Check for updates

OPEN ACCESS

EDITED BY Elena Ciaglia, University of Salerno, Italy

REVIEWED BY You Lang Zhou, Affiliated Hospital of Nantong University, China

Jo A. Van Ginderachter, Vrije University Brussel, Belgium

*CORRESPONDENCE Kuntang Shen shen.kuntang@zs-hospital.sh.cn

[†]These authors have contributed equally to this work

SPECIALTY SECTION

This article was submitted to Cancer Immunity and Immunotherapy, a section of the journal Frontiers in Immunology

RECEIVED 26 March 2022 ACCEPTED 11 October 2022 PUBLISHED 27 October 2022

CITATION

Zhang Q, Sun X, Sun J, Lu J, Gao X, Shen K and Qin X (2022) RNA m⁵C regulator-mediated modification patterns and the cross-talk between tumor microenvironment infiltration in gastric cancer. *Front. Immunol.* 13:905057. doi: 10.3389/fimmu.2022.905057

COPYRIGHT

© 2022 Zhang, Sun, Sun, Lu, Gao, Shen and Qin. 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.

RNA m⁵C regulator-mediated modification patterns and the cross-talk between tumor microenvironment infiltration in gastric cancer

Qiang Zhang^{1,2†}, Xiangfei Sun^{1†}, Jianyi Sun^{1†}, Jiangshen Lu¹, Xiaodong Gao¹, Kuntang Shen^{1*} and Xinyu Qin¹

¹Department of General Surgery, Zhongshan Hospital, Fudan University, Shanghai, China, ²Department of Gastrointestinal Surgery, the Second People's Hospital of Lianyungang Affiliated to Kangda College, Nanjing Medical University, Lianyungang, Jiangsu, China

The effect of immunotherapy strategy has been affirmed in the treatment of various tumors. Nevertheless, the latent role of RNA 5-methylcytosine (m⁵C) modification in gastric cancer (GC) tumor microenvironment (TME) cell infiltration is still unclear. We systematically explore the m⁵C modification patterns of 2,122 GC patients from GEO and TCGA databases by 16 m⁵C regulators and related these patterns to TME characteristics. LASSO Cox regression was employed to construct the m⁵Cscore based on the expression of regulators and DEGs, which was used to evaluate the prognosis. All the GC patients were divided into three m⁵C modification clusters with distinct gene expression characteristics and TME patterns. GSVA, ssGSEA, and TME cell infiltration analysis showed that m⁵C clusters A, B, and C were classified as immune-desert, immune-inflamed, and immuneexcluded phenotype, respectively. The m⁵Cscore system based on the expression of eight genes could effectively predict the prognosis of individual GC patients, with AUC 0.766. Patients with a lower m⁵Cscore were characterized by the activation of immunity and experienced significantly longer PFS and OS. Our study demonstrated the non-negligible role of m⁵C modification in the development of TME complexity and inhomogeneity. Assessing the m⁵C modification pattern for individual GC patients will help recognize the infiltration characterization and guide more effective immunotherapy treatment.

KEYWORDS

5-methylcytosine (m^sC), RNA methylation modification, tumor microenvironment, immune, prognosis, m^sCscore

Introduction

As a global disease, gastric cancer (GC) is the fifth most diagnosed malignancy and the third most common cause of cancer-related death, with 784,000 deaths worldwide in 2018 (1). Although the incidence and mortality rates of GC have declined in several countries, regions seriously threatened by GC, such as China and other East Asian countries, still bear severe health and economic burden. In China, 562,000 newly diagnosed GC patients were recorded, accounting for nearly half of the new cases worldwide (2). The 5-year survival rate of GC is 35.9% in China due to the late stage at diagnosis, notably lower than 71.5% in South Korea and 65% in Japan (3, 4). Due to the complexity of the pathogenic mechanism and the lack of specific biomarkers of GC, the effects of treatment strategies such as surgery, chemotherapy, and radiotherapy are not satisfactory.

RNA 5-methylcytosine (m⁵C) is an important kind of RNA methylation modification; there have been more than 150 RNA modifications identified to date (5). Traditionally, DNA m⁵C has been proven to be the most dominant DNA modification in mammals and functions by adding a methyl group at the carbon-5 position of the cytosine base (6). RNA m⁵C modification, as the third layer of epigenetics, can be found in but not limited to mRNA, noncoding RNA, and tRNA (7-13). Like other RNA epigenetic modifications, such as N6methyladenosine (m⁶A), m⁵C is a dynamic reversible process that can be regulated by "writers", "erasers", and "readers", namely, the methyltransferases, demethylases, and binding proteins (14). The methylation formation of m⁵C modification is catalyzed by methyltransferases composed of the SUN/NOL1/ NOP2 domain family of proteins (NSUN1, NSUN2, NSUN3, NSUN4, NSUN5, NSUN6, and NSUN7) and DNA methyltransferase homologues (DNMT1, DNMT2, DNMT3A, and DNMT3B) (15-17). At the same time, the demethylation process is regulated by the erasers consisting of enzymes of the ten-eleven translocation (TET) family members, including TET1, TET2, and TET3 (18, 19). Additionally, the whole methylation process is mediated by a cluster of special RNAbinding proteins, including ALYREF and YBX1 (8, 20). An increasing number of studies validate that the dynamic modification of m⁵C and its regulators is involved in a series of physiological and pathological processes, including RNA stability, gene expression, and protein synthesis. As for tumor malignant biological behaviors, it has been reported that m⁵C and its regulators play essential roles in the pathogenesis of leukemia (21), hepatocellular carcinoma (HCC) (22), glioblastoma multiforme (GBM) (12), and bladder cancer (20), indicating the promising prospect of m⁵C modification in cancer treatment.

Recently, immunotherapy, anti-PD-L1 antibody, and anti-PD-1 antibody have increased the overall survival rate of some advanced GC patients who were treated with two or more lines of

chemotherapy (23). The efficiency of immunotherapy depends on the status of EB virus infection, microsatellite instability (MSI)/ mismatch repair (MMR), and the expression of PD-L1. However, the dominant population of immunotherapy is still challenging to identify because of the heterogeneity of GC. Hence, to better analyze the heterogeneity and immunophenotype of patients with GC, it is essential to improve long-term survival. Consistently, epigenetic and genetic variations of malignant cells are the only factors participating in the tumor progression, which is a complex multistep process. Notwithstanding, numerous studies have proved that the tumor microenvironment (TME), where tumor cells survive and grow, is crucial in tumorigenesis and development. The composition of TME is rather complicated, including not only the tumor part but also the stromal cells, macrophages, bone marrow-derived cells (BMDCs), distant recruited cells, secreted factors, and neovascularization (24). The detailed types of cells and cytokines in the TME are complex, including cancer-associated fibroblasts (CAFs), myeloid cells, lymphocytes, chemokines, cytokines, and growth factors. Among these cells, tumorassociated macrophages (TAMs), tumor-associated neutrophils (TANs), myeloid-derived suppressor cells (MDSCs), Tie2expressing monocytes, and dendritic cells together constitute the tumor-associated myeloid cells (TAMCs) (25). The cross-talk between cancer cells and TME components promotes tumor proliferation and angiogenesis, avoids hypoxia, inhibits apoptosis, and mediates immune tolerance. With the gradual deepening of the understanding of the complexity and diversity of TME, increasing data depict its essential role in immune escape and immunotherapy. Moreover, the TME cell infiltration pattern can predict the response to the immune checkpoint blockade (ICB), which is promising in the tumor treatment strategies (26). Accordingly, particular tumor immunophenotypes are supposed to be validated via thoroughly parsing the TME landscape complexity and heterogeneity (27). As GC is characterized by tumor heterogeneity, it is urgent to identify the dominant population of immunotherapy by the landscape TME cell infiltration.

