M5C modulator mediated methylation pattern and TME cell infiltration in bladder cancer

Zhan Yang
The First Affiliated Hospital of Wenzhou Medical University
Zhichao Lang
The First Affiliated Hospital of Wenzhou Medical University
Qiqi Tao
The First Affiliated Hospital of Wenzhou Medical University
Bo Chen
The First Affiliated Hospital of Wenzhou Medical University
Xiaoqi Li
The First Affiliated Hospital of Wenzhou Medical University
Yating Zhan
The First Affiliated Hospital of Wenzhou Medical University
Rongrong Zhang
The First Affiliated Hospital of Wenzhou Medical University
Kai Zhu
The First Affiliated Hospital of Wenzhou Medical University
Zhengze Yao
The First Affiliated Hospital of Wenzhou Medical University
Yuxiang Gao
The First Affiliated Hospital of Wenzhou Medical University
Ji Chen
The First Affiliated Hospital of Wenzhou Medical University
Jianjian Zheng
The First Affiliated Hospital of Wenzhou Medical University
Yeping Li (liypwz@126.com)
The First Affiliated Hospital of Wenzhou Medical University

Research Article

Keywords: M5C, Tumor microenvironment, Immunotherapy, Immune cell, Stromal cell, Mutation burden

Posted Date: June 21st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1752939/v1

Abstract

Background:

In recent studies, the role of modification patterns of N6-methyladenosine (m6A) regulators in tumor microenvironment (TME) cell infiltration in the tumor microenvironment has been identified. However, the role of 5-methylcytosine (m5C) modification patterns in TME cell infiltration remains to be discovered.

Methods:

In this study, 660 bladder cancer patients were enrolled to comprehensively evaluate the m5C modification pattern based on 13 m5C regulators, and to explore the association between m5C modification pattern and TME cell infiltration status. Finally, PRINCIPAL component analysis (PCA) algorithm constructed m5Cscore, which was used to evaluate and quantify individual m5C modification patterns.

Results:

Our study identified three distinct m5C modification patterns in bladder cancer. Studies showed that TME cell infiltration characteristics under these three modification patterns were highly associated with three tumor immunophenotypes (immune-inflamed, immune-excluded, immune-desert). In addition, m5C methylation modification patterns were proved to be able to predict tumor disease progression, genetic variation and prognosis of patients. The individualized scoring system – m5Cscore, which was based on m5C methylation modification, showed great differences in tumor mutation burden (TMB), clinical characteristics, immune checkpoint and immunotherapy. The low m5Cscore group was characterized by matrix activation and lack of effective immune cell infiltration, indicating the immune-excluded phenotype. The higher m5Cscore group showed fewer PD-L1 mutations, which were associated with a promoted response to PD-L1/PD-1 immunotherapy and a better prognosis.

Conclusions:

This work confirmed the extensive regulatory mechanisms of the m5C modification pattern in the tumor microenvironment. The m5C modification pattern functioned as an important element of forming the heterogeneity and complexity of TME landscape. Understanding and in-depth evaluation of individual m5C modification patterns would help us understand the TME landscape and guide better immunotherapy strategies.

Introduction

5-methylcytosine (m5C) is an extensive RNA modification found in almost every RNA species, including cytoplasmic and mitochondrial ribosomal RNAs (rRNAs), enhancer RNAs (eRNAs), long noncoding RNAs (lncRNAs), transfer RNAs (tRNAs), and messenger RNAs (mRNAs) [1–3]. Similar to N6-methyladenosine (m6A), m5C modification is a reversible process in human cells that involves three dynamic control elements, including methyltransferases (writers), demethylases (erasers) and binding proteins (readers) [4, 5]. During the modification of m5C gene, NSUN1-7, DNMT1, DNMT2, DNMT3A, DNMT3B and TRDMT1 act as methyltransferases to catalyze m5C, while demethylases such as TET2 mediate its removal. In addition, m5C is specifically recognized by the mRNA export adaptor ALYREF, thus affecting m5C functions [6–8]. Recent studies have demonstrated that m5C regulators play a pivotal role in various biological functions such as malignant tumor invasion, abnormal immune function, cell developmental defects, and abnormal cell proliferation [9–11].

Immunotherapy immune checkpoint blockade (ICB, PD-L1/PD-1), a new immunotherapy, has been found to improve the efficacy of clinical drug therapy for cancer. Unfortunately, most patients show little or no benefit[12, 13]. Tumor microenvironment is the peripheral environment for tumor cells to survive. TME includes not only cancer cells, but also surrounding immune cells, lymphocytes, blood vessels, fibroblasts extracellular matrix (ECM), bone marrow-derived inflammatory cells and signaling molecules. Accumulating evidence suggests that the TME was crucial for tumor progression, immunosuppression and immunotherapy response. For example, YAP1 affects macrophages, bone marrow-derived inhibitory cells, and regulatory T cells to promote immunosuppressive TME[14]; Changes in tumor metabolic requirements lead to impaired functional responses of cytokine-activated natural killer cells, thereby inhibiting the toxic effects of NK cells[15].

A critical step in improving the efficacy of existing immune checkpoint therapies and develop new immunotherapy strategies is the individualized prediction of ICB response on the basis of TME cell infiltration signatures [16, 17]. Therefore, in-depth analysis of the heterogeneity and diversity of tumor landscape in TME contributes to determinate new tumor immunophenotypes and biomarkers for immune treatment response [18].

Many researches have proved that m5C modification plays a key role in TME. Pan et al. reported the increased expression of NSUN4, a m5C regulator, could promote the expression of CD4 + T cells and tumor-associated neutrophils, thus improving the anti-tumor effect [9]. In addition, DNMT3B and NSUN5-mediated m5C modification activates multiple cancer-related pathways including ERBB pathway, Wnt pathway and TGF-β pathway, associated with poor prognosis in Clear Cell Renal Cell Carcinoma [19]. However, the comprehensive understanding of the function of m5C modification-mediated TME in BC remains unknown.

In this study, gene expression profiles and clinical characteristics of 661 bladder cancer patients were included to comprehensively evaluate the m5C modification pattern, and to explore the correlation between m5C modification pattern and immune infiltration characteristics of TME. Three m5C modification patterns were found in bladder cancer patients, and the TME characteristics were highly

correlated with immunophenotypes. Our results demonstrated that m5C modification functions as a non-negligible factor in the formation of TME characteristics. Additionally, we established an individualized scoring system for TME immunophenotype in bladder cancer patients.

Materials And Methods

Patients Datasets and processing

Public gene-expression data and complete clinical annotation were retrieved from the Gene-Expression Omnibus (GEO) and the Cancer Genome Atlas (TCGA) databases. After excluding patients without survival information, this study collected eligible GC cohorts (GSE13507 and TCGA-BLCA (Cancer Genome Atlas-Urothelial Carcinoma of the Bladder)) for further analysis. We background-adjusted and quantile-normalized microarray data from Affymetrix® by "CEL" files and multiarray averaging methods and the affy and simpleaffy software packages. As for the microarray data from other platforms, we download the normalization matrix file directly. For the data set in TCGA, we used the R package TCGAbiolinks developed for comprehensive analysis using GDC data to download RNA sequencing data (FPKM value) of gene expression from Genomic Data Commons (GDC, https://portal.gdc.cancer.gov/). The FPKM values were then converted to a transcript of millions per kilobase (TPM) value. Batch effects that were not biotechnological biases were corrected for by the "ComBat" algorithm of the sva package. The TCGA-BLCA cohort was used for copy number variation (CNV) analysis. Then the R (version 3.6.1) and the R Bioconductor package were used for data analysis.

Unsupervised clustering for 21 m6A regulators

15 m5C regulators were extracted from published literature, including 12 writers (NSUN1-7, DNMT1, DNMT2, DNMT3A, DNMT3B, TRDMT1), 1 eraser (TET2) and 2 readers (ALYREF, YBX1) [20]. We combined TCGA-BLCA and GSE13507 patient cohorts, and five genes were not detected in the mixed cohorts. Thus, 10 m5C regulators were eventually incorporated into our study. Based on the expression of 10 m5C modulators, different m5C modification patterns were identified by unsupervised cluster analysis, and the classified patients were further analyzed. The consensus clustering algorithm determined the number of clusters and their stability. The ConsensuClusterPlus package performed the above steps and had been repeated 1000 times to ensure classification stability.

