

Received: 2020.05.07

Accepted: 2020.07.03

Available online: 2020.07.17

Published: 2020.07.24

Systematic Analysis of Alternative Splicing Landscape in Pancreatic Adenocarcinoma Reveals Regulatory Network Associated with Tumorigenesis and Immune Response

Authors' Contribution:
Study Design A
Data Collection B
Statistical Analysis C
Data Interpretation D
Manuscript Preparation E
Literature Search F
Funds Collection G

BCE 1,2,3,4 **Jiahua Lu**
BCD 5 **Shenyu Wei**
BF 6 **Jianying Lou**
DF 1,2,3,4 **Shengyong Yin**
CD 1,2,3,4 **Lin Zhou**
AEG 7,8 **Wu Zhang**
AEG 1,2,3,4 **Shusen Zheng**

1 Division of Hepatobiliary and Pancreatic Surgery, Department of Surgery, First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang, P.R. China
2 NHC Key Laboratory of Combined Multi-Organ Transplantation, Hangzhou, Zhejiang, P.R. China
3 Key Laboratory of the Diagnosis and Treatment of Organ Transplantation, Research Unit of Collaborative Diagnosis and Treatment for Hepatobiliary and Pancreatic Cancer, Chinese Academy of Medical Sciences (2019RU019), Hangzhou, Zhejiang, P.R. China
4 Key Laboratory of Organ Transplantation, Research Center for Diagnosis and Treatment of Hepatobiliary Diseases, Hangzhou, Zhejiang, P.R. China
5 First Clinical Medical College, Zhejiang Chinese Medical University, Hangzhou, Zhejiang, P.R. China
6 Department of Hepato-Pancreato-Biliary Surgery, Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, P.R. China
7 Shulan (Hangzhou) Hospital, Hangzhou, Zhejiang, P.R. China
8 School of Medicine, Zhejiang University, Hangzhou, Zhejiang, P.R. China

Corresponding Authors: Shusen Zheng, e-mail: shusenzheng@zju.edu.cn, Wu Zhang, e-mail: zhangwu516@zju.edu.cn

Source of support: This work was supported by the National S&T Major Project (No. 2018ZX10301201; No. 2017ZX10203205), Innovative Research Groups of National Natural Science Foundation of China (No. 81721091), the National Science Foundation of China (No. 81721091, No. 2017ZX10203205), Zhejiang International Science and Technology Cooperation Project (No. 2016C04003), Research Unit Project of Chinese Academy of Medical Sciences (2019-I2M-5-030), Major Program of National Natural Science Foundation of China (No. 91542205)

Background: Pancreatic ductal adenocarcinoma (PDAC) is one of the most aggressive gastrointestinal tumors and has an extremely high mortality rate. Recent studies indicate that alternative splicing (AS), a common post-transcriptional process, has important roles in tumor biological behaviors and may provide novel immunotherapeutic targets. This study systematically analyzes AS profiles in PDAC and reveals their potential regulatory effects on cancer immune response.

Material/Methods: AS event, RNA sequencing, and splicing factor (SF) data were extracted from SpliceSeq, The Cancer Genome Atlas, and SpliceAid2, respectively. Overall survival (OS)-associated AS events and SFs were identified with univariate analysis. The LASSO method and multivariate Cox regression analysis were used to construct predictive signatures for the prediction of patient prognosis. The proportions of immune cells within PDAC samples were evaluated using the CIBERSORT algorithm. The correlations among AS events, SFs, and immune cell proportions were calculated using Spearman correlation analysis. Consensus clustering and immune classification were performed on the PDAC cohort.

Results: A total of 4812 OS-related AS events from 3341 parent genes were identified, and 8 AS-based predictive models were constructed for PDAC. An OS-related SF-AS regulatory network was constructed. The AS events regulated by ELAVL4 exhibited strong correlations with CD8 T cells and regulatory T cells. In addition, AS-based clusters demonstrated distinct OS outcomes and immune features.

Conclusions: AS-based predictive models with high accuracy were constructed to facilitate prognosis prediction and treatment of PDAC. An SF-AS regulatory network was constructed, revealing the potential relationships among SF, AS, and immune response.

MeSH Keywords: **Alternative Splicing • Carcinoma, Pancreatic Ductal • Computers, Molecular • Decision Support Techniques • Immunity, Cellular**

Full-text PDF: <https://www.medscimonit.com/abstract/index/idArt/925733>



4875 2 10 66

Background

Pancreatic ductal adenocarcinoma (PDAC) is the most aggressive and common malignancy of the pancreas, and is estimated to be the second leading cause of cancer-related deaths in developed countries [1,2]. The overall 5-year survival rate of PDAC is less than 8% worldwide, primarily owing to late diagnosis, early metastasis, and a high recurrence rate [3]. The stromal-rich tumor microenvironment (TME) of PDAC is considered to be a great barrier for the infiltration of cytotoxic immune cells [4]. The interactions between pancreatic cancer cells and stromal cells like cancer-associated fibroblasts work in concert to induce the malignant phenotypes of PDAC, such as angiogenesis, proliferation, and epithelial to mesenchymal transition (EMT) [5,6]. Moreover, owing to the highly heterogeneous molecular characteristics of the TME, PDAC patients on chemotherapies frequently experience drug resistance and tumor recurrence [4,7]. With advances in understanding the diversity and complexity of the PDAC TME, aberrantly expressed signaling pathways have been increasingly found to be significant in the regulation of the immune microenvironment. For instance, it was reported that the aberrantly activated WNT pathway is a critical factor driving the central gene expression signatures that fuel lymph node metastasis in PDAC by shaping the tumor milieu [8]. In addition, immune-based therapies targeting such signaling show increased potency in enhancing tumor cytotoxicity. Thus, elucidating the molecular mechanisms underlying the PDAC TME with data analysis could be crucial for improving diagnostic, prognostic, and therapeutic techniques in PDAC.

Pre-messenger RNA (pre-mRNA) transcripts require the selective exclusion of introns and the inclusion of specific exons to form mature mRNA in a process known as alternative splicing (AS), one of the most researched forms of mRNA processing [9]. AS enables a single gene to be transcribed into various mRNA isoforms at the post-transcriptional level, thereby diversifying the translation of proteome [9,10]. AS is a ubiquitous RNA processing pattern across diverse cells and tissues, and its dysregulation is involved in multiple pathophysiological processes, including cancer [11,12]. Accumulating evidence implicates aberrant AS events as the hallmarks of carcinogenesis, including tumor growth, angiogenesis, immune evasion, and metastasis [13–15]. AS promotes tumor progression not only by activating oncogene isoforms such as *VEGFA* but also through facilitating the degradation of suppressor genes such as *TP53* [16,17]. AS events are also accurate predictive markers for various cancer types, including breast, lung, colorectal, and liver [18–21]. Peptides generated from tumor-specific AS events may serve as neopeptides (similar to neoantigens derived from somatic mutations), which are potential therapeutic targets for emerging immunotherapies [22,23]. The prognostic value of AS events and their potential use in novel therapies have made them a hot topic in research.

Splicing factors (SFs) play an important role in the regulation of AS. SFs promote or inhibit certain splicing events by binding to specific sequence motifs. Somatic mutations and altered expression of SFs, together with dysregulated AS events, have been found in several cancers, indicating a possible role in tumorigenesis [24–26]. However, because the regulation of aberrant AS events by SFs in cancers is only partially understood, it is imperative to investigate this relationship and its potential effects on PDAC.

Few studies have focused on both the complex genomic splicing events and the interplay of cancer cells with their surrounding tumor stroma, which together result in poor outcomes for PDAC patients. Therefore, we systematically characterized cancer-specific AS events in PDAC patients using The Cancer Genome Atlas (TCGA) data and constructed AS-based prognostic models. Furthermore, we established a regulatory network of SFs and their target AS events and analyzed their associations with immune-infiltrating cells. Finally, we clustered PDAC cohorts based on clinical and molecular characteristics, evaluated the immune landscape of each cluster, and explained the corresponding clinical phenotypes.

Material and Methods

Ethical compliance

All the clinical and genomic data in this study were acquired from open-access public databases, including TCGA, SpliceSeq, and SpliceAid2. Our acquisition procedure fully complied with the requirements of these databases. Therefore, no ethics committee approval or consent process was needed.

Data acquisition and pre-processing

Level-3 RNA-seq data and the corresponding clinical information for 177 PDAC patients were downloaded from TCGA (<https://portal.gdc.cancer.gov/>). To explicitly quantify the inclusion level of each splice junction and exon, SpliceSeq was used to profile the AS patterns of each PDAC sample. The percent spliced in (PSI) value, which ranges from 0 to 1, was calculated to assess the incidence of splicing patterns in each specific AS event [27]. A total of 7 AS event patterns, including exon skipping (ES), alternate donor site (AD), alternate acceptor site (AA), alternate promoter (AP), alternate terminator (AT), mutually exclusive exons (ME), and retained intron (RI), were included in the study. To obtain a set of reliable AS events, we used the R package *impute* to replenish the missing PSI data and applied strict filtering criteria (average PSI value >0.05, percentage of samples with PSI value ≥75%). To distinguish different AS events, each event was assigned a unique identifier code. For instance, in code “SMAD4|45557|AP”, SMAD4 is the gene symbol, 45557 is the ID number from SpliceSeq, and AP is

the splicing subtype. The R package UpSet (version 1.3.3) was used to visualize the intersections of the 7 AS event subtypes.

Survival analysis and functional annotation

Overall survival-associated AS (OS-AS) events were initially identified by univariate Cox analysis based on PSI values. A false discovery rate <0.05 was considered significant. A protein–protein interaction (PPI) network was constructed based on the parent genes of the top 500 significant OS-AS events from the String database (<https://string-db.org/>) with an interaction score of 0.7. Cytoscape (version 3.7.1) was used to visualize the network. We performed gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) functional enrichment analysis on parent genes of the prognostic AS events using the R package clusterProfile.

Construction of predictive models based on AS events

The 20 most significant OS-AS events for each subtype were used to construct prognostic models. The 20 most significant OS-AS events in all subtypes were used in the final model. The LASSO method is widely used to identify predictive rules and avoid data overfitting in multi-dimensional data [28]; therefore, we used LASSO implemented with the R package glmnet to decrease redundancy and irrelevance. Survival-associated AS events of each subtype were selected according to the optimal penalty λ value. Multivariate Cox regression analysis was used to optimize the prognostic model and to calculate the regression coefficients of identified predictor events in each model. Risk scores for each sample were calculated using the PSI value and regression coefficient of each predictor event. PDAC patients were further divided into high- and low-risk groups according to the median risk score value, and Kaplan-Meier curves were used to compare survival differences between the 2 groups. Receiver operator characteristic (ROC) curves and the area under curve (AUC) were used to evaluate and compare the predictive accuracies of each model.

Survival related SF-AS regulatory network and immune correlation analysis

A total of 71 SFs were extracted from SpliceAid2 (www.introni.it/spliceaid.html) [29]. Univariate Cox regression analysis was performed to identify OS-SFs based on mRNA expression data. Spearman correlation analysis was performed to evaluate the relationship between the PSI values of AS events and mRNA expression of the OS-SFs. An SF-AS regulatory network was generated using Cytoscape, with correlation coefficients greater than $|0.6|$.

CIBERSORT is a robust deconvolution algorithm that predicts the proportions of 22 human immune-infiltrating cell

subsets based on the mRNA expression profiles in cancer patients [30]. It was run using the LM22 signature (downloaded from website <https://cibersort.stanford.edu>) and 1000 permutations. Samples with output values of $P<0.05$ were preserved. The results were generated in violin plot using the R package ggplot2. We used CIBERSORT to evaluate and quantify the infiltration levels (fractions) of various immune cell types in 177 PDAC samples with $P<0.05$. Spearman correlation analysis determined the relationships between the PSI values of AS events and infiltrating immune cell subtypes. To illustrate the complex regulatory correlation, we visualized the SFs, AS events, and immune cell proportions in a dual axis plot using the R package ggplot2.