Lately, m⁵C modification is related to the TME-infiltrating immune cells, and the mechanisms are more complicated than expected. In systemic lupus erythematosus (SLE), abnormal m⁵C mRNAs were identified as relevant to critical immune pathways in CD4+ T cells (28). Another study reported that the eraser TET1 is downregulated *via* NF-κB signaling pathway activation in breast cancer cells (29). Interestingly, Andries and colleagues found that m⁵C-modified mRNA promoted protein expression by the increased ability of the mRNA to elude downstream innate immune signaling and activation of endosomal Toll-like receptor 3 (TLR3) (30). During virus infection, m⁵C RNA methyltransferases, such as NSUN family proteins, were employed to modify viral RNA and change antiviral host responses (31). All these latest findings reveal the fact that m5C modification and regulators may have a further effect on the TME, and previous studies focus only on one or two m⁵C regulators due to the limitation of technologies.

In the present study, the genomic and clinical data of 1,983 GC samples were employed to thoroughly estimate the m^5C modification patterns and the correlation between m^5C modification and TME features. Three different m^5C modification patterns and the specific TME cell infiltration peculiarities were identified. Three immunophenotypes, immune-inflamed, immune-excluded, and immune-desert phenotype, were related to the three m^5C clusters. Subsequently, a scoring system based on the m^5C modification pattern was established for individual GC patients.

Materials and methods

The detailed materials and methods can be found in the supplementary files (32–37).

Results

Blueprint of genetic variation of m⁵C regulators in GC

In the process of dynamic modification, methyltransferases and demethylases work together to keep the balance of the RNA m⁵C modification with the help of the readers. The ideograph of RNA m⁵C modification is shown in Figure 1A. Firstly, the characteristics of somatic mutations and copy number variations (CNVs) of the 16 m⁵C regulators were summarized in GC. Among all the 433 samples from TCGA, 83 (19.17%) patients experienced mutations of m⁵C regulators. We found that the three demethylases exhibited the highest mutation rates, while the readers (YBX1 and ALYREF) hardly showed any mutations (Figure 1B). Moreover, a significant mutation cooccurrence pattern was identified between NSUN2 and NSUN3 (Figure S1B). For CNV analysis, the most prevalent CNV alternation in the regulators was the amplification in copy number, except for NSUN3, TET2, and NSUN7, which were characterized by a high frequency of CNV deletion (Figure 1C). In Figure 1D, the detailed locations of CNV alternation of each m⁵C regulator are recorded on the chromosomes. Notably, we could thoroughly determine GC patients from normal samples based on the expression of the 16 m⁵C regulators (Figure 1E). To further ascertain the relation between the above genetic alternations and the expression of m⁵C regulators, we explored the expression of regulators in both GC and normal tissues. We found that CNVs might be the main factors leading to the abnormal expression of the m5C regulators. Regulators with amplificated CNV tended to highly expressed in tumor samples (e.g., DNMT1, ALYREF, and NSUN5), and vice versa (e.g., NSUN7 and NSUN6) (Figures 1C, F). The assessment disclosed the heterogeneity of expressional and genetic alternation patterns in m⁵C regulators between GC and

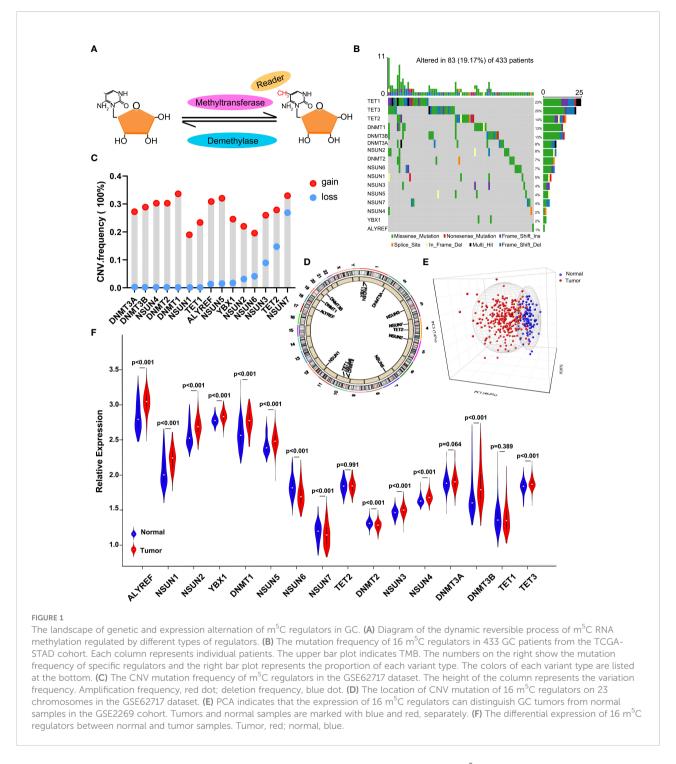
normal tissues, hinting that the aberrant expression of m^5C regulators played an essential role in the tumorigenesis and development of GC.

m⁵C methylation modification patterns mediated by 16 regulators

A meta-cohort including five GEO datasets (GSE57303, GSE84437, GSE34942, GSE62254, and GSE15459) with full OS and other clinical data was used to identify the expression pattern of 16 regulators. The prognostic values of 16 m⁵C regulators were analyzed through a univariate Cox regression model (Figure S1C and Figure 2A). We found that the readers ALYREF and YBX1 were favorable prognosis factors for GC patients. The cross-talk between 16 regulators and prognostic significance for patients was revealed in the m⁵C regulator network (Figures 2A, B). We noticed that a significant correlation was shown in both the same and different functional category regulators. Interestingly, the correlation of expression is consistent in regulators from the same functional category. However, we found that the relationship in writers is much complicated, such as DNMT1, which is remarkably negatively correlated with NSUN6 and NSUN7 (Figure 2B). In addition, the expression of the readers ALYREF and YBX1 was almost significantly correlated with other regulators. According to the expression of 16 m⁵C regulators, we further explored the m⁵C modification patterns via the ConsensusClusterPlus package, and identified three different modification patterns by the unsupervised clustering method, including 308 patients in m⁵C cluster A, 334 patients in m⁵C cluster B, and 417 patients in m⁵C cluster C (Figures S2A–D and Table S3). The heatmap of the 16 m⁵C regulators in 1,059 GC patients is depicted in Figure 2C. The expression of 16 regulators in three m⁵C clusters was remarkably different. LogRank analysis showed that the prognosis of patients in m⁵C cluster B was better than the other two clusters (Figure 2D).

