



OPEN ACCESS

EDITED BY
Nadine Bley,
University of Halle, Germany

REVIEWED BY
Danny Misiak,
Martin Luther University of Halle-
Wittenberg, Germany
Simon Müller,
Martin Luther University of Halle-
Wittenberg, Germany

*CORRESPONDENCE
Nathalie Britzen-Laurent,
nathalie.britzen-laurent@uk-
erlangen.de

SPECIALTY SECTION
This article was submitted to Cancer
Cell Biology,
a section of the journal
Frontiers in Cell and Developmental
Biology

RECEIVED 22 June 2022
ACCEPTED 28 July 2022
PUBLISHED 19 August 2022

CITATION
Yun D, Yang Z, Zhang S, Yang H, Liu D,
Grützmann R, Pilarsky C and
Britzen-Laurent N (2022), An m5C
methylation regulator-associated
signature predicts prognosis and
therapy response in pancreatic cancer.
Front. Cell Dev. Biol. 10:975684.
doi: 10.3389/fcell.2022.975684

COPYRIGHT
© 2022 Yun, Yang, Zhang, Yang, Liu,
Grützmann, Pilarsky and Britzen-
Laurent. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

An m5C methylation regulator-associated signature predicts prognosis and therapy response in pancreatic cancer

Duo Yun¹, Zhirong Yang¹, Shuman Zhang¹, Hai Yang¹,
Dongxue Liu¹, Robert Grützmann², Christian Pilarsky¹ and
Nathalie Britzen-Laurent^{1*}

¹Division of Surgical Research, Department of Surgery, Universitätsklinikum Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany, ²Department of Surgery, Universitätsklinikum Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

Pancreatic ductal adenocarcinoma (PDAC) is the most aggressive digestive malignancy due to frequent late-stage diagnosis, rapid progression and resistance to therapy. With increasing PDAC incidence worldwide, there is an urgent need for new prognostic biomarkers and therapy targets. Recently, RNA methylation has emerged as a new tumorigenic mechanism in different cancers. 5-methylcytosine (m5C) is one of the most frequent RNA modifications and occurs on a variety of RNA species including mRNA, thereby regulating gene expression. Here we investigated the prognostic role of m5C-regulator-associated transcriptional signature in PDAC. We evaluated m5C-regulator status and expression in 239 PDAC samples from publicly available datasets. We used unsupervised consensus clustering analyses to classify PDACs based on m5C-regulator expression. From the resulting signature of differentially expressed genes (DEGs), we selected prognosis-relevant DEGs to stratify patients and build a scoring signature (m5C-score) through LASSO and multivariate Cox regression analyses. The m5C-score represented a highly significant independent prognostic marker. A high m5C-score correlated with poor prognosis in different PDAC cohorts, and was associated with the squamous/basal subtype as well as activated cancer-related pathways including Ras, MAPK and PI3K pathways. Furthermore, the m5C-score correlated with sensitivity to pathway-specific inhibitors of PARP, EGFR, AKT, HER2 and mTOR. Tumors with high m5C-score were characterized by overall immune exclusion, low CD8⁺ T-cell infiltration, and higher PD-L1 expression. Overall, the m5C-score represented a robust predictor of prognosis and therapy response in PDAC, which was associated with unfavorable molecular subtypes and immune microenvironment.

KEYWORDS

5-methylcytosine (m5C), pancreatic cancer, prognosis, targeted drug therapy, tumor microenvironment (TME), immune cells infiltration

Introduction

Pancreatic ductal adenocarcinoma (PDAC) is one of the deadliest human malignant tumors, with an overall 5-year survival rate below 11%. It represents the seventh leading cause of cancer death globally for both sexes. In high-development-index countries, its mortality rank is estimated to increase from fifth to second or third by 2030 (Rahib et al., 2014; Siegel et al., 2022). The high mortality of PDAC has been notably imputed to the absence of obvious symptom at early stages, resulting in the majority of patients being diagnosed with locally advanced or metastatic tumors (Edwards et al., 2005; Arslan et al., 2010). While late diagnosis reduces the number of treatment options, PDAC is also highly recalcitrant to most therapies, including immunotherapy (Sung et al., 2021). Thus, it is crucial to develop new prognostic and predictive tools to improve therapeutic intervention in a personalized manner.

Post-transcriptional RNA methylation has emerged as an important epigenetic regulatory mechanism, which has recently gathered increasing interest, notably for its association with human cancer progression (Haruehanroengra et al., 2020; Han et al., 2021). 5-Methylcytosine (m5C) is one of the most prevalent RNA modifications, affecting mRNAs, tRNAs and rRNAs (Chen et al., 2021). The methylation of RNA cytosines is reversibly and specifically modulated by regulators including methyltransferases (writers) and demethylases (erasers), and is specifically recognized by binding proteins (readers) (Bohnsack et al., 2019). The m5C-modification of mRNAs has been shown to regulate their splicing, transport, translation, stabilization or degradation (Boo and Kim, 2020; Chen et al., 2021). Dysregulation of m5C-regulator gene expression is involved in pathophysiological processes such as cell proliferation, cell death or immune modulation, and has been observed in various cancers (Zhang et al., 2021a; Xu et al., 2021b). Changes in m5C regulators expression or m5C-regulator-associated signatures have been found to correlate with prognosis, immune infiltration and/or therapy response in different solid tumor entities (Zhang et al., 2021b; Wood et al., 2021).

The purpose of the present study was to investigate the prognostic relevance of tumor-associated dysregulation of the m5C-modification pathway for pancreas cancer patients. We combined genomic, transcriptional and clinical data from 239 PDAC patients to test the prognostic power of m5C-regulators and m5C-associated transcriptional signatures. We investigated the impact of different m5C-associated patterns on survival, gene/pathway expression, molecular subtypes, therapy response and immune infiltration in order to provide a comprehensive evaluation of the pathophysiological consequences of the m5C pathway dysregulation in PDAC.

Material and methods

Collection and pre-processing of pancreatic ductal adenocarcinoma gene expression public datasets

The gene and clinical information data were searched from the databases of Gene Expression-Omnibus (GEO) (<https://www.ncbi.nlm.nih.gov/geo/>), the Cancer Genome Atlas (TCGA) (<https://portal.gdc.cancer.gov/>) and the supplementary data of Bailey's cohort (<https://www.nature.com/articles/nature16965#Sec11>). For TCGA datasets, the fragments per kilobase of transcript per million mapped reads (FPKM) values were transformed into transcripts per kilobase million (TPM) values by using the R package "limma" (v. 3.50.3). For building and internal testing of prognostic model, we combined TCGA PAAD cohort and GSE57495 (RMA value) cohort data ($n = 239$). Batch effects were corrected using the "ComBat" function (batchType, par.prior = TRUE) of R package "sva" (v. 3.42.0) (Leek et al., 2012). Principal components analysis (PCA) algorithm was used to confirm the results of the batch effects correction based on the normalized sequencing data using the "prcomp" (scale. = TRUE) function in R package "limma" (v. 3.50.3). For external validation of the prognostic model, data from the GSE21501 (data were Lowess-normalized and then log₂ ratio was taken) cohort and Bailey's cohort (normalized RSEM data were converted to counts per million (c.p.m.) and log₂ transformed) were used independently. For comparison with other molecular classifications, Moffitt's cohort data (GSE71729: data were non-negative normalized log₂ Cy5 signal), Bailey's cohort data, Collision's cohort data (GSE17891; RMA log₂ data) were used. The copy number variation data of PDAC patients were downloaded from the UCSC Xena database (<https://xena.ucsc.edu/>) and the somatic mutation data were downloaded from the TCGA database. The genetic landscape of 12 m5C-regulators in 23 pairs of chromosomes and the variation were visualized using the R-package "Rcirco" (v. 1.2.2) (Zhang et al., 2013). The mutation status was visualized using the R-package "maftools" (v. 2.10.5) (Mayakonda et al., 2018). The expression of m5C regulators in pancreatic tumors and normal tissues was analyzed with the online tools GEPIA2 (<http://gepia2.cancer-pku.cn/#analysis>) (Tang et al., 2019) using data from TCGA and GTEx (<https://gtexportal.org/home/>) (Carithers and Moore, 2015). Patients without survival data were removed from further survival-related analyses. The cell line RNA-seq data were obtained from the Moffitt's cohort data (GSE71729) and GSE165949 (Alfarano et al., 2022), the latter as data normalized by "summarizeOverlaps" from the R package GenomicRanges (v. 1.38). Basic information of these datasets is listed in Supplementary Table S1, and the study workflow is presented in Supplementary Figure S1.

Unsupervised consensus cluster analysis

The “Consensusclusterplus” function (maxK = 9, reps = 50, clusterAlg = “pam”, distance = “euclidean”) was used to perform unsupervised consensus clustering analysis in order to classify PDAC patients based on normalized RNA sequencing data in R package “Consensusclusterplus” (v. 1.58.0) (Wilkerson and Hayes, 2010). The best k-value was selected based on the clustering effect and the cumulative distribution function (CDF). The result of classification was confirmed by Principal Component Analysis (PCA) based on the normalized sequencing data using the “prcomp” (scale. = TRUE) function in R basic package.

Gene set variation analysis

The “gsva” function (geneSets, min.sz = 10, max.sz = 500,) was used to analyze the transcriptome gene enrichment (Hänzelmann et al., 2013) using the gene document named “c2.cp.biocarta.v7.4.symbols.gmt” from the MSigDB database (<https://www.gsea-msigdb.org>). We used the R-packages of “gene set variation analysis (GSVA)” (v. 1.42.0) “limma” (v. 3.50.3), “GSEABase” (v. 1.56.0) to analyze the different pathways between the different m5C-clusters.

Construction and validation of the m5C-related prognostic risk scoring signature (m5C-score)

The differentially expression genes (DEGs) between three m5C-clusters were analyzed by using “lmFit”, “contrasts.fit” and “eBayes” in the R package “limma” (v.3.50.3) with a *p*-value filter adjusted to 0.001. We determined the differentially expressed genes (DEGs) between m5C-cluster A and B, A and C, or B and C. The potential functions of these DEGs were then analyzed by Gene Ontology (GO) annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment pathway analysis by using the “enrichKEGG” (organism = “hsa”, pvalueCutoff = 1, qvalueCutoff = 1) and “enrichGO” (OrgDb = org.Hs.eg.db, pvalueCutoff = 1, qvalueCutoff = 1) function in the R-package “clusterProfiler” (v. 4.2.2) and “org.Hs.eg.db”(version 3.14.0). An FDR < 0.05 was considered as significant. To build the m5C-score, we first randomly attributed the 239 patients to a training set (120 patients) and a testing set (119 patients). Univariate Cox regression analysis was applied on the training set to screen prognostic related DEGs with a *p*-value < 0.05 by using the function “coxph” in the R-package “survival” (v. 3.3.1). Then the least absolute shrinkage and selection operator (LASSO) regression analysis was used to avoid over-fitting by using the “glmnet” function (family = “cox”, maxit = 1000) function in the R-package “glmnet” (v. 4.1.4) where the value of the lambda is

chosen by the smallest likelihood deviance. Finally, multivariate Cox regression analysis was used to build the multivariate Cox proportional hazards regression model by using the “coxph” function in the R package “survival” (v. 3.3.1). The resulting m5C-score was constructed by four genes differentially regulated between clusters A, B and C (HIPK3, ZFAND4, DPP8, HIPK2). The signature could be represented using the formula: m5C-score = $\sum_n^i (Coef_i * X_i)$, in which *X* represents the expression level of each m5C-score gene and Coef represents the coefficient of each m5C-score gene in the multivariate Cox proportional hazards regression model. The m5C-score was then validated with the testing set, as well as two independent datasets (GSE21501 and Bailey cohort).

