



Uncovering N4-Acetylcytidine-Related mRNA Modification Pattern and Landscape of Stemness and Immunity in Hepatocellular Carcinoma

Sicheng Liu^{1†}, Yaguang Zhang^{1†}, Lei Qiu^{1†}, Su Zhang¹, Yang Meng¹, Canhua Huang¹, Zhixin Chen^{1,2}, Bo Zhang^{1,2} and Junhong Han^{1*}

¹Research Laboratory of Cancer Epigenetics and Genomics, Department of General Surgery, Frontiers Science Center for Disease-Related Molecular Network, State Key Laboratory of Biotherapy and National Clinical Research Center for Geriatrics, West China Hospital, Sichuan University, Chengdu, China, ²Department of Gastrointestinal Surgery, Frontiers Science Center for Disease-Related Molecular Network, West China Hospital, Sichuan University, Chengdu, China

OPEN ACCESS

Edited by:

Michael Liebman,
IPQ Analytics, United States

Reviewed by:

Ding Xinchun,
Indiana University, Purdue University
Indianapolis, United States

Lei Gu,
Max Planck Institute for Heart and
Lung Research, Germany

*Correspondence:

Junhong Han
hjunhong@scu.edu.cn

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 24 January 2022

Accepted: 04 April 2022

Published: 14 April 2022

Citation:

Liu S, Zhang Y, Qiu L, Zhang S,
Meng Y, Huang C, Chen Z, Zhang B
and Han J (2022) Uncovering N4-
Acetylcytidine-Related mRNA
Modification Pattern and Landscape of
Stemness and Immunity in
Hepatocellular Carcinoma.
Front. Cell Dev. Biol. 10:861000.
doi: 10.3389/fcell.2022.861000

N4-acetylcytidine (ac4C) is an ancient and conserved RNA modification. Previously, ac4C mRNA modification has been reported promoting proliferation and metastasis of tumor cells. However, it remains unclear whether and how ac4C-related mRNA modification patterns influencing the prognosis of hepatocellular carcinoma (HCC) patients. Hereby, we constructed an ac4Cscore model and classified patients into two groups and investigated the potential intrinsic and extrinsic characteristics of tumor. The ac4Cscore model, including COL15A1, G6PD and TP53I3, represented ac4C-related mRNA modification patterns in HCC. According to ac4Cscore, patients were stratified to high and low groups with distinct prognosis. Patients subject to high group was related to advanced tumor stage, higher TP53 mutation rate, higher tumor stemness, more activated pathways in DNA-repair system, lower stromal score, higher immune score and higher infiltrating of T cells regulatory. While patients attributed to low group were correlated with abundance of T cells CD4 memory, less aggressive immune subtype and durable therapy benefit. We also found ac4Cscore as a novel marker to predict patients' prognosis with anti-PD1 immunotherapy and/or mTOR inhibitor treatment. Our study for the first time showed the association between ac4C-related mRNA modification patterns and tumor intrinsic and extrinsic characteristics, thus influencing the prognosis of patients.

Keywords: hepatocellular carcinoma, ac4C, mRNA modification, tumor microenvironment, tumor stemness, prognosis predictor

Abbreviations: ac4C, N4-acetylcytidine; CNV, copy number variants; CSCs, cancer stem cells; DEGs, differentially expressed genes; GEO, gene expression omnibus; GISTIC, genomic identification of significant targets in cancer; GO, gene ontology; GSEA, gene set enrichment analysis; GSVA, gene set variant analysis; HCC, hepatocellular carcinoma; ICGC, international cancer genome consortium; INDEL, insertion-deletion; KEGG, kyoto encyclopedia of genes and genomes; LASSO, least absolute shrinkage and selection operator; LUAD, lung adenocarcinoma; MATH, mutant-allele tumor heterogeneity; mRNA_{asi}, mRNA expression-based stemness index; MSigDB, molecular signatures database; OCLR, one-class logistic regression; PAAD, pancreatic adenocarcinoma; ROC, receiver operating characteristic; SNV, single nucleotide variants; TCGA, the cancer genome atlas; TIDE, tumor immune dysfunction and exclusion; TMB, tumor mutation burden; TME, tumor microenvironment; TPM, transcripts per kilobase million.

INTRODUCTION

Worldwide, liver cancer has become a prominent public health problem with high morbidity and mortality. According to the latest statistics, liver cancer causes more than 9,00,000 cases, ranks sixth in terms of incidence and is the third leading cause of cancer-related death in 2020, leaving patients few treatment alternatives (Sung et al., 2021). Hepatocellular carcinoma (HCC) is a major histological subtype of liver cancer, accounting for 90% of primary liver cancer cases (Llovet et al., 2021). Notably, patients with HCC in China have a poor five-year survival rate of only 12% (Zheng et al., 2018). Various risk factors contribute to HCC tumorigenesis, including alcohol abuse, hepatitis B and C viral infection and metabolic disease (Yang J. D. et al., 2019). Traditional treatment options for HCC mainly rely on radiotherapy, chemotherapy, surgical resection and transplantation (Gao et al., 2021). Currently, targeted therapy, such as anti-mammalian target of rapamycin (mTOR) inhibitors, and immune checkpoint blockade therapy, such as anti-programmed cell death protein 1 (PD1) inhibitors, have great potential effects against HCC (Ruiz de Galarreta et al., 2019; Wang et al., 2019). However, due to the complicated pathogenesis, rapidly proliferating activities, unrevealed mechanism of invasion, migration and insensitive response, tumor relapse and metastasis subsequently occur after treatment, resulting in limited clinical benefit and poor prognosis of patients. Therefore, it is particularly urgent to discover and identify novel HCC markers for early diagnosis, prediction of prognosis, and to explore the molecular mechanisms behind the process of tumor initiation and progression. Thus, healthcare burden should be relieved, and the life quality of patients should be improved.

Cancer stem cells (CSCs) represent a rare subpopulation of cancer cells that exhibit extensive features of normal stem cells, such as limitless division, self-renewal and differentiation. Growing evidence has demonstrated that tumor recurrence, metastasis and the development of drug resistance can be explained by the presence of CSCs, leading to poor prognosis of patients (Visvader and Lindeman, 2008; Schulenburg et al., 2015). Recent studies have identified liver CSCs with several marker genes, including prominin 1 (PROM1), CD44, epithelial cell adhesion molecule (EPCAM) and others (Sun et al., 2016). Moreover, researchers have managed to identify tumor stemness characteristics using machine learning algorithm through transcriptome sequencing, promoting our understanding of CSCs and their roles and connections to specific oncogenic signaling pathways that sustain tumor growth and proliferation (Malta et al., 2018). In addition to CSCs, the tumor microenvironment (TME) consists of various stromal and immune cells, which together coregulate the progression of tumors and therefore influence the therapeutic response and prognosis of patients (Binnewies et al., 2018). For instance, differences in the composition of macrophages, neutrophils and T cells are correlated with clinical outcomes (Gonzalez et al., 2018). Moreover, the abundance of infiltrating T regulatory cells suppresses anticancer immunity, resulting in drug resistance (Togashi et al., 2019). With computational

tools such as ESTIMATE and CIBERSORT, the TME of tumor samples can be reasonably measured at low cost, providing new insight for clinical research (Yoshihara et al., 2013; Newman et al., 2015).

Cancer is well accepted as a heterogeneous disease caused by the accumulation of genetic and epigenetic variations. For example, BAZ2A directly interacts with EZH2 to maintain epigenetic silencing at genes repressed in prostate cancer metastasis (Gu et al., 2015). Additionally, down-regulated expression level of IRS-1 mediated by miRNA-203 leads to significant decrease of metastasis and proliferation capacities of prostate cancer cells (Meng et al., 2020). And new technologies had advanced the study of associations between diseases and epigenetics (Wang et al., 2013). Nevertheless, emerging evidence has revealed that posttranscriptional RNA modification, also termed the epitranscriptome, participates in the complex biological process of tumorigenesis, invasion, metastasis and even drug resistance (Delaunay and Frye, 2019). N⁶-methyladenosine (m⁶A), accounting for the most abundant mRNA modification, has been widely studied since debut. Based on the large number of m⁶A regulators being determined thus far, comprehensive bioinformatic analyses revealed that m⁶A modification patterns were involved in TME infiltration of gastric cancer (Zhang et al., 2020), clear cell renal cell carcinoma (Zhong et al., 2021) and pancreatic cancer (Zhou et al., 2021) and therefore are associated with the response and prognosis of patients treated with immunotherapy. In contrast, there are more than 170 RNA modifications whose functions in cancer are barely known yet (Esteve-Puig et al., 2020). N⁴-acetylcytidine (ac⁴C) modification is an ancient and highly conserved RNA modification previously reported in rRNAs and tRNAs (Jin et al., 2020), whereas the latest research found that ac⁴C modifications catalyzed by N-acetyltransferase 10 (NAT10) in human could also occur on a broad range of mRNAs in a dynamic fashion, promoting the translation process and enhancing mRNA stability (Arango et al., 2018; Sas-Chen et al., 2020). In Arango's study, NAT10-ablated HeLa cells showed reduced proliferative ability in an ac⁴C-dependent manner. Zhang et al. found that ac⁴C of COL5A1 mediated by NAT10 promoted metastasis of gastric cancer (Zhang et al., 2021). These results again confirmed the pivotal roles of ac⁴C mRNA modification in cancer progression. Therefore, comprehensive recognition of the association between intrinsic and extrinsic characteristics of tumor and ac⁴C modification patterns will contribute to our understanding of tumor management and treatment.

In this study, we identified differentially expressed genes between tumor and normal liver samples that are modified by NAT10-mediated ac⁴C and determined 21 genes. Combined with multivariate Cox regression and least absolute shrinkage and selection operator (LASSO) algorithms, COL15A1, G6PD and TP53I3 were selected to construct the risk model, termed ac⁴Cscore. We also demonstrated that ac⁴Cscore could classify patients into two groups with distinct prognoses, pathologies and somatic mutation profiles. The association between ac⁴Cscore and tumor stemness as well as TME infiltration was further confirmed using bioinformatic tools. Since ac⁴C is a conserved

modification, we proved that ac4Cscore was applicable not only in the prediction of stemness, TME distribution and survival outcome of patients with HCC, but also in those with lung and pancreatic cancers. Finally, we validated the practical value of ac4Cscore using clear cell renal cell carcinoma cohorts treated with anti-PD1 and anti-mTOR inhibitors. These results illustrated that ac4Cscore was a promising biomarker to predict the prognosis of patients.