Gene set variation analysis (GSVA) and functional annotation

GSVA enrichment analysis was performed by the "GSVA" R software package to explore different biological processes mediated by different m5C modification patterns. As a nonparametric and unsupervised method, GSVA is commonly performed to estimate changes in the pathway and biological process activity in expression dataset samples [21]. We downloaded the gene set of "c2.cp.kegg. V6.2" from MSigDB database for GSVA analysis. After adjustment, $P < 0.05$ was considered statistically significant. R package ClusterProfiler functionally annotated of m5C-related genes with cutoff $FDR < 0.05$.

Comprehensive assessment of TME cell infiltration status

We utilized the ssGSEA algorithm to quantify the relative abundance of various types of cell infiltrates in TME. A gene set resulting from the work of Charoentong, marking each TME-infiltrating immune cell type, stores multiple human immune cell subtypes, including activated activated B cells, CD4 T cells, macrophages, neutrophile, regulatory T cells, etc [22, 23]. The Stromal Score, Immune Score ESTIMATE Score and TumorPurity in each modified pattern was calculated by ESTIMATE analysis, and they were used to characterize cell infiltration. The infiltrating abundance of immune cells in each modification pattern was analyzed by the CIBERSOFT algorithm. Additionally, the expression levels of human leukocyte antigen (HLA) and immune checkpoints in the three modification patterns were further analyzed.

Identification of differentially expressed genes (DEGs) between m5C distinct phenotypes

Based on the expression levels of 10 m5C regulators, we identified M5C-related genes by classifying patients into three different m5C modification patterns. The differences among the different m5C modification modes were determined by Limma R package empirical Bayesian method. The significance criteria for determining DEGs was set to an adjusted P value < 0.001.

Construction of m5C gene signature

We constructed an evaluation system for evaluating the m5C modification pattern of a single bladder cancer patient, the m5C gene signature, and quantified the m5C modification pattern of a single tumor, which we called m5Cscore. The procedure for establishing m5C gene characteristics was as follows:

We first normalized the DEGs identified by different m5Cclusters in mixed cohort samples and extracted overlapping genes. The overlapping DEG was analyzed by unsupervised clustering method. Subsequently, the patients were divided into groups. Consensus clustering algorithm was used to define the number and stability of gene clustering. Univariate Cox regression model was used to analyze the prognosis of each gene in the trait, and the genes with significant prognosis were further analyzed. In order to construct m5C-related gene features, we conducted principal component analysis (PCA), and principal components 1 and 2 were both selected as signature scores. This approach focused the score on the group with the largest well-associated (or anti- associated) gene block in the group, while reducing the contribution of genes not tracked with other group members. The m5Cscore is defined by a method similar to GGI[24, 25]:

$$m5Cscore = \Sigma(PC1i + PC2i)$$

where i is the expression of m5C phenotype-related genes.

Clinical prognostic significance of m5C gene signature

After constructing the m5Cscore, we included clinical data (age, fustat, gender, T, N) from the mixed cohort to evaluate the clinical prognostic significance of m5Cscore. Survival analysis identified the value of m5Cscore in predicting the prognosis of patients with different clinical characteristics. The Kruskal Wallis test determined whether the m5Cscore was different in patients with different clinical traits. Bar graph shows the distribution of patients with different clinical characteristics in both m5Cscore group.

Collection of immune-checkpoint blockade genomic and clinical information

We systematically searched gene expression profiles for immune checkpoint blockade, which were publicly available and reported with complete clinical information. Our study eventually included two immunotherapy cohorts: Metastatic melanoma treated with pembrolizumab an, anti-PD-1 antibody (GSE78220 cohort from GEO) [27], and advanced uroepithelial carcinoma treated with Atezolizumab, an anti-PD-L1 antibody (IMvigor210 cohort) [26]. According to the Creative Commons license agreement, 3.0 IMvigor210 queue full expression data can be obtained from http://research-pub.Gene.com/imvigor210corebiologies and detailed clinical notes. The raw technical data were standardized using the DEseq2 package and the calculated values were converted into TPM values. For the GSE78220 cohort, FPKM data for gene expression profiles were also transformed into TPM values with greater inter-sample comparability after normalization with the limma package.

Statistical analysis

Correlation coefficients between m5C regulator expression and TME infiltrating immune cells were calculated by Spearman and distance correlation analysis [28]. One-way ANOVA and Kruskal-Wallis test were performed to compare differences between three or more groups [29]. The survminer R software package determined the cut-off points for each subgroup of data sets based on the correlation between patient survival and m5Cscore. The "surv-cutpoint" function splited the m5Cscore in two. In the light of the chosen maximum log-rank statistic, we divided patients into high and low m5Cscore groups so as to reduce the batch effects in the calculations. Kaplan-meier method was used to analyze the survival curve of patient prognosis, and log-rank test was used to determine the significance of the differences. Univariate Cox regression models were used to calculate hazard ratios (HRs) for m5C regulators and genes associated with m5C-modified phenotypes. Waterfall plots, drawn by the maftools package, showed the mutational profile of patients with high and low m5Cscore in the mixed TCGA-BLCA and GSE13507 cohorts. Copy number variation maps of 13 m5C regulons on 23 pairs of chromosomes were plotted by the R package RCircos, where all statistical values are bilateral. $P < 0.05$ was known as statistically significance. All data processing was completed in R3.6.1 software.

Result

CNV, mutation and prognostic characteristics of m5C regulators in bladder cancer

Figure 1 showed the technical route of this research. 13 regulators were included in this study. Figure 2A revealed the incidence of CNV of m5C regulators in bladder cancer. The results revealed that a prevalent amplificated CNV alteration of these 13 regulators, while TET2 and DNMT1 had an obvious CNV deletion frequency. The altered position of CNV on the chromosome of the m5C regulators were shown in Fig. 2B. Between BC samples and normal samples, we found that m5C regulators were differential expressed (Fig. 2C). Then, an investigation of somatic mutations in the m5C regulatory factor in BC revealed that 49 of the 408 samples had undergone mutations in the m5C regulatory factor. It was found that the mutation frequency of NSUN2, NSUN6, NSUN7, DNMT1, DNMT3A (writers) was 2%, and the mutation frequency of DNMT3B, the eraser (TET2) and the reader (ALYREF, YBX1) was 1% (Fig. 2D). The low expression of TRDMT1 and DNMT1 corresponded to a better prognosis, while the low expression of NSUN4-7 corresponded to a poor prognosis (Fig. 2E-J**)**. The above studies shown that the expression of m5C regulatory factors in normal and bladder cancer tissues had obvious heterogeneity, and its abnormal expression functioned as a key element in the occurrence and development of tumors and the progression of bladder cancer..

TME and immune infiltration characteristics in three distinct m5C modification patterns

According to the expression of 10 m5C regulatory factors in patients, we used R package ConensusClusterPlus to qualitatively classify patients with different m5C modification patterns. Unsupervised clustering method was adopted to determine three different modification patterns, including 195 cases of cluster A, 140 cases of cluster B, and 231 cases of cluster C.

On account of the correlation between these three different clusters is low, we refer to each of these patterns as m5Ccluster A-C (Fig. 3A-D). The survival analysis of these three m5C modification patterns revealed that patients in m5CclusterB had greater survival advantages (Fig. 3E). Principal component analysis (PCA) showed that the three patient groups were clearly distinguishable from each other (Fig. 3F). Figure 3G indicated that the expression of m5C regulatory factors in the three modification patterns and the clinical information between the three components. The m5C regulatory factors were found to be highly expressed in m5Ccluster A and low expressed in m5Ccluster C. Interestingly, the expression of NSUN6 and NSUN7 in m5Ccluster B was significantly lower.

