.1007/s00262-014-1524-0

27. Jochum W, Tucker JA, Tsang KY, Madan RA, Dahut WL, Liehr DJ, et al. A Combination Trial of Vaccine Plus Ipilimumab in High-Grade Prostate Cancer in the Placebo Arm of the Prostate Cancer Immunotherapy, Semin Cancer Biol (2022). doi: 10.1016/j.semcancer.2022.03.014

28. Porto-Filho JS, de Miranda S, de Souza MM, de Carvalho CP, de Melo RD. Chronic Inflammation in Prostate Carcinogenesis: Novel Therapeutic Opportunities. Mol Cancer (2021) 20:18. doi: 10.1186/s12952-020-01263-w

29. Gurel B, Lucia MS, Thompson IM Jr, Goodman PJ, Tangen CM, Kristal AR, et al. Preexisting Levels of CD4 T Cells Expressing PD-1 Are Related to Overall Survival in Prostate Cancer Patients Treated With Ipilimumab. Cancer Immunol Immunother (2015) 3:1008–16. doi: 10.1158/2326-6066.CIR-14-0227

30. Wagner GP, Kin K, Lynch VJ. Measurement of mRNA Abundance Using RNA-Seq Data: RPKM Measure is Inconsistent Among Samples. Theory Biosci (2012) 131:281–5. doi: 10.1007/s12920-012-0162-3

31. Zhao S, Ye Z, Stanton R. Misuse of RPKM or TPM Normalization When Comparing Across Samples and Sequencing Protocols. RNA (2020) 26:903–9. doi: 10.1261/rna.074922.120

32. Mermal CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 Facilitates Sensitive and Confident Localization of the Targets of Focal Somatic Copy Number Alteration in Human Cancers. Genome Biol (2011) 12:841. doi: 10.1186/gb-2011-12-4-r41

33. Mayakonda A, Lin DC, Assenov Y, Plass C, Koehler HP, Mafioff Efficient and Comprehensive Analysis of Somatic Variants in Cancer. Genome Res (2018) 28:1747–56. doi: 10.1101/gr.239244.118

34. Wilkerson MD, Hayes DN. ConsensusClusterPlus: A Class Discovery Tool With Confidence Assessments and Item Tracking. Bioinformatics (2010) 26:1572–3. doi: 10.1093/bioinformatics/btp170

35. Chen H, Ge XL, Zhang ZY, Liu M, Wu RY, Zhang XF, et al. M(5)C-Regulated Mediated Methylation Modification Patterns and Tumor Microenvironment Infiltration Characterization in Lung Adenocarcinoma. Transl Lung Cancer Res (2021) 10:2172–92. doi: 10.21037/tlcr-21-351

36. Zhang X, Shi M, Chen T, Zhang B. Characterization of the Immune Cell Infiltration Landscape in Head and Neck Squamous Cell Carcinoma to Aid Immunotherapy, Mol Ther Nucleic Acids (2020) 22:298–309. doi: 10.1016/j.omtn.2020.08.030

37. Hanzelmann S, Castelo R, Guinney J, GSVa: Gene Set Variation Analysis for Microarray and RNA-Seq Data. BMC Bioinformatics (2013) 14:7. doi: 10.1186/1471-2105-14-7

38. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-Cancer Immunogenomic Analyses Reveal Genotype-Inflammation relationships and Predictors of Response to Checkpoint Blockade. Cell Rep (2017) 18:248–62. doi: 10.1016/j.celrep.2016.12.019

39. Zhou P, Lu Y, Xun Y, Xu J, Liu C, Xia Q, et al. Ubiquitin Modification Patterns of Clear Cell Renal Cell Carcinoma and the Ubiquitin Score to Aid Immunotherapy and Targeted Therapy. Front Cell Dev Biol (2021) 9:659294. doi: 10.3389/fcell.2021.659294

40. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, et al. Signatures of T Cell Dysfunction and Exclusion Predict Cancer Immunotherapy Response. Nat Med (2018) 24:1510–8. doi: 10.1038/s41591-018-0136-1

41. Geeleher P, Cox N, Huang RS. Prrophetic: An R Package for Prediction of Clinical Chemotherapeutic Response From Tumor Gene Expression Levels. PLoS One (2014) 9:e107468. doi: 10.1371/journal.pone.0107468

42. Simon N, Friedman J, Hastie T, Tibshirani R. Regularization Paths for Cox’s Proportional Hazards Model With Application to DNA Microarray Data. J Stat Softw (2006) 13:1–26. doi: 10.18637/jss.v013.i01