MATERIALS AND METHODS

Dataset and Preprocess

Three publicly available databases were collected in this study, including The Cancer Genome Atlas (TCGA, <https://portal.gdc.cancer.gov/>), International Cancer Genome Consortium (ICGC, <https://dcc.icgc.org/>) and Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/geo/>). For TCGA database, normalized gene expression data of liver hepatocellular carcinoma (LIHC), pancreatic adenocarcinoma (PAAD) and lung adenocarcinoma (LUAD) by $\log_2(\text{FPKM}+1)$ were downloaded from UCSC Xena (<https://xenabrowser.net/datapages/>) along with clinical information (Goldman et al., 2020). The “ComBat” function from the R package sva was used to alleviate batch effects and to keep biological differences (Leek et al., 2012). Matched single nucleotide variants (SNVs), insertion-deletion variants (INDELs) and copy number variants (CNVs) of patients were extracted from the Genomic Data Commons (GDC). For the ICGC database, normalized expression data in FPKM format along with survival information from liver cancer patients (LIRI) were extracted (ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium, 2020). For the GEO database, the R package GEOquery was used to acquire Affymetrix microarray gene expression data normalized by the RMA method and matched phenoData of human hepatocellular carcinoma (GSE14520) (Davis and Meltzer, 2007; Roessler et al., 2010). Immune blockade therapy (anti-PD1) and mTOR inhibitor treatment (anti-mTOR) in a clear cell renal cell carcinoma cohort were also enrolled in this study. Whole-exome and transcriptome data were obtained from supplementary tables of the published article (Braun et al., 2020). To make it more comparable among multiple cohorts and similar to microarray data, RNA-sequencing data were transformed into transcripts per kilobase million (TPM) values based on the protein-coding gene annotations in GENCODE v22 and then logarithmically scaled. Dataset summaries are listed in **Supplementary Table S1**. Immunohistochemistry was obtained from The Human Protein Atlas (<https://www.proteinatlas.org/>). The tumor immune dysfunction and exclusion (TIDE) method, a well-known biomarker for modeling two primary mechanisms of tumor immune evasion and predicting the response of patients, was used in the anti-PD1 cohort (Jiang et al., 2018). TIDE scores were calculated with an online website (<http://tide.dfci.harvard.edu>).

Identification of Differentially Expressed ac4C-Modified Genes

The empirical Bayesian method from the R package limma (v3.42.2) was performed to identify differentially expressed genes of tumor and normal livers in LIHC with the threshold $|\log_2\text{FC}| > 2$ and $p\text{-value} < 0.05$ (Ritchie et al., 2015). Considering that ac4C is an ancient and highly conserved modification in eukaryotic mRNA, and NAT10 is the only known acetyltransferase at present, a list of ac4C-modified genes was selected manually from supplementary tables of the published article, which conducted ac4C-RIP-seq to screen modified peaks using wild-type and NAT10-ablated HeLa cells (Arango et al., 2018). Raw sequencing data were downloaded from the Gene Expression Omnibus dataset (GSE102113). Reads were first trimmed to remove low-quality bases and adaptor sequences using Trimmomatic (version 0.39) (Bolger et al., 2014). Hisat2 (version 2.2.1) was then used to map the reads to the human genome (hg38) (Kim et al., 2019). Samtools (version 1.9) was used to sort the coordinates of reads (Li et al., 2009). The BamCoverage function from deepTools (version 3.5.0) was applied to normalize aligned data (Ramirez et al., 2016). Finally, sample files were exported and visualized in IGV (version 2.8.10) (Thorvaldsdottir et al., 2013). Gene lists were intersected and defined as ac4C-DEGs for further analysis.

Construction and Validation of ac4C Gene Signature

Both multivariate Cox regression model and LASSO with 10-fold cross-validation algorithm were used to select the ac4C-DEGs that were significantly associated with prognosis in LIHC (Simon et al., 2011). The results were determined and visualized by the R package glmnet (v 4.1-1) and forestplot (1.10.1). Ac4C-DEGs with a $p\text{-value} < 0.05$ in the multi-Cox model and included in LASSO were considered key genes and validated using univariate Cox regression model based on the R packages survival (v3.2-7) and survminer (v0.4.8). Key genes were first scaled and standardized by Z-score. Then, ac4Cscore was calculated by the following formula: $ac4Cscore = \sum Exp_i * Coef_i$, where Exp_i is the scaled expression and $Coef_i$ is the coefficient of each key gene. This method ensures the comparability of ac4Cscore in different cohorts measured by various platforms. In addition, the maximum selected rank statistics method using the surv_cutpoint function from the R package survminer was applied to stratify patients into high and low ac4Cscore groups, thus reducing the batch effect of different cohorts. The prognostic value of ac4Cscore was further examined in other datasets using the Kaplan-Meier method with the log-rank test.

Somatic Mutation Analysis

SNVs and INDELs of ac4C-DEGs in LIHC and high-ac4Cscore and low-ac4Cscore groups in the mentioned datasets were analyzed to profile the variant characteristics. The capture size of the whole exon was deliberately set as 40 for calculation of tumor mutation burden (TMB) because it would not affect the result. Mutant-allele tumor heterogeneity (MATH) was

calculated by univariate density and cluster classification based on the variant allele frequency of individuals. All approaches above were performed by R package *maftools* (v2.6.05) (Mayakonda et al., 2018). Genomic Identification of Significant Targets in Cancer 2 (GISTIC2) software was used to detect CNV changes in the two ac4Cscore groups of LIHC (Mermel et al., 2011). The parameters were set as *-brlen* 0.5 *-conf* 0.90. UCSC hg38 was used as reference genome. Significantly aberrant regions were determined by *q*-value < 0.25.

Functional Enrichment Analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses of integrated ac4C-DEGs were performed using the R package *clusterProfiler* (v3.14.3) (Yu et al., 2012). Cutoffs of *p*-value < 0.05 and *q*-value < 0.2 indicated statistical significance.

Gene set enrichment analysis (GSEA) was performed to determine the concordant biological differences between the ac4Cscore-high and ac4Cscore-low groups in three HCC datasets (Subramanian et al., 2005). The R package *limma* was used to calculate the fold changes. Gene list was sorted by decreasing value. Gene sets *c2.all.v7.4.symbols.gmt* downloaded from the Molecular Signatures Database (MSigDB) and six categories of pathway information from KEGG (<https://www.genome.jp/kegg/>) were utilized as references in the R package *clusterProfiler*. A strict criterion with *p*-value < 0.01 and adjusted *p*-value < 0.05 represented statistical significance.

Estimation of Tumor Stemness

Cancer stem cells possess the abilities of self-renewal and differentiation, thus potentially giving rise to tumor progression and metastasis. A list of embryonic and liver cancer stem cell markers was obtained by comprehensive literature retrieval. Tathiane M. Malta et al developed an innovative one-class logistic regression (OCLR) machine learning algorithm to reveal characteristics of tumor stemness (Malta et al., 2018). The estimated cancer mRNA expression-based stemness index (mRNAsi) for TCGA datasets were extracted from the supplementary table of the paper. For other datasets, mRNAsi for each patient was predicted by the workflow described in this paper.

Estimation of Tumor Microenvironment

A variety of tumor and nontumor cells were mixed within the tumor microenvironment, which influenced tumor progression. In this context, the ESTIMATE algorithm was used to predict tumor purity and the presence of stromal and immune cell infiltration (Yoshihara et al., 2013). In addition, the CIBERSORT algorithm, which is based on a novel application of nu-support vector regression, was performed to estimate the relative fraction of 22 immune cell types with bulk transcriptome data from tumor tissues with parameter permutation test set as 1,000 times (Newman et al., 2015). The immune subtypes of individual LIHC patients were obtained from a previously published paper (Thorsson et al., 2018). To further examine the correlation of ac4Cscore and immune-related pathways, gene set variant analysis from the R package *GSEA* was

applied to evaluate the variation in pathway activity over the ac4C-high and ac4C-low groups. A list of immune signatures was downloaded from the Immunology Database and Analysis Portal (ImmPort, <https://www.immport.org/shared/genelists>).

Cell Culture and Gene Knockdown

HepG2, Huh7 and 293T cell lines verified by STR profiling were purchased from National Collection of Authenticated Cell Cultures (Shanghai, China) and maintained in DMEM (HyClone, AG29629575) supplemented with 10% FBS (ExCell Bio, FSP500) and 100 U/ml penicillin/streptomycin (Beyotime, C0222) at 37°C in a humidified incubator (Thermo) with 5% CO₂.

Gene knockdown was achieved using the short hairpin RNA system. The pLKO.1 vector was digested with AgeI (NEB, R3552) and EcoRI (NEB, R3101). Oligos were synthesized as follow: shNAT10#1, forward primer: CCGGTTGCTGTTCCACCCAGATTATCCTCGAGGATAATCTGGGTGAACAGCAATTTTGTG, reverse primer: AATT-CAAAAATTGCTGTTCCACCCAGATTATCCTCGAGGATAATCTGGGTGAACAG-CAA; shNAT10#2, forward primer: CCGGGAGATGTATTCACGGAATATGCTCG-AGCATATTCGGTGAATACATCTCTTTT TG, reverse primer: AATTCAAAAAGAG-ATGTATTCACGGAATATGCTCGAGCATATTCGGTGAATACATCTC. The forward oligo and reverse oligo were phosphorylated, annealed and further ligated to the linearized pLKO.1 vector by T4 DNA ligase (NEB, M0202). 293T cells were transfected with pLKO.1-shNAT10 with packing plasmids psPAX2 and PMD2.G. The resulting viral supernatant was used to infect target cells. Infected cells were selected with 2 µg/ml puromycin (InvivoGen).

Cell Proliferation Assay

Cell proliferation assays were performed as previously described (Liu et al., 2021). For CCK-8, a total of 2,000 cells were seeded in 96-well plates, and Cell Counting Kit-8 (CCK-8, TargetMol, C0005) was added and incubated for 3 h at days 0, 1, 2, 3, and 4 after seeding. The absorbance was measured with a microplate reader (BioTek) at 450 nm.

For colony formation, a total of 2,000 cells were seeded in 24-well plates and cultured for 10–14 days to form macroscopic clones. The colony was fixed with 4% paraformaldehyde for 20 min, washed twice with PBS buffer, and stained with crystal violet solution at room temperature for 30 min. Colony number was counted after washing with ddH₂O.

EdU Staining Assay

EdU was detected using a Cell-Light Edu Apollo488 *in Vitro* Kit (RiboBio) according to manufacturers' protocol. Briefly, control (shNT) and NAT10 knockdown (shNAT10_1# and shNAT10_2#) cells were seeded onto coverslips in 24-well. Cells were incubated with EdU solution (50 µM) for 2 h. 4% paraformaldehyde was fixed for 30 min, neutralized with 2 mg/ml glycine solution, incubated with penetrant (0.5% Triton X-100 in PBS). Apollo solution stained for 30 min in dark, washed twice with penetrant and methanol, counterstained the cell with Hoechst33342. Images were captured using fluorescence microscope (Nikon).