Gene set enrichment analysis

The gene set enrichment analysis (GSEA) software (version 4.1.0) was used to determine which pathways were significantly different between high- and low-m5C-score groups using the GSEA software (<https://www.broadinstitute.org/gsea/>). The KEGG-database “c2.cp.kegg.v7.5.symbols.gmt” was chosen as the reference database. The number of random permutations was set to 1000. The gene expression data in the TCGA-PAAD database were grouped via the m5C-score and recorded as the high- and low-m5C-score groups. The phenotype labels were set as high-m5C-score group versus low-m5C-score group patients. Collapse/Remap to gene symbols was set as “No_Collapse”. The other parameters were the default parameters in the GSEA software. A pathway with FDR *q* < 0.05 was defined as statistically significant.

Target drugs sensitivity prediction

The sensitivity to target drugs was determined by estimating the half-maximal inhibitory concentration (IC50) with the “pRRopheticPredict” (selection = 1) function in the R-package “pRRophetic” (version 0.5) (Geeleher et al., 2014). Common targets (Liu et al., 2021) being used in clinic or clinical trials were analyzed.

Immune cells infiltration and exclusion analyses

Cell type Identification By Estimating Relative Subsets Of RNA Transcripts (CIBERSORT) (Newman et al., 2015) (<https://cibersort.stanford.edu>) was used to quantify the proportions and distributions of tumor-infiltrating immune cells (TIICs) based on the RNA-seq data. The single-sample gene set enrichment analysis (ssGSEA) (Barbie et al., 2009) method was used to calculate the relative enrichment of immune cells for each

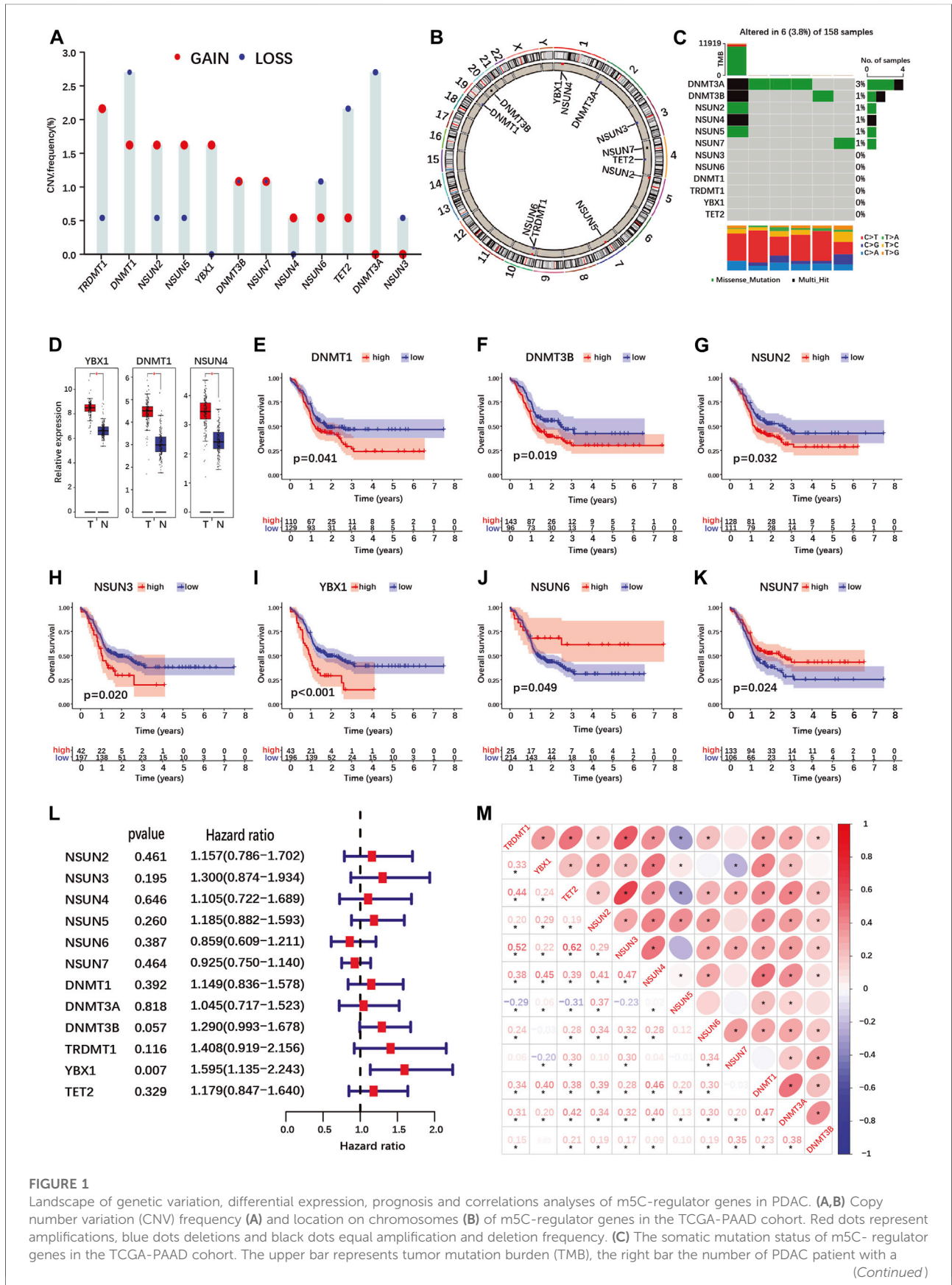


FIGURE 1

Landscape of genetic variation, differential expression, prognosis and correlations analyses of m5C-regulator genes in PDAC. (A,B) Copy number variation (CNV) frequency (A) and location on chromosomes (B) of m5C-regulator genes in the TCGA-PAAD cohort. Red dots represent amplifications, blue dots deletions and black dots equal amplification and deletion frequency. (C) The somatic mutation status of m5C-regulator genes in the TCGA-PAAD cohort. The upper bar represents tumor mutation burden (TMB), the right bar the number of PDAC patient with a

(Continued)

FIGURE 1

mutation in one or more m5C-regulator gene, and the bar below the distribution of conversions in each sample. (D) Differential expression of m5C-methylation regulators between PDAC and normal pancreas tissue in the TCGA-PAAD and GTEx cohorts. Tumor, red; Normal, blue. (E–K) K-M curves illustrating OS between high- and low- expression of m5C-regulator genes analyzed by log-rank test in the TCGA-PAAD and GSE57495 cohorts. (L) Univariate Cox regression analyses of OS of m5C-regulator genes in the TCGA-PAAD and GSE57495 cohorts. (M) Correlation analyses between m5C-regulator gene expression by Spearman analyses. * p -value < 0.05.

PDAC patient by using the “gsva” (method = “ssgsea”, kcdf = “Gaussian”, abs.ranking = TRUE) function in the R-package “GSVA” (v. 1.42.0). The marker genes for each infiltrating immune cell type were based on the results of Charoentong and colleagues (Charoentong et al., 2017). The Tumor Immune Dysfunction and Exclusion (TIDE) (<http://tide.dfci.harvard.edu/>) algorithm was used to predict PDAC immune exclusion of samples from the TCGA-PAAD cohort (Jiang et al., 2018).

Statistical analysis

For data analyses, we used the R-Studio (4.1.2) suite. For the comparison of data between two groups, the Student’s t -test was used for analyzing the quantitative statistics of normally distributed variables and the Wilcoxon rank sum test was used for analyzing the quantitative statistics of non-normally distributed variables. If the comparisons were between more than two groups, one-way ANOVA was used for parametric methods and the Kruskal–Wallis test was used for the non-parametric methods. The Kaplan–Meier (K-M) method and log-rank test were used for survival analyses. The fulfillment of the proportional hazards assumption was verified for all survival analyses with a significant survival difference (log-rank test p < 0.05) by hierarchical regression using the SPSS software. The “surv_cutpoint” function was used to identify the best cutoff in the R-package “survminer” (v. 0.4.9) for the K-M analyses of single genes and the external validation of the m5C-score. The specificity and sensitivity of factors were analyzed by receiver operating characteristic (ROC) curve, and the area under the curve (AUC) were calculated with the “timeROC” (weighting = “aalen”) function in the R-package “timeROC” (v.0.4). The degree of correlation of the m5C-score with gene expression and immune infiltration was determined by Spearman analyses. A p -value < 0.05 was regarded as statistically significant.

Results

Genetic variation and prognostic value of m5C-regulator expression in pancreatic ductal adenocarcinoma

Twelve known m5C-methylation regulators including ten writers, one eraser and one reader were selected for analysis

(Supplementary Table S2, Supplementary Figure S1) (Ghanbarian et al., 2016; Bohnsack et al., 2019; Chen et al., 2019; DeNizio et al., 2019). Copy number variations (CNV) and point mutations of m5C-regulator genes were rare in the TCGA-PAAD cohort with a frequency below 3 and 4%, respectively (Figures 1A–C). In addition, the presence of mutations in m5C-regulator genes was not associated with the survival of PDAC patients (Supplementary Figure S2A).

Using data from the TCGA-PAAD and GTEx cohorts, we compared m5C-regulator expression between tumor and normal tissues. Only three genes (YBX1, DNMT1 and NSUN4) out of twelve were significantly differentially expressed in tumors, all showing an up-regulation (Figure 1D; Supplementary Figure S2G). We then assessed the prognostic value of m5C-methylation regulators expression for PDAC patients in our cohort (TCGA-PAAD + GSE57495). Seven out of twelve m5C-methylation regulators showed a significant correlation between gene expression and overall patient survival (Figures 1E–K and Supplementary Figures S2B–F). High expression of DNMT1, DNMT3B, NSUN2, NSUN3 and YBX1 was associated with a worse prognosis (Figures 1E–I). On the contrary high expression of NSUN6 and NSUN7 correlated with a better prognosis (Figures 1J,K). However, only one gene, the m5C-reader YBX1, was significantly associated with PDAC prognosis in the Univariate Cox regression analysis (Figure 1L).

Given that the expression of several m5C-methylation regulators correlated positively in PDAC (Figure 1M), we investigated whether pancreatic cancers could be classified based on m5C-regulator gene expression. Unsupervised consensus clustering identified 3 subgroups named m5C-clusters A, B and C in our PDAC cohort (Supplementary Figures S3A–E; Supplementary Table S3). GSEA pathway analysis revealed differentially regulated cancer pathways between the three m5C-clusters, including KRAS-related pathways (MAPK-, mTOR-, EGF- and ERK-pathways), cell death pathways (TNFR1-, FAS- and death-pathways), and cancer-related pathways (Wnt-, Gleevec-, MET and CREB-pathways) (Supplementary Figure S3F). The expression of each m5C-regulator was also significantly different between three m5C-clusters (Supplementary Figure S3G). However, the different clusters did not correlate with the survival of PDAC patients in our cohort (Supplementary Figure S3H).