We used GSVA enrichment analysis to explore underlying biological behaviors. As shown in Fig. 3I, m5Ccluster A was significantly enriched in homologous recombination, aminoacyl tRNA biosynthesis, ubiquitin-mediated proteolysis, spliceosomes, and RNA degradation. M5Ccluster B was associated with mesenchymal and oncogenic activation pathways, such as cell and focal adhesion, ECM receptor interaction, cell transendothelial migration, and chemokine signaling pathway (Fig. 3H). And m5Ccluster C appeared to be related to a variety of biological pathways of acid metabolism and immune clearance (Fig. 3I-J). Surprisingly, subsequent analysis of TME cell infiltration revealed a large number of immune cell infiltrates in m5Ccluster B, including activated T and B lymphocytes, dendritic cells, natural killer cells, MDSC, macrophages, neutral cells, granulocytes and helper cells (Fig. 4G). However, patients in this m5C

modification pattern had the worst prognosis. The cell biology phenomenon of acid metabolism enrichment observed in m5Ccluster C corresponded to the Warburg effect, which meant that metabolic reprogramming in m5Ccluster C cancer cells was much more extensive [30]. The activation of acid metabolism in the tumor indicated enhanced tumorigenesis, which also corresponded to the worse survival of type C patients compared to type A in Fig. 3E.

In order to fully understand the immune infiltration status of these three m5Cclusters, we next scored the three m5C modification patterns in 29 immune genomes to obtain their immune score, matrix score, ESTIMATE score and tumor purity, and plotted a heatmap (Fig. 4A-E). Then we explored the expression of human leukocyte antigen (HLA) among m5Cclusters, and the generally high expression of HLA genes still verified the previous conclusions (Fig. 4F). Then, we used the CIBERSOFT algorithm to analyze the strength of cellular immunity between the three clusters, which demonstrated that the expression of Tregs and naive B cells in m5C cluster A was much higher than in the other two clusters (Fig. 4H). M5Ccluster C was more highly expressed in cells related to humoral immunity. In our research on immune checkpoints, the expression of m5Ccluster B related genes was still much higher than the other two clusters (Fig. 4I). The immune checkpoint study showed that the expression level of PD-L1 in 5CclusterB was the highest among the three groups, and the expression levels of the remaining two groups were about the same (Fig. 4J). In summary, m5Ccluster A was summarized as an immune desert type, which was characterized by immunosuppression and early activation of anti-tumor cytokines; m5Ccluster B was summarized as an immune rejection phenotype, which was characterized by numerous immune cells infiltrate and activated stroma; m5Ccluster C was summarized as a phenotype of immune inflammation and metabolic reprogramming, which was characterized by strong humoral immunity and activation of acid metabolism.

Generation of m5C gene clusters and functional annotation

Next, we attempted to gain insight into the biological behavior of each m5C modification pattern. The R package limma identified 902 m5C phenotype-related DEGs (Fig. 5A). Then GO and KEGG enrichment analysis were utilized to perform these 902 DEGs. It was worth noting that in the GO enrichment analysis, these DEGs were concentrated in matrix activation pathways and activation immune cells, such as activation of T cells and cell adhesion factors. While in KEGG enrichment analysis, these DEGs were enriched in cytokine and chemokine activation and pathways associated with abnormal immune inflammation. These conclusions indicated that m5C modification had a significant correlation with the biological process of immune cells, and once again confirmed that m5C modification functioned as a significant element in immune regulation in TME (Fig. 5B-C). To elucidate the regulatory mechanism in depth, we first obtained 377 prognostic genes based on 902 DEGs, and then performed unsupervised cluster analysis according to 377 prognostic genes associated with m5C phenotype to divide patients into distinct genomic subtypes. In keeping with the clustering grouping of m5C modification patterns, we also found three different m5C modified genome phenotypes, named m5C gene cluster A-C (Fig. 5D-F). This result revealed three different m5C modification patterns in bladder cancer. The tumors in m5C gene cluster B and m5C gene cluster C were poorly differentiated. We also found that patients in m5Ccluster B

were focused on gene cluster B mostly, which further proved the regulatory mechanism of m5C methylation modification (Fig. 5G). In addition, the m5C regulatory factor was found to be significantly differentially expressed in the three gene clusters, which proved the effectiveness of m5C methylation modification pattern at the gene level (Fig. 5H). And 192 cases were concentrated in gene cluster A, which was related to better prognosis. 160 cases were concentrated in gene cluster B and 214 cases were concentrated in gene cluster C. These two clusters were associated with worse prognosis (Fig. 5I).

Generation of individualized scoring criteria and associated transcriptional and TMB features

Based on the above conclusions, we built a scoring system to quantify m5C modification patterns in individual bladder cancer patients according to the understanding of m5C methylation and phenotypic modification patterns. The m5Cscore is calculated from 377 phenotype-related genes. We first calculated the m5Cscore of all samples, and then divided all patients into two groups with high AND low m5Cscore according to the median value of m5Cscore, and conducted survival analysis. It was found that the survival rate of the low m5C group was significantly lower than that of the high m5C group (Fig. 6A). In order to better explain the biological characteristics of m5Cscore, the relationships between m5Cscore and TME were also discussed. Through spearman analysis, we found that m5Cscore was significantly negatively correlated with immune cell activation, which meant there was more immune cell activation in the low m5C group (Fig. 6B). Next, we found matrix and immune activation related genes were highly expressed in the low m5C group (**Fig.S1A-B**). Alluvial diagrams showed changes in individual patient attributes. The results showed that patients with m5CclusterB characteristics were mainly focused on gene cluster B, and a few were focused on gene cluster C; patients with m5CclusterA characteristics were mainly focused on gene cluster A and gene cluster C; M5CclusterC was found in all three gene clusters. Almost all m5CclusterA patients were focused on the high m5C group, and almost all m5CclusterB patients were focused on the low m5C group. Patients with a large number of m5CclusterC acid metabolism characteristics were distributed in two m5C subgroups. However, we did not find any difference in gender and survival status between these two m5C groups (Fig. 6C-D). At the same time, Kruskal Wallis test proved that there were statistical differences of three modification patterns (m5Ccluster) and three gene clusters among m5Cscore groups **(**Fig. 6E-F**)**.

Subsequently, R package maftools analyzed the differences in the distribution of somatic mutations between low m5Cscore and high m5Cscore in the mixed cohort. Figure 6G-H showed that the high m5C group had a wider mutation load than the low m5C group, but for TP53, the mutation frequency of the low m5C group was much higher than that of the high m5C group. Previous literature suggested that the expression of TP53 could affect the expression level of PD-L1[31]. Therefore, we also detected the expression level of PD-L1 in the two m5C groups and found that the expression level of PD-L1 was higher in the low m5C group (Fig. 6L). And then, we also detected the expression level of PD-1. Results showed that the expression level of PD-1 was still higher in the low m5C group (**Fig.S1C**). Second, the prognostic analysis of patients with high and low mutation subtypes showed that the high mutation group had a

greater survival advantage (P < 0.001) (Fig. 6I). Specifically, we examined the patient survival rate under the combined TMB and m5Cscore. We again found that patients in the high mutation group had a higher survival rate, while in the high mutation group, patients with higher m5Cscore had slightly better survival advantages (Fig. 6J). In externally validated cohorts GSE48276 and GSE3184, the high m5Cscore group also showed better survival advantages **(FigS1.D-E**). In addition, tumors with low m5Cscore in the GSE31684 cohort were found to have stronger local recurrence activity (**FigS1.F-G**). The expression of PD-L1 in the high m5C group was lower, and the expression of PD-L1 in m5Cclusters was different. M5Cscore and TMB also showed a positive correlation (Fig. 6K). The accumulated evidence suggests that patients in the low mutation group may have a stronger resistance to anti-PD-L1 immunotherapy. These results indirectly suggested that the differences in different m5C modification patterns in bladder cancer patients may be a key factor mediating PD-L1 /PD-1 immunotherapy response.

Comprehensive evaluation of clinical significance of m5Cscore

Next, we attempted to further investigate the value of m5Cscore in assessing the clinical characteristics of patients. Results revealed that the m5Cscores of different age groups were different, and patients with high m5Cscores had a higher survival rate in the 65-year-old group (Fig. 7A, E). The survival analysis of m5Cscore in the survival state showed that there was no difference in m5Cscore between survival groups, but there were obvious differences in m5Cscore between death groups (Fig. 7B, F). At the same time, the corresponding m5Cscore of survival group was much higher than that of death group, and more patients survived in the high m5Cscore group (Fig. 7J, O). Similarly, survival analysis of the gender group of m5Cscore showed that there were significant differences in both groups (Fig. 7C, G). The corresponding m5Cscore was significantly higher in male patients, and there were more male patients with high m5Cscore (Fig. 7K, P). Survival analysis of the prognosis of patients with different tumor progression predicted by m5Cscore showed that the difference between tumor stages N1-3 and T3-4 was not obvious. However, in the N0 and T1-2 stages, m5Cscore showed a difference (Fig. 7D, H-I, N). In addition, the Kruskal Wallis test showed that patients with stage N0 or T1-2 had higher m5Cscore, which meant that patients with shorter stages had higher scores (Fig. 7L-M, Q-R). In short, m5Cscore had important prognostic significance for the clinical characteristics of patients.