Clustering and immune landscape analysis

Unbiased classification of the TCGA PDAC cohort was performed using the R package CancerSubtypes based on identified OS-AS events [31]. To ensure a robust and unsupervised classification, we used the consensus cluster algorithm with the following settings: $\text{maxk}=5$, $\text{pltem}=0.8$, $\text{innerlinkage}=\text{average}$, $\text{distance}=\text{pearson}$. The optimal number of clusters was determined by the cumulative distribution curve, and then survival analysis comparisons were made among these clusters. The immune scores and stromal scores of each PDAC cohort were calculated using the ESTIMATE algorithm to reflect the TME constitution [32]. Six disparate immune subgroups were referenced from the research of Thorsson et al. [33]. Distributions of clinicopathological phenotypes, including TNM stage, survival status, grade, gender, microsatellite instability, and driver gene mutations, including *TP53*, *KARAS*, *SMAD4*, and *CDKN2A* (encoding p16), were compared using chi-squared tests (2-sided, $P<0.05$). A flowchart of this study is shown in Figure 1.

Statistical analysis

Statistical analyses were carried out using R version 3.5.1 with multiple publicly available packages. Independent *t* test and analysis of variance were used to determine differences among groups. Survival curves were generated using the R package survival. Spearman's correlation analysis was used to evaluate the correlations of AS events, SFs, and immune infiltration levels. Pearson's chi-squared test was employed for comparison of the distributions of clinicopathological phenotypes and somatic mutation phenotypes. $P<0.05$ was considered statistically significant.

Results

Systematic characterization of AS events in PDAC cohort

A total of 177 PDAC patients with corresponding clinical information and AS patterns were included. The 7 types of AS events

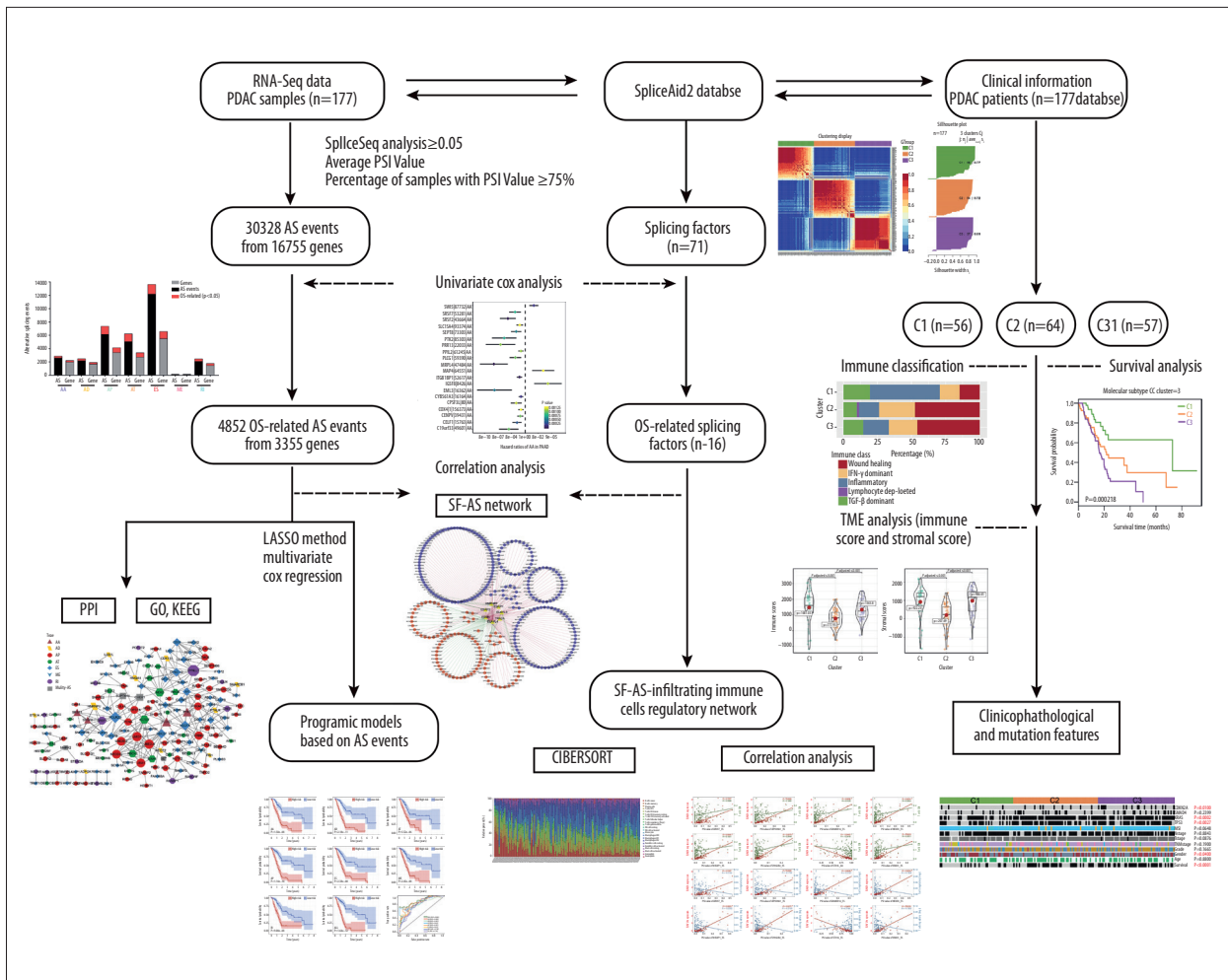


Figure 1. Experimental design of the study. RNA-seq data and clinical information of pancreatic ductal adenocarcinoma (PDAC) patients were extracted from the Cancer Genome Atlas database. Splicing factors were obtained from SpliceAid2 database. Percent spliced in (PSI) value of each AS event was calculated with SpliceSeq analysis, and subjected to stringent filters: average of PSI value ≥ 0.05 and percentage of samples with PSI value $\geq 75\%$. Univariate Cox analysis was used to identify overall survival-related alternate splicing (AS) events and splicing factors. Then, a splicing factor (SF)-AS regulatory network was constructed based on correlation analysis. In terms of OS-related AS events, enrichment analysis was performed on their parent genes for functional exploration, and LASSO multivariate Cox regression was used to build an AS-based prognostic model. CIBERSORT was used to further explore the correlation between SFs, AS events, and immune-infiltrating cells. Finally, to explore the relationship between AS events and patient survival in immune infiltration, consensus clustering was performed on the PDAC cohort. Three clusters exhibited distinct characteristics in survival analysis, immune classification, tumor microenvironment analysis, and clinicopathological features.

included in the study are demonstrated in the schematic diagram in Figure 2A. After rigorous filtering of the original data, 30 328 AS events were detected from 16 755 genes, including 2597 AAs in 1962 genes, 2184 ADs in 1655 genes, 6130 APs in 3388 genes, 5040 ATs in 2697 genes, 12 177 ESs in 5479 genes, 131 MEs in 130 genes, and 2069 RIs in 1444 genes (Figure 2B). As 1 gene could have multiple splicing events, UpSet plots were used to determine the intersections of the 7 AS subtypes with gene numbers (Figure 2C). ES splicing constituted the majority of AS patterns and accounted for more than one-third of AS events in PDAC; 60 genes had up to 5 types of AS events.

Prognostic AS events in PDAC and functional enrichment analysis

To explore the prognostic value of splicing events in PDAC, univariate Cox regression analysis was performed to acquire OS-AS events. We detected a total of 4812 OS-AS events from 3341 parent genes, including 263 AAs in 249 genes, 290 ADs in 263 genes, 1218 APs in 736 genes, 1173 ATs in 673 genes, 1455 ESs in 1071 genes, 24 MEs in 24 genes, and 389 RIs in 325 genes (Figure 2B, $P < 0.05$). Three genes occurred in up to 4 AS event types, although most genes underwent only 1 type

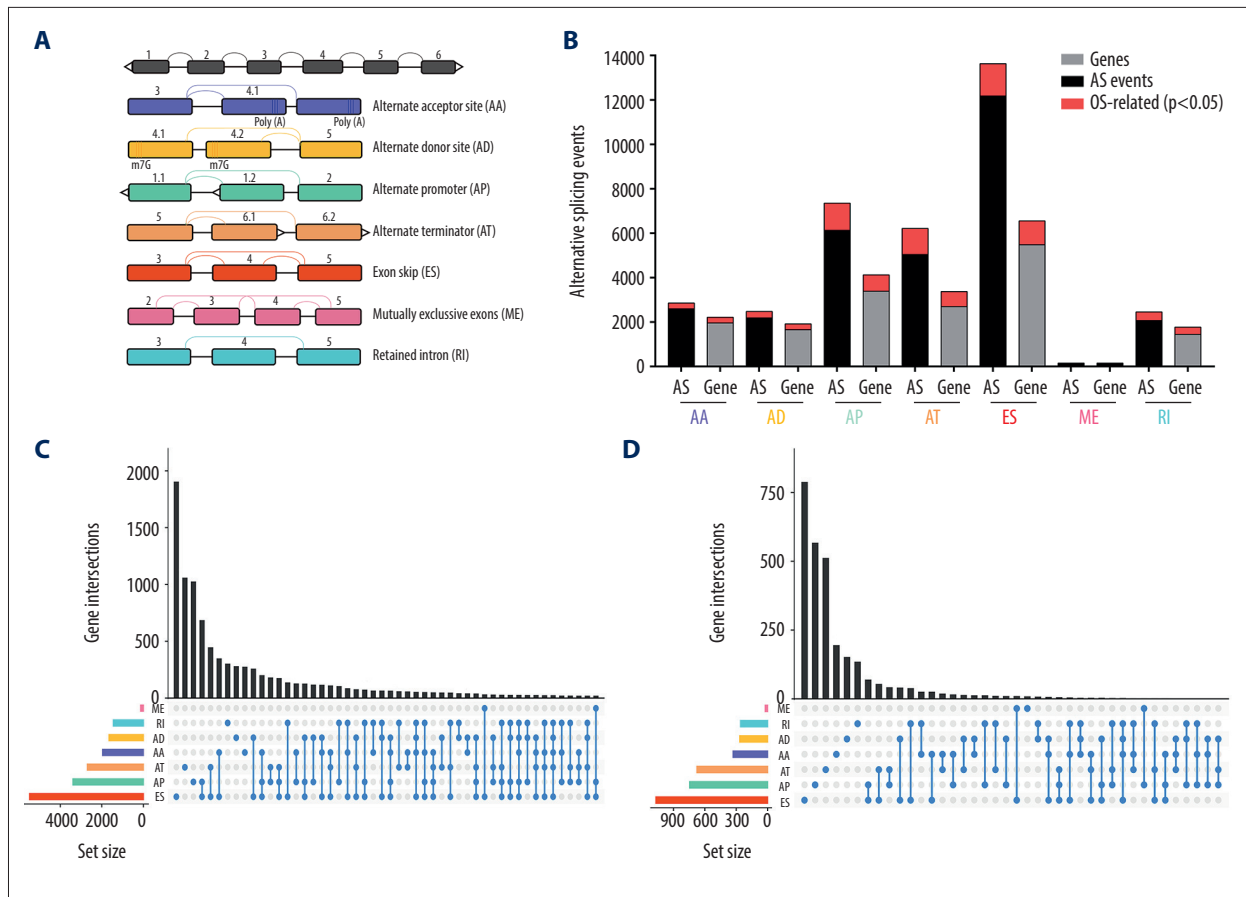


Figure 2. Profiles of 7 alternate splicing (AS) patterns and overall survival (OS)-related AS events in pancreatic ductal carcinoma (PDAC). **(A)** Schematic illustration of 7 AS patterns: exon skipping (ES), alternate donor site (AD), alternate acceptor site (AA), alternate promoter (AP), alternate terminator (AT), mutually exclusive exons (ME), and retained intron (RI). **(B)** Bar plot of counts of AS events and parent genes for the 7 AS patterns. Red bars represent OS-AS events and parent genes from univariate Cox analysis. Black and gray bars represent all AS events and corresponding parent genes in PDAC, respectively. **(C)** UpSet plot of intersections of the 7 patterns of AS events. One gene may have up to 4 types of AS events in the dot panel, as shown in the next to last line, which connects 4 AS events (RI, AD, AT and ES). **(D)** UpSet plot of intersections of the 7 patterns of OS-AS events.

of AS event (Figure 2D). The 20 most significant prognostic AS events in each pattern are shown in Figure 3. The majority of the mRNA splicing events acted as favorable prognostic predictors (hazard ratio [HR] <1) in the AA, AD, ES, and RI splicing patterns.