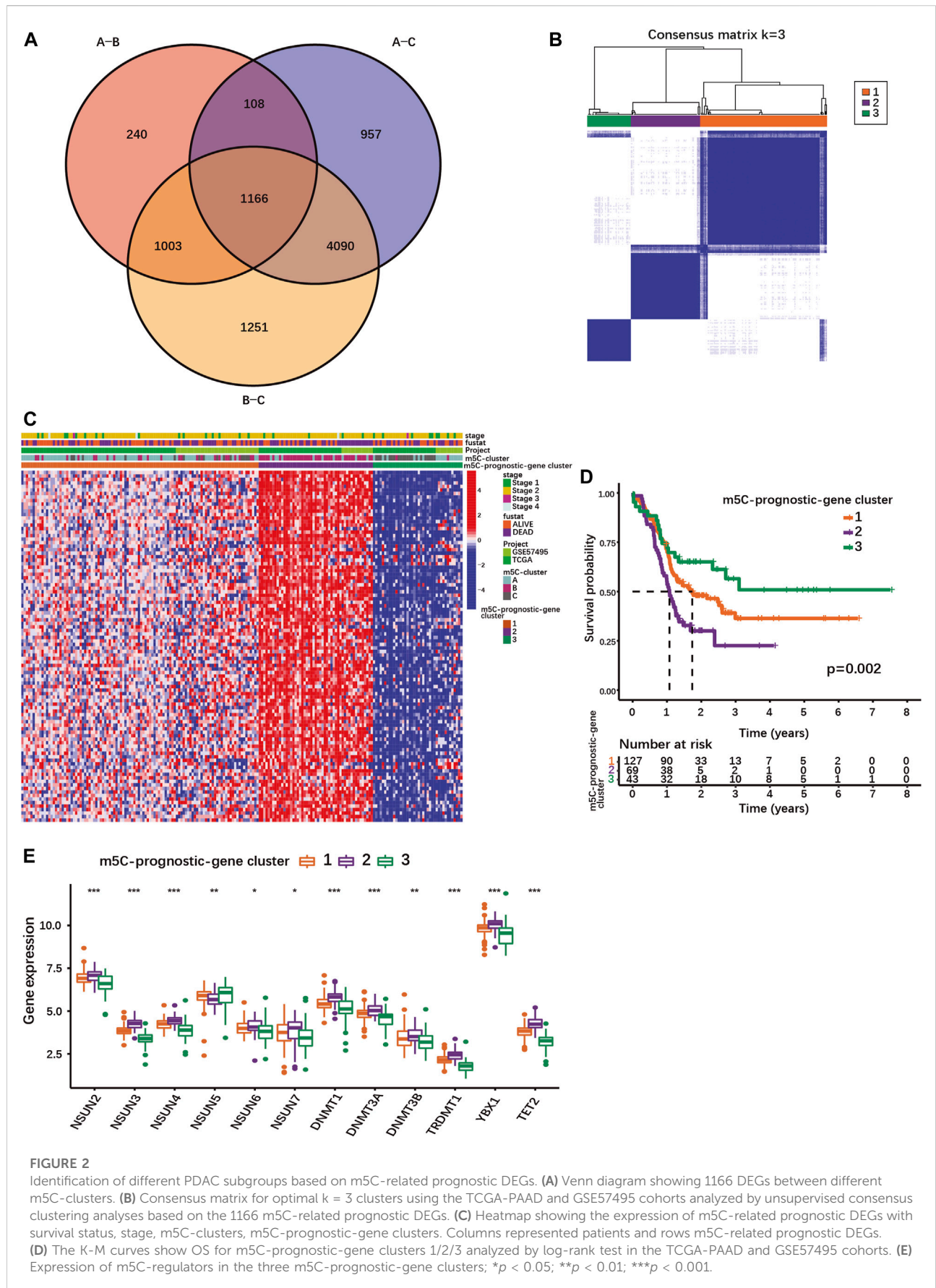


FIGURE 2

Identification of different PDAC subgroups based on m5C-related prognostic DEGs. (A) Venn diagram showing 1166 DEGs between different m5C-clusters. (B) Consensus matrix for optimal k = 3 clusters using the TCGA-PAAD and GSE57495 cohorts analyzed by unsupervised consensus clustering analyses based on the 1166 m5C-related prognostic DEGs. (C) Heatmap showing the expression of m5C-related prognostic DEGs with survival status, stage, m5C-clusters, m5C-prognostic-gene clusters. Columns represented patients and rows m5C-related prognostic DEGs. (D) The K-M curves show OS for m5C-prognostic-gene clusters 1/2/3 analyzed by log-rank test in the TCGA-PAAD and GSE57495 cohorts. (E) Expression of m5C-regulators in the three m5C-prognostic-gene clusters; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Identification of prognosis-relevant m5C-regulated gene subtypes

Since patient clustering according to m5C-regulator expression could not predict survival but was associated with transcriptional changes affecting several pancreatic cancer pathways, we investigated whether PDAC samples could be clustered according to m5C-related gene expression. In total, 1166 genes were differentially expressed between all three m5C-clusters (Figure 2A; Supplementary Table S4-4). Both KEGG and GO enrichment analyses showed that m5C-related differentially expressed genes (DEGs) were enriched in pathways regulating RNA regulation and processing (Supplementary Figures S4A,B; Supplementary Table S5-1/2), and in cancer-related pathways (Supplementary Figures S4A,B; Supplementary Table S5-1/2). This indicated that the changes in m5C-regulator expression in cancer have functional consequences impacting RNA regulation and cancer-pathway gene expression.

Using univariate Cox regression analysis, 100 m5C-related DEGs were identified as statistically significant prognostic markers in PDAC (Supplementary Table S6). Based on their expression, PDAC patients of our cohort could be stratified into 3 subgroups named m5C-prognostic-gene cluster 1, 2 and 3 (Figures 2B,C; Supplementary Figures S4C-E). The m5C-prognostic-gene cluster 3 was characterized by an overall low expression of the prognostic DEGs, and the highest overall survival rate (Figures 2C,D). The m5C-prognostic-gene cluster 2 was associated with a high expression of the prognostic DEGs and the lowest survival rate, while the m5C-prognostic-gene cluster 1 exhibited a mixed phenotype with intermediate survival rate (Figures 2C,D). Of note, the expression of each m5C-regulator was also significantly different between the three m5C-prognostic-gene clusters (Figure 2E). Except for NSUN5, all m5C-regulator genes were up-regulated in cluster 2 (Figure 2E), suggesting that the overexpression of m5C-regulators induces the expression of multiple genes involved in tumorigenesis.

Construction and validation of a prognostic model based on an m5C-related gene expression signature (m5C-score)

Next, we build a scoring signature (m5C-score) based on the 100 m5C-related prognostic DEGs to be used as prognostic model for PDAC patients. First, the patients of our cohort were randomly distributed into a training set and a testing set (Supplementary Table S7). LASSO regression analysis was performed on the training set using the m5C-related prognostic genes to avoid over-fitting (Supplementary Figures S5A,B). The multivariate Cox proportional hazards regression model was used to obtain the scoring signature (m5C-score), which is composed of four prognostic DEGs: Homeodomain Interacting Protein Kinase 3

(HIPK3), Zinc Finger AN1-Type Containing 4 (ZFAND4), Dipeptidyl Peptidase 8 (DPP8) and Homeodomain Interacting Protein Kinase 2 (HIPK2). The correlation coefficients of each gene was computed by multivariate Cox regression analysis and the m5C-score was calculated as follows: $m5C\text{-score} = (1.0722 \times \exp[\text{HIPK3}]) + (-1.168 \times \exp[\text{ZFAND4}]) + (0.9293 \times \exp[\text{DPP8}]) + (-0.6607 \times \exp[\text{HIPK2}])$, where “exp” means the expression of each m5C-score gene. Positive and negative coefficients indicated positive or negative correlation between single gene expression and the m5C-score, respectively.

Then, patients were classified into high-m5C-score and low-m5C-score groups using the median value of the m5C-score as cut-off (Figure 3A; Supplementary Table S8-1). Patients with low m5C-score survived significantly longer than patients with high m5C-score (Figure 3B). The area under the curve (AUC) of time-dependent receiver operating characteristic (ROC) curve was 0.821 for the 5-year overall survival (OS), showing that the m5C-score is an excellent predictor of survival (Figure 3C). The prognostic value of the m5C-score was then validated in the testing set and in two independent cohorts (GSE21501, Bailey cohort). Here again, patients with low m5C-score showed significantly longer survival than patients with high m5C-score, and the m5C-score could predict survival with good accuracy (Figures 3D-I; Supplementary Tables S8-2/3/4).

Univariate and multivariate Cox regression analyses performed using clinical data available for the TCGA-PAAD cohort showed that the m5C-score is an independent risk factor for PDAC patients, while age, gender, tumor grade and stage were not significant (Figures 4A,B). In addition, each individual m5C-score gene and common PDAC serum tumor markers were analyzed with the m5C-score by univariate Cox regression analysis in our cohort (Figures 4C,D). Each m5C-score gene taken individually (Figure 4C; Supplementary Figures S5C-F) and the marker genes MUC16, MUC1 and KRT19 (Figure 4D) were significantly associated with prognosis. All the statistically significant parameters ($p < 0.05$) were then selected to draw ROC curves, which showed that the m5C-score had the highest AUC at 1-, 3- and 5-year OS compared with either the m5C-score individual genes (Figures 4E-G), the m5C-methylation reader YBX1 or common PDAC serum tumor markers (Figures 4H-J). Overall, these results established the m5C-score is a robust predictive tool for survival of PDAC patients, that was superior to m5C-regulators or common tumor markers.

