MEDICAL SCIENCE MONITOR

DATABASE ANALYSIS

e-ISSN 1643-3750
© Med Sci Monit, 2020; 26: e924126
DOI: 10.12659/MSM.924126

Received: 2020.03.09
Accepted: 2020.04.28
Available online: 2020.05.22
Published: 2020.07.19

Systematic Profiling of Alternative Splicing for Sarcoma Patients Reveals Novel Prognostic Biomarkers Associated with Tumor Microenvironment and Immune Cells

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

ABCEF 1 Chuan Hu
ABC 1 Yuanhe Wang
DEF 2 Chuan Liu
DE 1 Rui Shen
F 3 Bo Chen
B 1 Kang Sun
C 4 Huili Rao
E 3 Lin Ye
C 3 Jianjun Ye
ABCDEF 1 Shaoqi Tian

1 Department of Joint Surgery, The Affiliated Hospital of Qingdao University, Qingdao, Shandong, P.R. China
2 Department of Medical Oncology, The First Hospital of China Medical University, Shenyang, Liaoning, P.R. China
3 Wenzhou Medical University, Wenzhou, Zhejiang, P.R. China
4 Sir Run Run Shaw Hospital, Zhejiang University, School of Medicine, Hangzhou, Zhejiang, P.R. China

Corresponding Author: Shaoqi Tian, e-mail: shaoqi99@aliyun.com
Source of support: Departmnetal sources

Background: Alternative splicing (AS) events is a novel biomarker of tumor prognosis, but the role of AS events in sarcoma patients remains unclear.

Material/Methods: RNA-seq and clinicopathologic data of the sarcoma cohort were extracted from the TCGA database and data on AS events were downloaded from the TCGASpliceSeq database. Univariate Cox analysis, LASSO regression analysis, and multivariate Cox analysis were performed to determine the overall survival (OS)- and disease-free survival (DFS)-related AS events. Two nomograms were developed based on the independent variables, and subgroup analysis was performed. The area under the curve (AUC), calibration curve, and decision curve analysis (DCA) were used to evaluate the nomograms. Then, we used the CIBERSORT and ESTIMATE package to determine the immune cell proportion and tumor microenvironment (TME) score, respectively. The associations between AS events-based clusters and TME and immune cells were studied.

Results: We identified 1945 and 1831 AS events as OS- and DFS-related AS events, respectively. Two nomograms based on the AS events and clinical data were established and the AUCs of nomograms ranged from 0.807 to 0.894. The calibration curve and DCA showed excellent performance of nomograms. In addition, the results indicated the distinct relationships between AS events-based clusters and OS, DFS, immune score, stromal score, and 10 immune cells.

Conclusions: Our study indicated that AS events are novel prognostic biomarkers for sarcoma patients that may be associated with the TME and immune cells.

MeSH Keywords: Alternative Splicing • Immune System • Nomograms • Prognosis • Sarcoma • Tumor Microenvironment

Full-text PDF: https://www.medscimonit.com/abstract/index/idArt/924126

3450 4 11 51

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

Background

Sarcoma is a heterogeneous group of malignancies originating from mesenchymal tissues, which has significant histological diversity. Among the many histological types of sarcoma, osteosarcoma, leiomyoma, lymphosarcoma, and synovial sarcoma are common histological types [1]. Surgery is the first-line treatment for sarcoma patients, and some patients receive radiotherapy if needed [2]. However, metastasis can occur even in early-stage sarcoma patients, and more than 50% of patients have high risk of metastasis or death [1]. Hence, it is essential to find contributing markers for diagnosis, risk stratification, and prediction of prognosis for sarcoma patients. Intensive efforts have been made to determine the prognostic biomarkers of sarcoma patients in previous studies, such as clinicopathologic variables, lncRNA, gene-signature, miRNA, and plasmacytoma variant translocation 1 [3-7]. However, although previous research has contributed sarcoma research, mostly studies have focused on the transcriptional level, while genome-wide profiling of splicing variant is lacking.