Given that AS might affect the expression and translation of the parent RNA and protein, we constructed a PPI network of the 500 most significant prognostic AS events to provide an overview of the cellular interactions at a protein level (Figure 4A). Hub genes including *POLR2D*, *DYNLL1*, *SRSF1*, *VEGFA*, and *PRKACB* were identified in the PPI network. GO and KEGG functional enrichment analyses were performed on all the parent genes of OS-AS events (Figure 4B). Most of the significantly enriched GO terms were in “RNA splicing”, “focal adhesion”, “cell adhesion molecule binding”, “tumor necrosis

factor-mediated signaling pathway”, and “cadherin binding”; whereas the KEGG analysis showed enrichment in pathways including “guanyl-nucleotide exchange factor activity”, “protein binding, bridging”, and “EGFR tyrosine kinase inhibitor resistance”.

Construction of predictive models in PDAC cohort

To obtain predictive models with high accuracy, we performed LASSO logistic regression to select the 20 most powerful prognostic predictor OS-AS events in each pattern. Multivariate Cox analysis was used to determine the independent predictors. Seven prognostic models were constructed, with 11 AS-events in AAs, 8 in ADs, 7 in APs, 5 in ATs, 7 in ESs, 7 in MEs, and 5 in RIs (Supplementary Table 1). The final predictive model was constructed using the top 20 prognostic events from all

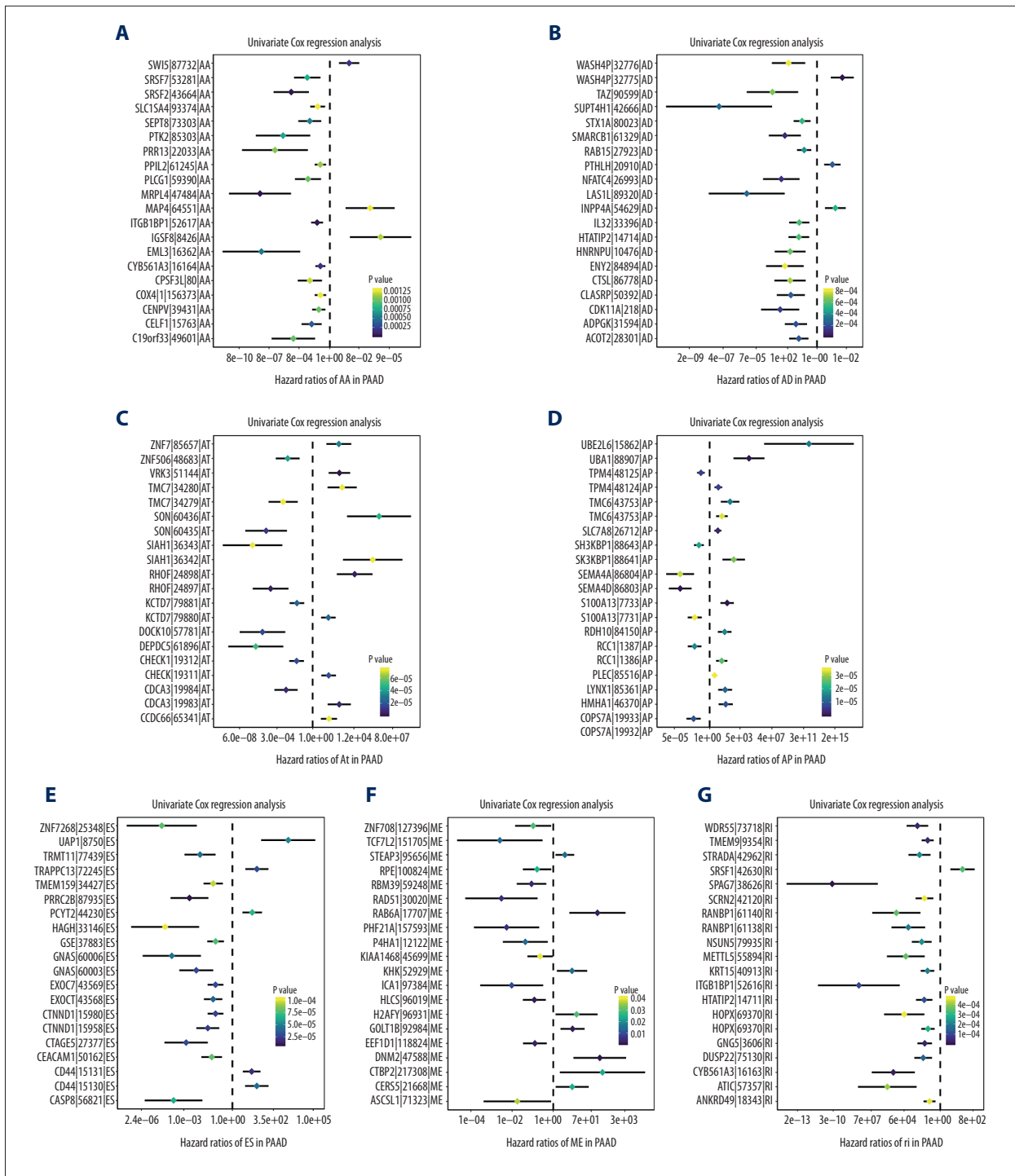


Figure 3. Forest plots for subgroup analyses of the 7 patterns of alternative splicing (AS) events associated with overall survival (OS) in pancreatic ductal carcinoma. (A–G) Forest plots of hazard ratios for top 20 OS-AS events: exon skipping (ES), alternate donor site (AD), alternate acceptor site (AA), alternate promoter (AP), alternate terminator (AT), mutually exclusive exons (ME), and retained intron (RI). P-values are indicated by the color scale; horizontal bars represent 95% CIs.

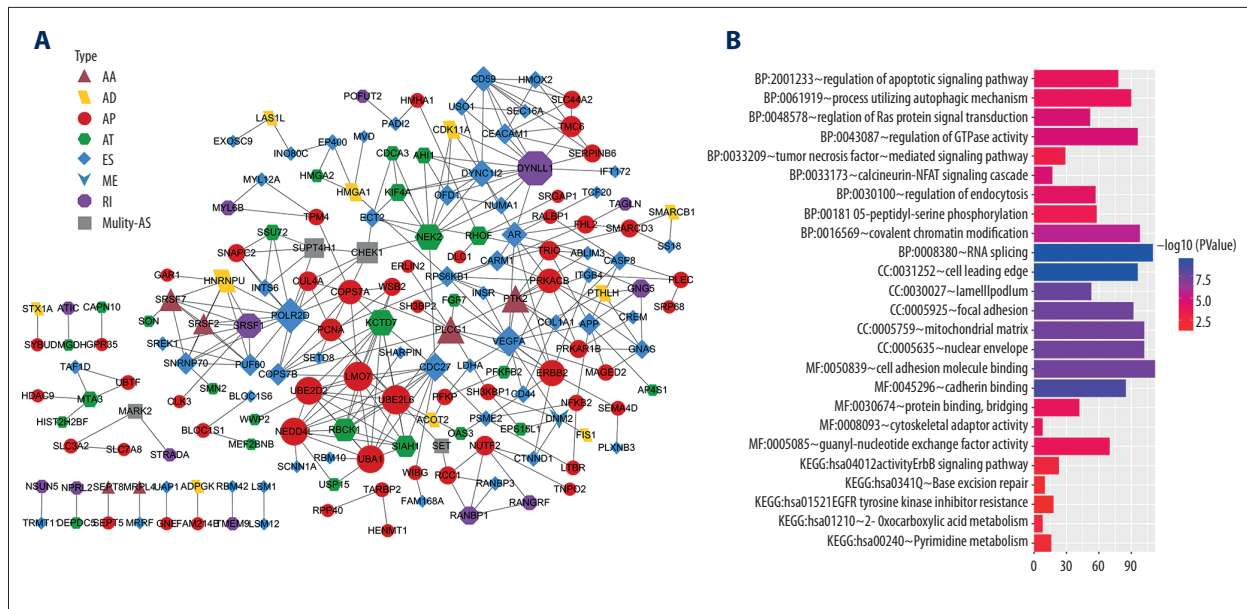


Figure 4. Protein-protein interaction network and functional annotations for parent genes of overall survival (OS)-associated alternative splicing (AS) events in pancreatic ductal carcinoma. **(A)** Interaction network of top 500 significant OS-associated AS events. Different shapes and colors represent different AS types of parent genes; node size corresponds to the number of neighboring genes. **(B)** Functional enrichment pathways from gene ontology and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis. Top 10 significant biological process pathways, and top 5 significant molecular function (MF), cellular component (CC), and KEGG pathways. Color and length of each pathway represent false discovery rate and number of enriched pathways, respectively. AA – alternate acceptor site; AD – alternate donor site; AP – alternate promoter; AT – alternate terminator; ES – exon skipping; HR – hazard ratio; ME – mutually exclusive exons; RI – retained intron.

AS patterns, and 6 events were eventually included in the final model (Supplementary Table 1). Risk scores were calculated using the coefficients of each event's PSI value. The PDAC cohort was divided into high- and low-risk groups according to the median risk score. Supplementary Figure 1 shows the risk score curve and distribution of survival outcomes, as well as a PSI heatmap for each model. Kaplan-Meier survival analysis showed significant differences in survival outcomes between the high- and low-risk groups in all 8 prognostic models, suggesting excellent predictive performance ($P < 0.001$, Figure 5A–5H). The ROC curve confirmed the superiority of the final prognostic model, with an AUC of 0.884 for 3-year survival (Figure 5I).

To assess the performance of the final prognostic model, we performed univariate and multivariate Cox regression analyses on the TCGA PDAC cohort. Associations between survival and risk scores, as well as clinicopathological features including age, gender, grade, stage, and T and N stage were analyzed (Supplementary Table 2). In the univariate Cox analysis, age, N stage, and risk score were all significantly correlated with poor prognosis (Figure 6A). Multivariate Cox analysis showed that risk score (HR=1.149, 95% CI [1.097–1.202], $P < 0.001$), age, and N stage were independent prognostic factors for PDAC (Figure 6B).