A high m5C-score is associated with mutation and transcriptomic regulation of pancreatic cancer-related genes

To further characterize the pathobiological significance of the m5C-score, somatic mutations were compared in PDACs with high and low m5C-score. The frequency of somatic mutations in the TCGA cohort was higher in the high-m5C-score group (87.5%) than in the low-m5C-score group (68.83%, Figure 5A). In particular,

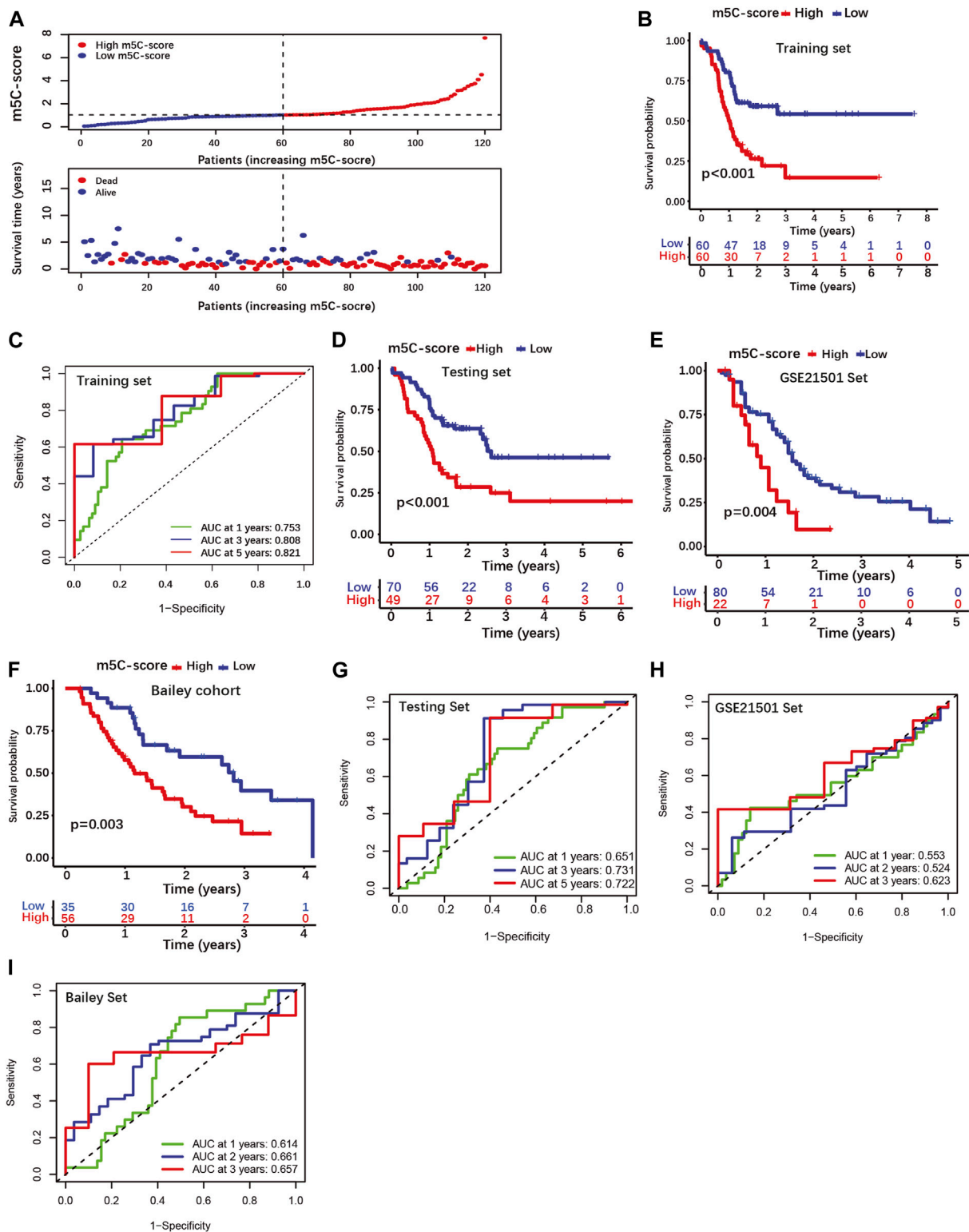
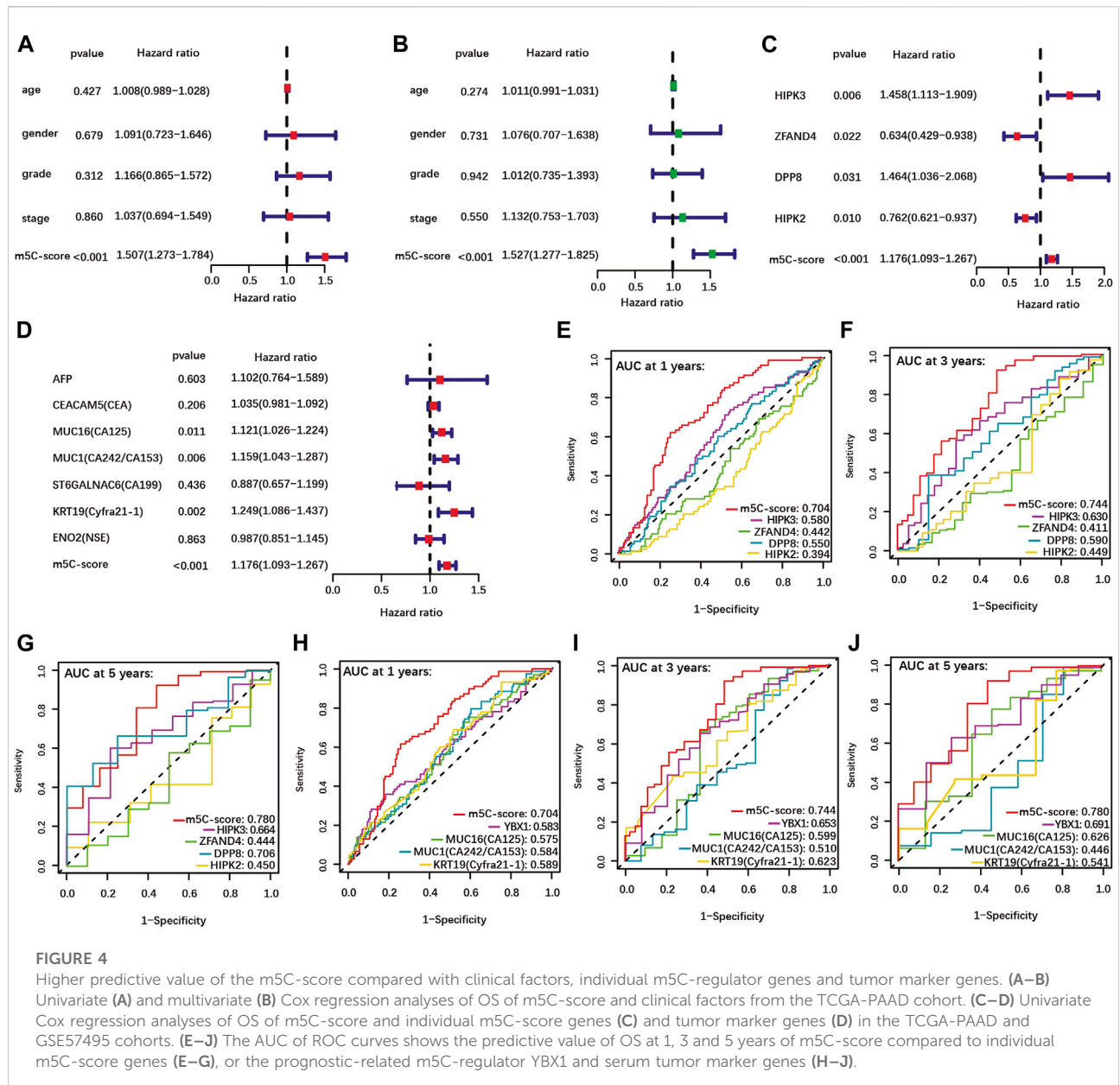


FIGURE 3 Construction and validation of a four-gene scoring signature (m5C-score) prognostic model. (A) Distributions of PDAC patients and survival status according to m5C-score in the training set. (B) K-M curves illustrating OS according to the m5C-score analyzed by log-rank test in the training set. (C) ROC curves for m5C-score prediction of OS at 1, 3 and 5 years in the training set. (D–F) K-M curves illustrating OS according to the m5C-score analyzed by log-rank test in the testing set and external cohorts (GSE21501 set, Bailey cohorts). (G–I) ROC curves of m5C-score prediction of OS at 1, 3 and 5 years in the testing and external cohorts.



KRAS, TP53, SMAD4 and CDKN2A were more frequently mutated in the high-m5C-score group. In agreement with these results, a higher expression of KRAS was observed in tumors with high-m5C-score compared to low-m5C-score (Figure 5B).

Then, GSEA pathway analysis was performed to investigate whether specific pathways could be related to the m5C-score. PDAC-related pathways were enriched in the high-m5C-score group, which included “pathways in cancer”, “pancreatic cancer”, and several signaling pathways (TGF-beta, MAPK, p53, hedgehog, ERBB and Wnt; Supplementary Table S9; Figures 5C,D). In addition, the genes differentially expressed between high-m5C-score and low-m5C-score groups (Figure 5E; Supplementary Table S10) were enriched in many

malignancy-related pathways (Figure 5F; Supplementary Table S11). Taken together, these results showed an association between a high-m5C-score, mutation rates and transcriptional regulation of pancreatic cancer-related pathways, including the MAPK pathway downstream of KRAS, and the TGFβ pathway.

A high m5C-score is associated with the squamous/basal pancreatic ductal adenocarcinoma subtype

We then evaluated whether the m5C-score correlated with the common molecular subtypes of PDAC. For this, we used