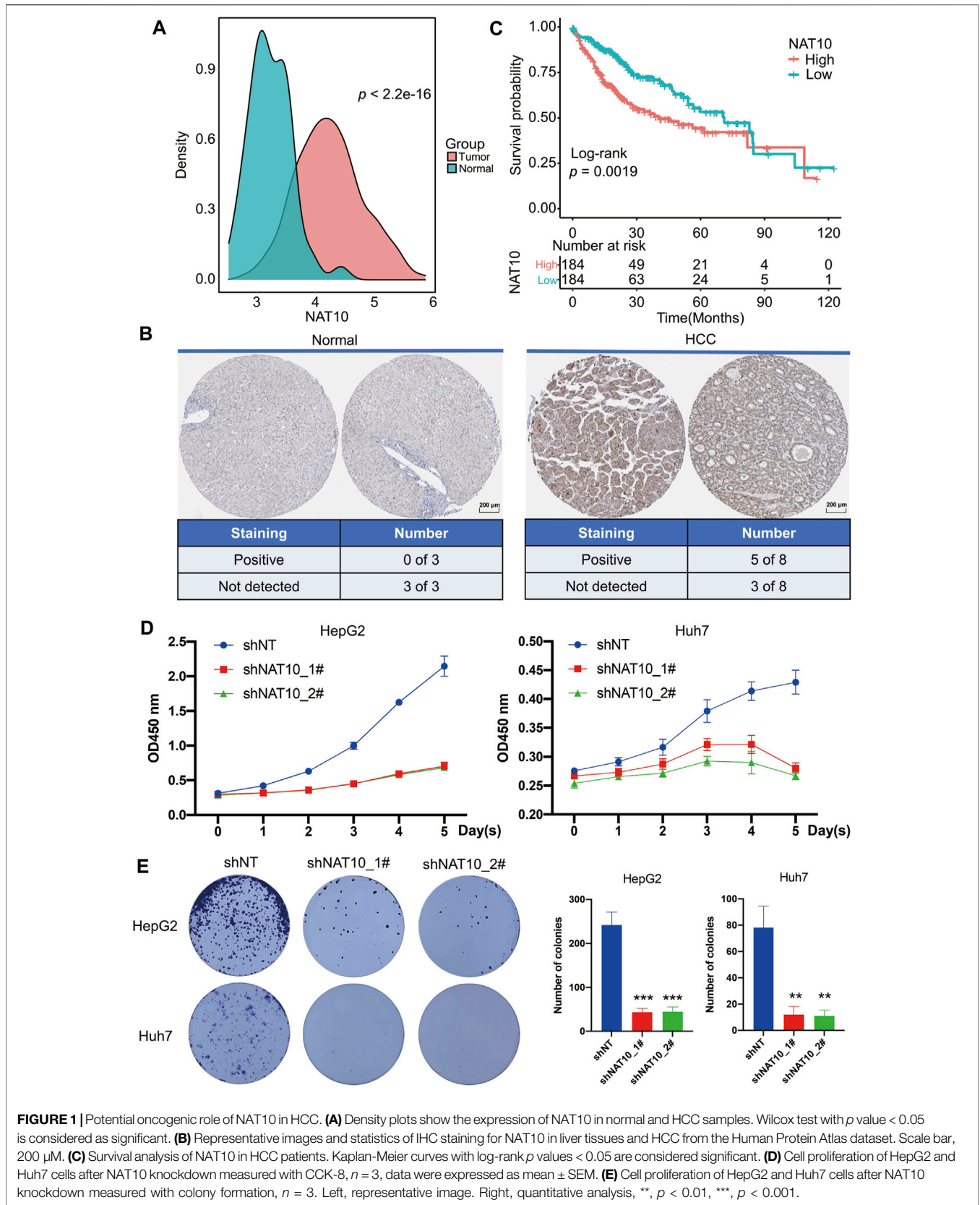


FIGURE 1 | Potential oncogenic role of NAT10 in HCC. **(A)** Density plots show the expression of NAT10 in normal and HCC samples. Wilcox test with p value < 0.05 is considered as significant. **(B)** Representative images and statistics of IHC staining for NAT10 in liver tissues and HCC from the Human Protein Atlas dataset. Scale bar, 200 μ M. **(C)** Survival analysis of NAT10 in HCC patients. Kaplan-Meier curves with log-rank p values < 0.05 are considered significant. **(D)** Cell proliferation of HepG2 and Huh7 cells after NAT10 knockdown measured with CCK-8, $n = 3$, data were expressed as mean \pm SEM. **(E)** Cell proliferation of HepG2 and Huh7 cells after NAT10 knockdown measured with colony formation, $n = 3$. Left, representative image. Right, quantitative analysis, **, $p < 0.01$, ***, $p < 0.001$.

Apoptosis Assay

Apoptosis assay was detected using an Annexin V-FITC/PI Kit (KeyGen BioTECH) according to manufacturers' protocol. Briefly, cells (shNT, shNAT10_1# and shNAT10_2#) were collected after trypsinization, washed twice with pre-cooled PBS, and resuspended in binding buffer to adjust the cell concentration to 5×10^6 /ml. Take 100 μ l of cell suspension and add 5 μ l of Annexin V/FITC to incubate for 5 min, add 10 μ l of PI and 400 μ l of PBS, and then use flow cytometer (Beckman) for detection.

Statistical Analysis

R software 3.6.3 (<https://www.R-project.org/>) and GraphPad Prism five were used for statistical analysis. Student's t test was applied to calculate the *p*-value of continuous variables. When data were not normally distributed, the Wilcoxon test was used. Fisher's exact test was performed to compare categorical variables. Log-rank test was used for survival analysis. Pearson correlation coefficients were used to compare the relationship of two variables. A two-sided *p*-value < 0.05 was considered statistically significant unless otherwise specified. The R package pheatmap (v1.0.12) was used to scale and visualize the expression of ac4C-DEGs between tumor and normal tissues, and stemness-related genes and score of immune-related pathways between the ac4Cscore-high and ac4Cscore-low groups. The R package ggalluvial (v0.12.3) was used to visualize the attribution of immune subtype to individual patients in different ac4Cscore groups. The R package corrplot (v0.84) was applied to draw the correlation of ac4Cscore and immune signatures. The R package circlize (v0.4.14) was utilized for ring heatmap (Gu et al., 2014). The R package ggradar (v0.2) was used to display the distribution of immune cells. The R package timeROC (v0.4) was used for receiver operating characteristic analysis and to quantify the area under the curve.

RESULTS

N-Acetyltransferase 10 Is Highly Expressed in Hepatocellular Carcinoma and Functions as a Potential Oncogene

The expression of NAT10 in HCC and normal liver samples were downloaded from TCGA. NAT10 was aberrantly higher expressed in HCC compared with normal samples (Figure 1A). The protein expression of NAT10 was also examined using the Human Protein Atlas database, and the same trend was observed. Among eight HCC samples, five were staining-positive, while all three out of three normal liver samples were negative (Figure 1B). For those who diagnosed with HCC, higher expression of NAT10 indicated a shorter survival outcome, especially within the first 5 years (Figure 1C). To investigate cell proliferation in HCC cells lacking NAT10, CCK-8, colony formation and EdU staining assay were performed. Results showed that cell growth was significantly inhibited in HepG2 and Huh7 with NAT10 knocked down by shRNAs (Figures 1D,E and Supplementary Figure S1A–C).

Interestingly, we also found that NAT10 might play a role in regulating cell apoptosis (Supplementary Figure S1 D,E). These results suggest NAT10 exerts a potential oncogenic role in HCC.

Identification of N-Acetyltransferase 10-Related ac4C-DEGs in Liver Hepatocellular Carcinoma

NAT10 was known as an acetyltransferase to many molecular components. Recent study revealed the ability of NAT10 for mRNA acetylcytidine, bringing us new insight into epitranscriptome. A list of 2,156 genes that are modified by NAT10-mediated ac4C was obtained (Supplementary Table S2). A total of 250 genes showing significantly different expression patterns between tumor and normal liver tissues were identified (Supplementary Table S3). In these two gene sets, 21 genes intersected and were defined as ac4C-DEGs (Figure 2A). Most of these genes were upregulated in tumors compared with normal samples, while only *LY6E* was downregulated (Supplementary Figure S2A). Therefore, distinct expression patterns of these two groups could be observed with principal component analysis (Figure 2B). GO analysis demonstrated that ac4C-DEGs were associated with DNA replication and the cell cycle, while KEGG annotated that these genes were involved in the p53 signaling pathway and multiple tumor types (Supplementary Figure S2B). Somatic mutation analysis revealed that ac4C-DEGs had little SNV or INDEL (Fig S3A), but CNV results showed that the amplification frequencies of *FAM189B*, *LY6E*, *RECQL4* and *TK1*, and the deletion frequencies of *CDKN2A* and *SFN* were greater than 10% (Supplementary Figure S3B). Then, a multivariate Cox regression model and LASSO algorithm were constructed to determine the key genes that affected the prognosis of patients, and *COL15A1*, *G6PD* and *TP53I3* were selected (Figure 2C and Supplementary Figure S4A–C). Interestingly, while three key genes were significantly upregulated in tumor samples (Figure 2D), *COL15A1* played a role as a favorable factor, while the other two served as risk factors (Figure 2C). The independent prognostic value of the three key genes was further confirmed by K-M curve analysis (Figures 2E–G). Finally, the existence of ac4C modification mediated by NAT10 within *COL15A1*, *G6PD* and *TP53I3* was confirmed (Figure 2H). These results imply that ac4C-related modification genes are potentially involved in tumor initiation and are related to different prognoses of patients.

Construction and Clinical Relevance of ac4Cscore

To profile the dynamic mRNA ac4C modification is costly, however, based on the knowledge that NAT10 is the only identified mRNA ac4C writer, the modification level should be positive correlated with the expression of NAT10, and the coefficient should not perfectly near to one because there might be other ac4C regulators not determined yet. Therefore, we constructed a model, termed ac4Cscore, using the expression and coefficient of key genes. Patients were divided into high and