Regulatory network of OS-associated SFs and AS events

SFs regulate AS by binding to specific sequence motifs of pre-mRNA, which in turn might contribute to tumor progression. Existing evidence suggests that prognostic AS events are determined by a limited number of SFs [34]. Hence, we constructed a potential regulatory network to explore this underlying relationship. Based on level-3 RNA-seq data from TCGA, we performed univariate Cox analysis on 71 SFs and acquired 16 SFs whose expression was significantly associated with OS in PDAC patients. Spearman's correlation analysis identified 7 SFs strongly ($|cor| > 0.6$, $P < 0.001$) correlated with 319 AS events, among which 105 were adverse AS events (HR > 1) and 214 were favorable AS events (HR < 1, Figure 7A). More than half the AS events were positively correlated with SF expression and were mainly favorable prognostic events. Furthermore, the small proportion of AS events negatively regulated by SFs mainly predicted unfavorable prognosis. Thus, the AS events and SFs in the regulatory network may inhibit the development of pancreatic cancer. Moreover, PDAC patients were divided into 2 groups according to the median expression value of OS-related SFs. Kaplan-Meier analysis showed significant prognostic differences in 4 SFs: TRA2A, ESRP1, ELAVL3, and ELAVL4. High expression of ELAVL4 (HR=0.57, $P=6.1e-0.3$), RBM5 (HR=0.53, $P=1.5e-0.2$), and TRA2A (HR=0.58, $P=4.6e-0.3$) predicted favorable OS in

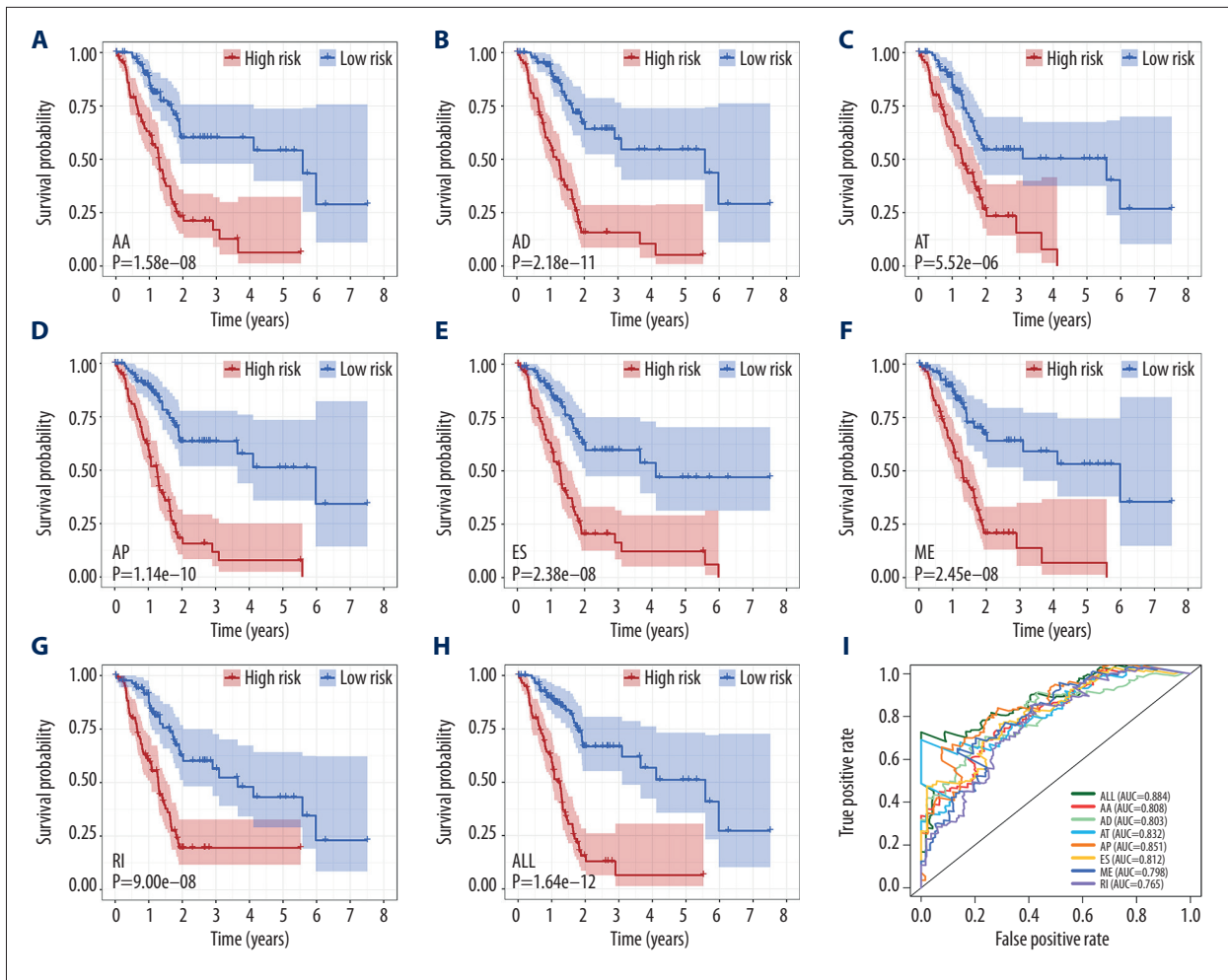


Figure 5. Kaplan-Meier survival analysis and receiver operating characteristic (ROC) curves of 8 prognostic models based on overall survival (OS)-related alternative splicing events for pancreatic ductal carcinoma (PDAC) cohort. (A–H) High- and low-risk groups in each prognostic model show OS difference in Kaplan-Meier plot. (I) ROC curves and area under the curve of each prognostic model for 3-year survival in PDAC cohort. AA – alternate acceptor site; AD – alternate donor site; AP – alternate promoter; AT – alternate terminator; ES – exon skipping; HR – hazard ratio; ME – mutually exclusive exons; RI – retained intron.

PDAC, whereas high expression of ESRP1 (HR=1.9, $P=1.3e-03$) predicted adverse OS (Figure 7B–7E). These results indicate the overall tumor inhibitory effect of the SF-AS regulatory network on PDAC. The most significantly correlated AS events with 4 SFs are shown in Figure 7F–7M.

Correlation between AS events, SFs, and infiltrating immune cells

Accumulating evidence demonstrates the important role of AS in TME remodeling and that AS alterations may affect immune cell infiltrations [35]. To explore potential mechanisms, we used CIBERSORT to evaluate the proportion of each infiltrating immune cell type in the PDAC samples (Figure 8A). Correlation analysis showed that expression of ELAVL4 was

strongly positively correlated with the PSI values of ARVCF, ATP6V0A1, GRAMD1A, MEAF6, NCKAP1, SH3GLB2, and VDACC3 in the ES pattern (Figure 8B). Moreover, these AS events were significantly positively correlated with the fraction of CD8 cytotoxic T cells in PDAC patients, indicating a positive regulatory relationship between ELAVL4 and CD8 cytotoxic T cells. In addition, these AS events were negatively associated with the regulator T cell (Treg) subtype (Figure 8C). Thus, we identified a potential regulatory network in which ELAVL4 not only positively regulates the activation of CD8 T cells but also negatively regulates Tregs, via alterations in AS events.

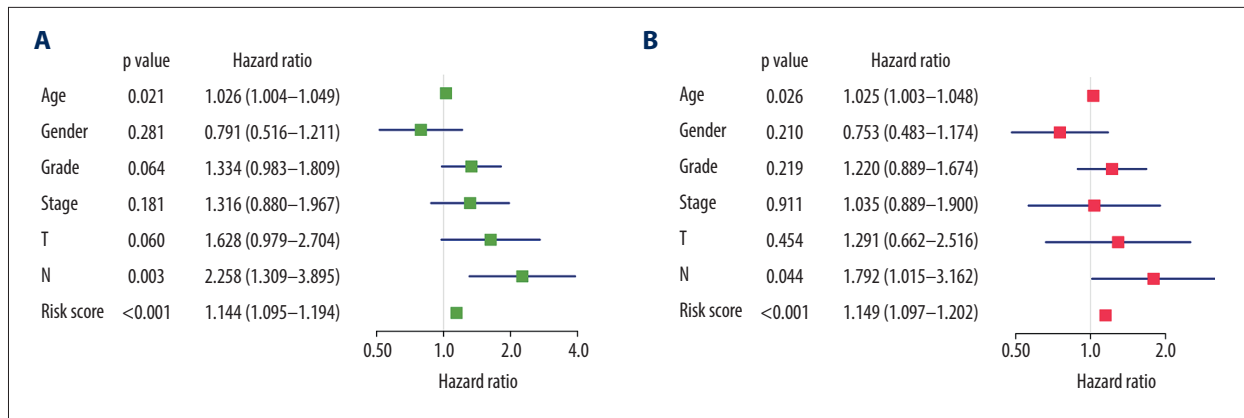


Figure 6. Forest plots of clinicopathological features and risk score (all) in pancreatic ductal carcinoma patients. **(A)** Univariate Cox regression analysis of the association between survival and clinicopathological features as well as alternative splicing (AS)-based risk score. **(B)** Multivariate Cox analysis of the association between survival and clinicopathological features as well as AS-based risk score.

Clustering based on AS events significantly correlated with immune features

Previous experiments have demonstrated the heterogeneity of AS events between various PDAC cohorts, and the identification of aberrant AS events has helped to elucidate patterns among the PDAC cohorts. To better understand the molecular heterogeneity underlying AS events in PDAC, we performed unsupervised consensus clustering on the 177 PDAC samples. As shown in Figure 9A, the clusters had significant independence and acquired a balanced partition with $k=3$. Thus, the tumor-specific clustering based on AS events classified the PDAC cohorts into 3 subgroups: C1 ($n=56$, 31.6%), C2 ($n=64$, 36.2%), and C3 ($n=57$, 32.2%) (Figure 9B). Kaplan-Meier survival analysis of the 3 subgroups demonstrated distinct survival outcomes: C1 showed the best prognosis, followed by C2, whereas C3 exhibited poor prognosis ($P<0.001$) (Figure 9C).

Given the close relationship between AS events and infiltrating immune cells identified in previous steps, we systematically evaluated the immune landscapes of the 3 clusters. Immune classification was performed on each sample within the clusters, based on a previous study by Thorsson et al. [33]. The majority of C1 samples were classified as inflammatory subtypes (51.2%), characterized by elevated Th17 and Th1 expression and inhibited tumor growth. In addition, C1 had lower proportions of wound healing subtypes (14.6%) compared with the other clusters, indicating reduced angiogenesis and lower proliferation of tumor cells. C2 and C3 had higher proportions of wound healing (C2, 47.5%; C3, 45.8%) and IFN- γ -dominant subtypes (C2, 26%; C3, 20.8%), suggesting an immunosuppressive state, whereas C2 contained fewer TGF- β dominant samples and more IFN- γ samples compared with C3, indicating higher levels of inflammatory response (Figure 9D).

We also calculated the immune scores and stromal scores of each PDAC sample to assess the potential constitution of TME in the 3 clusters. C1 had the highest immune scores, followed by C3 and C2 ($P<0.001$), consistent with the immune classifications (Figure 9E). C3 had slightly higher stromal scores than C1, and C2 had the lowest ($P<0.001$) (Figure 9F). Thus, the majority of immune-infiltrating cells in C3 may have immunosuppressive status and facilitate tumor progression. Figure 9G shows clinicopathological information and driver gene mutations for each cluster. The prevalent mutated genes in pancreatic cancer, including *KRAS* ($P=0.0002$), *CDKN2A* ($P=0.0100$), and *TP53* ($P=0.0027$), were preferentially distributed in the C2 and C3 clusters, indicating a close relationship between AS-based clusters and somatic mutation.

Overall, these findings indicate that clusters based on AS events differ significantly in terms of immune molecular patterns; thus, they represent potential immunotherapeutic targets as well as prognostic tools for PDAC.