TME cell infiltration characteristics in distinct m⁵C modification patterns

To better understand the biological characteristics among the distinct m⁵C modification clusters, the GSVA enrichment method was conducted. In Figure 2E, m⁵C cluster A is related to the immune suppression process, while m⁵C cluster B is notably enriched in immune full activation pathways, including cytokine–cytokine receptor interaction, natural killer cellmediated cytotoxicity, antigen processing and presentation, Toll-like receptor signaling pathway, and chemokine signaling pathway. m⁵C cluster C is enriched in carcinogenic and stromal activation pathways, such as ECM receptor interaction, TGF beta signaling pathway, adhesion and gap junction, mTOR, and



MAPK signaling pathways (Figure 2F). Interestingly, TME immune cell infiltration analysis subsequently showed that m⁵C cluster C was rich in resting and naïve immune cells, such as dendritic cells, CD4 memory T cells, mast cells, B cells, and other innate immune cells, by the CIBERSORT method. On the contrary, m⁵C cluster B is characterized by specific immune cell enrichment (Figure 3A, Figure S2E, and Table S4). The

correlation of specific m⁵C regulators and immune cell is shown in Figure S2F. To further reveal the TME features, the singlesample GSEA (ssGSEA) analysis of all the 1,059 cases was conducted. In addition to immune cells, more details about immune functions and pathways can be summarized *via* the ssGSEA method. As shown in Figure 3B, three distinct immune patterns under three m⁵C clusters are identified (Table S5).

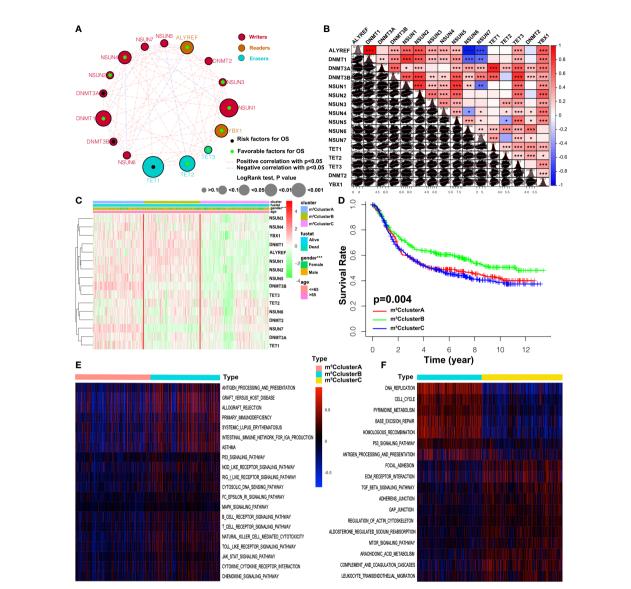
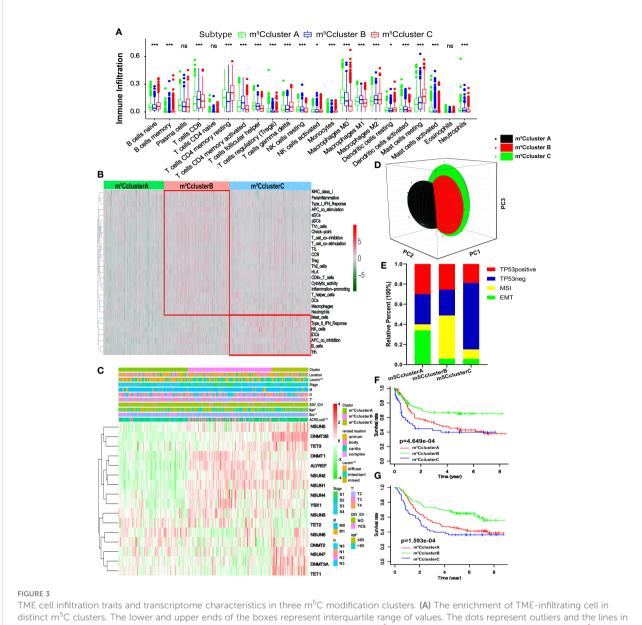


FIGURE 2

 m^5C methylation modification patterns and biological characteristics. (A) The interaction between 16 m^5C regulators in GC. The size of the circles represents the effect of a specific regulator on the prognosis, and the values calculated *via* the Log-rank method were p > 0.1, p < 0.0, p < 0.05, p < 0.01, and p < 0.001, respectively. Black dots, favorable factors; green dots, risk factors. The lines between two regulators show the interaction. Negative correlation and positive correlation are marked with blue and red. (B) The expression correlation of the 16 m^5C regulators in GC. Each cell in the lower left corner represents the expression of two regulators, and the red line is the trend line. Cells in the upper right corner show the statistic result. The asterisks represent the *p*-value (*p < 0.05, **p < 0.01, and ***p < 0.001). (C) Unsupervised clustering of 16 m^5C regulators in the five independent gastric cancer cohorts. The m⁶A clusters and basic clinical information are used as patient annotations. Each column represents patients and each row represents m^5C regulators. (D) Overall survival of three m^5C modification clusters based on the GC patients from five GEO cohorts. Log-rank *p*-value 0.004 indicates a significant prognostic difference among three m^5C clusters. (E, F) GSVA enrichment analysis revealing the activation states of biological pathways in distinct m^5C clusters. The heatmap is employed to visualize the biological processes, and blue represents inhibited pathways and yellow represents activated pathways.

Combined with the survival results above, we were surprised to find that m^5C cluster A belonged to the immune-desert phenotype, characterized by immunological suppression; m^5C cluster B was classified as immune-inflamed phenotype, which features immune activation and immune cell infiltration; m^5C cluster C was labeled as immune-excluded phenotype, characterized by stromal activation and innate immune cell infiltration (Figures 2D and 3A, B). These results demonstrated that the interaction among the writers, erasers, and readers might play fundamental roles in distinct m5C modification patterns and TME cell infiltration characteristics of individual GC patients.