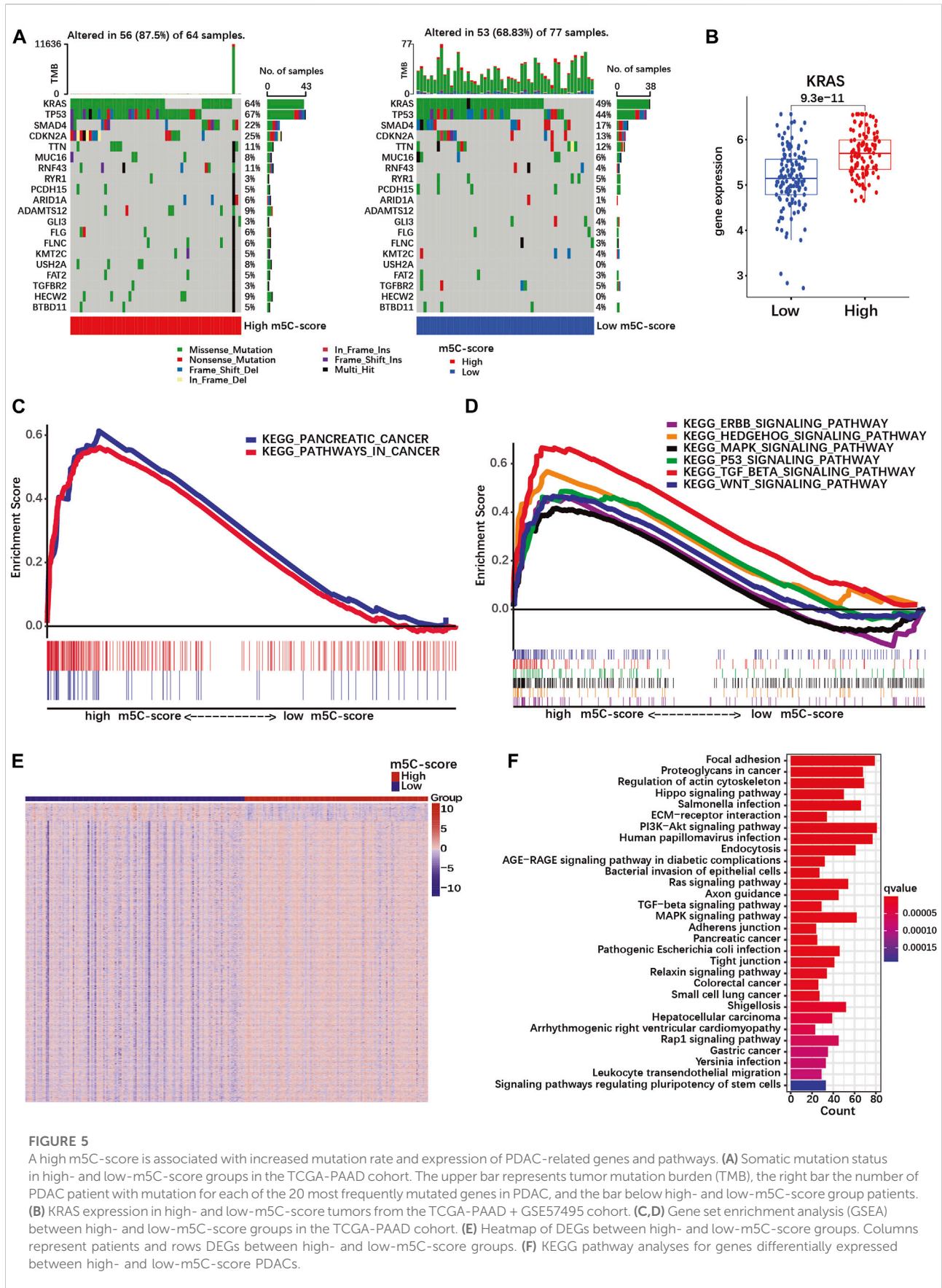
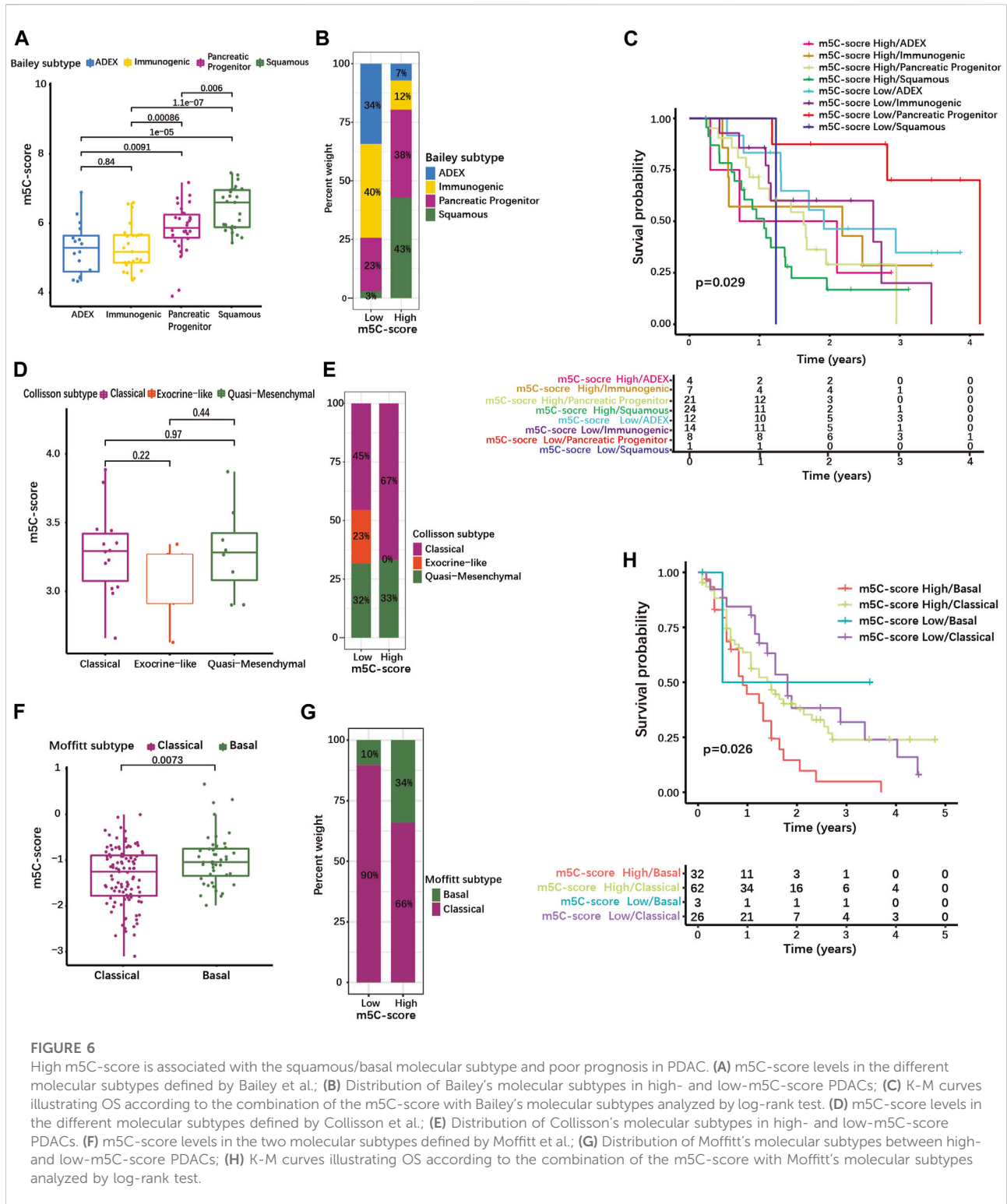


FIGURE 5

A high m5C-score is associated with increased mutation rate and expression of PDAC-related genes and pathways. **(A)** Somatic mutation status in high- and low-m5C-score groups in the TCGA-PAAD cohort. The upper bar represents tumor mutation burden (TMB), the right bar the number of PDAC patient with mutation for each of the 20 most frequently mutated genes in PDAC, and the bar below high- and low-m5C-score group patients. **(B)** KRAS expression in high- and low-m5C-score tumors from the TCGA-PAAD + GSE57495 cohort. **(C,D)** Gene set enrichment analysis (GSEA) between high- and low-m5C-score groups in the TCGA-PAAD cohort. **(E)** Heatmap of DEGs between high- and low-m5C-score groups. Columns represent patients and rows DEGs between high- and low-m5C-score groups. **(F)** KEGG pathway analyses for genes differentially expressed between high- and low-m5C-score PDACs.



classifications published by Bailey, Collisson, and Moffitt et al. (Collisson et al., 2011; Moffitt et al., 2015; Bailey et al., 2016). The Bailey classification distinguishes 4 molecular subtypes (Bailey et al., 2016). The Squamous subtype showed a significantly higher

m5C-score compared to the Progenitor, the Immunogenic and the Aberrantly Differentiated Endocrine Exocrine (ADEX) subtypes (Figure 6A, Supplementary Table S8-4). The proportion of tumor of the Squamous and Progenitor subtypes was higher in the high-

m5C-score group, while Immunogenic and ADEX subtypes were over-represented in the low-m5C-score group (Figure 6B). Survival analysis revealed that patients with high-m5C-score combined with Squamous subtype had the worst prognosis, while patients with low-m5C-score combined with Progenitor subtype showed the best prognosis (Figure 6C).

The Collison classification differentiates Classical, Exocrine-like and Quasi-mesenchymal subtypes (Supplementary Table S8-5) (Collisson et al., 2011). The m5C-score in the Exocrine-like subgroup was marginally lower than in the other two subtypes (Figure 6D), and the proportion of Exocrine-like tumors was drastically reduced in the high-m5C-score group (Figure 6E).

In the Moffitt cohort (Moffitt et al., 2015), tumors are classified as either Basal-like or Classical (Supplementary Table S8-6). The Basal-like subtype was associated with a higher m5C-score (Figure 6F) and the high-m5C-score group displayed a higher proportion of Basal-like tumors (Figure 6G). In the Moffitt cohort, Basal-like tumors with high-m5C-score had the worst prognosis, while low-m5C-score combined with Classical subtype showed the best prognosis (Figure 6H). A strong overlap has been described between the Classical/Pancreatic Progenitor subtypes on one side, and the Squamous/Basal-like/Quasi-mesenchymal subtypes associated with poor a prognosis on the other side (Xu Z. et al., 2021). Taken together, this means that a high m5C-score is strongly associated with the Squamous/Basal-like subtype, which represents a more aggressive phenotype with a poor prognosis.

The m5C-score correlates with targeted therapy response in pancreatic ductal adenocarcinoma

Targeted therapy of PDAC remains limited, with the EGFR-inhibitor Erlotinib being the only approved drug for treatment of advanced inoperable PDAC in combination with gemcitabine. However, multiple drugs targeting the KRAS signaling pathway, either upstream or downstream, as well as DNA repair are being tested in preclinical and clinical trials (Liu et al., 2021). We used the R package “pRRophetic” to predict the sensitivity of tumors with high or low m5C-score towards drugs targeting these pathways.

First, we investigated drugs targeting the signaling downstream of KRAS. The mTOR inhibitor Temsirolimus had a higher predicted IC50, meaning a decreased sensitivity, in the high-m5C-score group (Figure 7A). On the contrary, inhibition of AKT by A.443654, of MEK1/2 by PD.0325901 and RDEA119, and of PI3K β by AZD6482 was predicted to be more efficient in the high-m5C-score group in agreement with the high gene expression of AKT1-3, MEK1 and PIK3CB (Figures 7B–D). This indicated that tumors with high m5C-score, which exhibit higher KRAS mutation and expression (see Figures 5A,B), might be eligible for targeted therapy with inhibitors of AKT, MEK or PI3K, but not mTOR.

PDAC is characterized by frequent overexpression of ERBB receptors upstream of KRAS, notably EGFR and HER2 (ERBB2). Targeted therapy against EGFR or HER2 has yielded contrasted results in clinical trials, but combined targeting showed a synergistic effect in preclinical models (Maron et al., 2013; Liu et al., 2021). The expression of both EGFR and ERBB2 was increased in the high-m5C-score group compared to the low-m5C-score group (Figure 7E). In the high-m5C-score group, the predicted sensitivity was reduced for the EGFR inhibitor Gefitinib but increased for Lapatinib, an inhibitor of both EGFR and HER2, indicating that a combined inhibition might be effective in tumors with a high m5C-score (Figure 7E).

A subset of sporadic PDAC is characterized by DNA repair efficiency due to mutations in the homologous recombination (HR) pathway genes BRCA1/2 and/or PALB2. The inhibition of PARP, a downstream HR gene, has proven to be efficient for the treatment of BRCA1/2-mutated PDAC. The expression of BRCA1/2 and PALB2 was reduced and the predicted sensitivity to the PARP-inhibitor ABT.888 was accordingly increased in tumors with a low m5C-score compared to the high-m5C-score group (Figure 7F).

Overall, tumors with a high m5C-score were predicted to be more sensitive to AKT, MEK and PI3K inhibitors, as well as combined EGFR/HER2. In contrast, tumors with low m5C-score might be more sensitive to PARP inhibition.

Subsequently, we determined the m5C-score in different PDAC cell lines. We compared the m5C-score in 17 different pancreatic cell lines included in the Moffitt classification study (Moffitt et al., 2015). The Panc1005 cell line showed highest m5C-score, followed by CAPAN1, while PANC1 showed the lowest m5C-score (Figure 7H and Supplementary Table S8-7). In addition, the sequencing data from four pancreatic cancer cell lines (CAPAN1, PANC1, MIAPACA2 and CFPAC1) was acquired from GSE165949 (Alfarano et al., 2022). CAPAN1 showed the highest m5C-score, the PANC1 showed the lowest m5C-score, while MIAPACA2 cells had an almost negative m5C-score, confirming the results above (Figure 7G; Supplementary Table S8-8). These results might serve as basis for further studies using cell lines to compare the biological effects of high and low m5C-score.