Discussion

PDAC is one of the most malignant tumors of the gastrointestinal system and exhibits high molecular heterogeneity and a complex immune microenvironment [7]. Recent research into aberrant AS events has shed some light on the complex biological behaviors of such tumors [23]. In this study, we systematically collected clinical and genomic data from TCGA and SpliceSeq to analyze the role of AS in PDAC. We used univariate analysis to identify 7 subtypes of AS events associated with survival, many of which are involved in PDAC development. CD44|15131 and CD44|15130 were identified as top AS events significantly correlated with survival outcomes ($P<0.001$). Gansauge et al. found that the CD44 variant 6 (CD44v6) and

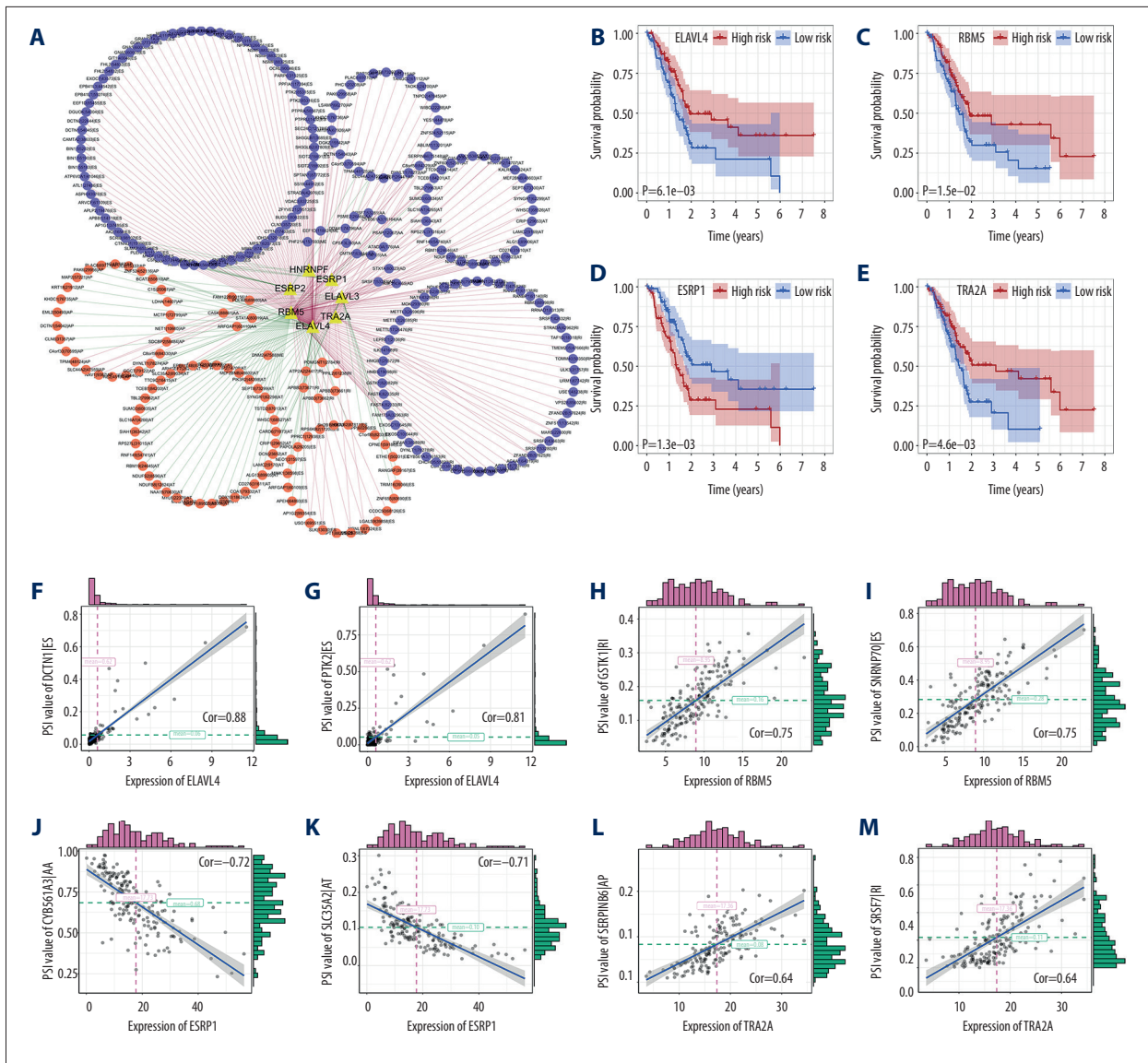


Figure 7. (A) Regulatory splicing factor (SF)-alternative splicing (AS) network based on strong correlations between percent spliced in (PSI) value of splicing events and mRNA expression of SFs ($|cor| > 0.6$). Favorable (purple circles) or unfavorable (orange circles) AS events were positively (red lines) or negatively (green lines) regulated by SFs (yellow triangles). (B-E) Kaplan-Meier survival analysis validated the prognostic value of SFs. Four SFs (ELAVL4, ESRP1, RBM5, and TRA2A) were significantly associated with overall survival of pancreatic ductal carcinoma patients ($P < 0.05$). (F-M) Representative dot plots showing the correlation between expression of prognostic SFs (ELAVL4, ESRP1, RBM5, and TRA2A) and PSI values of splicing events.

CD44 standard (CD44s) isoforms were significantly reduced in PDAC patients and soluble CD44v6 is a potential biomarker for unfavorable outcomes [36]. Another study showed that exon v7-containing CD44 isoforms could mediate tumor stroma formation through activating PI3K/AKT signaling, which is essential for tumor metastasis [37]. Differentially expressed splice variants of CD44 also serve as biomarkers in PDAC [38,39]. The diversity of splicing-derived CD44 isoforms makes them promising targets for immunotherapies.

IL32|33396|AD was identified as an unfavorable prognostic event in this study. Interleukin-32 (IL32) was previously reported to be a novel cytokine comprised of 8 small exons that has 9 AS isoforms [40]. In PDAC, the IL-32 α isoform can revert IL-6-induced EMT, invasion, and metastasis through inactivation of JAK2/STAT3 signaling [41]. Many other prognostic AS events identified here were also potentially associated with PDAC, including SMAD4|45557|AP, SMAD4|45559|AP, TP53AIP1|19438|AT, and SRSF2|43664|AA, whose parent genes are reported to regulate PDAC progression [42,43]. In addition

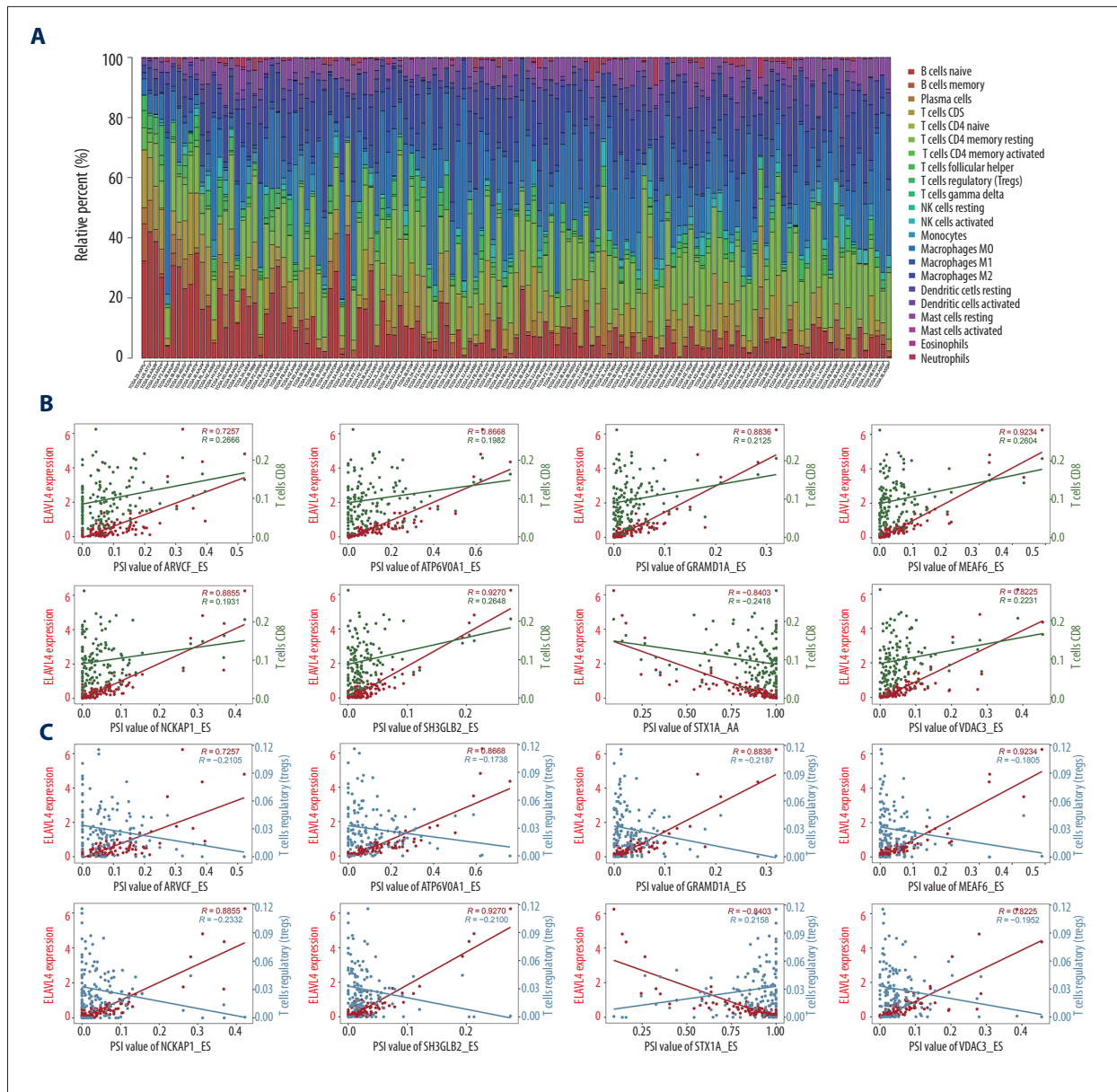


Figure 8. Correlation analysis of prognostic splicing factors (SFs), alternative splicing (AS) events, and infiltrating immune cells. **(A)** Proportion of the 22 human immune cell subtypes in each pancreatic ductal carcinoma sample calculated with CIBERSORT. **(B, C)** Dual axis correlation dot plot of SFs, AS events, and infiltrating immune cells. The x-axis represents percent spliced in (PSI) values of AS events, the left y-axis represents SF mRNA expression, and the right y-axis represents relative proportions of immune cell subtypes.

to RNA-splicing pathways, significantly enriched functions included “focal adhesion”, “cell adhesion molecule binding”, “cadherin binding”, and “EGFR tyrosine kinase inhibitor resistance”, which have been implicated in tumor metastasis, invasion, and drug resistance [44,45]. However, the molecular mechanisms associated with these AS events in PDAC require more research.

Clinically, pancreatic cancer is difficult to diagnose at an early stage, which usually leads to poor prognosis. Many diagnostic biomarkers and prognostic models for pancreatic cancer have been proposed based on mRNA and noncoding RNA profiles [46]; however, gene expression signatures are infrequently used in clinical practice. In this study, the final prognostic model based on AS subtypes showed efficient prognostic prediction. Moreover, ROC curves and multivariate analysis confirmed that the final model, consisting of 6 AS events, could