TME cell infiltration traits and transcriptome characteristics in three m⁵C modification clusters. (A) The enrichment of TME-infiltrating cell in distinct m⁵C clusters. The lower and upper ends of the boxes represent interquartile range of values. The dots represent outliers and the lines in the boxes show the median value. (B) The heatmap of TME-infiltrating cells in three m⁵C patterns. (C) Unsupervised clustering of 16 m⁵C regulators in the ACRG cohort. The m⁵C cluster, location, Lauren types, tumor stage, EBV status, sex, age, and ACRG molecular subtypes are used as patient annotations. High and low expression of regulators is marked with red and green, respectively. (D) PCA shows that the expression of 16 m⁵C clusters. (F) Recurrence-free survival of three m⁵C modification clusters based on the GC patients from the ACRG cohort. (G) Overall survival of three m⁵C modification clusters based on the GC patients from the ACRG cohort. **p* < 0.05; ***p* < 0.01; ****p* < 0.001).

m⁵C methylation modification patterns in the ACRG cohort

We focused on the ACRG cohort, a group of 300 GC participants with complete clinicopathological information, to further reveal the biological behaviors and the features of m^5C modification patterns. Like the meta-cohort datasets, the ACRG cohort is divided into three distinct m^5C modification clusters as

well by the unsupervised clustering method (Figures S3A–D and Figures 3C, D). The heatmap based on the expression of 16 m⁵C regulators shows that m⁵C cluster A exhibits a high expression of TET2 and NSUN6 and is downregulated in other regulators; m⁵C cluster B is characterized by the upregulated readers and five writers including NSUN1–4 and DNMT1; m⁵C cluster C shows high levels of two erasers and four writers (Figure 3C). We found that patients in m⁵C cluster A were rich in the diffuse

subtype and tended to have poor differentiation. Intestinal subtype was more likely to be observed in m⁵C cluster B and C modification pattern. GC patients with diffuse histological and EMT molecular subtypes were related to poorer survival; on the contrary, MSI was linked to a better prognosis. Consequently, patients with an m⁵C cluster pattern were markedly linked to high malignancy, stromal activation, and poor prognosis (Figure 3C). Moreover, Figure 3E shows that patients in m^5C cluster A are also significantly related to EMT molecular subtypes, while the MSI subtypes are featured by m⁵C cluster B modification. The survival results revealed that patients in m⁵C cluster B are related to a favorable prognosis, while m⁵C clusters A and C show a shorter survival time (Figure 3F). Notably, we also found that the relapse-free survival (RFS) of m^5C cluster B is better than the other two clusters (Figure 3G). The findings above demonstrate that most GC patients with EMT molecular subtypes were divided into m⁵C cluster A and related to stromal activation; most patients with MSI instead of the EMT subtype were in m⁵C cluster B and characterized by immune activation.

Immunomodulatory effect of m^5C modification on the TME

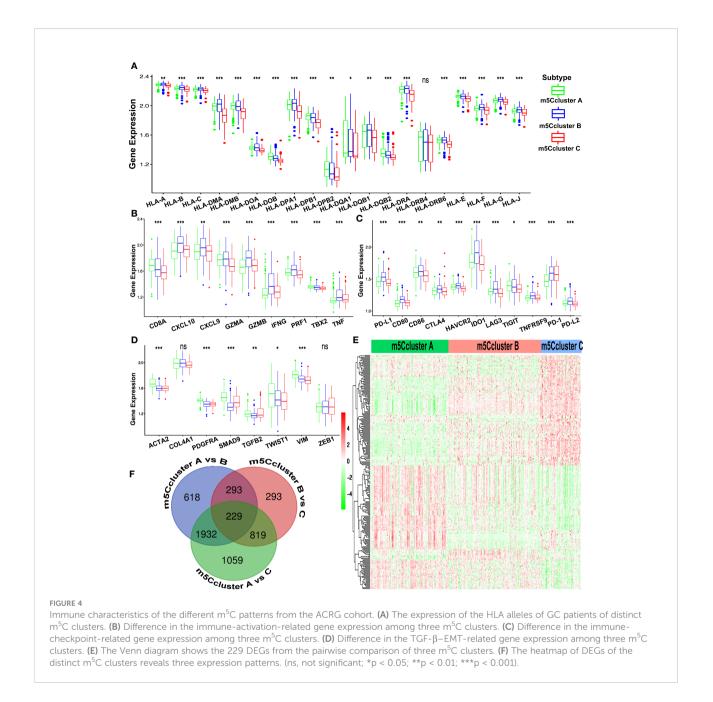
Subsequently, four gene clusters belonging to distinct immune processes were used to reveal the role of m5C modification on the immune regulation of the TME. Chemokines and cytokines with different functions were selected from the published literature. The essential members of human leukocyte antigen (HLA), the major histocompatibility complex (MHC) of human beings, present antigen and mediate immune response. CD8A, CXCL9, CXCL10, GZMA, GZMB, IFNG, PRF1, TBX2, and TNF are related to immune activation. CD80, CD86, HAVCR2, CTLA-4, LAG3, IDO1, PD-1, PD-L1, PD-L2, TNFRSF9, and TIGIT are supposed to correlate with immune checkpoints. ACTA2, CLDN3, COL4A1, SMAD9, TGRB1, TGFBR2, TWIST1, VIM, and ZEB1 are considered to associate with TGF-β and EMT pathways (24, 38). In Figure 4A, HLA-I molecules, including HLA-A, B, C, E, F, and G, are remarkably highly expressed in m⁵C cluster B, which means stronger antigen presentation and tumor-killing ability. We noted that HLA-II molecules, such as HLA-DPB2, HLA-DQA1, HLA-DQB2, and HLA-DQA1, were upregulated in m⁵C cluster A. HLA-G is reported to suppress the immune response and leads to long-term immune escape and tolerance (39). Meanwhile, we also found that the expression of genes related to TGF-B and EMT pathways was remarkably upregulated in m⁵C cluster A, which added the evidence of stromal activation, while m5C cluster B exhibited overexpression of mRNAs relevant to immune activation (Figures 4B-D). Immune checkpoint analysis showed that all the genes, including CTLA-4, PD-1, and PD-L1, were upregulated in m⁵C cluster B (Figure 4C). The results above demonstrate that m⁵C modification patterns are significantly related to TME immune regulation and may play crucial roles in

immunotherapy. However, these findings were only based on the $16 \text{ m}^5\text{C}$ modification regulators.

Considering the heterogeneity and complexity of m^5C methylation modification, we tried to identify the DEGs under different m^5C clusters using the limma package. Finally, 229 m^5C phenotype-related DEGs were found and showed a distinct expression pattern under three m^5C clusters (Figures 4E, F). The GO and KEGG enrichment analysis of the 229 DEGs showed that (Figures S3E, F) the DEGs were rich in immune-related biological processes and pathways, including CD8+ $\alpha\beta$ T cell activation, negative regulation of the immune system process, NOD-like receptor signaling pathway, and TNF signaling pathways.

Generation of m⁵Cscore and capability to predict prognosis

We established a scoring system that depended on the expression of DEGs and m⁵C regulators to quantify the individual m⁵C modification pattern; we termed this m⁵Cscore. The univariate Cox regression method was employed to determine the DEGs that were significantly related to the survival of GC patients in ACRG (Table S6). Ninety-nine genes were involved in the LASSO Cox regression algorithm to generate the m5Cscore signature, and eight genes were selected, including seven DEGs (RBPMS2, TNFRSF11A, NBEA, INHBB, RGN, DFNA5, and TPD52L1) and one writer (DNMT3A) (Figures 5A, B). The m5Cscore of each GC patient and prognostic information is summarized in Table S7. The alluvial diagram shows the attribute changes of individual GC patients (Figure 5C). Log-rank results depict that the OS of patients with a low m⁵Cscore is remarkably higher than patients with a high m⁵Cscore under the cutoff value of 9.92 (Figures 5D, E). The area under the curve (AUC) is 0.766, quantified by the pROC package (Figure 5F). Univariate and multivariate analysis demonstrates that age, N stage, M stage, and m⁵Cscore are the independent factors of prognosis (Figures 5G, H). Meanwhile, we found that m⁵Cscores significantly differed in distinct ACRG molecular subtypes. Patients in the EMT subgroup showed the highest m5Cscore compared to the other molecular groups (Figure 6A). Additionally, patients in m⁵C cluster B showed the lowest m⁵Cscore compared to other clusters (Figure 6B). In Figure S3G-I, GC patients with a high m⁵Cscore show a significantly higher stromal score and a lower tumor purity score. The results added to the evidence that a low m5Cscore was significantly related to immune activation and a high m⁵Cscore was correlated with stromal activation. m5Cscore could be a better marker to estimate the m⁵C modification of individual GC patients. Notably, patients with a low m5Cscore and who received adjuvant chemotherapy showed significant treatment advantages (Figure 6C). The result also demonstrated that the prediction value of m⁵Cscore was not affected by chemotherapy, and a low m⁵Cscore showed obvious survival advantage, regardless of whether patients received chemotherapy or not (Figure 6C).