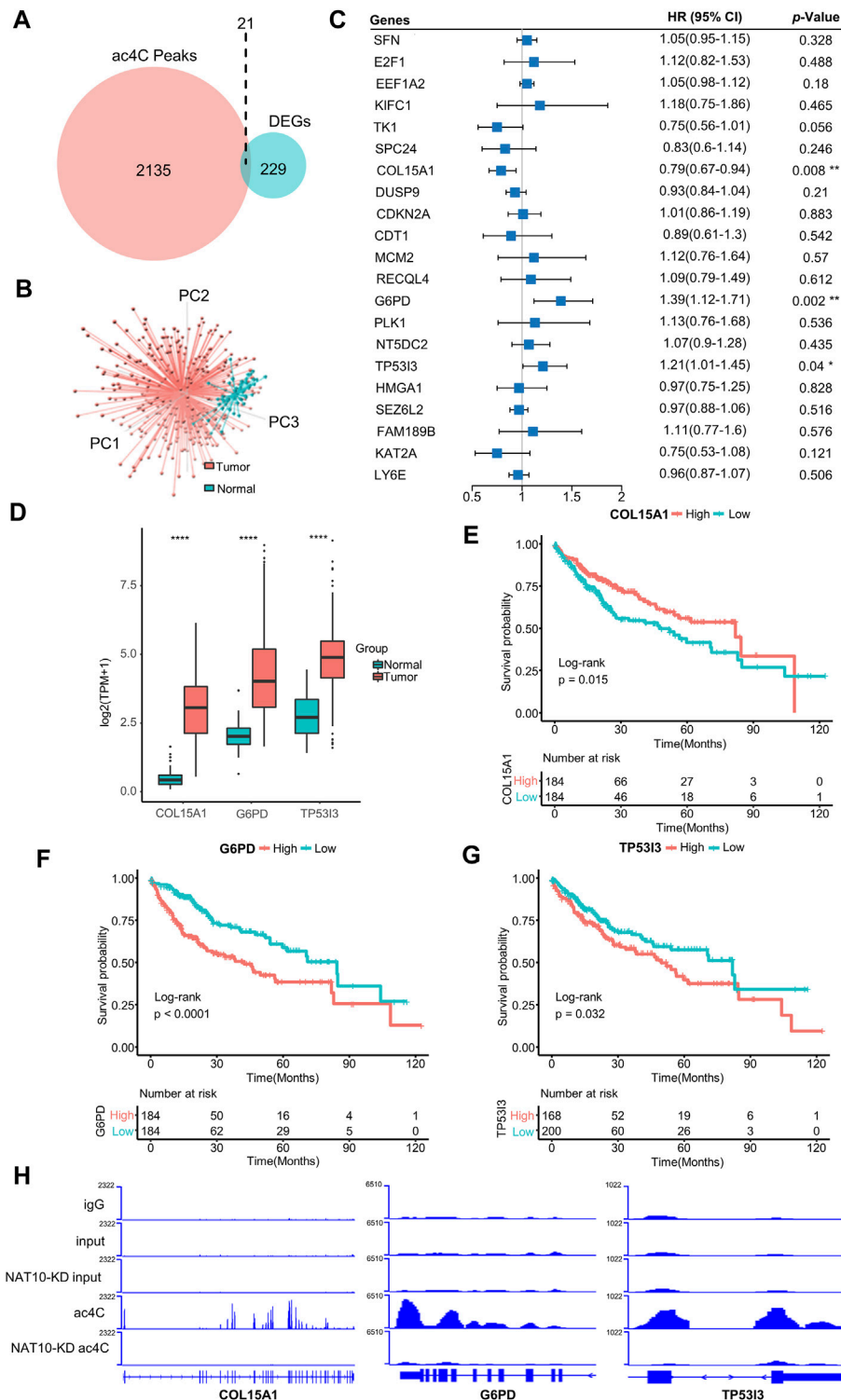
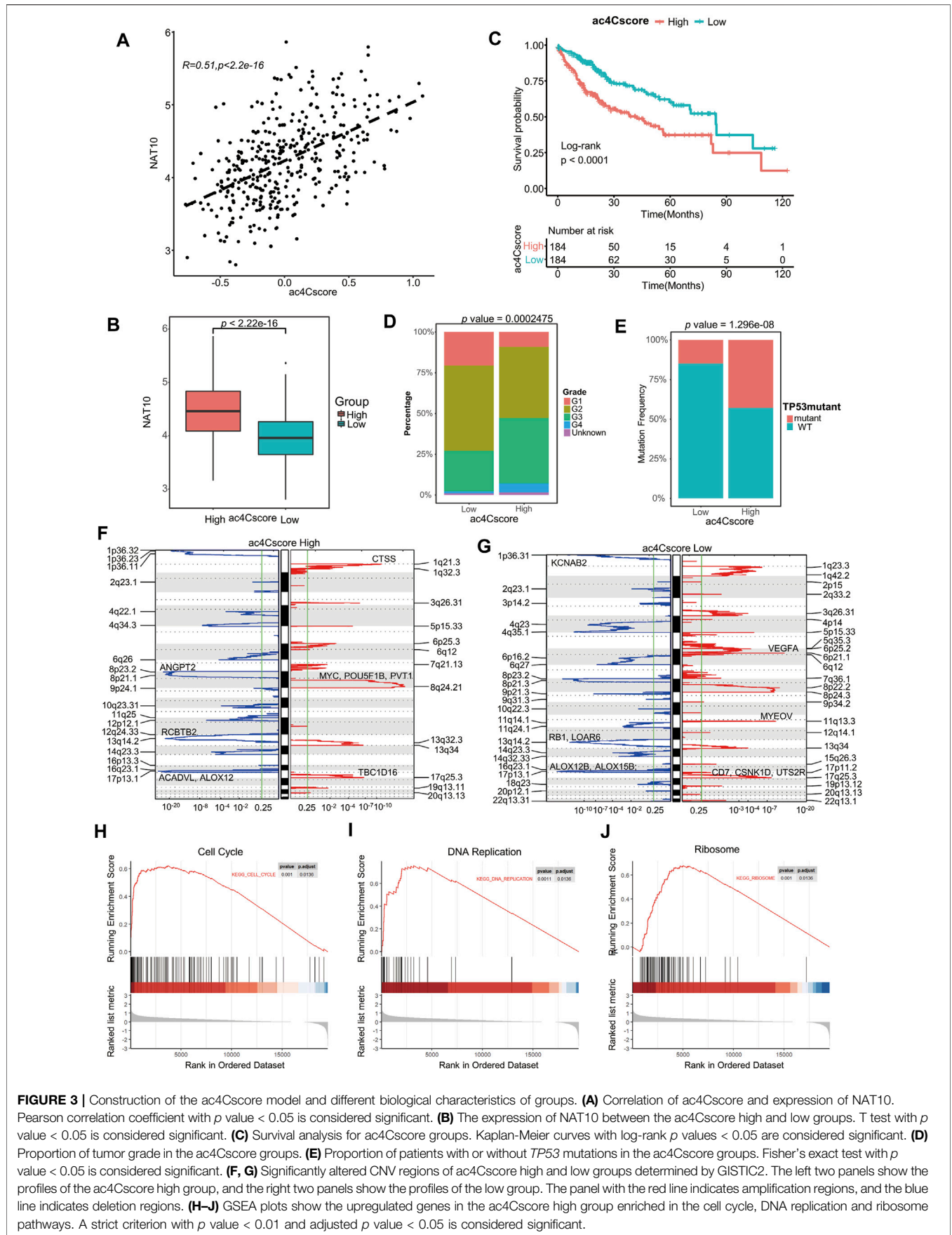
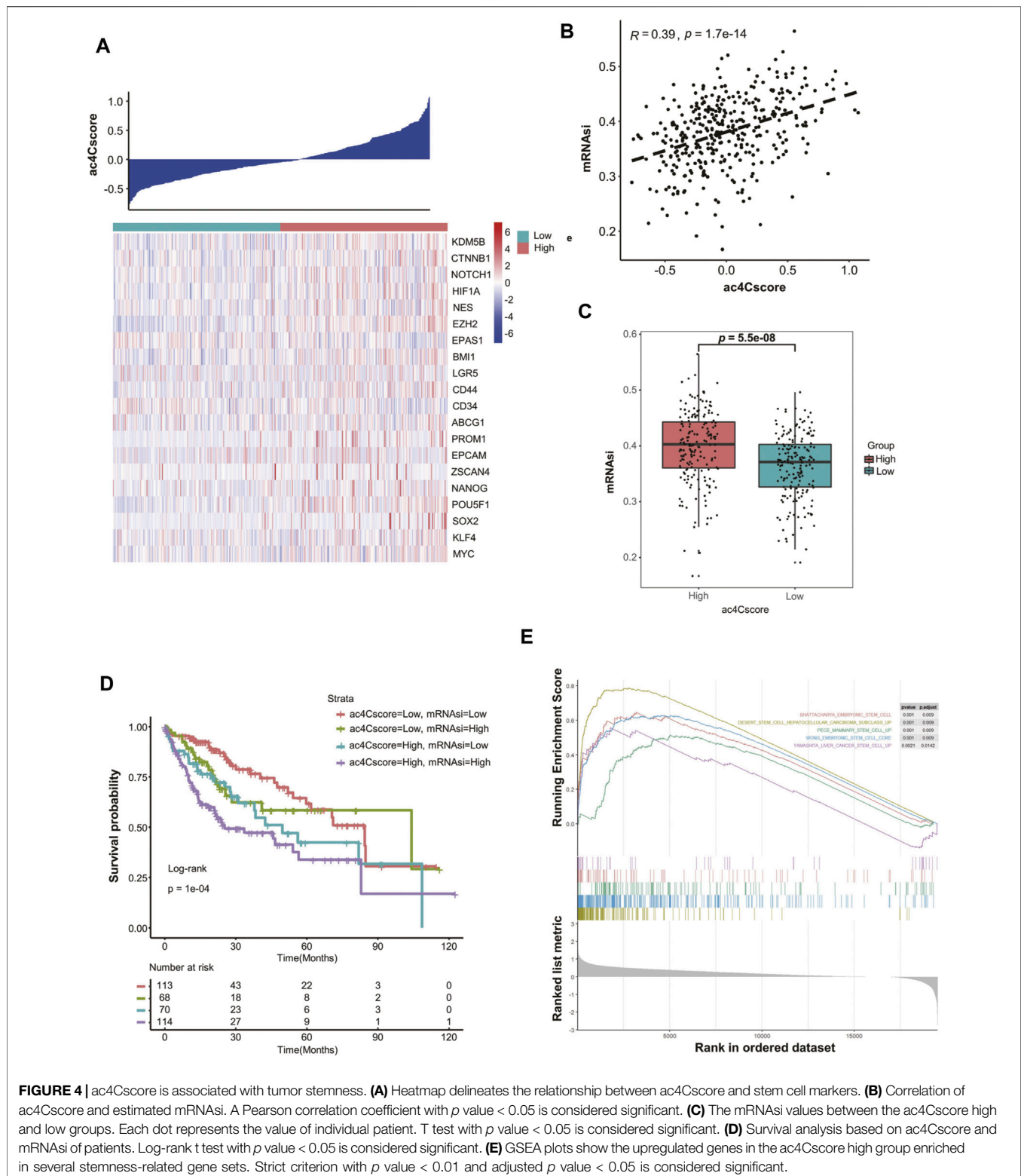


FIGURE 2 | Identification of ac4C-DEGs associated with prognosis in HCC. **(A)** Venn diagram shows the number of intersected genes between ac4C peaks and DEGs. **(B)** Principal component analysis of the expression patterns of 21 ac4C-DEGs distinguishes tumor and normal samples. **(C)** Forest plot of multivariate Cox analysis determines three key genes significantly associated with overall survival. HRs are shown with 95% confidence intervals. * indicates $p < 0.05$ and ** $p < 0.01$. **(D)** The expression of three key genes between normal and tumor samples. The top and bottom of the boxes represent the 75th and 25th percentiles, respectively. The middle lines in the boxes represent median values. The black dots indicate outliers. **** indicates $p < 0.0001$. **(E–G)** Survival analysis for the three key genes. Kaplan-Meier curves with log-rank p values < 0.05 are considered significant. **(H)** Bar graphs show differential ac4C modification of COL15A1, G6PD, and TP5313 between wild-type and NAT10-ablated cells.





low groups according to their ac4Cscore (Supplementary Table S4). As expected, ac4Cscore was moderately positive correlated with the expression of NAT10 (Figure 3A), and high ac4Cscore group had significantly higher expression of NAT10 than that of

low ac4Cscore group (Figure 3B). Unsurprisingly, those who had higher ac4Cscore suffered from significantly worse prognosis (Figure 3C). No age or gender differences were observed in these two groups since this was an unbiased model based only on

transcriptome expression data (**Supplementary Figures S5A,B**). However, compared with the ac4C-low group, the ac4C-high group had a larger proportion of patients with advanced tumor grade (G3-4), suggesting the predictive power of ac4Cscore in clinical and pathological diagnosis (**Figure 3D**). There was no significant difference between the two groups of T, N and M stages (**Supplementary Figures S5C–E**).

The profile of somatic mutations between the low and high ac4Cscore groups was also analyzed. The high ac4Cscore group was notably associated with a higher prevalence of *TP53* mutations (**Figure 3E**). However, the TMB and MATH scores of these two groups were not significantly different (**Supplementary Figures S5F,G**). Therefore, GISTIC2 was used to detect most different aberrant CNV regions. The high ac4Cscore group contained fewer regions of both amplification and deletion than the low ac4Cscore group (**Figures 3F,G**). For the high ac4Cscore group, the most significant peak of amplification fell in the cytoband of 8q24.21. Genes located in this region, including *MYC*, *POU5F1B* and *PVT1*, were notorious for tumor progression. Other peaks fell in 1q21.3 and 17q25.3. The deletion peak fell in the cytobands of 17p13.1, 13q14.2 and 8p23.2 (**Figure 3F** and **Supplementary Table S5, 6**). For the low ac4Cscore group, the amplification peaks included 6p21.1, 11q13.3 and 17q25.3, involving *VEGFA*, *CD7*, *CSNK1D* and *UTS2R*. The deletion peaks also contained 17p13.1, 1p36.31 and 13q14.2 (**Figure 3G** and **Supplementary Table S7, 8**). The above results show a different distribution of somatic mutations between the two groups.