An m5C-score is associated with immune exclusion and immune resistance

Recent studies have suggested a role of RNA modification in cell infiltration within the tumor microenvironment (TME) (Xu F. et al., 2021; Zhang M. et al., 2021). Therefore, we investigated the overall abundance of tumor infiltrating immune cells (TIICs) in the TCGA-PAAD cohort. Immune exclusion was significantly higher in tumors with high m5C-score compared to the low-m5C-score group (Figure 8A; Supplementary Table S12). Next, we evaluated the relative abundance of 25 different types of TIICs in every PDAC sample in our cohort (Supplementary Table S13),

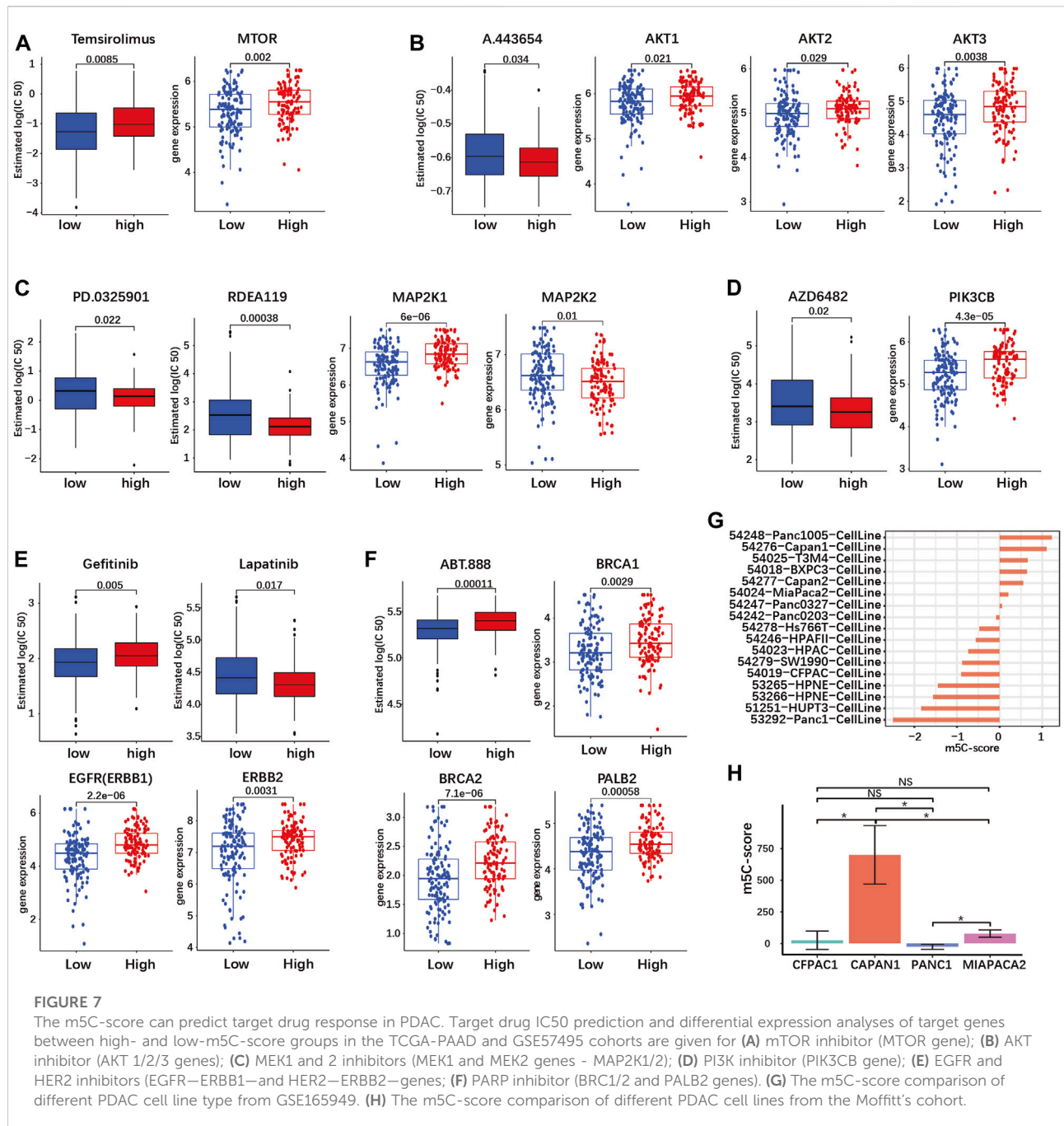
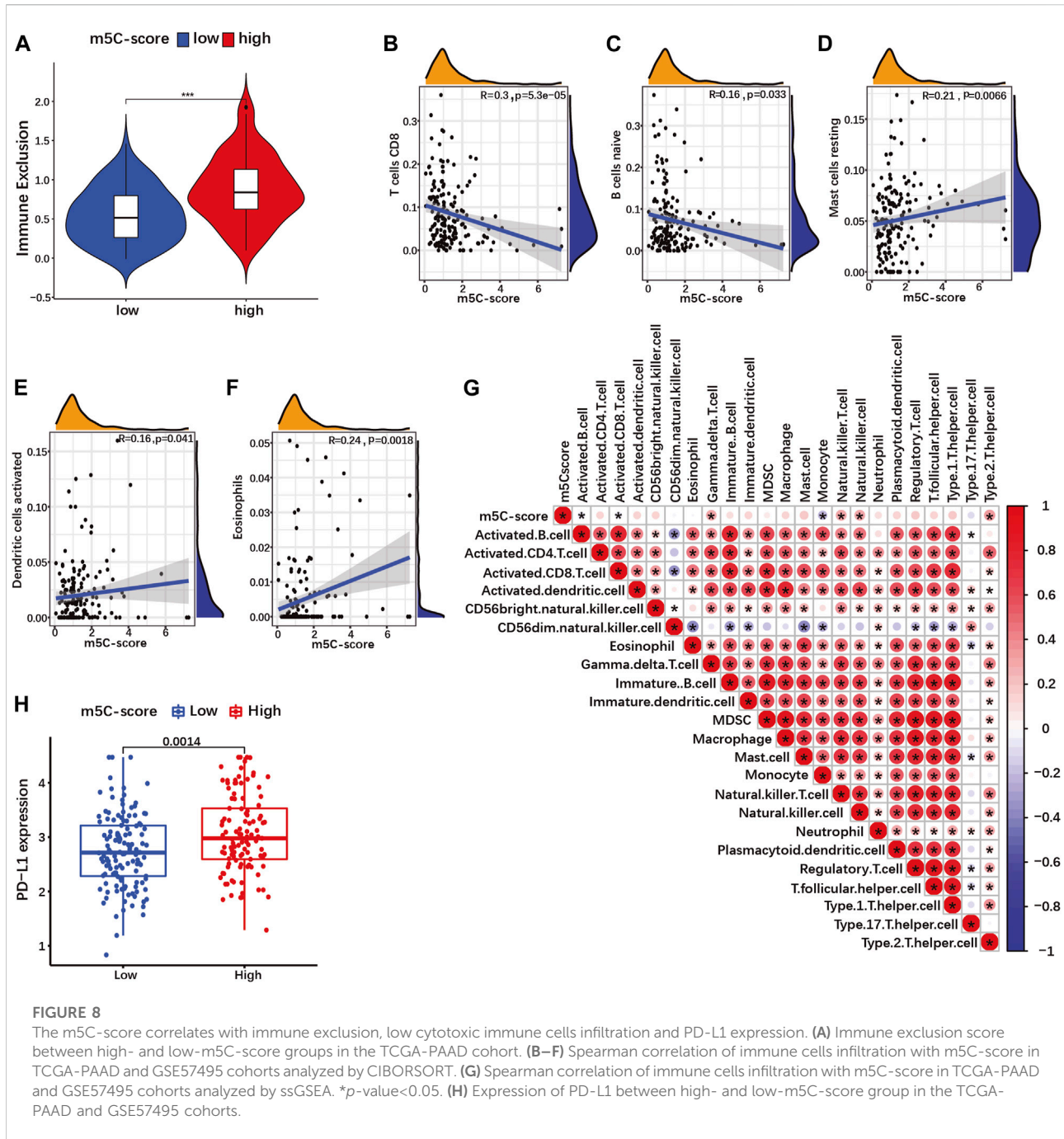


FIGURE 7

The m5C-score can predict target drug response in PDAC. Target drug IC50 prediction and differential expression analyses of target genes between high- and low-m5C-score groups in the TCGA-PAAD and GSE57495 cohorts are given for (A) mTOR inhibitor (MTOR gene); (B) AKT inhibitor (AKT 1/2/3 genes); (C) MEK1 and 2 inhibitors (MEK1 and MEK2 genes - MAP2K1/2); (D) PI3K inhibitor (PIK3CB gene); (E) EGFR and HER2 inhibitors (EGFR—ERBB1— and HER2—ERBB2—genes); (F) PARP inhibitor (BRCA1/2 and PALB2 genes). (G) The m5C-score comparison of different PDAC cell line type from GSE165949. (H) The m5C-score comparison of different PDAC cell lines from the Moffitt's cohort.

and determined the correlation between m5C-score and the various TIICs. The m5C-score correlated negatively with the presence of CD8⁺ T cells (Figure 8B) and naive B cells (Figure 8C), and was positively with to the presence of resting mast cells (Figure 8D), activated dendritic cells (Figure 8E) and eosinophils (Figure 8F). Then, ssGSEA was used to calculate the relative enrichment of immune cells in tumor microenvironment (TME) of PDAC patients in our cohort (Supplementary Table S14). Here, the m5C-score correlated positively with the

infiltration of gamma delta T cells, natural killer T cells, natural killer cells and type 2 T helper (Th2) cells, and negatively with the infiltration of activated B cells, activated CD8⁺ T cells, and monocytes (Figure 8G). Taken together, these results showed that a high m5C-score was associated with more frequent immune exclusion, a type 2 immune response (mast cells, eosinophils, Th2 T cells), and a low cytotoxic T cell infiltration and activity (CD8⁺ T cells), all indicative of an immune-deprived microenvironment with low anti-tumorigenic activity.



Overexpression by tumor cells of immune checkpoint ligands such as the programmed death ligand 1 (PDL1) represents a mechanism of immune escape, but also a therapy target. In our cohort, PD-L1 expression was significantly more expressed in high-m5C-score tumors compared to low-m5C-score tumors (Figure 8H). Altogether, tumors with a high m5C-score displayed characteristics of immunologically “cold” cancers and immune evasion.

YBX1 expression correlates with m5C-score, target drug IC50, immune exclusion and immune infiltration.

Since YBX1 had the best predictive power for survival among the m5C-regulators, albeit lower than the m5C-score (see Figure 4E), we analyzed the relationship between YBX1 expression level and the m5C-score, predicted targeted drug response and immune infiltration. The m5C-score was

significantly higher in the high-YBX1 expression level group compared to the low-YBX1 expression group (Supplementary Figure S6A). Next, we estimated the IC50 of target drugs between high and low YBX1 expression groups (for drugs significantly related to the m5C-score). Only two target drugs showed a significant difference. Similar to the m5C-score, high YBX1 expression was associated with a higher sensitivity to the PI3K β inhibitor AZD6482, and a lower sensitivity to the EGFR inhibitor Gefitinib (Supplementary Figures S6B,C). Contrary to the m5C-score, YBX1 was not associated with the response to inhibitors of AKT, MEK, combined HER1/HER2 or PARP. Then, we analyzed the correlation of YBX1 with immune exclusion and immune infiltration. Tumors with high YBX1 expression had more immune exclusion compared to the low YBX1 expression group (Supplementary Figure S6D). CIBERSORT analyses showed that YBX1 expression correlated negatively with mast cells, and positively with CD4⁺ memory T cells in a significant manner. Using ssGSEA analyses, we found a positive correlation between YBX1 expression and many kinds of immune cells such as activated B cell, activated CD8⁺ T cell, activated dendritic cell, natural killer cell and type 1 and 2 T helper cell (Supplementary Figure S6G). PD-L1 expression was significantly up-regulated in the high YBX1 expression group compared to the low YBX1 expression group (Supplementary Figure S6H). Overall, these results showed that the YBX1 expression correlates with the m5C-score, but the m5C-score was a better predictor for prognosis and drug sensitivity. Contrary to the m5C-score, YBX1 expression was associated with an active anti-tumor immune response. However, both factors were equally associated with PDL-1 expression and immune exclusion.