Moreover, older patients, intestinal histological subtype, and early GC patients were notably related to a low m⁵Cscore, which demonstrated that these GC patients were characterized by the m⁵C cluster B and immune-inflamed phenotype, with a better prognosis (Figure 6D).

Validation of m^5C modification in TCGA and other datasets

Data from the TCGA-STAD cohort and GEO were used for external and internal validation to determine the role of $\rm m^5C$

modification and the prognostic value of m⁵Cscore. m⁵Cscore was employed to evaluate the individual m⁵C modification of the single patients in the TCGA dataset, among which 267 patients have a low m⁵Cscore and 69 patients have a high m⁵Cscore. Combined with the prognosis information, we revealed that patients with a low m⁵Cscore had a better prognosis (Figure 6E). ROC curve analysis showed that the AUC was 0.898, which was even higher than that in the training cohort (Figure 6F). We also noticed that patients with a high m⁵Cscore and without chemotherapy experienced the worst prognosis, while patients with a low m⁵Cscore and chemotherapy showed a favorable prognosis (Figure 6G). As shown in Figures 6H, I, patients in the

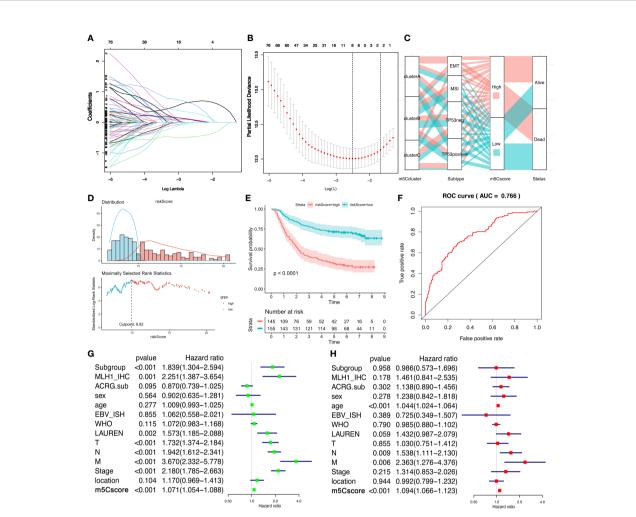
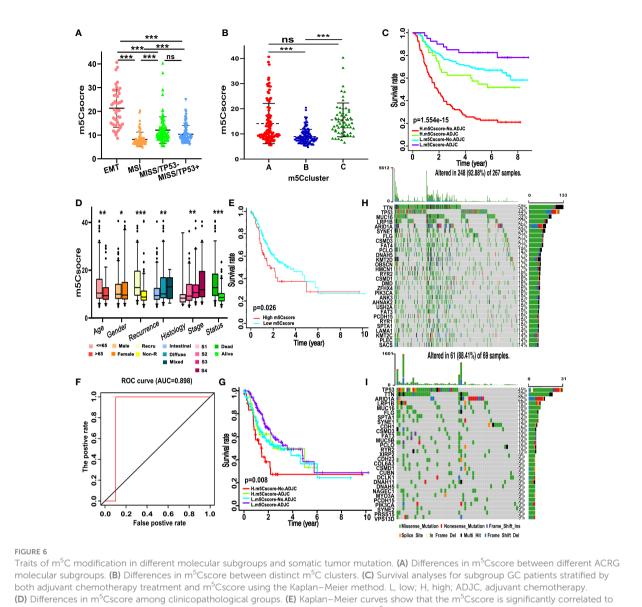


FIGURE 5

The generation of the m⁵Cscore system in the ACRG cohort. (**A**, **B**) The LASSO regression model and verification. (**A**) The longitudinal, lower transverse, and upper transverse coordinates are the correlation coefficient, Log Lambda (penalty coefficient), and the number of non-zero coefficients in the model. The lines with different colors show the trajectories of related variables changing with Log Lambda in the model. (**B**) The upper transverse, lower transverse, and longitudinal coordinates represent the format of the factor, the Log Lambda (penalty coefficient), and the error of cross-verification. The point with the smallest cross-verification error corresponds to the number of factors involved in the LASSO regression model. (**C**) Alluvial diagram demonstrates the changes of m⁶A clusters, molecular subtypes, m⁵Cscores, and status. (**D**) The distribution of the m⁵Cscore of each patient from the ACRG cohort. The cutoff point of m⁵Cscore is 9.92. (**E**) Kaplan–Meier curves show that the m⁵Cscore is significantly related to the overall survival of GC patients in the ACRG cohort, of which 145 patients were in the high-m⁵Cscore group and 155 patients were in the low-m⁵Cscore group (p < 0.0001, Log-rank test). (**F**) The predictive value of m⁵Cscore in the ACRG cohort. AUC, 0.766. (**G**) Univariate Cox regression analysis for m⁵Cscore in the ACRG cohort shown by the forest plot. (**H**) Multivariate Cox regression analysis for m⁵Cscore in the ACRG cohort shown by the forest plot.

high-m⁵Cscore group exhibit less extensive tumor mutation burden than patients in the low-m⁵Cscore group, with alternation rates of 88.41% and 92.88%, respectively. TMB analysis demonstrated that a high m⁵Cscore was significantly related to lower TMB, and showed a notable negative correlation (Figures 6H, I). Furthermore, the mean TMB of patients with a high or low m5Cscore was 2.31 and 1.26 mutations per MB. The violin plot also demonstrated that the TMB of patients in the highm5Cscore group was significantly higher than that of patients in the low-m5Cscore group, and the *p*-value was 0.012 (Figure S3J). Next, to further validate the stability of the m⁵Cscore system, the m⁵Cscore model was applied to other independent GC cohorts to confirm the prognostic value. Figures 7A–C show that GC patients with a low m⁵Cscore have a better prognosis in GSE57303, GSE84437, and GSE 15459. Moreover, we combined all the five GEO datasets together and found that the m⁵Cscore model was validated (Figure 7D). The ROC curve was drawn, and all AUCs were over 0.6 (Figure 7E). In addition, GSE26253, a new GEO dataset, was used to evaluate the predictive value of recurrence-free survival. Figure 7F confirms the ability of



the overall survival of patients in the TCGA-STAD cohort. (F) The predictive value of m^5 Cscore in the TCGA-STAD cohort. AUC, 0.898. (G) Survival analyses for subgroup GC patients divided by adjuvant chemotherapy treatment and m^5 Cscore using Kaplan–Meier curves in the TCGA-STAD cohort. (H,I) The waterfall plot of tumor somatic mutation based on those with high m^5 Cscore (H) and low m^5 Cscore (I) in the TCGA-STAD cohort. Each column represents each GC patient. The upper and right bar plots show TMB and the proportion of each variant type. The number on the right indicated the mutation frequency in each gene. (ns, not significant; *p < 0.05; **p < 0.01; ***p < 0.001).

 m^5 Cscore to predict RFS, which means the underlying potential mechanisms exist between m^5 C modification and tumor relapse to be elucidated.