Meanwhile, differentially expressed genes between the two ac4Cscore groups were determined (**Supplementary Table S9**). GSEA results illustrated that upregulated genes in the high ac4Cscore group were mainly enriched in the cell cycle and DNA replication pathways (**Figures 3H,I**), suggesting that patients' poor prognosis might be caused by the rapid proliferation of tumor cells. The activated ribosome pathway in the high ac4Cscore group was consistent with previous findings that ac4C modification of mRNA might promote translation (**Figure 3J**), again confirmed the efficacy of our model.

ac4Cscore is Associated With Tumor Stemness

To elicit the underlying connection of ac4Cscore with tumor stemness, the expression of several stemness-related genes was extracted. With increased ac4Cscore among patients, the expression of most markers, such as *KDM5B*, *EZH2*, *CD44* and *POU5F1*, tended to be elevated (**Figure 4A**). A significant correlation between ac4Cscore and these genes was observed (**Supplementary Figure S6A–D**). However, several markers exhibited no significant ascent or even an opposite trend (**Supplementary Figure S6E,F**). Considering that single gene expression might be confounded by tumor heterogeneity and tissue sampling bias in bulk analysis, we adopted a machine learning method, termed mRNAsi, to comprehensively analyze the association of ac4Cscore and tumor stemness (**Supplementary Table S10**). The Pearson correlation

coefficient results showed that a significantly positive correlation between ac4Cscore and mRNAsi existed (**Figure 4B**). In addition, the ac4Cscore-high group had a remarkably higher mRNAsi than the low ac4Cscore group (**Figure 4C**). We also noticed that the combination of mRNAsi and ac4Cscore could be a prognostic indicator to predict patient outcome. Patients with high ac4Cscore and high mRNAsi had the worst prognosis, while those with low ac4Cscore and low mRNAsi had a better prognosis (**Figure 4D**). The predictive power was strong within 60 months. Finally, GSEA was performed, and the results demonstrated that various stem cell gene sets were enriched in ac4Cscore-high group (**Figure 4E**). Moreover, diverse mechanisms, including the base excision repair, mismatch repair and nucleotide excision repair pathways, which could help CSCs survive in extreme environments, were also enriched in ac4Cscore-high group (**Supplementary Figure S6G**). These findings uncover the association of ac4Cscore and tumor stemness.

Characteristics of the Tumor Microenvironment in Different ac4Cscore Groups

To quantitatively profile the landscape of the tumor microenvironment in different ac4Cscore groups, both ESTIMATE and CIBERSORT algorithms were performed (**Figure 5A** and **Supplementary Table S11, 12**). Upon comparison with the low ac4Cscore group, stromal scores were significantly lower in the ac4Cscore-high group, while immune scores were higher (**Supplementary Figure S7A**). There were no obvious differences in ESTIMATE scores or tumor purity between the two groups (**Supplementary Figure S7B**). GSEA results indicated that the primary immunodeficiency pathway was enriched in ac4Cscore-high group, suggesting that the poor prognosis of this group could be caused by immune cell dysregulation (**Figure 5B**). In this context, we further investigated the differences in immune infiltration of these two groups. The proportions of plasma cells, T cells follicular helper, T cells regulatory, macrophage M0 and neutrophils were more abundant in the high ac4Cscore group, while the proportions of T cells memory resting, NK cells resting, monocytes, macrophage M2 and mast cells resting were more adequate in the low ac4Cscore group (**Figure 5C**). To better illustrate the role of ac4Cscore in immunity, we extracted available immune subtypes of LIHC. Subtype IFN-gamma Dominant and Wound Healing were mainly consisted of the ac4Cscore-high group, while the low ac4Cscore group was mostly made up of subtypes Inflammatory and TGF-beta Dominant (**Figure 5D**). Even though subtype Lymphocyte Depleted comprised approximately half of each group, the survival curve showed that the prognosis of patients with higher ac4Cscore was worse than that of the low group (**Supplementary Figure S7C**). Next, GSVA was conducted to investigate the correlation of ac4Cscore and immune signatures (**Supplementary Table S13**). The high ac4Cscore group had higher variation of pathway activities in interferon

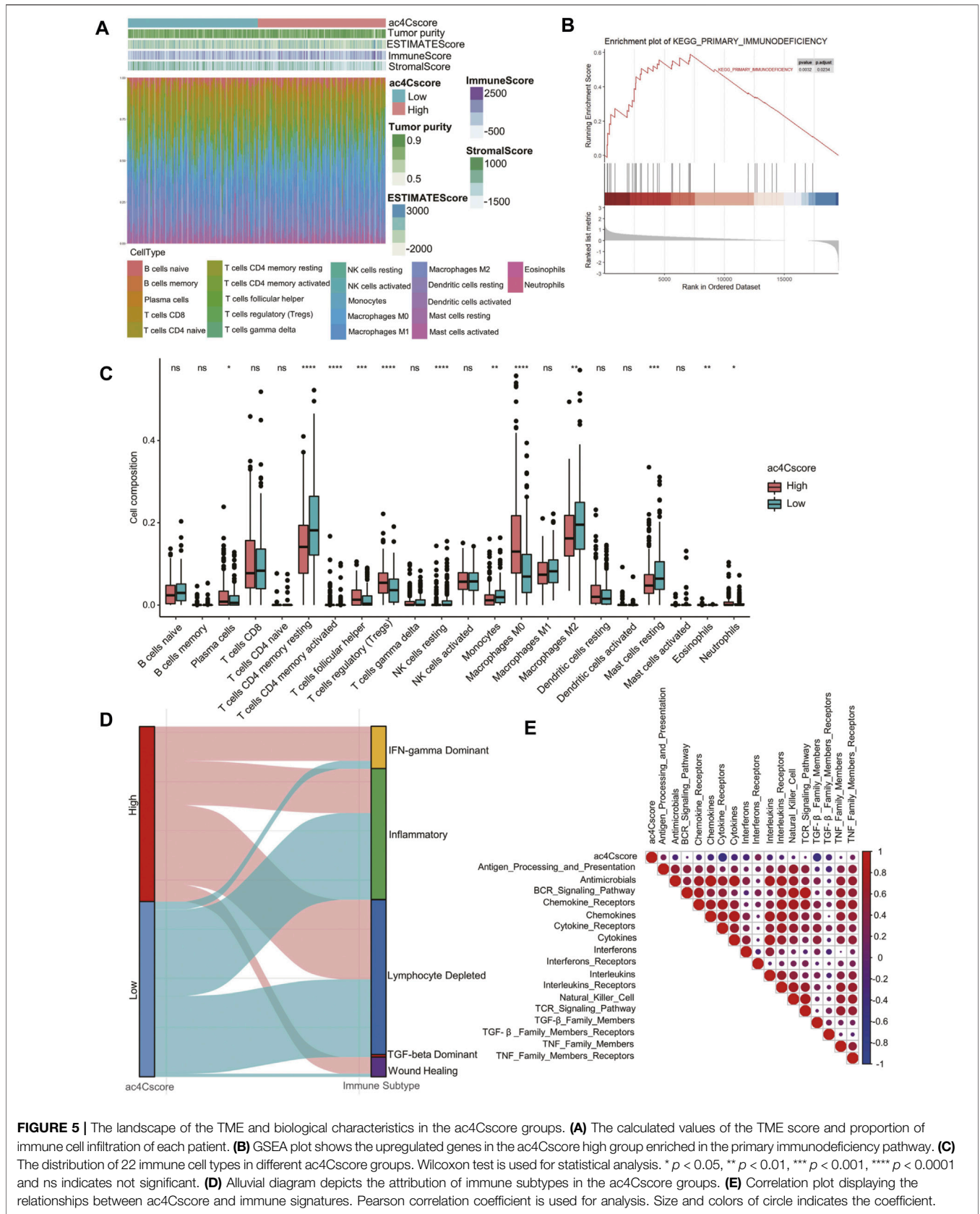


FIGURE 5 | The landscape of the TME and biological characteristics in the ac4Cscore groups. **(A)** The calculated values of the TME score and proportion of immune cell infiltration of each patient. **(B)** GSEA plot shows the upregulated genes in the ac4Cscore high group enriched in the primary immunodeficiency pathway. **(C)** The distribution of 22 immune cell types in different ac4Cscore groups. Wilcoxon test is used for statistical analysis. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$ and ns indicates not significant. **(D)** Alluvial diagram depicts the attribution of immune subtypes in the ac4Cscore groups. **(E)** Correlation plot displaying the relationships between ac4Cscore and immune signatures. Pearson correlation coefficient is used for analysis. Size and colors of circle indicates the coefficient.

receptors as well as antigen processing and presentation, while higher activity variation of cytokine receptors and TGF- β family members was observed in the ac4Cscore-low group (**Supplementary Figure S7D**). Corrplot results showed a positive or negative correlation between ac4Cscore and immune signatures (**Figure 5E**). The results above suggest that ac4Cscore might have an impact on the tumor microenvironment, especially immune infiltration.

Validation of ac4Cscore in Other Datasets

To validate the practical value of ac4Cscore, four other datasets were involved in this study. Patient information is listed in **Supplementary Table S14**. According to the ac4Cscore calculated by a previously described workflow, patients in each dataset were dichotomized into high and low groups. The prognosis of patients from the high ac4Cscore group was significantly worse than that of the low ac4Cscore group ($p = 0.044$ for GSE14520, $p = 0.00022$ for LIRI, $p = 0.049$ for LUAD and $p = 0.034$ for PAAD) (**Figures 6A–D**). Consistent GSEA results from three HCC cross-datasets for KEGG pathway analysis were extracted. The ac4Cscore-high group showed common downregulation of various metabolic pathways. Of note, cross-dataset results confirmed the upregulation of cell proliferation-related pathways, including DNA replication, cell cycle, Hippo signaling pathway and p53 signaling pathway (**Figure 6E**). Then, mRNasi was estimated to represent the tumor stemness of each patient (**Supplementary Table S15**). The mRNasi of the high ac4Cscore group was significantly higher than that of the low ac4Cscore group in the four other datasets respectively (**Figure 6F**). In addition, compared with the low ac4Cscore group, the high ac4Cscore group had a higher mean mRNasi, and a higher positive correlation with ac4Cscore was also observed (**Supplementary Figure S8A**). Interestingly, pathways such as nucleotide excision repair, mismatch repair and homologous recombination, which boost CSC survival, were generally upregulated in the ac4Cscore-high group (**Figure 6E**). Next, the fraction of 22 immune cells of every patient from the four other datasets was predicted (**Supplementary Table S16**) and merged into high and low ac4Cscore groups. Distinct compositions of T cell types, including CD4 memory resting, CD4 memory activated, follicular helper, regulatory and gamma delta were observed between ac4Cscore high and low groups (**Figure 6G**). Detailed information about the correlation of the immune cells with ac4Cscore in each group and each dataset is also displayed (**Supplementary Figure S8B**). In addition, the cross-dataset results also showed consistent upregulation of various immune pathways, such as the IL-17 signaling pathway and Fc gamma R-mediated phagocytosis (**Figure 6E**). In conclusion, the predictive power of ac4Cscore in patient prognosis, tumor stemness and immune infiltration was confirmed using multiple datasets. The results suggest that ac4Cscore also work in other cancer types, including lung cancer and pancreatic cancer.