Role of m5C modification patterns in anti-PD-1 /L1 immunotherapy

To analyze the role of m5C modification patterns in anti-PD-L1 / PD-1 immunotherapy in depth, we included two immunotherapy cohorts to investigate whether m5C modification patterns can predict patients' response to ICB therapy. In the anti-PD-L1 cohort (IMvigor210), high m5Cscore group was associated with better clinical benefit and longer survival (Fig. 8A). And patients in high m5Cscore group had significant therapeutic advantages and anti-PD-L1 clinical efficacy (Fig. 8B-C). However, in the anti-PD-1 cohort (GSE78220), patients with a high m5C score had a smaller survival advantage (Fig. 8D). We speculated that the possible reason was that the sample size of the queue is too small, leading to

sampling error (n = 27). Another possible reason was that tumor mutations produce new PD-L1 epitopes, so anti-PD-L1 blocking therapy did not show significant effect. It was worth mentioning that in anti-PD-L1 immunotherapy, patients with high m5Cscore still had better immunotherapy effects (Fig. 8E-F). In addition, we also studied the therapeutic effect of patients in the anti-PD-L1 cohort (IMvigor210) evaluated by the combination of m5Cscore and tumor neoantigen load, and found that patients with high m5Cscore and tumor neoantigen load had a greater survival advantage (Fig. 8G). In view of the tumor immunophenotypes detected in the IMvigor210 cohort, we also studied the differences in m5Cscore between different phenotypes. The results revealed that high m5Cscore was significantly correlated with immune desert subtypes, while low m5Cscore was significantly correlated with immune rejection and immune inflammation types (Fig. 8H). In conclusion, the quantification of m5C modification patterns can be used as potential and reliable biomarker for evaluating the prognosis and clinical response to immunotherapy.

Secondly, by detecting the immune cell infiltration status of both m5Cscore groups, we found that the low m5Cscore group generally had high expression of immune cells (Fig. 8I). The results of CIBERSOFT analysis revealed that the expression levels of Tregs and dendritic cells were higher in the high m5Cscore group, which meant that the high m5Cscore group and m5CclusterA had similar immune characteristics (Fig. 8J). The study on the expression levels of immune checkpoints in both groups showed that the expression levels of immune checkpoints in the low m5Cscore group increased widely, which also showed that the low m5C group was more capable of resisting immune checkpoint treatment from the side (Fig. 8K). These results once again confirmed that there were different m5C methylation patterns in patients with bladder cancer, and these m5C methylation patterns were significantly related to tumor immunophenotypes and anti-PD-L1/PD-1 immunotherapy responses.

Based on our elaboration and quantification of the m5C methylation modification mode, m5Cscore has a strong ability to evaluate clinical characteristics such as age, gender, and T stage, and m5Cscore has a strong evaluation ability in anti-PD-L1 /PD-1 immunotherapy effect. These findings strongly indicate that patients with bladder cancer have corresponding m5C methylation patterns. The established m5Cscore can help predict the response to anti-PD-1/L1 immunotherapy and detect the clinicopathological characteristics of patients.

Discussion

Many previous studies have examined the TME landscape from the perspective of multiple m6A regulators. This provided an idea for our study. Multiple m5C regulatory factors were included in this study to investigate the effects of m5C modification on tumor growth, tumor immunity and anti-tumor under the interaction of multiple m5C regulatory factors. To clarify the different roles of different m5C modification patterns in TME cell infiltration will be helpful to understand the anti-tumor immune response effect of TME under m5C modification pattern, and to guide more effective immunotherapy strategies.

In this study, three different m5C modification patterns were revealed based on 10 m5C regulators, all of which had significantly different characteristics of tumor microenvironment cell infiltration. M5Ccluster A was summarized as immune-desert type, characterized by immunosuppression and early activation of anti-tumor cytokines; m5Ccluster B was summarized as immune-excluded phenotype, characterized by amount innate immune cell infiltration and stromal activation; m5Ccluster C was summarized as immune-inflammatory and metabolic reprogramming phenotype, characterized as strong humoral immune and acid metabolism activation. The immune-desert and immune-excluded phenotypes can be considered as non-inflammatory tumors, with little or no immune cell infiltration in the TME. The immune-inflammatory phenotype was known as a thermo tumor and was characterized by extensive immune cell infiltrates in the TME[32–35]. Although a large number of immune cells also existed in the immune-excluded phenotype, they only existed in the stroma around the tumor cell nest and failed to break through the tumor matrix to kill tumor cells or inhibit tumor cell proliferation. In some tumors, immune cells may be confined to the stroma of the tumor, or they may be present in the tumor itself, resulting in the appearance of immune cells within the tumor [36, 37]. In fact, tumor active stroma is thought to inhibit immune cells [38]. Therefore, we inferred that the activated matrix inhibits the anti-tumor ability of immune cells.

In our study, m5CcluserA showed little or no infiltration of immune cells, with high expression of Tregs and multiple anti-tumor cytokines. Previous studies have found that Tregs can help suppress immune surveillance in tumors, thereby promoting cancer progression[39]. At the same time, it had also been proved that naive B cells can inhibit the growth of cancer cells by secreting four factors that negatively regulate cell growth, and naive B cells inhibited the growth of cancer cells in the early stage and promoted the growth of cancer cells in the late stage [40]. M5CclusterB showed significant mesenchymal activation, and its mesenchymal score was much higher than that of the other two clusters, including high expression of adhesion factors and ECM receptors, which was considered as T cell inhibition. M5CclusterC showed an extensive immune cell infiltration in TME. These findings substantiated the reliability of our phenotypic classification of m5Ccluster, which may help us understand the m5C modification pattern in bladder cancer.

Moreover, mRNA transcriptional differences between different m5C modification patterns further confirmed that m5C was significantly correlated with immune-related biological pathways. These DEGs were considered to be characteristic genes associated with m5C. The prognostic characteristic genes were selected by univariate Cox regression analysis, and then three gene clusters were determined based on prognostic characteristic genes, which were also related to matrix and immune activated pathways. This finding proved that the alteration of m5C is significantly related to the heterogeneity of TME landscape again. Therefore, comprehensive evaluation of the m5C modification pattern would contribute to understanding the characteristics of TME cell infiltration. Considering the individual differences of m5C modification, we needed to quantify the m5C modification pattern of a single tumor. To this end, we constructed a scoring system to assess m5C modification patterns in bladder cancer individuals – m5C score. The m5C modification pattern characterized by immune-desert phenotype showed a higher m5Cscore, while the m5C modification pattern characterized by immune-excluded phenotype exhibited a

lower m5Cscore, and higher m5Cscore was associated with a greater survival advantage. Additionally, the immunophenotyping cohort well validated the above results (IMvigor210). These results suggested that m5Cscore could be used as a reliable tool to assess the individual m5C modification patterns comprehensively and to further determine tumor immunophenotypes. Our data also showed a positive correlation between m5Cscore and TMB and the expression of immune checkpoint PD-L1 and PD-1 was also significantly different in the high and low groups of m5Cscore. In our evaluation of clinical traits of patients assessed by m5Cscore, m5Cscore also showed strong predictive ability. M5Cscore can be used to evaluate the survival rate of patients at age, sex and stage of tumor progression. Patients with different clinical traits had significant differences in high and low m5CScore, and there were significant stratifications.

Previous studies showed that matrix activation played a critical role in response and resistance to checkpoint immunotherapy[41, 42]. In addition, activation of Tregs, expression of multiple anti-tumor growth factors, TGF-β and other pathways were associated with activation of EMT and matrix[43, 44]. It had also been proved that the activation of TGF-β and Tregs leads to a decrease in the amount of T cells entering the tumor and a decrease in the killing ability of the tumor[45]. In conclusion, matrix activation had the ability to weaken TME immunity and mediate resistance to immune checkpoint blocking therapy, thus influencing the individualized immunotherapy of bladder cancer. In our study, the methylation pattern of m5C played an important role in the formation of different substrates and the immune TME landscape, suggesting that m5C modification may influence the therapeutic effect of ICB immunotherapy. M5C gene markers include multiple biomarkers such as TMB, PD-L1 expression, neoantigen load, immune TME, Tregs expression and matrix, suggesting more effective immunotherapy prediction strategies. Furthermore, we substantiated the predictive value of m5Cscore in immunotherapy for anti-PD-L1 and anti-PD-1 cohorts. There were significant differences in patients' response to treatment and survival prognosis between high and low m5Cscores.