Discussion

Growing evidence revealed that aberrant RNA m^5C methylation modification played a crucial role in tumorigenesis, progression, and patient prognosis by means of dynamic RNA epigenetic modification. In the current study, we analyzed that m⁵C regulators in GC explored the correlation between TME and m⁵C modification, as well as established an m⁵Cscore system to evaluate the prognosis of GC patients *via* the data from GEO datasets and the TCGA-STAD cohort. The m⁵Cscore model was further validated by internal and external datasets. These findings added clues for understanding the m⁵C modification of individual GC patients.

Sixteen m⁵C methylation regulators were involved in the analysis, including methyltransferases, demethylases, and RNA

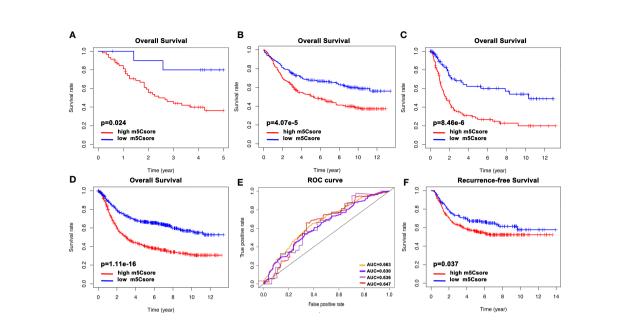


FIGURE 7

The validation of the m⁵Cscore system of individual GC patients from internal and external cohorts. (A) Overall survival analysis of m⁵Cscore in the GSE57303 cohort. (B) Overall survival analysis of m⁵Cscore in the GSE5459 cohort. (D) Overall survival analysis of m⁵Cscore in all five GEO cohorts. (E) The predictive value of m⁵Cscore in the four validation cohorts. The orange line represents all five GEO datasets with AUC 0.636; the purple line represents the GSE57303 cohort with AUC 0.636; the red line represents the GSE57303 cohort with AUC 0.647. (F) Recurrence-free survival analysis of m⁵Cscore in the GSE26253 cohort.

binding proteins. Although the exact number of m⁵C regulators and detailed mechanisms of m⁵C methylation are far from clear, the existing evidence has validated the essential function of m⁵C modification on different types of RNA, physiological, and pathological processes (7, 14). Among all the regulators, 13 regulators are significantly aberrantly expressed with 10 genes upregulated and 3 downregulated in GC samples. NSUN7 and DNMT2 are the only low-expression regulators out of the 11 methyltransferases. Sato et al. reported that NSUN7 was upregulated in low-grade glioma with an unknown mechanism (40). However, in GC, we suppose that the low expression of NSUN7 is caused by the loss of CNV frequency. Mei and colleagues found that NSUN2 was overexpressed in GC, which is consistent with our results, and they further validated that NSUN2 promotes GC cell proliferation via repressing p57(Kip2) in an m⁵C-dependent manner (41). In correlation analysis, we noticed that the methyltransferases tended to be related to each other, indicating the underlying interaction of mediating the m⁵C methylation modification. As for the readers, ALYREF and YBX1 were remarkably overexpressed in GC patients. Research on bladder cancer, breast cancer, HCC, and oral squamous cell carcinoma revealed that ALYREF and YBX1 were upregulated as well (22, 42-44). Intriguingly, high expression of ALYREF and YBX1 are also significantly correlated with the favorable prognosis of GC patients. All three erasers are notably related

to the OS of GC patients despite the fact that only TET3 is significantly abnormally expressed in tumor samples.

Based on the expression of 16 m⁵C methylation regulators, three m⁵C modification patterns were distinguished. The three m⁵C modification clusters were characterized by different TME cell infiltration patterns. m⁵C cluster A was included in the immunedesert phenotype, characterized by immunosuppression; m5C cluster B belonged to the immune-inflamed phenotype, showing the activation of adaptive immunity; m5C cluster C was classified as immune-excluded phenotype, characterized by stroma and immunity activation. The GSVA analysis also revealed that m⁵C cluster B is enriched in cytokine-cytokine receptor interaction, natural killer cell-mediated cytotoxicity, antigen processing and presentation, Toll-like receptor signaling pathway, and chemokine signaling pathway. These results added to the evidence that the immune-inflamed phenotype, also known as a hot tumor, is characterized by immune cell infiltration and immune-related signal pathway stimulation in TME (45, 46). Additionally, we found that the immune checkpoints in m⁵C cluster B were notably overexpressed than the other two m5C clusters, which indicated that patients in m⁵C cluster B might benefit from immunotherapy. In the immune-excluded phenotype, TGF- β and EMT pathways are activated and abate the efficiency of immunotherapy (47). However, we observed the activation of TGF-β and EMT pathways in m⁵C cluster A, which was

10.3389/fimmu.2022.905057

classified as the immune-desert phenotype. The anomaly may be due to the limited number of TGF- β and EMT pathway-related genes, which requires more data analysis and illustrates the complexity of m⁵C methylation modification. In survival analysis, m⁵C cluster B showed the most favorable prognosis, which is consistent with the above-mentioned immune-inflamed phenotype.

The m⁵Cscore system was established based on the expression of eight genes via the LASSO Cox regression method, namely, DNMT3A, RBPMS2, TNFRSF11A, NBEA, INHBB, RGN, DFNA5, and TPD52L1. Among all the genes calculated in the m⁵Cscore system, only DNMT3A is an m⁵C modification regulator; TNFRSF11A, INHBB, and DFNA5 are involved in TNF-related pathways (48-50); TPD52L1 participates in cell proliferation and calcium signaling; and RBPMS2, as an RNA binding protein, is involved in the regulation of cell differentiation and proliferation (51, 52). m⁵Cscore is a reliable marker to evaluate the prognosis of GC patients with an AUC of 0.766 in the ACRG training set and 0.898 in the TCGA validation set. m⁵Cscore was verified by other GEO datasets as well. Inevitably, m⁵Cscore is distinct in different m5C clusters, in which m5C cluster B had the lowest m⁵Cscore. We noticed that GC patients with the EMT molecular subtype show the highest m⁵Cscore, demonstrating poor prognosis. Furthermore, GC patients with a high m5Cscore tend to have a shorter RFS, indicating that m⁵C methylation may play an essential role in tumor recurrence.