Role of ac4Cscore in Anti-PD1/mTOR Treatment Cohorts

Existing evidence has shown that the NAT10-mediated ac4C modification pattern is associated with tumor stemness and immune cell infiltration. Therefore, we investigated whether ac4Cscore could indicate patient response to immune checkpoint blockade therapy or targeted treatment with anti-PD1 and anti-mTOR cohorts. Patient information is listed in **Supplementary Table S14**. In the anti-PD1 cohort, patients with low ac4Cscore presented significant clinical benefits and notably prolonged overall and progression-free survival (**Figure 7A**). Compared with the low ac4Cscore group, the high ac4Cscore group had elevated mRNasi (**Supplementary Figure S9A** and **Supplementary Table S15**). Different distributions of immune cells could also be observed in these groups (**Supplementary Figure S9B** and **Supplementary Table S16**). In addition, the low ac4Cscore group exhibited a more intensified TMB than the high ac4Cscore group (**Supplementary Figure S9C**), with a mutation rate of 14 versus 6% for the 20th most mutated gene (**Figure 7B**). Moreover, the TIDE method was adopted to assess tumor immune evasion. Notably, although the two ac4Cscore groups showed distinct outcomes, there were no significant differences in either TIDE score or exclusion score between these groups except dysfunction score (**Figure 7C** and **Supplementary Figure S9D,E**), indicating that ac4C mRNA modification should be a novel mechanism involved in T cell dysfunction-related tumor immune evasion and has not been fully considered thus far. The predictive power of ac4Cscore on patient survival was also examined. ROC curves showed that ac4Cscore could be a better indicator than *PDCD1* expression or TIDE score to predict patient outcome at 12, 36 and 60 months (**Figure 7D**). For those treated with mTOR inhibitors, high ac4Cscore patients suffered from worse overall and progression-free survival than low ac4Cscore patients (**Figure 7E**). The gene mutation landscape of these two groups was also drawn (**Supplementary Figure S9F**), illustrating that the high ac4Cscore group had a higher mutation rate of *PBRM1* and *BAP1* than the low ac4Cscore group. The results above suggest that ac4Cscore could be a novel marker to predict patient prognosis with immunotherapy and targeted therapy.

DISCUSSION

Increasing evidence suggests that posttranscriptional mRNA modification plays an essential role in tumor initiation and progression, raising our attention to the epitranscriptome at an unprecedented level (Bauer et al., 2016; Gu H. et al., 2020; Gu L. et al., 2020; Blanco et al., 2021). As an ancient and highly conserved RNA modification in all domains of life, the process, distribution and function of ac4C in mRNAs has been unraveled in the past few years (Arango et al., 2018; Sas-Chen et al., 2020). However, ac4C remains relatively unexplored, and few studies have focused on how ac4C regulates and influences the intrinsic and extrinsic characteristics of tumors. Here, we reported an ac4Cscore model that can stratify patients into two groups with

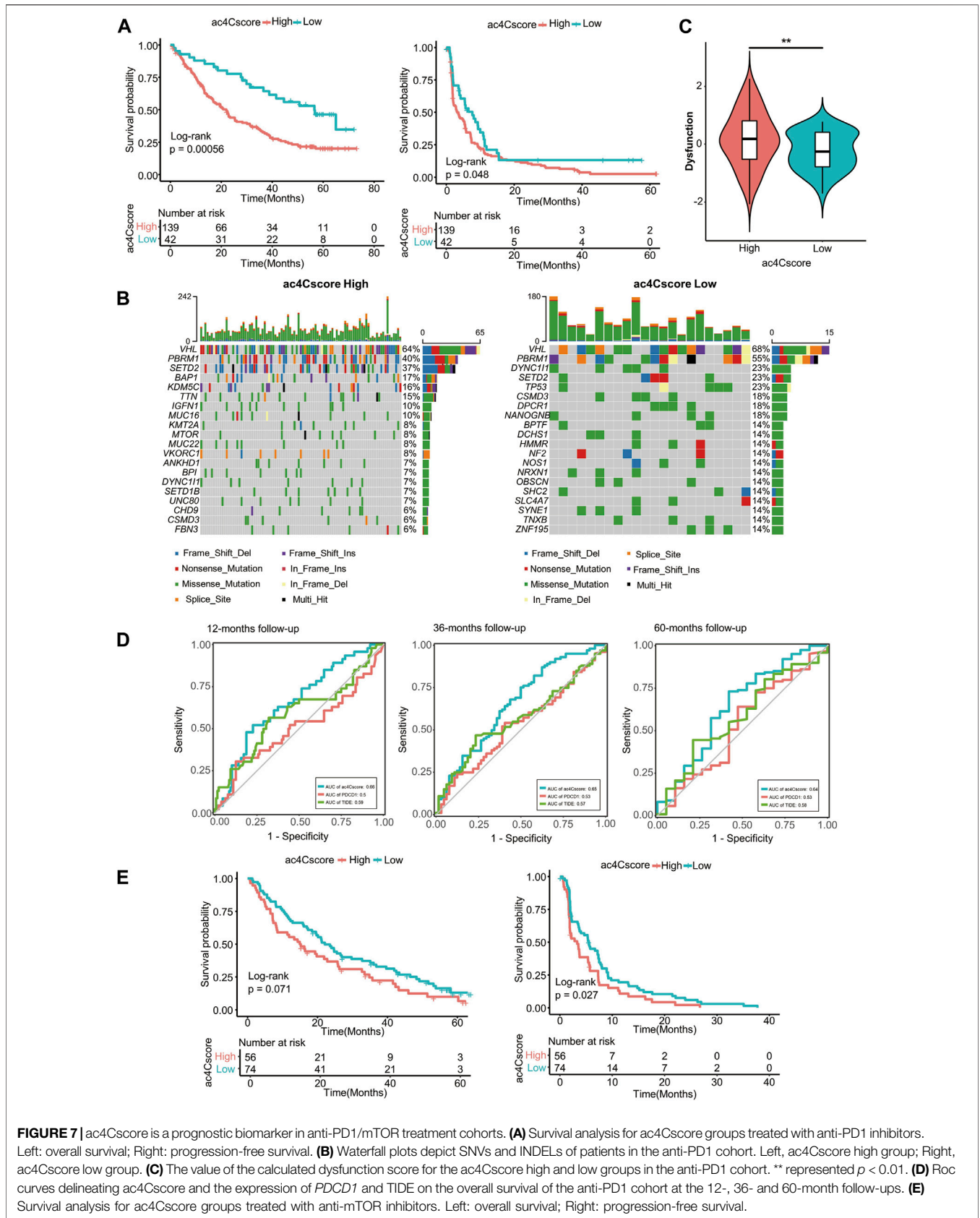


FIGURE 7 | ac4Cscore is a prognostic biomarker in anti-PD1/mTOR treatment cohorts. **(A)** Survival analysis for ac4Cscore groups treated with anti-PD1 inhibitors. Left: overall survival; Right: progression-free survival. **(B)** Waterfall plots depict SNVs and INDELS of patients in the anti-PD1 cohort. Left, ac4Cscore high group; Right, ac4Cscore low group. **(C)** The value of the calculated dysfunction score for the ac4Cscore high and low groups in the anti-PD1 cohort. ** represented $p < 0.01$. **(D)** Roc curves delineating ac4Cscore and the expression of *PDCD1* and TIDE on the overall survival of the anti-PD1 cohort at the 12-, 36- and 60-month follow-ups. **(E)** Survival analysis for ac4Cscore groups treated with anti-mTOR inhibitors. Left: overall survival; Right: progression-free survival.

distinct prognoses caused by different somatic mutation landscapes, stemness features and TME infiltration. Through multiple validation cohorts, we confirmed the conserved efficacy of ac4Cscore in predicting the prognosis of patients (**Supplementary Figure S10**).

In this study, we obtained ac4C-DEGs comparing HCC with normal liver samples by comprehensive data mining and bioinformatic analysis. In addition, we constructed a predictive model named ac4Cscore by multivariate Cox regression and LASSO algorithms to determine genes with the greatest effect on patient outcome. Multivariate analysis has the advantage of gleaning authentic associations between variables rather than focusing on any specific one. LASSO regression computes the average error and standard deviation over each fold and fits a generalized linear model by maximizing the partial likelihood with an elastic net penalty. These two methods enabled the selection of optimal genes, including a favorable factor, *COL15A1*, and two risk factors, *G6PD* and *TP53I3*. In other words, ac4C-related modification genes are a double-edged sword that inhibit and promote tumors at the same time. The ac4Cscore model managed to dichotomize patients into two groups. Of note, the low ac4Cscore group had a significantly prolonged survival outcome compared with the high ac4Cscore group. We also observed that pathologically, the high ac4Cscore group consisted of patients with advanced tumor grade. In this context, we sought to determine the characteristics of tumors behind ac4C modification patterns. For somatic mutation profiles, the SNVs and INDELS rate of *TP53* ranked the most frequently in the high ac4Cscore group, much higher than that in the low ac4Cscore group. The loss of this suppressor gene activation is strong evidence for the poor prognosis of patients (Staib et al., 2003). Another proof supporting our findings was the amplification regions of the high ac4Cscore group, involving cytoband 8q24.21, where oncogenes *MYC*, *POU5F1B* and lncRNA *PVT1* are located. The promoting roles of these genes in HCC have been widely studied and revealed (Qu et al., 2014; Pan et al., 2018; Jiang et al., 2020). We then performed GSEA to uncover the concordant biological difference between the two groups and observed that the cell cycle and DNA replication were enriched in the high ac4Cscore group, which is responsible for rapid tumor growth. We also found that the ribosome pathway was highly activated in the high ac4Cscore group. This made sense because Arango et al proved that ac4C modification promoted mRNA translation (Arango et al., 2018).

The ac4Cscore model was made up of three genes, *COL15A1*, *G6PD* and *TP53I3*. Collagen type XV alpha one chain (*COL15A1*) is a protein-coding gene mainly localized at basement membrane zones and is essential for maintaining the structure of the extracellular matrix. Emerging evidence has demonstrated that it is secreted from a fibroblast-tumor cell hybrid and acts as a tumor suppressor. Re-expression of *COL15A1* in the HeLa cell line hindered tumor formation in a xenograft model (Mutolo et al., 2012), and the ability of *COL15A1* to stabilize the extracellular matrix also prevented distal tumor metastasis (Clementz and Harris, 2013). Glucose-6-phosphate dehydrogenase (*G6PD*) is a housekeeping gene located on the X chromosome. It encodes a cytosolic enzyme whose main

function is to produce nicotinamide adenine dinucleotide phosphate (NADPH) and maintain redox homeostasis. Evidence has shown that *G6PD* is essential for cell growth, survival and embryonic development through redox-sensitive mechanisms and is therefore upregulated in several tumor types (Yang H. C. et al., 2019). In addition, the G6PD-NADPH redox system was shown to be important for stabilizing the metabolism of T cells and regulating antitumor immunity (Gu et al., 2021). It was reported as a prognostic marker to predict the immunotherapy response of Merkel cell carcinoma (Nakamura et al., 2020). For tumor protein p53 inducible protein 3 (*TP53I3*), it functions as reactive oxygen species and involves in DNA damage response pathway (Kotsinas et al., 2012). There were experiments illustrating oncogenic role of *TP53I3* in papillary thyroid cancer through activating the PI3K/AKT/PTEN pathway (Xu et al., 2015) and in non-small cell lung cancer by promoting mitotic progression (Li et al., 2017). In summary, the ac4Cscore model was formed from three key genes that individually had proven physiological functions in tumors, thereby representing ac4C mRNA modification patterns.