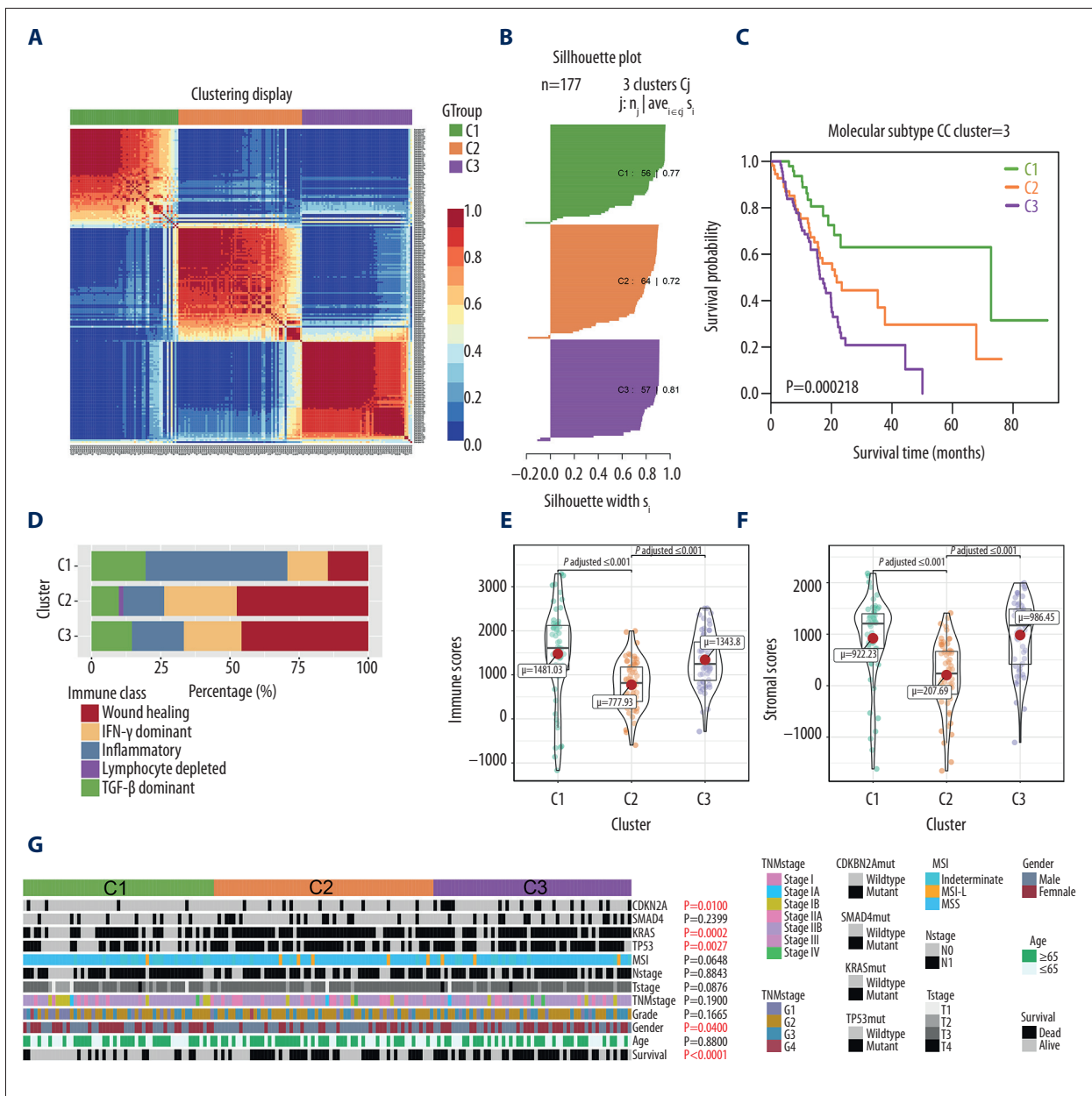


Figure 9. Consensus clustering based on prognosis-related alternative splicing (AS) events and immune landscape analysis. (A, B) Consensus clustering matrix and silhouette plot identifying 3 clusters. (C) Kaplan-Meier survival plot showing significant differences among the 3 clusters (log-rank *t* test, $P<0.05$). (D) Classification of pancreatic ductal carcinoma samples based on 5 immune subtypes in 3 clusters: wound healing, IFN- γ dominant, inflammatory, lymphocyte depleted, and TGF- β dominant. (E, F) Comparison of immune and stromal scores of the 3 clusters with analysis of variance. (G) Heat map of clinicopathological characteristics and gene phenotypes in the 3 clusters. Driver gene mutations, including *KRAS*, *CDKN2A*, and *TP53*, showed distinct distribution in the 3 clusters with chi-squared test.

accurately predict patient outcomes and independently stratify high-risk PDAC patients. Therefore, our final prognostic model is suitable for clinical use.

The extensive occurrence of aberrant AS events in the TME may be driven by a small number of SFs. A regulatory network

was constructed based on SFs showing strong correlation ($|cor| > 0.6$) with OS-related AS events; these included ESRP1, ESRP2, RBM5, hnRNP, ELAVL4, ELAVL3, and TRA2A, most of which have been reported to affect the biological behaviors of tumors [47–49]. For example, heterogenous nuclear ribonucleoprotein family F (hnRNP) inhibits EMT in breast cancer

by regulating AS events of CD44 in a G-quadruplex-dependent manner [50]. Another ESRP1-CD44 regulatory axis, not shown in the correlation network, has an important role in PDAC development. EMT reduced the expression of ESRP1, thereby regulating the splicing of CD44, turning variant CD44v into CD44s isoforms [51]. ZEB1 represses ESRP1 to regulate AS events of CD44s, and CD44s activate ZEB1, forming a regulatory loop between CD44s, ESRP1, and ZEB1. This feedback loop promotes EMT and sustains stemness in pancreatic cancer, reflecting the crucial role of both SFs and OS-related AS events in PDAC [51]. SRSF1 enhances chemoresistance and inhibits drug-induced apoptosis in PDAC by regulating the AS events of target genes such as MNK2b [52]. In the current study, SRSF1 was positively regulated by RBM5 in event SRSF1|42632|RI ($P < 0.001$), which indicated favorable prognosis (HR < 1), which is contrary to previous reports. This discrepancy suggests that the same gene may have completely different functions in the regulation of different AS events. Furthermore, it highlights the complexity of the SF-AS regulatory network underlying PDAC progression.

The majority of favorable prognostic AS events (HR < 1) were upregulated by SFs, whereas adverse AS events (HR > 1) were downregulated. In contrast to observations in other cancers, this suggests that most of the SFs in PDAC inhibit tumor progression through the regulation of AS events [53,54]. However, owing to the lack of normal samples in TCGA, we were unable to identify differentially expressed SFs and draw comprehensive conclusions about the effects of SFs on PDAC. Nevertheless, we identified 4 SFs (ELAVL4, RBM5, ESRP1, and TRA2A) significantly related to OS in PDAC patients. The high expression of ELAVL4, RBM5, and TRA2A predicted favorable outcomes.

Recent studies indicated that aberrant AS events may create a predisposition to the regulation of immune response of cancers [22,23,35]. Proteins derived from aberrant AS events in tumors can be processed into residual peptides by proteasome. The peptides are then transported into the endoplasmic reticulum through the transporter associated with processing, after which they could be loaded onto major histocompatibility complex class I molecules, which can be recognized by T lymphocytes to induce anti-tumor immune responses [55]. Therefore, we evaluated the correlation of AS events with immune cell subtypes within tumor samples. AS events regulated by ELAVL4 were significantly associated with upregulated CD8 cytotoxic T cells and downregulated Tregs within the PDAC TME, indicating ELAVL4's potential role in inhibiting PDAC tumor progression. ELAVL4 (also known as HuD) was initially identified as a neural antigen specifically expressed in small cell lung cancer or neuroblastoma [56]. Later, some studies reported that HuD could elicit specific cytotoxic T lymphocytes to inhibit tumor growth in small cell lung cancer patients, and it was studied as a potential vaccination for lung cancers [57,58]. The present study suggests that ELAVL4 elicits the activation

of effector T cells within the PDAC TME in an AS-dependent manner, making it a promising new immunotherapeutic target in PDAC. Moreover, its inhibition of immunosuppressive T cells and activation of CD8 T cells indicate that the SF-AS axis could reverse the immunosuppressive TME. However, more detailed experiments are needed to validate this effect.

To the best of our knowledge, this is the first study to comprehensively analyze AS-based clusters in PDAC from an immune landscape perspective. Using integrated algorithms, 3 prognostic AS event-based clusters were identified, with distinct clinical outcomes according to survival analysis. Clinical characteristics and the genetic phenotypes of *KRAS*, *TP53*, *SMAD4*, and *CDKN2A* were unevenly distributed within the clusters. In a recent study, Thorsson et al. identified 6 novel immune subtypes across 33 cancer types and 10 000 TCGA tumors, which could help to explain the biological behaviors of tumors and tailor targeted immunotherapies [33]. Notably, the 3 AS clusters identified here exhibited distinct immune features after reclassification. The C1 cluster included inflammatory subtypes, suggesting an immune-activated status, featuring higher expression of pro-inflammatory genes and low to moderate tumor growth [59]. Furthermore, the inflammatory subtype showed lower levels of aneuploidy and somatic mutations compared with the other subtypes [60]. Somatic mutations of 4 driver genes were preferentially distributed in the C2 and C3 clusters, supporting the efficacy of this immune classification. C2 showed higher proportions of IFN- γ -dominant subtypes as well as wound healing subtypes. IFN- γ subtypes have high M1/M2 macrophage polarization and expression of CD8 genes and also demonstrate significant tumor cell proliferation, which may override the evolving inflammatory response to promote tumor progression [61]. The wound healing subtype was characterized by the transformation into immunosuppressive states and promotion of tumor angiogenesis [62]. Overall, the C2 cluster represents a transition from inflammatory-dominant to immunosuppressive status. In C3, increasing TGF- β subtypes suppressed immune response and promoted tumor development [63]. This AS-based clustering and immune classification may facilitate prediction of prognosis and patient response to different immunotherapies, as has been demonstrated in squamous cell carcinoma [64].

The abundant fibrotic stroma surrounding pancreatic tumors could impede the infiltration of effector T cells, greatly reducing the efficacy of immunotherapies [65]. We evaluated TME components in each cluster: C3 had significantly higher stromal and immune scores than those of C2, whereas C1 had the highest immune scores. Previous studies indicated that stromal compartments such as cancer-associated fibroblasts and mast cells in the PDAC microenvironment can greatly promote tumor invasion and drug resistance [66]. Moreover, these effects were mediated through the TGF- β pathways, consistent

with our previous clustering results. Thus, it is suggested that the majority of infiltrating immune cells within the C3 TME are inhibited by tumor stroma and stromal cells, leading to tumor progression, and that the highly active immune response in C1 could reduce tumor growth. The interaction of tumor stroma and AS events suggests a new way to bypass immune barriers in PDAC and directly regulate immune response.

This study had some limitations which will be addressed in further experiments. First, owing to the lack of normal samples for comparison with PDAC, we selected the most significant AS events instead of differentially expressed AS events in the univariate analysis for further model construction. Also, comparison between databases was not feasible because TCGA is the only available source for patient AS information. Second, the accuracy of the infiltration cell levels calculated by the CIBERSORT algorithm was restricted by the fidelity of the reference profile, especially in the heterogenous PDAC TME. To lower estimation bias and increase accuracy, larger sample sizes and additional experiments are required for future validation. Third, the retrospective nature of this study makes it difficult to rule out the impact of some clinical features such as R0 or R1 surgical resection and adjuvant therapies on HR, though multivariate Cox analysis was performed.

Conclusions

We have systematically evaluated the splicing profiles of PDAC patients and established a prognostic model with high accuracy

Supplementary Data

Supplementary Table 1. Multivariate Cox regression analysis for alternative splicing (AS) events in each predictive model.

Type	AS event	Coef	HR	95% lower	95% Up	P-value
AA	MRPL4 47484 AA	-8.02	3.30E-04	2.68E-08	4.04	0.095098
	ITGB1BP1 52617 AA	-1.51	0.22	0.04	1.21	0.081335
	CYB561A3 16164 AA	-2.16	0.12	0.03	0.49	0.003402
	SWI5 87732 AA	2.51	12.27	0.69	216.66	0.087002
	CELF1 15763 AA	-2.94	0.05	3.59E-03	0.78	0.032441
	EML3 16362 AA	-9.34	8.82E-05	4.39E-09	1.77	0.064770
	PLCG1 59390 AA	-6.08	2.29E-03	5.51E-05	0.09	0.001375
	C19orf33 49601 AA	-8.14	2.91E-04	4.23E-07	0.20	0.014598
	SRSF7 53281 AA	5.00	149.15	1.82	1.22E+04	0.026008
	IGSF8 8426 AA	12.85	3.83E+05	91.05	1.61E+09	0.002530
	SLC15A4 93374 AA	-1.59	0.20	0.03	1.51	0.119692

based on AS events to facilitate clinical decision making. We also constructed an SF-AS regulatory network, which could regulate the immune response in PDAC. We stratified PDAC patients based on AS events and found a significant relationship between AS clusters and the immune landscape, identifying a novel direction for immunotherapeutic research. Further studies are needed to explore the regulation of immune responses by AS events and identify possible tumor-specific antigens derived from AS events.