Conclusion

In summary, we revealed the potential regulatory mechanisms of m^5C methylation modification on the GC TME. The characteristics of distinct m^5C modification patterns might lead to the complexity and heterogeneity of individual GC TME. A farreaching understanding of specific m^5C modification patterns in GC will contribute to identifying TME cell infiltration and guiding clinical immunotherapy treatments.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Ethics statement

The study complied with the principles set forth in the Declaration of Helsinki. Access to the deidentified linked dataset was obtained from the TCGA and GEO databases in accordance with the database policy. For analyses of deidentified data from the TCGA and GEO databases, institutional review board approval and informed consent were not required.

Author contributions

All authors searched the literature, designed the study, interpreted the findings, and revised the manuscript. QZ, JS, XS and KS carried out data management and statistical analysis and drafted the manuscript. JL, XG, KS and XQ helped with cohort identification and data management. QZ, JS, XS and KS performed project administration. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by grants from the National Natural Science Foundation of China (82003184) and the Health Science and Technology Project of Lianyungang (201914).

Acknowledgments

The authors thank all individuals who participated in this study and donated samples.

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/ fimmu.2022.905057/full#supplementary-material

References

1. Smyth EC, Nilsson M, Grabsch HI, van Grieken NC, Lordick F. Gastric cancer. Lancet (2020) 396(10251):635-48. doi: 10.1016/S0140-6736(20)31288-5

2. Collaborators GBDSC. The global, regional, and national burden of stomach cancer in 195 countries, 1990-2017: a systematic analysis for the global burden of disease study 2017. *Lancet Gastroenterol Hepatol* (2020) 5(1):42–54. doi: 10.1016/S2468-1253(19)30328-0

3. Matsuda T, Ajiki W, Marugame T, Ioka A, Tsukuma H, Sobue T, et al. Population-based survival of cancer patients diagnosed between 1993 and 1999 in Japan: a chronological and international comparative study. *Japanese J Clin Oncol* (2011) 41(1):40–51. doi: 10.1093/jjco/hyq167

4. Hong S, Won YJ, Park YR, Jung KW, Kong HJ, Lee ES, et al. Cancer statistics in Korea: Incidence, mortality, survival, and prevalence in 2017. *Cancer Res Treat* (2020) 52(2):335–50. doi: 10.4143/crt.2020.206

5. Boccaletto P, Machnicka MA, Purta E, Piatkowski P, Baginski B, Wirecki TK, et al. MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res* (2018) 46(D1):D303–D7. doi: 10.1093/nar/gkx1030

6. Bestor TH. Cloning of a mammalian DNA methyltransferase. *Gene* (1988) 74 (1):9–12. doi: 10.1016/0378-1119(88)90238-7

7. Garcia-Vilchez R, Sevilla A, Blanco S. Post-transcriptional regulation by cytosine-5 methylation of RNA. *Biochim Biophys Acta Gene Regul Mech* (2019) 1862(3):240–52. doi: 10.1016/j.bbagrm.2018.12.003

8. 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(5):606–25. doi: 10.1038/cr.2017.55

9. Van Haute L, Lee SY, McCann BJ, Powell CA, Bansal D, Vasiliauskaite L, et al. NSUN2 introduces 5-methylcytosines in mammalian mitochondrial tRNAs. *Nucleic Acids Res* (2019) 47(16):8720–33.

10. Sharma S, Yang J, Watzinger P, Kotter P, Entian KD. Yeast Nop2 and Rcm1 methylate C2870 and C2278 of the 25S rRNA, respectively. *Nucleic Acids Res* (2013) 41(19):9062–76. doi: 10.1093/nar/gkt679

11. Xue S, Xu H, Sun Z, Shen H, Chen S, Ouyang J, et al. Depletion of TRDMT1 affects 5-methylcytosine modification of mRNA and inhibits HEK293 cell proliferation and migration. *Biochem Biophys Res Commun* (2019) 520(1):60–6. doi: 10.1016/j.bbrc.2019.09.098

12. Cheray M, Etcheverry A, Jacques C, Pacaud R, Bougras-Cartron G, Aubry M, et al. Cytosine methylation of mature microRNAs inhibits their functions and is associated with poor prognosis in glioblastoma multiforme. *Mol Cancer* (2020) 19 (1):36. doi: 10.1186/s12943-020-01155-z

13. David R, Burgess A, Parker B, Li J, Pulsford K, Sibbritt T, et al. Transcriptomewide mapping of RNA 5-methylcytosine in arabidopsis mRNAs and noncoding RNAs. *Plant Cell* (2017) 29(3):445-60. doi: 10.1105/tpc.16.00751

14. Xue C, Zhao Y, Li L. Advances in RNA cytosine-5 methylation: detection, regulatory mechanisms, biological functions and links to cancer. *biomark Res* (2020) 8:43. doi: 10.1186/s40364-020-00225-0

15. Reid R, Greene PJ, Santi DV. Exposition of a family of RNA m(5)C methyltransferases from searching genomic and proteomic sequences. *Nucleic Acids Res* (1999) 27(15):3138-45. doi: 10.1093/nar/27.15.3138

16. Goll MG, Kirpekar F, Maggert KA, Yoder JA, Hsieh CL, Zhang X, et al. Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. *Sci* (*New York NY*) (2006) 311(5759):395–8. doi: 10.1126/science.1120976

17. Schaefer M, Pollex T, Hanna K, Tuorto F, Meusburger M, Helm M, et al. RNA Methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. *Genes Dev* (2010) 24(15):1590–5. doi: 10.1101/gad.586710

18. Traube FR, Carell T. The chemistries and consequences of DNA and RNA methylation and demethylation. *RNA Biol* (2017) 14(9):1099–107. doi: 10.1080/15476286.2017.1318241

19. Huang Y, Wang G, Liang Z, Yang Y, Cui L, Liu CY. Loss of nuclear localization of TET2 in colorectal cancer. *Clin Epigenet* (2016) 8:9. doi: 10.1186/s13148-016-0176-7

20. 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(8):978–90. doi: 10.1038/s41556-019-0361-y

21. Cheng JX, Chen L, Li Y, Cloe A, Yue M, Wei J, et al. RNA Cytosine methylation and methyltransferases mediate chromatin organization and 5-azacytidine response and resistance in leukaemia. *Nat Commun* (2018) 9 (1):1163. doi: 10.1038/s41467-018-03513-4

22. Li D, Li K, Zhang W, Yang KW, Mu DA, Jiang GJ, et al. The m6A/m5C/m1A regulated gene signature predicts the prognosis and correlates with the immune status of hepatocellular carcinoma. *Front Immunol* (2022) 13:918140. doi: 10.3389/fimmu.2022.918140

23. Smyth EC, Cervantes A. Immunotherapy is not for all comers in chemotherapy-refractory advanced gastric cancer. better predictive biomarkers are needed. *Ann Oncol* (2018) 29(10):2027–8. doi: 10.1093/annonc/mdy331