In conclusion, m5Cscore can comprehensively assesse the methylation modification pattern of m5C and the corresponding characteristics of immune cell infiltration in clinical practice, so as to further evaluate the immunophenotype of tumor. Meanwhile, m5Cscore can evaluate clinicopathological characteristics of patients, including clinical stage, tumor differentiation level, genetic variation, tumor mutation load and neoantigen load, etc. In addition, m5Cscore can be used as an effective independent prognostic biomarker to evaluate the prognosis of patients, and it can also be used to predict the clinical response to anti-PD-L1 /PD-1 immunotherapy. Moreover, our study has provided some novel insights into the mechanisms of TME landscape change and tumor immunotherapy. For example, modifying the mode of modification targeted at m5C regulation and genes associated with the m5C phenotype to further reverse the TME immune cell invasion or stromal invasion, which transforms the "cold tumor" into a "hot tumor". Contribute to the development of new immunotherapy drugs for cancer in the future. Our results provide new ideas for identifying different tumor immune TME phenotypes, improving patients' clinical response to immunotherapy, and promoting personalized treatment of tumors.

Conclusion

To sum up, this work confirmed the extensive regulatory mechanisms of the m5C modification pattern in the tumor microenvironment. The m5C modification pattern was a crucial factor to form the heterogeneity and complexity of TME landscape. Understanding and in-depth evaluation of individual m5C modification patterns would help us understand the TME landscape and guide better immunotherapy strategies.

Declarations

Acknowledgments

Thanks to my girlfriend for supporting me all the time.

Author contribution

Z Y, ZC L, and B C conceived of and designed the study. QQ T, XQ L and YT Z

performed literature search. Z Y, B C, and RR Z generated the figures and tables. K

Z, ZZ Y and YX G analyzed the data. Z Y and J C wrote the manuscript and JJ Z

critically reviewed the manuscript. YP L supervised the research. All authors have read

and approved the manuscript

Fund

The project was supported by the Project of Wenzhou Medical University (wyx2021101102) and National College Students Innovation and Entrepreneurship Training Program (No.202110343006)

Availability of data and materials

The datasets generated and/or analyzed during the current study are available in the

TCGA repository (https://portal.gdc.cancer.gov/) and GEO repository

(https://www.ncbi.nlm.nih.gov/).

Ethics approval and consent to participate

The research didn’t involve animal experiments and human specimens, no ethics related

issues.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Xue M, Shi Q, Zheng L, Li Q, Yang L, Zhang Y: **Gene signatures of m5C regulators may predict prognoses of patients with head and neck squamous cell carcinoma**. *Am J Transl Res* 2020, **12**(10):6841-6852.
2. Chen X, Li A, Sun BF, Yang Y, Han YN, Yuan X, Chen RX, Wei WS, Liu Y, Gao CC *et al*: **5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs**. *Nat Cell Biol* 2019, **21**(8):978-990.
3. Bohnsack KE, Höbartner C, Bohnsack MT: **Eukaryotic 5-methylcytosine (m□C) RNA Methyltransferases: Mechanisms, Cellular Functions, and Links to Disease**. *Genes (Basel)* 2019, **10**(2).
4. 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**(1):53.
5. Chi L, Delgado-Olguín P: **Expression of NOL1/NOP2/sun domain (Nsun) RNA methyltransferase family genes in early mouse embryogenesis**. *Gene Expr Patterns* 2013, **13**(8):319-327.
6. Van Haute L, Lee SY, McCann BJ, Powell CA, Bansal D, Vasiliauskaitė L, Garone C, Shin S, Kim JS, Frye M *et al*: **NSUN2 introduces 5-methylcytosines in mammalian mitochondrial tRNAs**. *Nucleic Acids Res* 2019, **47**(16):8720-8733.
7. Zhang Y, Zhang X, Shi J, Tuorto F, Li X, Liu Y, Liebers R, Zhang L, Qu Y, Qian J *et al*: **Dnmt2 mediates intergenerational transmission of paternally acquired metabolic disorders through sperm small non-coding RNAs**. *Nat Cell Biol* 2018, **20**(5):535-540.
8. Yang X, Yang Y, Sun BF, Chen YS, Xu JW, Lai WY, Li A, Wang X, Bhattarai DP, Xiao W *et al*: **5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m(5)C reader**. *Cell Res* 2017, **27**(5):606-625.
9. 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.
10. Chellamuthu A, Gray SG: **The RNA Methyltransferase NSUN2 and Its Potential Roles in Cancer**. *Cells* 2020, **9**(8).
11. Trixl L, Lusser A: **The dynamic RNA modification 5-methylcytosine and its emerging role as an epitranscriptomic mark**. *Wiley Interdiscip Rev RNA* 2019, **10**(1):e1510.
12. Tang B, Yan X, Sheng X, Si L, Cui C, Kong Y, Mao L, Lian B, Bai X, Wang X *et al*: **Safety and clinical activity with an anti-PD-1 antibody JS001 in advanced melanoma or urologic cancer patients**. *J*

Hematol Oncol 2019, **12**(1):7.

13. Brahmer JR, Tykodi SS, Chow LQ, Hwu WJ, Topalian SL, Hwu P, Drake CG, Camacho LH, Kauh J, Odunsi K *et al*: **Safety and activity of anti-PD-L1 antibody in patients with advanced cancer**. *N Engl J Med* 2012, **366**(26):2455-2465.
14. Shibata M, Ham K, Hoque MO: **A time for YAP1: Tumorigenesis, immunosuppression and targeted therapy**. *Int J Cancer* 2018, **143**(9):2133-2144.
15. Chambers AM, Lupo KB, Matosevic S: **Tumor Microenvironment-Induced Immunometabolic Reprogramming of Natural Killer Cells**. *Front Immunol* 2018, **9**:2517.
16. Bader JE, Voss K, Rathmell JC: **Targeting Metabolism to Improve the Tumor Microenvironment for Cancer Immunotherapy**. *Mol Cell* 2020, **78**(6):1019-1033.
17. Subudhi SK, Vence L, Zhao H, Blando J, Yadav SS, Xiong Q, Reuben A, Aparicio A, Corn PG, Chapin BF *et al*: **Neoantigen responses, immune correlates, and favorable outcomes after ipilimumab treatment of patients with prostate cancer**. *Sci Transl Med* 2020, **12**(537).
18. 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.
19. Wu J, Hou C, Wang Y, Wang Z, Li P, Wang Z: **Comprehensive Analysis of m(5)C RNA Methylation Regulator Genes in Clear Cell Renal Cell Carcinoma**. *Int J Genomics* 2021, **2021**:3803724.
20. Xue S, Xu H, Sun Z, Shen H, Chen S, Ouyang J, Zhou Q, Hu X, Cui H: **Depletion of TRDMT1 affects 5-methylcytosine modification of mRNA and inhibits HEK293 cell proliferation and migration**. *Biochem Biophys Res Commun* 2019, **520**(1):60-66.
21. Hänzelmann S, Castelo R, Guinney J: **GSVA: gene set variation analysis for microarray and RNA-seq data**. *BMC Bioinformatics* 2013, **14**:7.
22. Xiao B, Liu L, Li A, Xiang C, Wang P, Li H, Xiao T: **Identification and Verification of Immune-Related Gene Prognostic Signature Based on ssGSEA for Osteosarcoma**. *Front Oncol* 2020, **10**:607622.
23. Yi M, Nissley DV, McCormick F, Stephens RM: **ssGSEA score-based Ras dependency indexes derived from gene expression data reveal potential Ras addiction mechanisms with possible clinical implications**. *Sci Rep* 2020, **10**(1):10258.
24. Kempowsky-Hamon T, Valle C, Lacroix-Triki M, Hedjazi L, Trouilh L, Lamarre S, Labourdette D, Roger L, Mhamdi L, Dalenc F *et al*: **Fuzzy logic selection as a new reliable tool to identify molecular grade signatures in breast cancer--the INNODIAG study**. *BMC Med Genomics* 2015, **8**:3.
25. Lu X, Lu X, Wang ZC, Iglehart JD, Zhang X, Richardson AL: **Predicting features of breast cancer with gene expression patterns**. *Breast Cancer Res Treat* 2008, **108**(2):191-201.
26. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, Kadel EE, III, Koeppen H, Astarita JL, Cubas R *et al*: **TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells**. *Nature* 2018, **554**(7693):544-548.