Discussion

An increasing body of studies suggests that m5C modification regulators are involved in tumorigenesis, tumor progression, and anti-tumor immune response in multiple malignancies including breast, ovarian, cervical or prostate cancers (Esteve-Puig et al., 2020; Nombela et al., 2021). In the present study, we aimed to comprehensively investigate the predictive value of m5C regulation in PDAC. Overall, we observed a low level of CNVs or SNVs for m5C-regulator genes in PDAC samples and the presence of genomic alterations did not correlate with patient survival, indicating that mutations of m5C-regulator genes were rather passenger than driver mutations. These findings were in accordance with a recent study by Yu et al. (2021), who uncovered a small number of m5C-regulator gene mutations in PDAC, that were without predictive statistical significance. In a pan-cancer study across 33 tumor types, He et al. (2021) also reported low mutation frequency among m5C-regulators. At the RNA level, we found that most of the m5C-regulators were differentially expressed

between normal and tumor tissues, or between different tumor samples, confirming that m5C-regulator gene expression is frequently dysregulated in tumors (He et al., 2021). The prognostic value of m5C-regulator gene expression has been assessed in several other tumor entities. Several m5C-regulators are associated with overall survival prognosis in cutaneous melanoma (DNMT2, NSUN1, NSUN3, NSUN6 and YBX1). In particular, low NSUN6 expression was shown to correlate with melanoma progression (Huang et al., 2021a). On the contrary, in lung squamous carcinoma, while most of the m5C-regulators showed significantly different expression between tumor and normal samples, only NSUN3 could significantly predict prognosis (and NSUN4 almost significantly) (Pan et al., 2021). NSUN6 expression could predict prognosis in both breast and colorectal cancer. However, NSUN6 expression was associated with a poor prognosis in breast cancer, and with a more favorable prognosis in colorectal cancer (Huang et al., 2021b; Fang et al., 2022). In the present study, NSUN6 expression was associated with a longer survival. Out of the seven genes whose expression correlated with survival, only YBX1, an m5C-binding protein, was significantly able to predict prognosis in PDAC. Hence, it appears that while the expression of m5C-regulators is often associated with prognosis in several tumor types, which specific genes are involved and the prognostic value of each single gene vary from entity to entity.

Since only one single m5C-regulator (YBX1) could predict prognosis in our study but a high level of correlation was found between the expression of the different m5C-regulators, we investigated whether m5C-regulator-associated expression profiles were related to prognosis in PDAC. Patient clustering according to m5C-regulator expression did not correlate with survival in our PDAC cohort. However, we could identify up to 100 genes differentially regulated between the three m5C-clusters, that showed prognostic value. Based on these prognostic DEGs, we identified three m5C-prognostic-gene clusters, which could stratify PDAC patients. Notably, the m5C-prognostic-gene cluster 2 was associated with high prognostic DEG expression, and poor survival. In addition, high expression of m5C-regulators was observed in the m5C-prognostic-gene cluster 2, confirming the validity of the clustering. These data are supported by the fact that the expression of m5C-related long noncoding RNAs (lncRNAs) were found to correlate with prognosis in PDAC (Yuan et al., 2021; Liu et al., 2022). They also indicated that, while m5C-regulators themselves could not predict prognosis for PDAC patients, the profound pro-tumorigenic transcriptional changes induced by m5C-dysregulation could be used to stratify patients and predict disease outcome. This effect might be attributed to the redundancy of function of some m5C-writers, and/or to compensatory mechanisms.

In order to predict the prognosis of patients with PDAC based on the expression of m5C-related prognostic DEG expression, we constructed a risk score (m5C-score) including four prognostic DEGs and classified patients into high- and low-risk groups. Overall, the m5C-score was well suitable for patient stratification. It significantly correlated with survival and successfully predicted PDAC patient prognosis in our cohort, as well as two additional independent cohorts. A similar approach was used by Li et al. (2021) in papillary thyroid carcinoma based on DEGs between three m5C-regulator clusters. In agreement with our results, a high m5C-score correlated with a lower survival rate. In our study, the m5C-score could independently predict prognosis in PDAC patients. It represented a stronger prediction factor than tumor grade, stage or classical PDAC markers such as CEA, MUC16, MUC1, CA199, KRT19 and NSE, and also performed better than any m5C-score gene taken individually. The comparison of high- and low-m5C-score groups offered a further validation by revealing that a high m5C-score was associated with a more aggressive tumor phenotype. For instance, tumors with high-m5C-score displayed more mutations, notably for KRAS, TP53, SMAD4 and CDKN2A, as well as a higher KRAS expression. In addition, cancer pathway expression including the Ras, p53, MAPK, ERBB and TGF-beta PI3K-AKT pathways was enriched in the high-m5C-score group. Furthermore, there was an association between high m5C-score and the prognostically unfavorable squamous/basal molecular pathway using all three PDAC classifications from Bailey, Collisson and Moffitt (Collisson et al., 2011; Moffitt et al., 2015; Bailey et al., 2016). Interestingly, therapy response prediction showed that the high m5C-score group was more likely to respond to therapy targeted at the PI3K-AKT, MEK, ERBB pathways. This was well in agreement with the elevated activation of these pathways in tumors with high m5C-score. On the contrary, tumors with a high m5C-score were predicted to have a lower sensitivity to PARP inhibition than low-m5C-score tumors. This might be explained by the fact that m5C-methylation of mRNA is involved in DNA damage repair, notably by promoting homologous recombination (HR). Chen et al. have recently shown that the RNA methyltransferase TRDMT1 is recruited to DNA damage sites to promote m5C-methylation. Loss of TRDMT1 compromised HR, increased cellular sensitivity to DNA double-strand breaks and conferred sensitivity to PARP inhibitors *in vitro* and *in vivo* (Chen et al., 2020). Finally, a high m5C-score correlated with more frequent immune exclusion, immune exhaustion (low activated cytotoxic T-cells), and immune evasion (high PDL-1 expression). These findings indicated that m5C phenotype-associated patterns also affect the TME, and that the m5C-score might be used as a marker for immunologically “cold” cancers and immune evasion. This is in line with the fact that m6A-regulators have been reported to correlate with anti-tumor immunity in PDAC, and have been proposed to regulate the immune microenvironment (Xu F. et al., 2021). Notably, a high m6A-risk signature associated with poor prognosis was found to correlate with lower naive B cells and CD8⁺ T cells infiltration, similarly to what we

observed for the high m5C-score (Xu F. et al., 2021). More generally, RNA methylation has been shown to inhibit RNA recognition by Toll-like receptors or dendritic cells, to regulate T-cell differentiation and expression of immune factors such as IL-17, or to modulate macrophages polarization (Zhang M. et al., 2021).

When comparing YBX1 expression with the m5C-score, we found that the m5C-score was a better predictor of prognosis and drug sensitivity. YBX1 codes for the Y-box binding protein-1, a multifunctional oncoprotein regulating cell proliferation, survival, drug resistance in cancer (Kuwano et al., 2019). YBX1 can act as transcription factor, but is also involved in DNA repair and RNA splicing. In agreement with our findings, overexpression of YBX1 has been reported in PDAC, where it regulates cell-cycle progression and proliferation through the expression of cell-cycle-related cyclins and GSK3B (Liu et al., 2020). However, YBX1 is not specific to m5C, can bind several other modifications or be activated by oncogenic pathways (Ban et al., 2020; Alkrekshi et al., 2021; Feng et al., 2021). In conclusion, even if YBX1 represents a good prognostic marker, it is not specific of m5C-methylation and could therefore complement but not substitute the m5C-score.

Taken together, our results demonstrated that the m5C-score represents a robust prognostic tool for patients with PDAC, which correlates with molecular subtype as well as an immune-deprived and immune-resistant tumor microenvironment.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: Gene Expression-Omnibus (GEO) (<https://www.ncbi.nlm.nih.gov/geo/>): GSE21501; GSE71729; GSE17891; the Cancer Genome Atlas (TCGA) (<https://portal.gdc.cancer.gov/>): TCGA-PAAD; <https://www.nature.com/articles/nature16965#Sec11>.

Author contributions

DY and NB-L analyzed the data and wrote the article. NB-L and DY conceived the study. NB-L and CP obtained financial support. CP and RG applied the guiding suggestions. ZY, SZ, HY, and DL prepared the dataset. All authors read and approved the final article.

Funding

This work was supported by grants from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation: BR 5196/2-1 to NB-L, SFB/TRR305-B07 to NB-L); the Interdisciplinary Center for Clinical Research (IZKF) of the Clinical Center Erlangen (ELAN-Fonds project P027 to NB-L and Project J19 to NB-L).

Acknowledgments

Thanks to all who have contributed to the process of writing this article. The authors also acknowledge financial support by Deutsche Forschungsgemeinschaft and Friedrich-Alexander-Universität Erlangen-Nürnberg within the funding programme “Open Access Publication Funding”.

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.