Acknowledgements

We thank TCGA and SpliceSeq for the availability of the data.

Conflict of interest

None.

Abbreviations

PDAC – pancreatic ductal adenocarcinoma; **TME** – tumor microenvironment; **EMT** – epithelial to mesenchymal transition; **TCGA** – The Cancer Genome Atlas; **AS** – alternative splicing; **SF** – splicing factor; **OS** – overall survival; **ES** – exon skipping; **AD** – alternate donor site; **AA** – alternate acceptor site; **AP** – alternate promoter; **AT** – alternate terminator; **ME** – mutually exclusive exons; **RI** – retained intron; **ROC** – receiver operator characteristic; **AUC** – area under curve; **GO** – Gene ontology; **KEGG** – Kyoto Encyclopedia of Genes and Genomes.

Supplementary Table 1 continued. Multivariate Cox regression analysis for alternative splicing (AS) events in each predictive model.

Type	AS event	Coef	HR	95% lower	95% Up	P-value
AD	CDK11A 218 AD	-3.19	0.04	2.25E-03	0.75	0.031458
	SUPT4H1 42666 AD	-14.97	3.16E-07	4.77E-11	2.09E-03	0.000853
	NFATC4 26993 AD	-4.08	0.02	9.09E-04	0.31	0.006114
	CLASRP 50392 AD	-2.22	0.11	9.80E-03	1.21	0.070738
	RAB15 27923 AD	-1.26	0.28	0.08	1.06	0.061960
	PTHLH 20910 AD	1.95	7.01	1.79	27.35	0.005089
	NDUFAF2 72172 AD	-6.18	2.08E-03	2.30E-05	0.19	0.007177
	CCDC58 66432 AD	-13.82	9.95E-07	5.96E-10	1.66E-03	0.000262
AP	S100A13 7733 AP	-5.05	6.40E-03	1.66E-04	0.25	0.006691
	UBA1 88907 AP	4.78	118.67	1.01	1.39E+04	0.049376
	COPS7A 19933 AP	2.10	8.16	1.16	57.64	0.035306
	UBE2L6 15862 AP	13.94	1.13E+06	1.59	8.09E+11	0.042642
	SH3KBP1 88641 AP	-1.36	0.26	0.06	1.20	0.083722
	SLC7A8 26712 AP	2.17	8.73	1.37	55.77	0.022006
	EIF5A 38908 AP	3.60	36.48	0.94	1414.49	0.053957
AT	VRK3 51144 AT	2.09	8.10	0.51	128.40	0.137897
	RHOF 24898 AT	5.72	305.26	3.21	2.90E+04	0.013794
	CDCA3 19984 AT	-4.26	0.01	4.28E-04	0.47	0.016996
	DOCK10 57781 AT	-8.96	1.28E-04	6.29E-07	0.03	0.000949
	ZNF506 48683 AT	-3.63	0.03	1.49E-03	0.47	0.013505
ES	CTNND1 15980 ES	-1.16	0.31	0.09	1.11	0.072179
	CTNND1 15958 ES	-2.19	0.11	0.02	0.71	0.020238
	CTAGE5 27377 ES	-5.08	6.20E-03	2.13E-04	0.18	0.003117
	UAP1 8750 ES	6.59	725.58	10.87	4.84E+04	0.002116
	NCAPH 54528 ES	-11.51	1.00E-05	2.19E-08	4.58E-03	0.000230
	GSE1 37883 ES	-1.46	0.23	0.06	0.86	0.028096
	CASP8 56821 ES	-6.82	1.09E-03	1.10E-05	0.11	0.003605
ME	RAD51 30020 ME	-3.24	0.04	4.58E-04	3.33	0.152682
	RBM39 59248 ME	-3.12	0.04	0.01	0.25	0.000410
	ICA1 97384 ME	-5.22	5.40E-03	1.80E-04	0.16	0.002600
	TCF7L2 151705 ME	-4.03	0.02	1.31E-04	2.43	0.108170
	AMT 64866 ME	7.26	1428.93	1.19	1.71E+06	0.044545
	CERS5 21668 ME	2.08	8.01	1.23	52.28	0.029715
	DNM2 47588 ME	4.75	115.11	4.19	3162.60	0.004994

Supplementary Table 1 continued. Multivariate Cox regression analysis for alternative splicing (AS) events in each predictive model.

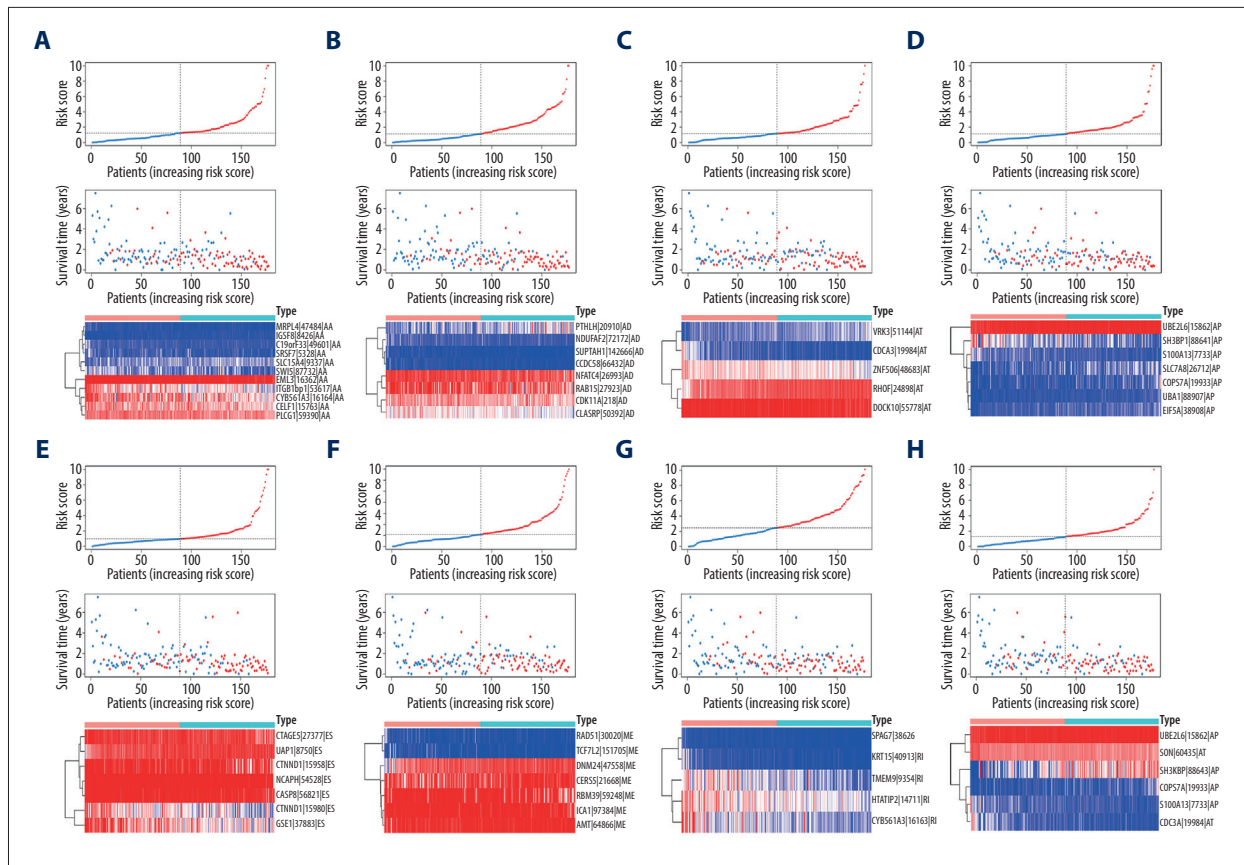
Type	AS event	Coef	HR	95% lower	95% Up	P-value
RI	SPAG7 38626 RI	-12.53	3.63E-06	4.86E-11	0.27	0.028648
	CYB561A3 16163 RI	-4.65	9.58E-03	7.33E-05	1.25	0.061473
	TMEM9 9354 RI	-1.32	0.27	0.08	0.92	0.036773
	HTATIP2 14711 RI	-1.93	0.14	0.02	0.84	0.031768
	KRT15 40913 RI	-1.52	0.22	0.05	0.90	0.035230
ALL	S100A13 7733 AP	-4.68	9.29E-03	2.66E-04	0.32	0.009853
	SH3KBP1 88643 AP	1.05	2.87	0.91	9.07	0.073418
	CDCA3 19984 AT	-2.95	0.05	1.82E-03	1.52	0.086090
	COPS7A 19933 AP	2.42	11.20	1.61	77.90	0.014633
	UBE2L6 15862 AP	15.43	5.01E+06	2.16	1.16E+13	0.039109
	SON 60435 AT	-6.49	1.52E-03	5.80E-06	0.40	0.022398

AA – alternate acceptor site; AD – alternate donor site; AP – alternate promoter; AT – alternate terminator; ES – exon skipping; HR – hazard ratio; ME – mutually exclusive exons; RI – retained intron.

Supplementary Table 2. Clinical features of PDAC patients in TCGA database.

Characteristics	Patients, n (%)	Characteristics	Patients, n (%)
Age		Pathologic T	
>65	83 (47.9)	T1	7 (4.0)
≤65	94 (53.1)	T2	24 (13.5)
Sex		T3	141 (79.7)
Male	97 (54.8)	T4	3 (1.7)
Female	80 (45.2)	Unknow	2 (1.1)
Grade		Pathologic N	
G1	30 (16.9)	N0	49 (27.7)
G2	95 (53.7)	N1	123 (69.5)
G3	48 (27.2)	Unknow	5 (2.8)
G4	2 (1.1)	Median follow-up, OS (days)	440 (4–2,741)
GX	2 (1.1)	OS event (0/1)	88/89
Pathologic stage			
I	21 (11.9)		
II	146 (82.4)		
III	3 (1.7)		
IV	4 (2.3)		
Unknow	3 (1.7)		

OS – overall survival.



Supplementary Figure 1. Risk score calculation based on overall survival (OS)-related alternative splicing (AS) events. Risk score curves showing that pancreatic ductal carcinoma patients were divided into high- and low-risk groups based on the median risk score, as shown at the top of each assembly. Survival status and OS of each patient are shown in the middle, and the percent spliced in (PSI) heatmaps at the bottom. The 8 assemblies of risk scores were based on (A) alternate acceptor site, (B) alternate donor site, (C) alternate terminator, (D) alternate promoter site, (E) exon skipping, (F) mutually exclusive exons, (G) retained intron, and (H) all patterns of OS-related AS events.