27. Zeng D, Li M, Zhou R, Zhang J, Sun H, Shi M, Bin J, Liao Y, Rao J, Liao W: **Tumor Microenvironment Characterization in Gastric Cancer Identifies Prognostic and Immunotherapeutically Relevant Gene Signatures**. *Cancer Immunol Res* 2019, **7**(5):737-750.
28. Liu Q, Li C, Wanga V, Shepherd BE: **Covariate-adjusted Spearman's rank correlation with probability-scale residuals**. *Biometrics* 2018, **74**(2):595-605.
29. Hazra A, Gogtay N: **Biostatistics Series Module 3: Comparing Groups: Numerical Variables**. *Indian J Dermatol* 2016, **61**(3):251-260.
30. Li Z, Zhang H: **Reprogramming of glucose, fatty acid and amino acid metabolism for cancer progression**. *Cell Mol Life Sci* 2016, **73**(2):377-392.
31. Thiem A, Hesbacher S, Kneitz H, di Primio T, Heppt MV, Hermanns HM, Goebeler M, Meierjohann S, Houben R, Schrama D: **IFN-gamma-induced PD-L1 expression in melanoma depends on p53 expression**. *J Exp Clin Cancer Res* 2019, **38**(1):397.
32. Huang X, Qiu Z, Li L, Chen B, Huang P: **m6A regulator-mediated methylation modification patterns and tumor microenvironment infiltration characterization in hepatocellular carcinoma**. *Aging (Albany NY)* 2021, **13**(16):20698-20715.
33. Gao X, Huang H, Wang Y, Pan C, Yin S, Zhou L, Zheng S: **Tumor Immune Microenvironment Characterization in Hepatocellular Carcinoma Identifies Four Prognostic and Immunotherapeutically Relevant Subclasses**. *Front Oncol* 2020, **10**:610513.
34. Li J, Wang W, Zhou Y, Liu L, Zhang G, Guan K, Cui X, Liu X, Huang M, Cui G *et al*: **m6A Regulator-Associated Modification Patterns and Immune Infiltration of the Tumor Microenvironment in Hepatocarcinoma**. *Front Cell Dev Biol* 2021, **9**:687756.
35. Yang Q, Yu XL, Wang Y, Cheng ZG, Han ZY, Liu FY, Qian TG, Yu J, Liang P: **Predictive effects of a combined indicator in patients with hepatocellular carcinoma after thermal ablation**. *J Cancer Res Ther* 2020, **16**(5):1038-1050.
36. Peranzoni E, Rivas-Caicedo A, Bougherara H, Salmon H, Donnadieu E: **Positive and negative influence of the matrix architecture on antitumor immune surveillance**. *Cell Mol Life Sci* 2013, **70**(23):4431-4448.
37. Savino W, Mendes-Da-Cruz DA, Smaniotto S, Silva-Monteiro E, Villa-Verde DM: **Molecular mechanisms governing thymocyte migration: combined role of chemokines and extracellular matrix**. *J Leukoc Biol* 2004, **75**(6):951-961.
38. Chen DS, Mellman I: **Elements of cancer immunity and the cancer-immune set point**. *Nature* 2017, **541**(7637):321-330.
39. Pacella I, Grimaldi A, Piconese S: **Assessment of lipid load in tumor-infiltrating Tregs by flow cytometry**. *Methods Enzymol* 2020, **632**:283-294.
40. Chen J, Tan Y, Sun F, Hou L, Zhang C, Ge T, Yu H, Wu C, Zhu Y, Duan L *et al*: **Single-cell transcriptome and antigen-immunoglobin analysis reveals the diversity of B cells in non-small cell lung cancer**. *Genome Biol* 2020, **21**(1):152.

41. Tamura R, Tanaka T, Akasaki Y, Murayama Y, Yoshida K, Sasaki H: **The role of vascular endothelial growth factor in the hypoxic and immunosuppressive tumor microenvironment: perspectives for therapeutic implications**. *Med Oncol* 2019, **37**(1):2.
42. Voron T, Marcheteau E, Pernot S, Colussi O, Tartour E, Taieb J, Terme M: **Control of the immune response by pro-angiogenic factors**. *Front Oncol* 2014, **4**:70.
43. de Streel G, Bertrand C, Chalon N, Liénart S, Bricard O, Lecomte S, Devreux J, Gaignage M, De Boeck G, Mariën L *et al*: **Selective inhibition of TGF-β1 produced by GARP-expressing Tregs overcomes resistance to PD-1/PD-L1 blockade in cancer**. *Nat Commun* 2020, **11**(1):4545.
44. Calon A, Espinet E, Palomo-Ponce S, Tauriello DV, Iglesias M, Céspedes MV, Sevillano M, Nadal C, Jung P, Zhang XH *et al*: **Dependency of colorectal cancer on a TGF-β-driven program in stromal cells for metastasis initiation**. *Cancer Cell* 2012, **22**(5):571-584.
45. Schmidt A, Éliás S, Joshi RN, Tegnér J: **In Vitro Differentiation of Human CD4+FOXP3+ Induced Regulatory T Cells (iTregs) from Naïve CD4+ T Cells Using a TGF-β-containing Protocol**. *J Vis Exp* 2016(118).