Except for cancer cells, the TME contains CSCs, stromal and immune cells, and other molecular components that interact with cancer cells and contribute to the development of tumors. A recent review suggested that CSC-immune cell crosstalk impacted tumor growth and the immune response, revealing their synergistic effect on immune escape, therapeutic resistance and recurrence (Bayik and Lathia, 2021). Despite the scarce evidence showing a direct association of ac4C with the TME, it is reasonable to hypothesize that ac4C regulates tumor stemness and immune infiltration. In this context, we examined ac4Cscore with such tumor characteristics. Cancer stem cells drive tumorigenesis, eventually leading to tumor progression and therapy resistance (Batlle and Clevers, 2017). We found that ac4Cscore was positively correlated with tumor stemness, and some stem cell-related gene sets were enriched in the high ac4Cscore group. The DNA repair system plays a key role in maintaining genome stability. However, they also confer the ability of CSCs to survive in extreme conditions by promoting the DNA damage sensor and repair machinery (Maugeri-Sacca et al., 2012). Additionally, we quantitatively calculated the infiltration of noncancerous cells. Since being proposed, the ESTIMATE and CIBERSORT algorithms have been proven effective in assessing the presence of stromal and immune cells. Evidence has shown the antitumor immunity of regulatory T helper cells, and diverse macrophages boost tumor progression and metastasis (Qian and Pollard, 2010; Togashi et al., 2019). We observed different stromal scores and immune scores along with the distribution of T cells and macrophages, which together assisted tumor evasion of immune destruction, accounting for the distinct prognosis of patients in the two ac4Cscore groups. Notably, the ac4Cscore high group predominated in the immune subtypes IFN-gamma Dominant and Wound Healing, which were classified as having increased expression of angiogenic genes and a higher proliferation rate, while the ac4Cscore low group was mainly attributed to subtype Inflammatory, defined as low to moderate tumor cell proliferation (Thorsson et al., 2018). These intrinsic and

extrinsic characteristics of tumors explained the different outcomes of patients in the two ac4Cscore groups. Finally, we validated the results in different primary tumor cohorts, suggesting the predictive power of the ac4Cscore model. Considering this, we next analyzed clinical cohorts treated with anti-PD1 and anti-mTOR inhibitors respectively. We also noted worse overall survival and progression-free survival in the high ac4Cscore group in both cohorts. Specifically, in the anti-PD1 cohort, we observed a more extensive TMB status in the low ac4Cscore group. Accumulating evidence has demonstrated that patients with high TMB correlate with an enhanced response, benefit from immunotherapy and therefore have prolonged survival (Goodman et al., 2017). Conventionally, the expression level of *PDCD1* predicts the outcome of anti-PD1 treatment. Recently, TIDE has emerged as a novel biomarker for predicting the outcome of melanoma and non-small lung cancer patients. The dysfunction score, instead of the TIDE score or exclusion score, which was calculated by TIDE, was significantly different between the ac4Cscore groups. This result suggested that ac4C mRNA modification related genes might impact T cell dysfunction with higher infiltration of cytotoxic T cells, such as T regulatory cells, but was not associated with the immune exclusion microenvironment. In comparison, time-dependent ROC curves showed that ac4Cscore was better than the expression of *PDCD1* or TIDE in predicting the OS of patients. These results indirectly supported the crucial role of the ac4Cscore model.

In conclusion, we developed an ac4Cscore model and proved that it was a promising biomarker for the prediction of patient outcome. The easy-to-use ability of ac4Cscore because it consists of only three genes broadens its clinical application. Our study showed for the first time that ac4C-related mRNA modification patterns might play an important role in regulating the intrinsic and extrinsic characteristics of tumors via a plethora of pathways, thus influencing the prognosis of patients. We complemented the role of transcriptional mRNA modification in tumor initiation and progression. However, limitations still remain. Our findings were based on bulk RNA sequencing data, further analysis at

single cell resolution could provide more rigorous information. Considering that tumors are a complicated system, the accuracy of this model is not outstanding, and how ac4C modification patterns directly impact the TME needs further experimental confirmation.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

Conceptualization: SL and JH. Original draft writing and editing: SL, YZ, LQ, YM, ZC, BZ and JH. Figure creating: SL, YZ, LQ and SZ. Resources, supervision, funding acquisition: CH, ZC and JH.

FUNDING

This work was supported by National Natural Science Foundation of China (31972884 to JH), the Foundation for Innovative Research Groups of the National Natural Science Foundation of China (81821002 to CH), National Clinical Research Center for Geriatrics (Z20201007 to JH), and 1-3-5 Project for Disciplines of Excellence, West China Hospital (ZYGD18003 to JH, ZYJC210021 to ZC), Sichuan University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2022.861000/full#supplementary-material>

REFERENCES

- Arango, D., Sturgill, D., Alhusaini, N., Dillman, A. A., Sweet, T. J., Hanson, G., et al. (2018). Acetylation of Cytidine in mRNA Promotes Translation Efficiency. *Cell* 175 (7), 1872–1886. e24. doi:10.1016/j.cell.2018.10.030
- Battle, E., and Clevers, H. (2017). Cancer Stem Cells Revisited. *Nat. Med.* 23 (10), 1124–1134. doi:10.1038/nm.4409
- Bayik, T., Trump, S., Ishaque, N., Thurmann, L., Gu, L., Bauer, M., et al. (2016). Environment-induced Epigenetic Reprogramming in Genomic Regulatory Elements in Smoking Mothers and Their Children. *Mol. Syst. Biol.* 12 (3), 861. doi:10.15252/msb.20156520
- Bayik, D., and Lathia, J. D. (2021). Cancer Stem Cell-Immune Cell Crosstalk in Tumour Progression. *Nat. Rev. Cancer* 21 (8), 526–536. doi:10.1038/s41568-021-00366-w
- Binnewies, M., Roberts, E. W., Kersten, K., Chan, V., Fearon, D. F., Merad, M., et al. (2018). Understanding the Tumor Immune Microenvironment (TIME) for Effective Therapy. *Nat. Med.* 24 (5), 541–550. doi:10.1038/s41591-018-0014-x
- Blanco, M. A., Sykes, D. B., Gu, L., Wu, M., Petroni, R., Karnik, R., et al. (2021). Chromatin-state Barriers Enforce an Irreversible Mammalian Cell Fate Decision. *Cell Rep* 37 (6), 109967. doi:10.1016/j.celrep.2021.109967
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* 30 (15), 2114–2120. doi:10.1093/bioinformatics/btu170
- Braun, D. A., Hou, Y., Bakouny, Z., Ficial, M., Sant' Angelo, M., Forman, J., et al. (2020). Interplay of Somatic Alterations and Immune Infiltration Modulates Response to PD-1 Blockade in Advanced clear Cell Renal Cell Carcinoma. *Nat. Med.* 26 (6), 909–918. doi:10.1038/s41591-020-0839-y
- Clementz, A. G., and Harris, A. (2013). Collagen XV: Exploring its Structure and Role within the Tumor Microenvironment. *Mol. Cancer Res.* 11 (12), 1481–1486. doi:10.1158/1541-7786.Mcr-12-0662
- Davis, S., and Meltzer, P. S. (2007). GEOquery: a Bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics* 23 (14), 1846–1847. doi:10.1093/bioinformatics/btm254
- Delaunay, S., and Frye, M. (2019). RNA Modifications Regulating Cell Fate in Cancer. *Nat. Cell Biol* 21 (5), 552–559. doi:10.1038/s41556-019-0319-0