43. Altmann L, Gutierrez-Osborn ML, Lee Y, Tangen CM, Belinson JL, et al. A Combination Trial of Vaccine Plus Ipilimumab in High-Grade Prostate Cancer in the Placebo Arm of the Prostate Cancer Immunotherapy, Semin Cancer Biol (2022). doi: 10.1016/j.semcancer.2022.03.014

44. Pardoll DM. The Blockade of Immune Checkpoints in Cancer Immunotherapy. Nat Rev Cancer (2012) 12(4):252–64. doi: 10.1038/nrc3239
46. Schaeffer E, Srinivas S, Antonarakis ES, Armstrong AJ, Bekelman JE, Cheng H, et al. NCCN Guidelines Insights: Prostate Cancer, Version 1.2021. *J Natl Compr Canc Netw* (2021) 19(2):134–43. doi: 10.6004/jnccn.2021.0008

47. Chen MK, Liang ZJ, Luo DS, Xue KY, Liao DY, Li Z, et al. Abiraterone, Orteronel, Enzalutamide and Docetaxel: Sequential or Combined Therapy? *Front Pharmacol* (2022) 13:843110. doi: 10.3389/fphar.2022.843110

48. Petrylak DP, Tangen CM, Hussain MH, Lara PN, Jones JA, Taplin ME, et al. Docetaxel and Estramustine Compared With Mitoxantrone and Prednisone for Advanced Refractory Prostate Cancer. *N Engl J Med* (2004) 351(15):1513–20. doi: 10.1056/NEJMoa041318

49. Tannock IF, de Wit R, Berry WR, Horti J, Pluzanska A, Chi KN, et al. Docetaxel Plus Prednisone or Mitoxantrone Plus Prednisone for Advanced Prostate Cancer. *N Engl J Med* (2004) 351(15):1502–12. doi: 10.1056/NEJMoa040720.50

50. Li H, Fu J, Zeng Z, Cohen D, Li J, Chen Q, et al. TIMER2.0 for Analysis of Tumor-Infiltrating Immune Cells. *Nucleic Acids Res* (2020) 48:W509–14. doi: 10.1093/nar/gkaa407

51. Chen H, Yang H, Zhu X, Yadav T, Ouyang J, Truesdell SS, et al. M(5)C Modification of mRNA Serves a DNA Damage Code to Promote Homologous Recombination. *Nat Commun* (2020) 11:2834. doi: 10.1038/s41467-020-16722-7

52. Tixid L, Lusser A. The Dynamic RNA Modification 5-Methylcytosine and its Emerging Role as an Epitranscriptomic Mark. *Wiley Interdiscip Rev RNA* (2019) 10:e1510. doi: 10.1002/wrna.1510

53. Su J, Wu G, Ye Y, Zhang J, Zeng L, Huang X, et al. NSUN2-Mediated RNA 5-Methylcytosine Promotes Esophageal Squamous Cell Carcinoma Progression via LIN28B-Dependent GRB2 mRNA Stabilization. *Oncogene* (2021) 40 (39):5814–28. doi: 10.1038/s41388-021-01978-0

54. Zeng H, Xu Y, Xu S, Jin L, Shen Y, Rajan KC, et al. Construction and Analysis of a Colorectal Cancer Prognostic Model Based on N6-Methyladenosine-Related lncRNAs. *Front Cell Dev Biol* (2021) 9:698388. doi: 10.3389/fcell.2021.698388

55. Liu XS, Zhou LM, Yuan LL, Gao Y, Kui XY, Liu XY, et al. NPM1 Is a Prognostic Biomarker Involved in Immune Infiltration of Lung Adenocarcinoma and Associated With M6a Modification and Glycolysis. *Front Immunol* (2021) 12:724741. doi: 10.3389/fimmu.2021.724741

56. Zhang B, Wu Q, Li B, Wang D, Wang L, Zhou YL. M(6)A Regulator-Mediated Methylation Modification Patterns and Tumor Microenvironment Infiltration Characterization in Gastric Cancer. *Mol Cancer* (2020) 19:53. doi: 10.1186/s12943-020-01170-0

57. Xue C, Zhao Y, Li G, Li L. Multi-Omic Analyses of the M(5)C Regulator ALYREF Reveal Its Essential Roles in Hepatocellular Carcinoma. *Front Oncol* (2021) 11:633415. doi: 10.3389/fonc.2021.633415

58. Pan J, Huang Z, Xu Y. M5c RNA Methylation Regulators Predict Prognosis and Regulate the Immune Microenvironment in Lung Squamous Cell Carcinoma. *Front Oncol* (2021) 11:657466. doi: 10.3389/fonc.2021.657466

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

**Copyright © 2022 Xu, Chen, Zhang, Liu and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.**