Figures

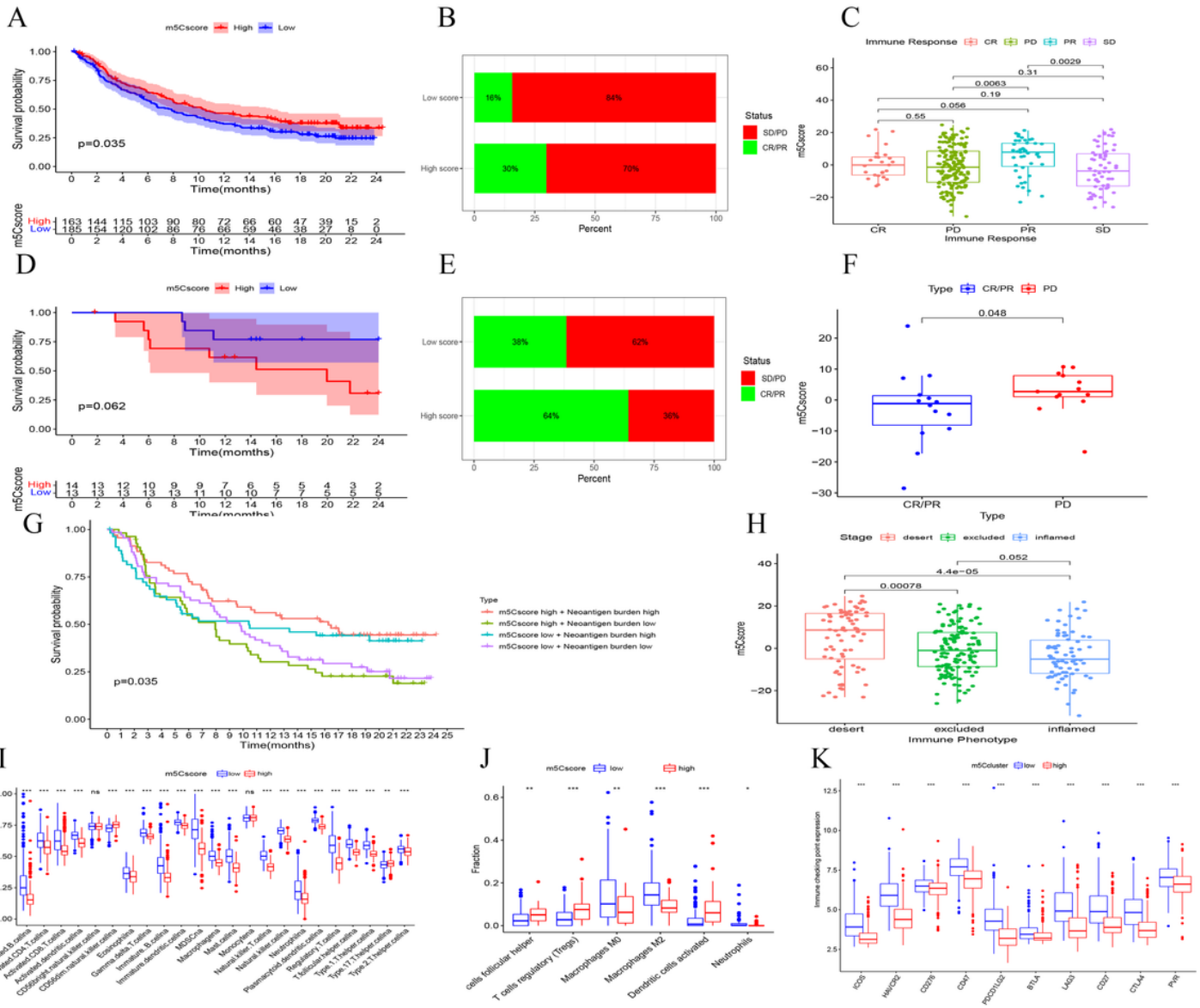

Figure 1

The technical route of this study

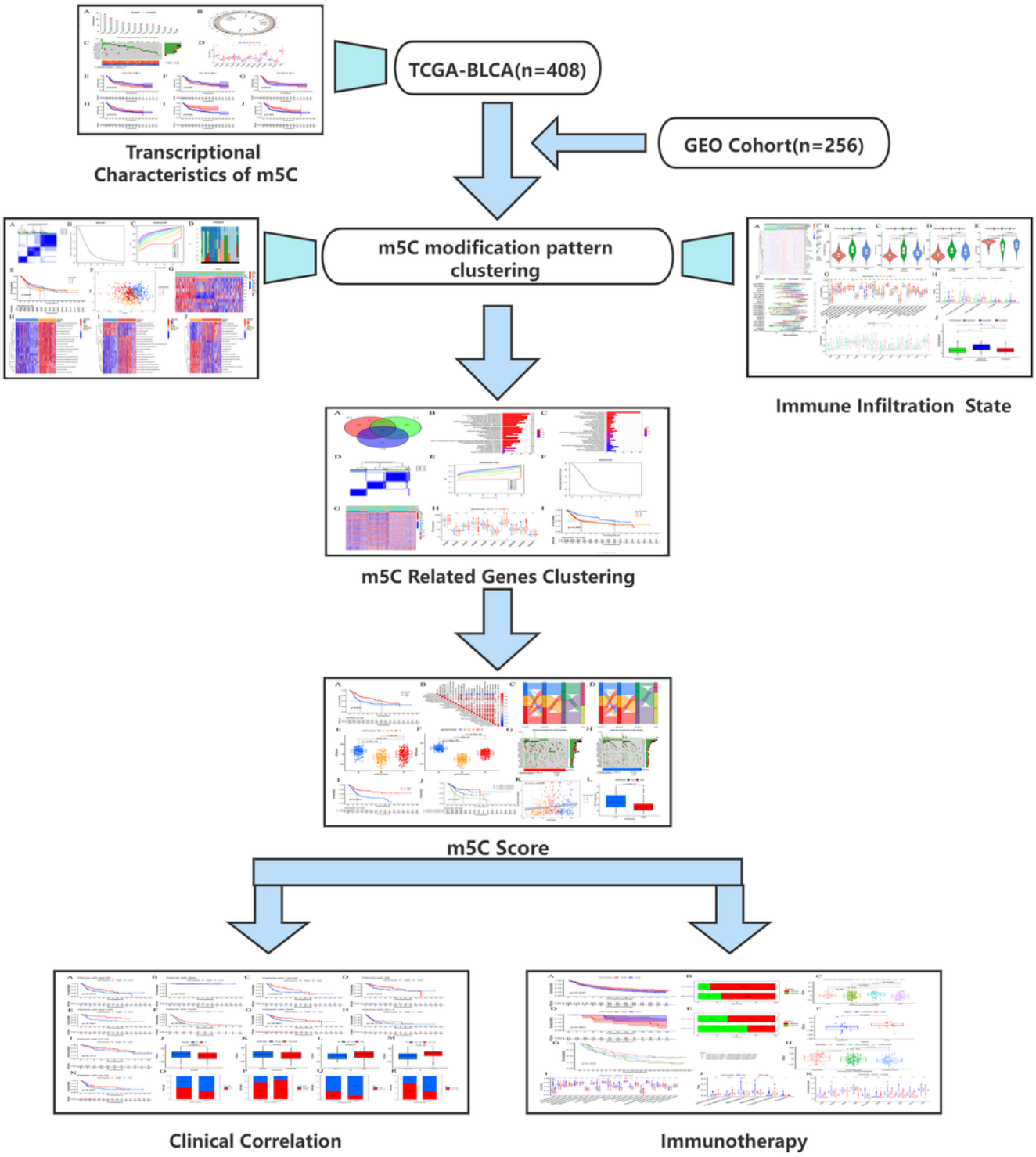

Figure 2

CNV, mutation and prognostic characteristics of m5C regulators in bladder cancer. **A** The CNV variation frequency of m5C regulators in TCGA-BLCA cohort. The height of the column represented the alteration frequency. The deletion frequency, blue dot; The amplification frequency, red dot. **B** The location of CNV alteration of m5C regulators on 23 chromosomes in TCGA-BLCA cohort. **C** The expression of 13 m5C regulators between normal and gastric tissues. Tumor, red; Normal, blue. The upper and lower ends of the

boxes represented interquartile range of values. The lines in the boxes represented median value, and black dots showed outliers. The asterisks represented the statistical P value (*P< 0.05; **P< 0.01; ***P< 0.001) **D** The mutation frequency of 13 m5C regulators in 409 patients with gastric cancer from TCGA-BLCA cohort. Each column represented individual patients. The upper barplot showed TMB, the number on the right indicated the mutation frequency in each regulator. The right barplot showed the proportion of each variant type. The stacked barplot below showed fraction of conversions in each sample. **E-G** Correlation analysis of mutation and expression of m5C regulatory factor in TCGA-BLCA cohort. The vertical axis represents the expression level of the target gene. Blue is the wild type of another gene, red is the mutant type of another gene. **H-M** Survival analysis for low and high m5C regulators expression groups of patients using Kaplan-Meier curves (P< 0.05 was considered significant).

Figure 3

Distinct modification patterns and biological characteristics of each m5C cluster. **A-D** Three m5C methylation modification patterns classification and verification. **E** Survival analysis of three different m5C modification patterns based on GEO cohort, including 195 cases in m5Ccluster A, 140 cases in m5Ccluster B, 231 cases in m5Ccluster C. The Log-rank p value of Kaplan-Meier was 0.026, which indicated that there are significant discrepancies among the three modification patterns. The survival advantage of m5Ccluster A was mas prominently higher than the other clusters. **F** Principal component analysis of the transcriptome spectra of the three modification patterns showed that there were significant differences among these three modification patterns. **G** Enrichment diagram of m5C regulators and clinical information in three modification patterns. Red indicated high expression and blue indicated low expression. **H-J** GSVA analysis for each two clusters, showing pathways in which genes were enriched in each m5Ccluster. Red indicated high expression and blue indicated low expression.

Figure 4

TME and immune cell infiltration in distinct m5C modification patterns. **A** Abundance of immune cell infiltration in three m5C modification patterns based on the expression level of 29 immune genes sets. Red indicated high expression and blue indicated low expression. **B-E** The immune score, stromal score, estimate score and tumor purity of each m5C modification pattern. **F** Differences in HLA gene family among three m5C modification patterns. The asterisks represented the statistical P value (*P< 0.05; **P< 0.01; ***P< 0.001). **G** Infiltrate abundance of immune cells under three m5C modification patterns by GSEA analysis method. The asterisks represented the statistical P value (*P< 0.05; **P< 0.01; ***P< 0.001). **H** Cibersoft method was used to analyze the strength of cellular immunity in three m5C modified patterns. The asterisks represented the statistical P value (*P< 0.05; **P< 0.01; ***P< 0.001). **I-J** Expression differences of immune checkpoint in three m5C modification patterns. The asterisks represented the statistical P value (*P< 0.05; **P< 0.01; ***P< 0.001).