- Esteve-Puig, R., Bueno-Costa, A., and Esteller, M. (2020). Writers, Readers and Erasers of RNA Modifications in Cancer. *Cancer Lett.* 474, 127–137. doi:10.1016/j.canlet.2020.01.021
- Gao, Y., Lyu, L., Feng, Y., Li, F., and Hu, Y. (2021). A Review of Cutting-Edge Therapies for Hepatocellular Carcinoma (HCC): Perspectives from Patents. *Int. J. Med. Sci.* 18 (14), 3066–3081. doi:10.7150/ijms.59930
- Goldman, M. J., Craft, B., Hastie, M., Repecka, K., McDade, F., Kamath, A., et al. (2020). Visualizing and Interpreting Cancer Genomics Data via the Xena Platform. *Nat. Biotechnol.* 38 (6), 675–678. doi:10.1038/s41587-020-0546-8
- Gonzalez, H., Hagerling, C., and Werb, Z. (2018). Roles of the Immune System in Cancer: from Tumor Initiation to Metastatic Progression. *Genes Dev.* 32 (19–20), 1267–1284. doi:10.1101/gad.314617.118
- Goodman, A. M., Kato, S., Bazhenova, L., Patel, S. P., Frampton, G. M., Miller, V., et al. (2017). Tumor Mutational Burden as an Independent Predictor of Response to Immunotherapy in Diverse Cancers. *Mol. Cancer Ther.* 16 (11), 2598–2608. doi:10.1158/1535-7163.MCT-17-0386
- Gu, H., Guo, Y., Gu, L., Wei, A., Xie, S., Ye, Z., et al. (2020). Deep Learning for Identifying Corneal Diseases from Ocular Surface Slit-Lamp Photographs. *Sci. Rep.* 10 (1), 17851. doi:10.1038/s41598-020-75027-3
- Gu, L., Frommel, S. C., Oakes, C. C., Simon, R., Grupp, K., Gerig, C. Y., et al. (2015). BAZZA (TIP5) Is Involved in Epigenetic Alterations in Prostate Cancer and its Overexpression Predicts Disease Recurrence. *Nat. Genet.* 47 (1), 22–30. doi:10.1038/ng.3165
- Gu, L., Wang, L., Chen, H., Hong, J., Shen, Z., Dhall, A., et al. (2020). CG14906 (Mettl4) Mediates M(6)A Methylation of U2 snRNA in *Drosophila*. *Cell Discov* 6, 44. doi:10.1038/s41421-020-0178-7
- Gu, M., Zhou, X., Sohn, J. H., Zhu, L., Jie, Z., Yang, J. Y., et al. (2021). NF-kappa-B-inducing Kinase Maintains T Cell Metabolic Fitness in Antitumor Immunity. *Nat. Immunol.* 22 (2), 193–204. doi:10.1038/s41590-020-00829-6
- Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). Circlize Implements and Enhances Circular Visualization in R. *Bioinformatics* 30 (19), 2811–2812. doi:10.1093/bioinformatics/btu393
- ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium (2020). Pan-cancer Analysis of Whole Genomes. *Nature* 578 (7793), 82–93. doi:10.1038/s41586-020-1969-6
- Jiang, B., Yang, B., Wang, Q., Zheng, X., Guo, Y., and Lu, W. (2020). lncRNA PVT1 Promotes Hepatitis B Viruspositive Liver Cancer Progression by Disturbing Histone Methylation on the cMyc Promoter. *Oncol. Rep.* 43 (2), 718–726. doi:10.3892/or.2019.7444
- Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of T Cell Dysfunction and Exclusion Predict Cancer Immunotherapy Response. *Nat. Med.* 24 (10), 1550–1558. doi:10.1038/s41591-018-0136-1
- Jin, G., Xu, M., Zou, M., and Duan, S. (2020). The Processing, Gene Regulation, Biological Functions, and Clinical Relevance of N4-Acetylcytidine on RNA: A Systematic Review. *Mol. Ther. Nucleic Acids* 20, 13–24. doi:10.1016/j.omtn.2020.01.037
- Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based Genome Alignment and Genotyping with HISAT2 and HISAT-Genotype. *Nat. Biotechnol.* 37 (8), 907–915. doi:10.1038/s41587-019-0201-4
- Kotsinas, A., Aggarwal, V., Tan, E. J., Levy, B., and Gorgoulis, V. G. (2012). PIG3: a Novel Link between Oxidative Stress and DNA Damage Response in Cancer. *Cancer Lett.* 327 (1–2), 97–102. doi:10.1016/j.canlet.2011.12.009
- Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., and Storey, J. D. (2012). The Sva Package for Removing Batch Effects and Other Unwanted Variation in High-Throughput Experiments. *Bioinformatics* 28 (6), 882–883. doi:10.1093/bioinformatics/bts034
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). Project Data Processing: The Sequence Alignment/Map Format and SAMtools. *Bioinformatics* 25 (16), 2078–2079. doi:10.1093/bioinformatics/btp352
- Li, M., Li, S., Liu, B., Gu, M. M., Zou, S., Xiao, B. B., et al. (2017). PIG3 Promotes NSCLC Cell Mitotic Progression and Is Associated with Poor Prognosis of NSCLC Patients. *J. Exp. Clin. Cancer Res.* 36 (1), 39. doi:10.1186/s13046-017-0508-2
- Liu, S., Zhang, Y., Zhang, S., Qiu, L., Zhang, B., and Han, J. (2021). Identification of Hub Genes Related to Liver Metastasis of Colorectal Cancer by Integrative Analysis. *Front. Oncol.* 11, 714866. doi:10.3389/fonc.2021.714866
- Llovet, J. M., Kelley, R. K., Villanueva, A., Singal, A. G., Pikarsky, E., Roayaie, S., et al. (2021). Hepatocellular Carcinoma. *Nat. Rev. Dis. Primers* 7 (1), 6. doi:10.1038/s41572-020-00240-3
- Malta, T. M., Sokolov, A., Gentles, A. J., Burzykowski, T., Poisson, L., Weinstein, J. N., et al. (2018). Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell* 173 (2), 338–354. e15. doi:10.1016/j.cell.2018.03.034
- Maugeri-Sacca, M., Bartucci, M., and De Maria, R. (2012). DNA Damage Repair Pathways in Cancer Stem Cells. *Mol. Cancer Ther.* 11 (8), 1627–1636. doi:10.1158/1535-7163.MCT-11-1040
- Mayakonda, A., Lin, D. C., Assenov, Y., Plass, C., and Koeffler, H. P. (2018). Maftools: Efficient and Comprehensive Analysis of Somatic Variants in Cancer. *Genome Res.* 28 (11), 1747–1756. doi:10.1101/gr.239244.118
- Meng, Y., Hu, X., Li, S., Zeng, X., Qiu, L., Wei, M., et al. (2020). miR-203 Inhibits Cell Proliferation and ERK Pathway in Prostate Cancer by Targeting IRS-1. *BMC Cancer* 20 (1), 1028. doi:10.1186/s12885-020-07472-2
- Mermel, C. H., Schumacher, S. E., Hill, B., Meyerson, M. L., Beroukhi, R., and Getz, G. (2011). GISTIC2.0 Facilitates Sensitive and Confident Localization of the Targets of Focal Somatic Copy-Number Alteration in Human Cancers. *Genome Biol.* 12 (4), R41. doi:10.1186/gb-2011-12-4-r41
- Mutolo, M. J., Morris, K. J., Leir, S. H., Caffrey, T. C., Lewandowska, M. A., Hollingsworth, M. A., et al. (2012). Tumor Suppression by Collagen XV Is Independent of the Restin Domain. *Matrix Biol.* 31 (5), 285–289. doi:10.1016/j.matbio.2012.03.003
- Nakamura, M., Nagase, K., Yoshimitsu, M., Magara, T., Nojiri, Y., Kato, H., et al. (2020). Glucose-6-phosphate Dehydrogenase Correlates with Tumor Immune Activity and Programmed Death Ligand-1 Expression in Merkel Cell Carcinoma. *J. Immunother. Cancer* 8 (2), e001679. doi:10.1136/jitc-2020-001679
- Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust Enumeration of Cell Subsets from Tissue Expression Profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337
- Pan, Y., Zhan, L., Chen, L., Zhang, H., Sun, C., and Xing, C. (2018). POU5F1B Promotes Hepatocellular Carcinoma Proliferation by Activating AKT. *Biomed. Pharmacother.* 100, 374–380. doi:10.1016/j.biopha.2018.02.023
- Qian, B. Z., and Pollard, J. W. (2010). Macrophage Diversity Enhances Tumor Progression and Metastasis. *Cell* 141 (1), 39–51. doi:10.1016/j.cell.2010.03.014
- Qu, A., Jiang, C., Cai, Y., Kim, J. H., Tanaka, N., Ward, J. M., et al. (2014). Role of Myc in Hepatocellular Proliferation and Hepatocarcinogenesis. *J. Hepatol.* 60 (2), 331–338. doi:10.1016/j.jhep.2013.09.024
- Ramirez, F., Ryan, D. P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., et al. (2016). deepTools2: a Next Generation Web Server for Deep-Sequencing Data Analysis. *Nucleic Acids Res.* 44 (W1), W160–W165. doi:10.1093/nar/gkw257
- Ritchie, M., Phipson, B., Wu, D., Hu, Y., Law, C., Shi, W., et al. (2015). LIMMA powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res.* 43, e47. doi:10.1093/nar/gkv007
- Roessler, S., Jia, H. L., Budhu, A., Forgues, M., Ye, Q. H., Lee, J. S., et al. (2010). A Unique Metastasis Gene Signature Enables Prediction of Tumor Relapse in Early-Stage Hepatocellular Carcinoma Patients. *Cancer Res.* 70 (24), 10202–10212. doi:10.1158/0008-5472.CAN-10-2607
- Ruiz de Galarreta, M., Bresnahan, E., Molina-Sanchez, P., Lindblad, K. E., Maier, B., Sia, D., et al. (2019). Beta-Catenin Activation Promotes Immune Escape and Resistance to Anti-PD-1 Therapy in Hepatocellular Carcinoma. *Cancer Discov.* 9 (8), 1124–1141. doi:10.1158/2159-8290.CD-19-0074
- Sas-Chen, A., Thomas, J. M., Matzov, D., Taoka, M., Nance, K. D., Nir, R., et al. (2020). Dynamic RNA Acetylation Revealed by Quantitative Cross-Evolutionary Mapping. *Nature* 583 (7817), 638–643. doi:10.1038/s41586-020-2418-2
- Schulenburg, A., Blatt, K., Cerny-Reiterer, S., Sadovnik, I., Herrmann, H., Marian, B., et al. (2015). Cancer Stem Cells in Basic Science and in Translational Oncology: Can We Translate into Clinical Application? *J. Hematol. Oncol.* 8, 16. doi:10.1186/s13045-015-0113-9
- Simon, N., Friedman, J. H., Hastie, T., and Tibshirani, R. (2011). Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent. *J. Stat. Softw.* 39 (5), 13. doi:10.18637/jss.v039.i05
- Staib, F., Hussain, S. P., Hofseth, L. J., Wang, X. W., and Harris, C. C. (2003). TP53 and Liver Carcinogenesis. *Hum. Mutat.* 21 (3), 201–216. doi:10.1002/humu.10176

- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-wide Expression Profiles. *Proc. Natl. Acad. Sci.* 102 (43), 15545–15550. doi:10.1073/pnas.0506580102
- Sun, J. H., Luo, Q., Liu, L. L., and Song, G. B. (2016). Liver Cancer Stem Cell Markers: Progression and Therapeutic Implications. *World J. Gastroenterol.* 22 (13), 3547–3557. doi:10.3748/wjg.v22.i13.3547
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660
- Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T. H., et al. (2018). The Immune Landscape of Cancer. *Immunity* 48 (4), 812–830. e14. doi:10.1016/j.immuni.2018.03.023
- Thorvaldsdottir, H., Robinson, J. T., and Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): High-Performance Genomics Data Visualization and Exploration. *Brief Bioinform* 14 (2), 178–192. doi:10.1093/bib/bbs017
- Togashi, Y., Shitara, K., and Nishikawa, H. (2019). Regulatory T Cells in Cancer Immunosuppression - Implications for Anticancer Therapy. *Nat. Rev. Clin. Oncol.* 16 (6), 356–371. doi:10.1038/s41571-019-0175-7
- Visvader, J. E., and Lindeman, G. J. (2008). Cancer Stem Cells in Solid Tumours: Accumulating Evidence and Unresolved Questions. *Nat. Rev. Cancer* 8 (10), 755–768. doi:10.1038/nrc2499
- Wang, C., Vegna, S., Jin, H., Benedict, B., Liefink, C., Ramirez, C., et al. (2019). Inducing and Exploiting Vulnerabilities for the Treatment of Liver Cancer. *Nature* 574 (7777), 268–272. doi:10.1038/s41586-019-1607-3
- Wang, Q., Gu, L., Adey, A., Radlwimmer, B., Wang, W., Hovestadt, V., et al. (2013). Tagmentation-based Whole-Genome Bisulfite Sequencing. *Nat. Protoc.* 8 (10), 2022–2032. doi:10.1038/nprot.2013.118
- Xu, J., Cai, J., Jin, X., Yang, J., Shen, Q., Ding, X., et al. (2015). PI3K Plays an Oncogenic Role in Papillary Thyroid Cancer by Activating the PI3K/AKT/PTEN Pathway. *Oncol. Rep.* 34 (3), 1424–1430. doi:10.3892/or.2015.4096
- Yang, H. C., Wu, Y. H., Yen, W. C., Liu, H. Y., Hwang, T. L., Stern, A., et al. (2019). The Redox Role of G6PD in Cell Growth, Cell Death, and Cancer. *Cells* 8 (9), 1055. doi:10.3390/cells8091055
- Yang, J. D., Hainaut, P., Gores, G. J., Amadou, A., Plymoth, A., and Roberts, L. R. (2019). A Global View of Hepatocellular Carcinoma: Trends, Risk, Prevention and Management. *Nat. Rev. Gastroenterol. Hepatol.* 16 (10), 589–604. doi:10.1038/s41575-019-0186-y
- Yoshihara, K., Shahmoradgoli, M., Martinez, E., Vegesna, R., Kim, H., Torres-Garcia, W., et al. (2013). Inferring Tumour Purity and Stromal and Immune Cell Admixture from Expression Data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612
- Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A J. Integr. Biol.* 16 (5), 284–287. doi:10.1089/omi.2011.0118
- Zhang, B., Wu, Q., Li, B., Wang, D., Wang, L., and Zhou, Y. L. (2020). m(6A Regulator-Mediated Methylation Modification Patterns and Tumor Microenvironment Infiltration Characterization in Gastric Cancer. *Mol. Cancer* 19 (1), 53. doi:10.1186/s12943-020-01170-0
- Zhang, Y., Jing, Y., Wang, Y., Tang, J., Zhu, X., Jin, W. L., et al. (2021). NAT10 Promotes Gastric Cancer Metastasis via N4-Acetylated COL5A1. *Signal. Transduct. Target. Ther.* 6 (1), 173. doi:10.1038/s41392-021-00489-4
- Zheng, R., Qu, C., Zhang, S., Zeng, H., Sun, K., Gu, X., et al. (2018). Liver Cancer Incidence and Mortality in China: Temporal Trends and Projections to 2030. *Chin. J. Cancer Res.* 30 (6), 571–579. doi:10.21147/j.issn.1000-9604.2018.06.01
- Zhong, J., Liu, Z., Cai, C., Duan, X., Deng, T., and Zeng, G. (2021). m(6A Modification Patterns and Tumor Immune Landscape in clear Cell Renal Carcinoma. *J. Immunother. Cancer* 9 (2), e001646. doi:10.1136/jitc-2020-001646
- Zhou, Z., Zhang, J., Xu, C., Yang, J., Zhang, Y., Liu, M., et al. (2021). An Integrated Model of N6-Methyladenosine Regulators to Predict Tumor Aggressiveness and Immune Evasion in Pancreatic Cancer. *EBioMedicine* 65, 103271. doi:10.1016/j.ebiom.2021.103271

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

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

Copyright © 2022 Liu, Zhang, Qiu, Zhang, Meng, Huang, Chen, Zhang and Han. 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.