References:

- Rahib L, Smith BD, Aizenberg R et al: Projecting cancer incidence and deaths to 2030: The unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res*, 2014; 74(11): 2913–21
- Serenari M, Ercolani G, Cucchetti A et al: The impact of extent of pancreatic and venous resection on survival for patients with pancreatic cancer. *Hepatobiliary Pancreat Dis Int*, 2019; 18(4): 389–94
- Vincent A, Herman J, Schulick R et al: Pancreatic cancer. *Lancet*, 2011; 378(9791): 607–20
- Kleeff J, Korc M, Apte M et al: Pancreatic cancer. *Nat Rev Dis Primers*, 2016; 2: 16022
- Ren B, Cui M, Yang G et al: Tumor microenvironment participates in metastasis of pancreatic cancer. *Mol Cancer*, 2018; 17(1): 108
- Ligorio M, Sil S, Malagon-Lopez J et al: Stromal microenvironment shapes the intratumoral architecture of pancreatic cancer. *Cell*, 2019; 178(1): 160–75.e27
- Collisson EA, Bailey P, Chang DK, Biankin AV: Molecular subtypes of pancreatic cancer. *Nat Rev Gastroenterol Hepatol*, 2019; 16(4): 207–20
- Argentiero A, De Summa S, Di Fonte R et al: Gene expression comparison between the lymph node-positive and -negative reveals a peculiar immune microenvironment signature and a therapeutic role for WNT targeting in pancreatic ductal adenocarcinoma: A pilot study. *Cancers (Basel)*, 2019; 11(7): 942
- Wang ET, Sandberg R, Luo S et al: Alternative isoform regulation in human tissue transcriptomes. *Nature*, 2008; 456(7221): 470–76
- Matlin AJ, Clark F, Smith CW: Understanding alternative splicing: Towards a cellular code. *Nat Rev Mol Cell Biol*, 2005; 6(5): 386–98
- Kozlovski I, Siegfried Z, Amar-Schwartz A, Karni R: The role of RNA alternative splicing in regulating cancer metabolism. *Hum Genet*, 2017; 136(9): 1113–27
- Baralle FE, Giudice J: Alternative splicing as a regulator of development and tissue identity. *Nat Rev Mol Cell Biol*, 2017; 18(7): 437–51
- Sveen A, Kilpinen S, Ruusulehto A et al: Aberrant RNA splicing in cancer; Expression changes and driver mutations of splicing factor genes. *Oncogene*, 2016; 35(19): 2413–27
- Sakuma K, Sasaki E, Kimura K et al: HNRNPLL, a newly identified colorectal cancer metastasis suppressor, modulates alternative splicing of CD44 during epithelial-mesenchymal transition. *Gut*, 2018; 67(6): 1103–11
- Mavrou A, Oltean S: SRPK1 inhibition in prostate cancer: A novel anti-angiogenic treatment through modulation of VEGF alternative splicing. *Pharmacol Res*, 2016; 107: 276–81
- Salton M, Voss TC, Misteli T: Identification by high-throughput imaging of the histone methyltransferase EHMT2 as an epigenetic regulator of VEGFA alternative splicing. *Nucleic Acids Res*, 2014; 42(22): 13662–73

17. Schittenhelm MM, Walter B, Tsintari V et al: Alternative splicing of the tumor suppressor ASP2 results in a stress-inducible, oncogenic isoform prevalent in acute leukemia. *EBioMedicine*, 2019; 42: 340–51
18. Ma S, Chen C, Ji X et al: The interplay between m6A RNA methylation and noncoding RNA in cancer. *J Hematol Oncol*, 2019; 12(1): 121
19. Li Y, Sun N, Lu Z et al: Prognostic alternative mRNA splicing signature in non-small cell lung cancer. *Cancer Lett*, 2017; 393: 40–51
20. Hou J, Zhang H, Liu J et al: YTHDF2 reduction fuels inflammation and vascular abnormalization in hepatocellular carcinoma. *Mol Cancer*, 2019; 18(1): 163
21. Muller S, Glass M, Singh AK et al: IGF2BP1 promotes SRF-dependent transcription in cancer in a m6A- and miRNA-dependent manner. *Nucleic Acids Res*, 2019; 47(1): 375–90
22. Kahles A, Lehmann KV, Toussaint NC et al: Comprehensive analysis of alternative splicing across tumors from 8,705 patients. *Cancer Cell*, 2018; 34(2): 211–24.e6
23. Jayasinghe RG, Cao S, Gao Q et al: Systematic analysis of splice-site-creating mutations in cancer. *Cell Rep*, 2018; 23(1): 270–81.e3
24. Cohen-Eliav M, Golan-Gerstl R, Siegfried Z et al: The splicing factor SRSF6 is amplified and is an oncoprotein in lung and colon cancers. *J Pathol*, 2013; 229(4): 630–39
25. Karni R, de Stanchina E, Lowe SW et al: The gene encoding the splicing factor SF2/ASF is a proto-oncogene. *Nat Struct Mol Biol*, 2007; 14(3): 185–93
26. David CJ, Chen M, Assanah M et al: HnRNP proteins controlled by c-Myc de-regulate pyruvate kinase mRNA splicing in cancer. *Nature*, 2010; 463(7279): 364–68
27. Schafer S, Miao K, Benson CC et al: Alternative splicing signatures in RNA-seq data: Percent spliced in (PSI). *Curr Protoc Hum Genet*, 2015; 87: 1161–64
28. Goeman JJ: L1 penalized estimation in the Cox proportional hazards model. *Biom J*, 2010; 52(1): 70–84
29. Piva F, Giulietti M, Burini AB, Principato G: SpliceAid 2: A database of human splicing factors expression data and RNA target motifs. *Hum Mutat*, 2012; 33(1): 81–85
30. Newman AM, Liu CL, Green MR et al: Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*, 2015; 12(5): 453–57
31. Xu T, Le TD, Liu L et al: CancerSubtypes: An R/Bioconductor package for molecular cancer subtype identification, validation and visualization. *Bioinformatics*, 2017; 33(19): 3131–33
32. Yoshihara K, Shahmoradgoli M, Martinez E et al: Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*, 2013; 4: 2612
33. Thorsson V, Gibbs DL, Brown SD et al: The immune landscape of cancer. *Immunity*, 2018; 48(4): 812–30.e14
34. Seiler M, Peng S, Agrawal AA et al: Somatic mutational landscape of splicing factor genes and their functional consequences across 33 cancer types. *Cell Rep*, 2018; 23(1): 282–96.e4
35. Frankiw L, Baltimore D, Li G: Alternative mRNA splicing in cancer immunotherapy. *Nat Rev Immunol*, 2019; 19(11): 675–87
36. Gansauge F, Gansauge S, Rau B et al: Low serum levels of soluble CD44 variant 6 are significantly associated with poor prognosis in patients with pancreatic carcinoma. *Cancer*, 1997; 80(9): 1733–39
37. Klingbeil P, Marhaba R, Jung T et al: CD44 variant isoforms promote metastasis formation by a tumor cell-matrix cross-talk that supports adhesion and apoptosis resistance. *Mol Cancer Res*, 2009; 7(2): 168–79
38. Abetamann V, Kern HF, Elsasser HP: Differential expression of the hyaluronan receptors CD44 and RHAMM in human pancreatic cancer cells. *Clin Cancer Res*, 1996; 2(9): 1607–18
39. Chaudhry A, Gobl A, Eriksson B et al: Different splice variants of CD44 are expressed in gastrinomas but not in other subtypes of endocrine pancreatic tumors. *Cancer Res*, 1994; 54(4): 981–86
40. Hong JT, Son DJ, Lee CK et al: Interleukin 32, inflammation and cancer. *Pharmacol Ther*, 2017; 174: 127–37
41. Chen J, Wang S, Su J et al: Interleukin-32alpha inactivates JAK2/STAT3 signaling and reverses interleukin-6-induced epithelial-mesenchymal transition, invasion, and metastasis in pancreatic cancer cells. *Onco Targets Ther*, 2016; 9: 4225–37
42. Cicenas J, Kvederaviciute K, Meskinyte I et al: KRAS, TP53, CDKN2A, SMAD4, BRCA1, and BRCA2 mutations in pancreatic cancer. *Cancers (Basel)*, 2017; 9(5): 42
43. Lee JH, Giovannetti E, Hwang JH et al: Loss of 18q22.3 involving the carboxypeptidase of glutamate-like gene is associated with poor prognosis in resected pancreatic cancer. *Clin Cancer Res*, 2012; 18(2): 524–33
44. Chong CR, Janne PA: The quest to overcome resistance to EGFR-targeted therapies in cancer. *Nat Med*, 2013; 19(11): 1389–400
45. Wong BS, Shea DJ, Mistriotis P et al: A direct podocalyxin-dynamin-2 interaction regulates cytoskeletal dynamics to promote migration and metastasis in pancreatic cancer cells. *Cancer Res*, 2019; 79(11): 2878–91
46. Wu B, Wang K, Fei J et al: Novel three lncRNA signature predicts survival in patients with pancreatic cancer. *Oncol Rep*, 2018; 40(6): 3427–37
47. Bechara EG, Sebestyen E, Bernardis I et al: RBM5, 6, and 10 differentially regulate NUMB alternative splicing to control cancer cell proliferation. *Mol Cell*, 2013; 52(5): 720–33
48. Bellavia D, Mearozzi M, Campese AF et al: Notch3 and the Notch3-upregulated RNA-binding protein HuD regulate Ikaros alternative splicing. *EMBO J*, 2007; 26(6): 1670–80
49. Kedzierska H, Piekliko-Witkowska A: Splicing factors of SR and hnRNP families as regulators of apoptosis in cancer. *Cancer Lett*, 2017; 396: 53–65
50. Huang H, Zhang J, Harvey SE et al: RNA G-quadruplex secondary structure promotes alternative splicing via the RNA-binding protein hnRNP. *Genes Dev*, 2017; 31(22): 2296–309
51. Preca BT, Bajdak K, Mock K et al: A self-enforcing CD44s/ZEB1 feedback loop maintains EMT and stemness properties in cancer cells. *Int J Cancer*, 2015; 137(11): 2566–77
52. Adesso L, Calabretta S, Barbagallo F et al: Gemcitabine triggers a pro-survival response in pancreatic cancer cells through activation of the MNK2/elf4E pathway. *Oncogene*, 2013; 32(23): 2848–57
53. Wu F, Chen Q, Liu C et al: Profiles of prognostic alternative splicing signature in hepatocellular carcinoma. *Cancer Med*, 2020; 9(6): 2171–80
54. Chen T, Zheng W, Chen J et al: Systematic analysis of survival-associated alternative splicing signatures in clear cell renal cell carcinoma. *J Cell Biochem*, 2019 [Online ahead of print]
55. Smart AC, Margolis CA, Pimentel H et al: Intron retention is a source of neo-epitopes in cancer. *Nat Biotechnol*, 2018; 36(11): 1056–58
56. Manley GT, Smitt PS, Dalmau J, Posner JB: Hu antigens: Reactivity with Hu antibodies, tumor expression, and major immunogenic sites. *Ann Neurol*, 1995; 38(1): 102–10
57. Ohwada A, Nagaoka I, Takahashi F et al: DNA vaccination against HuD antigen elicits antitumor activity in a small-cell lung cancer murine model. *Am J Respir Cell Mol Biol*, 1999; 21(1): 37–43
58. Roberts WK, Deluca IJ, Thomas A et al: Patients with lung cancer and paraneoplastic Hu syndrome harbor HuD-specific type 2 CD8+ T cells. *J Clin Invest*, 2009; 119(7): 2042–51
59. Calabro A, Beissbarth T, Kuner R et al: Effects of infiltrating lymphocytes and estrogen receptor on gene expression and prognosis in breast cancer. *Breast Cancer Res Treat*, 2009; 116(1): 69–77
60. Spira A, Disis ML, Schiller JT et al: Leveraging premalignant biology for immune-based cancer prevention. *Proc Natl Acad Sci USA*, 2016; 113(39): 10750–58
61. Wolf DM, Lenburg ME, Yau C et al: Gene co-expression modules as clinically relevant hallmarks of breast cancer diversity. *PLoS One*, 2014; 9(2): e88309
62. Chang HY, Sneddon JB, Alizadeh AA et al: Gene expression signature of fibroblast serum response predicts human cancer progression: Similarities between tumors and wounds. *PLoS Biol*, 2004; 2(2): E7
63. Teschendorff AE, Gomez S, Arenas A et al: Improved prognostic classification of breast cancer defined by antagonistic activation patterns of immune response pathway modules. *BMC Cancer*, 2010; 10: 604
64. Li ZX, Zheng ZQ, Wei ZH et al: Comprehensive characterization of the alternative splicing landscape in head and neck squamous cell carcinoma reveals novel events associated with tumorigenesis and the immune microenvironment. *Theranostics*, 2019; 9(25): 7648–65
65. Kleeff J, Beckhove P, Esposito I et al: Pancreatic cancer microenvironment. *Int J Cancer*, 2007; 121(4): 699–705
66. Porcelli L, Iacobazzi RM, Di Fonte R et al: CAFs and TGF-beta signaling activation by mast cells contribute to resistance to gemcitabine/nabpaclitaxel in pancreatic cancer. *Cancers (Basel)*, 2019; 11(3): 330