Figure 5

Three gene clusters and its corresponding modification pattern. **A** Venn plot exhibited 902 DEGs among three m5Cclusters. **B** Functional annotation for m5Ccluster related DEGs using GO analysis. The color depth of the plots represented the number of genes enriched. **C** Functional annotation for m5Ccluster related DEGs using KEGG analysis. The color depth of the plots represented the number of genes enriched. **D- F** Three m5C gene modification patterns classification and verification. **G** Heatmap of 377 prognostic genes associated with m5C modification phenotype, and unsupervised clustering based on these 377 prognostic genes classified patients into three gene genomic clusters, termed as m5C gene cluster A-C, respectively. We used fustat, N, T, gender, age for patient annotation. **H** The expression of m5C regulators in three distinct m5C gene clusters. The lines in the boxes represented median value, and black dots showed outliers. The asterisks represented the statistical P value (*P< 0.05; **P< 0.01; ***P< 0.001). **I** Kaplan-Meier curves for three m5C gene clusters. Patient prognosis of three distinct m5C modification phenotype were of remarkably difference with 192 cases in gene cluster A, 160 cases in gene cluster B and 214 cases in gene cluster C (P < 0.001).

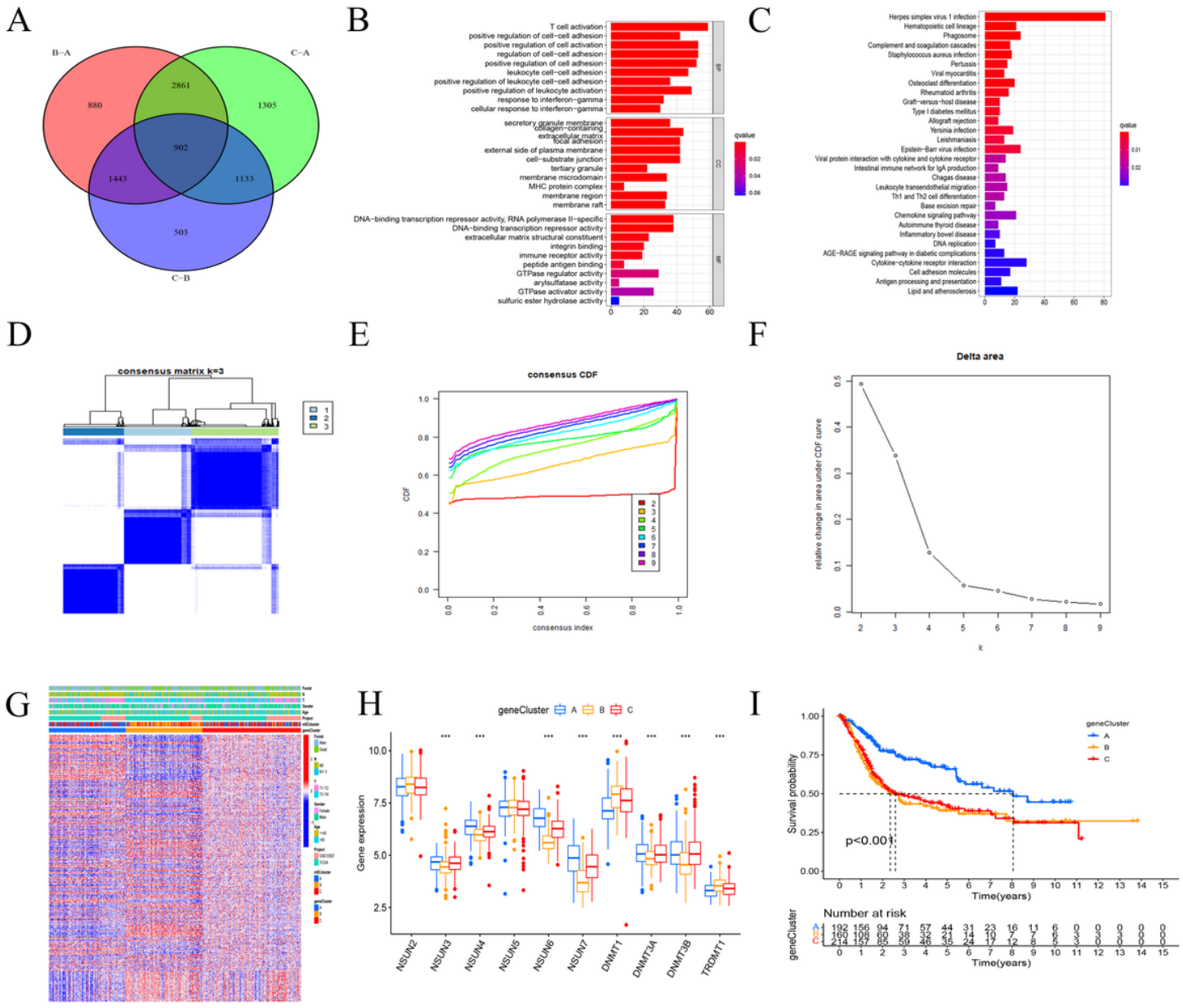

Figure 6

Transcriptional and TMB characteristics of two m5Cscore groups. **A** Survival analyses for low (304 cases) and high (262 cases) m5Cscore patient groups in the mixed cohort using Kaplan-Meier curves (P < 0.001). **B** Correlations between m5Cscore and known gene signatures. Red indicated positive and blue indicated negative. **C-D** Alluvial map showed the changes of m5C clusters, gene clusters, m5Cscore and clinical characteristics (gender and survival state). **E-F** Differences of m5Cscore in m5C cluster and gene cluster. The Kruskal Wallis test was used to compare the statistical difference between three m5C clusters (P < 0.001) and gene clusters (P < 0.001). **G-H** Waterfall plot depicted the characteristics of TMB in the low and high m5Cscore groups. **I** Survival analyses for low (268 cases) and high (126 cases) TMB patient groups using Kaplan-Meier curves (P < 0.001). **J** M5Cscore and TMB were combined to describe the survival status of patients, and there were differences in the four groups of survival analysis (P < 0.001). **K** The scatter plot showed that there was a negative correlation between TMB and m5Cscore. **L**

The boxplot showed that the expression of PD-L1 was significantly different in the high and low m5Cscore groups.

Figure 7

Comprehensive evaluation of clinical significance of m5Cscore. **A-I, N** Survival analysis for m5Cscore in different clinical groups (Age, survival state, gender, T stage, N stage) using Kaplan-Meier curves. $P < 0.05$ was considered as significant difference. **J-M** Kruskal Wallis test was used to determine whether m5Cscore was different in patients with different clinical traits (survival state, gender, T stage, N stage). **O-R** The histogram showed the distribution of patients with different m5Cscore in different clinical characteristics.

Figure 8

M5C modification patterns in the role of anti-PD-1/L1 immunotherapy and the revalidation of m5C modification patterns. **A** Survival analyses for low (185 cases) and high (163 cases) m6Ascore patient groups in the anti-PD-L1 immunotherapy cohort using Kaplan-Meier curves (IMvigor210 cohort; $P = 0.035$). **B** The proportion of patients with response to PD-L1 blockade immunotherapy in low or high m5Cscore groups. SD, stable disease; PD, progressive disease; CR, complete response; PR, partial response. **C** Distribution of m5Cscore in distinct anti-PD-L1 clinical response groups. **D** Survival analyses for low and high m5Cscore patient groups in the anti-PD1 immunotherapy cohort using Kaplan-Meier curves (GSE78220 cohort; $P=0.062$). **E** The proportion of patients with response to PD-1 blockade immunotherapy in low or high m5Cscore groups. **F** Distribution of m5Cscore in distinct PD-1 blockade clinical response groups. **G** Survival analyses for patients receiving anti-PD-L1 immunotherapy stratified by both m5Cscore and neoantigen burden using Kaplan-Meier curves. H, high; L, Low; NEO, neoantigen burden ($P < 0.0001$). **H** Differences in m6Ascore among distinct tumor immune phenotypes in IMvigor210 cohort. **I** Differences in expression of immune cells among high and low m5Cscore groups. **J** CIBERSOFT analysis was used to assess the cellular immunity strength among high and low m5Cscore groups. **K** Expression of immune checking point in high and low m5Cscore group.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- figureS1.jpg
- supplementary.docx