References

- Alfarano, G., Audano, M., Di Chiaro, P., Balestrieri, C., Milan, M., Polletti, S., et al. (2022). Interferon regulatory factor 1 (IRF1) controls the metabolic programmes of low-grade pancreatic cancer cells. *Gut*. Epub ahead of print, 2021-325811. doi:10.1136/gutjnl-2021-325811
- Alkrekshi, A., Wang, W., Rana, P. S., Markovic, V., and Sossey-Alaoui, K. (2021). A comprehensive review of the functions of YB-1 in cancer stemness, metastasis and drug resistance. *Cell. Signal.* 85, 110073. doi:10.1016/j.cellsig.2021.110073
- Arslan, A. A., Helzlsouer, K. J., Kooperberg, C., Shu, X. O., Stepnowski, E., Bueno-de-Mesquita, H. B., et al. (2010). Anthropometric measures, body mass index, and pancreatic cancer: a pooled analysis from the pancreatic cancer cohort consortium (PanScan). *Arch. Intern. Med.* 170 (9), 791–802. doi:10.1001/archinternmed.2010.63
- Bailey, P., Chang, D. K., Nones, K., Johns, A. L., Patch, A. M., Gingras, M. C., et al. (2016). Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* 531 (7592), 47–52. doi:10.1038/nature16965
- Ban, Y., Tan, P., Cai, J., Li, J., Hu, M., Zhou, Y., et al. (2020). LNCAROD is stabilized by m6A methylation and promotes cancer progression via forming a ternary complex with HSPA1A and YBX1 in head and neck squamous cell carcinoma. *Mol. Oncol.* 14 (6), 1282–1296. doi:10.1002/1878-0261.12676
- Barbie, D. A., Tamayo, P., Boehm, J. S., Kim, S. Y., Moody, S. E., Dunn, I. F., et al. (2009). Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* 462 (7269), 108–112. doi:10.1038/nature08460
- Bohnsack, K. E., Höbartner, C., and Bohnsack, M. T. (2019). Eukaryotic 5-methylcytosine (mC) RNA methyltransferases: mechanisms, cellular functions, and links to disease. *Genes (Basel)* 10 (2), E102. doi:10.3390/genes10020102
- Boo, S. H., and Kim, Y. K. (2020). The emerging role of RNA modifications in the regulation of mRNA stability. *Exp. Mol. Med.* 52 (3), 400–408. doi:10.1038/s12276-020-0407-z
- Carithers, L. J., and Moore, H. M. (2015). The genotype-tissue expression (GTEx) project. *Biopreserv. Biobank.* 13 (5), 307–308. doi:10.1089/bio.2015.29031.hmm
- Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D., et al. (2017). Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep.* 18 (1), 248–262. doi:10.1016/j.celrep.2016.12.019
- Chen, X., Li, A., Sun, B. F., Yang, Y., Han, Y. N., Yuan, X., et al. (2019). 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. *Nat. Cell Biol.* 21 (8), 978–990. doi:10.1038/s41556-019-0361-y
- Chen, H., Yang, H., Zhu, X., Yadav, T., Ouyang, J., Truesdell, S. S., et al. (2020). m5C modification of mRNA serves a DNA damage code to promote homologous recombination. *Nat. Commun.* 11 (1), 2834. doi:10.1038/s41467-020-16722-7
- Chen, Y. S., Yang, W. L., Zhao, Y. L., and Yang, Y. G. (2021). Dynamic transcriptomic m(5) C and its regulatory role in RNA processing. *Wiley Interdiscip. Rev. RNA* 12 (4), e1639. doi:10.1002/wrna.1639
- Collisson, E. A., Sadanandam, A., Olson, P., Gibb, W. J., Truitt, M., Gu, S., et al. (2011). Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nat. Med.* 17 (4), 500–503. doi:10.1038/nm.2344
- DeNizio, J. E., Liu, M. Y., Leddin, E. M., Cisneros, G. A., and Kohli, R. M. (2019). Selectivity and promiscuity in TET-mediated oxidation of 5-methylcytosine in DNA and RNA. *Biochemistry* 58 (5), 411–421. doi:10.1021/acs.biochem.8b00912
- Edwards, B. K., Brown, M. L., Wingo, P. A., Howe, H. L., Ward, E., Ries, L. A., et al. (2005). Annual report to the nation on the status of cancer, 1975–2002, featuring population-based trends in cancer treatment. *J. Natl. Cancer Inst.* 97 (19), 1407–1427. doi:10.1093/jnci/dji289
- 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
- Fang, X., Miao, C., Zeng, T., Chu, W., Zheng, Y., Sun, X., et al. (2022). Role of m5C RNA methylation regulators in colorectal cancer prognosis and immune microenvironment. *J. Clin. Lab. Anal.* 36, e24303. doi:10.1002/jcla.24303
- Feng, M., Xie, X., Han, G., Zhang, T., Li, Y., Li, Y., et al. (2021). YBX1 is required for maintaining myeloid leukemia cell survival by regulating BCL2 stability in an m6A-dependent manner. *Blood* 138 (1), 71–85. doi:10.1182/blood.202009676
- Geeleher, P., Cox, N., and Huang, R. S. (2014). pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLoS One* 9 (9), e107468. doi:10.1371/journal.pone.0107468
- Ghanbarian, H., Wagner, N., Polo, B., Baudouy, D., Kiani, J., Michiels, J. F., et al. (2016). Dnmt2/Trdm1 as mediator of RNA polymerase II transcriptional activity in cardiac growth. *PLoS One* 11 (6), e0156953. doi:10.1371/journal.pone.0156953
- Han, X., Wang, M., Zhao, Y. L., Yang, Y., and Yang, Y. G. (2021). RNA methylations in human cancers. *Semin. Cancer Biol.* 75, 97–115. doi:10.1016/j.semcancer.2020.11.007
- Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinform.* 14, 7. doi:10.1186/1471-2105-14-7
- Haruehanroengra, P., Zheng, Y. Y., Zhou, Y., Huang, Y., and Sheng, J. (2020). RNA modifications and cancer. *RNA Biol.* 17 (11), 1560–1575. doi:10.1080/15476286.2020.1722449
- He, Y., Yu, X., Zhang, M., and Guo, W. (2021). Pan-cancer analysis of m(5)C regulator genes reveals consistent epigenetic landscape changes in multiple cancers. *World J. Surg.* 19 (1), 224. doi:10.1186/s12957-021-02342-y
- Huang, M., Zhang, Y., Ou, X., Wang, C., Wang, X., Qin, B., et al. (2021a). m5C-Related signatures for predicting prognosis in cutaneous melanoma with machine learning. *J. Oncol.* 2021, 6173206. doi:10.1155/2021/6173206
- Huang, Z., Pan, J., Wang, H., Du, X., Xu, Y., Wang, Z., et al. (2021b). Prognostic significance and tumor immune microenvironment heterogeneity of m5C RNA methylation regulators in triple-negative breast cancer. *Front. Cell Dev. Biol.* 9, 657547. doi:10.3389/fcell.2021.657547
- 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
- Kuwano, M., Shibata, T., Watari, K., and Ono, M. (2019). Oncogenic Y-box binding protein-1 as an effective therapeutic target in drug-resistant cancer. *Cancer Sci.* 110 (5), 1536–1543. doi:10.1111/cas.14006

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/fcell.2022.975684/full#supplementary-material>

- 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, F., Deng, Q., Pang, X., Huang, S., Zhang, J., Zhu, X., et al. (2021). m5C regulator-mediated methylation modification patterns and tumor microenvironment infiltration characterization in papillary thyroid carcinoma. *Front. Oncol.* 11, 729887. doi:10.3389/fonc.2021.729887
- Liu, Z., Li, Y., Li, X., Zhao, J., Wu, S., Wu, H., et al. (2020). Overexpression of YBX1 promotes pancreatic ductal adenocarcinoma growth via the GSK3B/cyclin D1/cyclin E1 pathway. *Mol. Ther. Oncolytics* 17, 21–30. doi:10.1016/j.omto.2020.03.006
- Liu, X., Li, Z., and Wang, Y. (2021). Advances in targeted therapy and immunotherapy for pancreatic cancer. *Adv. Biol.* 5 (3), e1900236. doi:10.1002/adbi.201900236
- Liu, X., Wang, D., Han, S., Wang, F., Zang, J., Xu, C., et al. (2022). Signature of m5C-related lncRNA for prognostic prediction and immune responses in pancreatic cancer. *J. Oncol.* 2022, 7467797. doi:10.1155/2022/7467797
- Maron, R., Schechter, B., Mancini, M., Mahlknecht, G., Yarden, Y., and Sela, M. (2013). Inhibition of pancreatic carcinoma by homo- and heterocombinations of antibodies against EGF-receptor and its kin HER2/ErbB-2. *Proc. Natl. Acad. Sci. U. S. A.* 110 (38), 15389–15394. doi:10.1073/pnas.1313857110
- 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
- Moffitt, R. A., Marayati, R., Flate, E. L., Volmar, K. E., Loeza, S. G., Hoadley, K. A., et al. (2015). Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat. Genet.* 47 (10), 1168–1178. doi:10.1038/ng.3398
- 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
- Nombela, P., Miguel-López, B., and Blanco, S. (2021). The role of m(6)A, m(5)C and Ψ RNA modifications in cancer: Novel therapeutic opportunities. *Mol. Cancer* 20 (1), 18. doi:10.1186/s12943-020-01263-w
- Pan, J., Huang, Z., and Xu, Y. (2021). m5C RNA methylation regulators predict prognosis and regulate the immune microenvironment in lung squamous cell carcinoma. *Front. Oncol.* 11, 657466. doi:10.3389/fonc.2021.657466
- Rahib, L., Smith, B. D., Aizenberg, R., Rosenzweig, A. B., Fleshman, J. M., and Matrisian, L. M. (2014). Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res.* 74 (11), 2913–2921. doi:10.1158/0008-5472.Can-14-0155
- Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2022). Cancer statistics, 2022. *CA Cancer J. Clin.* 72 (1), 7–33. doi:10.3322/caac.21708
- 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
- Tang, Z., Kang, B., Li, C., Chen, T., and Zhang, Z. (2019). GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* 47 (W1), W556–w560. doi:10.1093/nar/gkz430
- Wilkerson, M. D., and Hayes, D. N. (2010). ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* 26 (12), 1572–1573. doi:10.1093/bioinformatics/btq170
- Wood, S., Willbanks, A., and Cheng, J. X. (2021). The role of RNA modifications and RNA-modifying proteins in cancer therapy and drug resistance. *Curr. Cancer Drug Targets* 21 (4), 326–352. doi:10.2174/1568009621666210127092828
- Xu, F., Zhang, Z., Yuan, M., Zhao, Y., Zhou, Y., Pei, H., et al. (2021a). M6A regulatory genes play an important role in the prognosis, progression and immune microenvironment of pancreatic adenocarcinoma. *Cancer Invest.* 39 (1), 39–54. doi:10.1080/07357907.2020.1834576
- Xu, L., Zhang, C., Yin, H., Gong, S., Wu, N., Ren, Z., et al. (2021b). RNA modifications act as regulators of cell death. *RNA Biol.* 18 (12), 2183–2193. doi:10.1080/15476286.2021.1925460
- Xu, Z., Hu, K., Bailey, P., Springfield, C., Roth, S., Kurilov, R., et al. (2021c). Clinical impact of molecular subtyping of pancreatic cancer. *Front. Cell Dev. Biol.* 9, 743908. doi:10.3389/fcell.2021.743908
- Yu, X., Zhang, Q., Gao, F., Zhang, M., Zheng, Q., He, Y., et al. (2021). Predictive value of m5C regulatory gene expression in pancreatic adenocarcinoma. *Sci. Rep.* 11 (1), 17529. doi:10.1038/s41598-021-96470-w
- Yuan, H., Liu, J., Zhao, L., Wu, P., Chen, G., Chen, Q., et al. (2021). Prognostic risk model and tumor immune environment modulation of m5C-related lncRNAs in pancreatic ductal adenocarcinoma. *Front. Immunol.* 12, 800268. doi:10.3389/fimmu.2021.800268
- Zhang, H., Meltzer, P., and Davis, S. (2013). RCircos: an R package for Circos 2D track plots. *BMC Bioinform.* 14, 244. doi:10.1186/1471-2105-14-244
- Zhang, M., Song, J., Yuan, W., Zhang, W., and Sun, Z. (2021a). Roles of RNA methylation on tumor immunity and clinical implications. *Front. Immunol.* 12, 641507. doi:10.3389/fimmu.2021.641507
- Zhang, Q., Liu, F., Chen, W., Miao, H., Liang, H., Liao, Z., et al. (2021b). The role of RNA m(5)C modification in cancer metastasis. *Int. J. Biol. Sci.* 17 (13), 3369–3380. doi:10.7150/ijbs.61439