24. Zhang B, Wu Q, Li B, Wang D, Wang L, Zhou YL. m(6)A regulatormediated methylation modification patterns and tumor microenvironment infiltration characterization in gastric cancer. *Mol Cancer* (2020) 19(1):53. doi: 10.1186/s12943-020-01170-0

 Pitt JM, Marabelle A, Eggermont A, Soria JC, Kroemer G, Zitvogel L. Targeting the tumor microenvironment: removing obstruction to anticancer immune responses and immunotherapy. *Ann Oncol* (2016) 27(8):1482–92. doi: 10.1093/annonc/mdw168

26. Ali HR, Chlon L, Pharoah PD, Markowetz F, Caldas C. Patterns of immune infiltration in breast cancer and their clinical implications: A gene-Expression-Based retrospective study. *PloS Med* (2016) 13(12):e1002194. doi: 10.1371/journal.pmed.1002194

27. Binnewies M, Roberts EW, Kersten K, Chan V, Fearon DF, Merad M, et al. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat Med* (2018) 24(5):541–50. doi: 10.1038/s41591-018-0014-x

28. Guo G, Wang H, Shi X, Ye L, Yan K, Chen Z, et al. Disease activity-associated alteration of mRNA m(5) c methylation in CD4(+) T cells of systemic lupus erythematosus. *Front Cell Dev Biol* (2020) 8:430. doi: 10.3389/fcell.2020.00430

29. Collignon E, Canale A, Al Wardi C, Bizet M, Calonne E, Dedeurwaerder S, et al. Immunity drives TET1 regulation in cancer through NF-kappaB. *Sci Adv* (2018) 4(6):eaap7309. doi: 10.1126/sciadv.aap7309

30. Andries O, Mc Cafferty S, De Smedt SC, Weiss R, Sanders NN, Kitada T. N (1)-methylpseudouridine-incorporated mRNA outperforms pseudouridine-incorporated mRNA by providing enhanced protein expression and reduced immunogenicity in mammalian cell lines and mice. *J Control Release* (2015) 217:337–44. doi: 10.1016/j.jcorrel.2015.08.051

31. Wnuk M, Slipek P, Dziedzic M, Lewinska A. The roles of host 5methylcytosine RNA methyltransferases during viral infections. *Int J Mol Sci* (2020) 21(21): 8176. doi: 10.3390/ijms21218176

32. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res* (2016) 44(8):e71. doi: 10.1093/nar/gkv1507

33. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* (2010) 26 (12):1572–3. doi: 10.1093/bioinformatics/btq170

34. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep* (2017) 18 (1):248–62. doi: 10.1016/j.celrep.2016.12.019

35. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* (2015) 43(7):e47. doi: 10.1093/nar/gkv007

36. Xiang ZJ, Wang Y, Ramadge PJ. Screening tests for lasso problems. *IEEE Trans Pattern Anal Mach Intell* (2017) 39(5):1008-27. doi: 10.1109/TPAMI.2016.2568185

37. Engebretsen S, Bohlin J. Statistical predictions with glmnet. Clin Epigenet (2019) 11(1):123. doi: 10.1186/s13148-019-0730-1

38. Zeng D, Li M, Zhou R, Zhang J, Sun H, Shi M, et al. Tumor microenvironment characterization in gastric cancer identifies prognostic and immunotherapeutically relevant gene signatures. *Cancer Immunol Res* (2019) 7 (5):737–50. doi: 10.1158/2326-6066.CIR-18-0436

39. Amiot L, Vu N, Samson M. Biology of the immunomodulatory molecule HLA-G in human liver diseases. *J Hepatol* (2015) 62(6):1430-7. doi: 10.1016/j.jhep.2015.03.007

40. Sato K, Tahata K, Akimoto K. Five genes associated with survival in patients with lower-grade gliomas were identified by information-theoretical analysis. *Anticancer Res* (2020) 40(5):2777–85. doi: 10.21873/anticanres.14250

41. Mei L, Shen C, Miao R, Wang JZ, Cao MD, Zhang YS, et al. RNA Methyltransferase NSUN2 promotes gastric cancer cell proliferation by repressing p57(Kip2) by an m(5)C-dependent manner. *Cell Death Dis* (2020) 11 (4):270. doi: 10.1038/s41419-020-2487-z

42. Yamashita T, Higashi M, Momose S, Morozumi M, Tamaru JI. Nuclear expression of y box binding-1 is important for resistance to chemotherapy including gemcitabine in TP53-mutated bladder cancer. *Int J Oncol* (2017) 51 (2):579–86. doi: 10.3892/ijo.2017.4031

43. Campbell TM, Castro MAA, de Oliveira KG, Ponder BAJ, Meyer KB. ERalpha binding by transcription factors NFIB and YBX1 enables FGFR2 signaling to modulate estrogen responsiveness in breast cancer. Cancer Res (2018) 78 (2):410–21. doi: 10.1158/0008-5472.CAN-17-1153

44. Saito Y, Kasamatsu A, Yamamoto A, Shimizu T, Yokoe H, Sakamoto Y, et al. ALY as a potential contributor to metastasis in human oral squamous cell carcinoma. *J Cancer Res Clin Oncol* (2013) 139(4):585–94. doi: 10.1007/s00432-012-1361-5

45. Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. *Nature* (2017) 541(7637):321–30. doi: 10.1038/nature21349

46. Turley SJ, Cremasco V, Astarita JL. Immunological hallmarks of stromal cells in the tumour microenvironment. *Nat Rev Immunol* (2015) 15(11):669–82. doi: 10.1038/nri3902

47. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, et al. TGFbeta attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* (2018) 554(7693):544–8. doi: 10.1038/nature25501

48. Luo J, Yang Z, Ma Y, Yue Z, Lin H, Qu G, et al. LGR4 is a receptor for RANKL and negatively regulates osteoclast differentiation and bone resorption. *Nat Med* (2016) 22(5):539–46. doi: 10.1038/nm.4076

49. Zou G, Ren B, Liu Y, Fu Y, Chen P, Li X, et al. Inhibin b suppresses anoikis resistance and migration through the transforming growth factor-beta signaling pathway in nasopharyngeal carcinoma. *Cancer Sci* (2018) 109(11):3416–27. doi: 10.1111/cas.13780

50. Hu L, Chen M, Chen X, Zhao C, Fang Z, Wang H, et al. Chemotherapyinduced pyroptosis is mediated by BAK/BAX-caspase-3-GSDME pathway and inhibited by 2-bromopalmitate. *Cell Death Dis* (2020) 11(4):281. doi: 10.1038/ s41419-020-2476-2

51. Byrne JA, Nourse CR, Basset P, Gunning P. Identification of homo- and heteromeric interactions between members of the breast carcinoma-associated D52 protein family using the yeast two-hybrid system. *Oncogene* (1998) 16(7):873–81. doi: 10.1038/sj.onc.1201604

52. Notarnicola C, Rouleau C, Le Guen L, Virsolvy A, Richard S, Faure S, et al. The RNA-binding protein RBPMS2 regulates development of gastrointestinal smooth muscle. *Gastroenterology* (2012) 143(3):687–97.e9. doi: 10.1053/j.gastro.2012.05.047