Alternative splicing (AS) event is one of the most important steps in post-transcription regulation for pre-RNA. About 92–94% of genes in humans are modified by AS event [8]. With the limited number of genes, AS event is very important for biological protein diversity by regulating the production of different mRNA subtypes [9–12]. Moreover, the close relationship between AS event and malignancy has been studied recently for the first time [13,14]. AS event is considered to be related with tumorigenesis, progression, metastasis, and drug resistance of tumors [15–18]. More importantly, with the development of high-throughput sequencing, many studies have focused on comprehensive genome-wide profiling of AS events in cancer and confirmed that AS events can be used as robust prognostic biomarkers. For example, Zhang et al. [19] studied 330 oral squamous cell carcinoma patients, and a AS event-based signature was established and validated. In addition, the AS-splicing factor (SF) regulatory network was also established. Notably, similar studies were also performed to study breast cancer [20], kidney renal clear cell carcinoma [21], colorectal cancer [22], and esophageal carcinoma [23]. However, our literature review found few studies focused on the association between sarcoma and AS events. Furthermore, The Cancer Genome Atlas (TCGA) data portal provides the AS events data of sarcoma patients, and the corresponding clinical data were also provided in the dataset, which facilitated the present study.

Therefore, in our study, comprehensive bioinformatics and statistical methods were used to identify the prognosis-associated AS events in the sarcoma cohort based on the TCGA data portal. We developed 2 nomograms based on the AS events and clinicopathologic data to predict the overall survival (OS) and disease-free survival (DFS) in sarcoma patients. To the best of our knowledge, this is the first published study exploring the potential relationship between AS events and immune features.

Material and Methods

Data acquisition and preprocessing

The RNA-seq data (Level 3) and clinicopathologic data of the sarcoma cohort were extracted from the TCGA data portal (*https://tcga-data.nci.nih.gov/tcga/*). We included patients with complete data and with OS more than 30 days in the TCGA-SARC. In addition, we used the percent spliced in (PSI) value to quantify the AS event, and the value of PSI in TCGA-SARC was downloaded from the TCGASpliceSeq [24]. Samples with less than 25% of lacking PSI values were included in the present study. To show the intersections between the 7 types of AS events – Exon Skip (ES), Mutually Exclusive Exons (ME), Retained Intron (RI), Alternate Promoter (AP), Alternate Terminator (AT), Alternate Donor site (AD), and Alternate Acceptor site (AA) – we generated UpSet plots and bar plots [25].

Identification of OS- and DFS-related AS events

Based on the inclusion criteria, 195 patients diagnosed as having sarcoma were included in our cohort. We performed the univariate Cox analysis to determine the prognostic AS events, including OS-related AS events and DFS-related AS events. The AS events with a p-value <0.05 in the univariate Cox analysis were considered as the prognostic AS events. A circular dendrogram was generated to show the most significant prognostic AS events (top 20, if possible).

Construction of a prognostic AS events signature

On the basis of the prognostic AS events, the least absolute shrinkage and selection operator (LASSO) regression analysis was performed to avoid overfitting [26]. Then, multivariate Cox analysis was performed and independent prognostic AS events were determined. Subsequently, based on the independent AS events, a prognostic model was constructed using multivariate Cox analysis, and the risk scores of each patient were calculated. The optimal cutoff value of risk score was identified using X-tile software, and 195 patients were stratified into the low-, middle-, and high-risk groups [27].

Identification of independent predictors and construction of a nomogram

The clinicopathologic data used were age, sex, neoplasm histologic type, disease multifocal indicator, surgical margin resection status, and tumor site. Firstly, we used the X-tile software to determine the optimal cutoff value of age in predicting

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

OS and DFS [27]. Then, univariate Cox analysis was performed and the prognostic clinicopathologic data were determined (p-value <0.05). The risk classification based on AS events and prognostic clinicopathologic data was incorporated into the multivariate Cox analysis, and independent predictors of prognosis in sarcoma patients were identified.

Based on the independent predictors, 2 nomograms were developed to predict the OS and DFS in sarcoma patients. The time-dependent receiver operating characteristic (ROC) curves with the area under the curve (AUC) of 3-, 5-, and 7-year survival were generated to quantify the discrimination of the prognostic nomogram [28], and the calibration curves of 3-, 5-, and 7-year survival were used to calibrate the nomogram. The net benefit of the nomogram was determined by decision curve analysis (DCA) [29].

Subgroup analysis of the nomograms

To further assess if the nomograms can serve as an effective tool in the various histological types of sarcoma, subgroup analysis was performed for 4 major histological types. First, the risk scores of each patient were calculated based on the nomogram. Then, the time-dependent ROC curves were plotted and the AUCs were calculated. According to the cutoff value (median) in each subgroup, we stratified patients into a high-risk group and a low-risk group. The Kaplan-Meier curves of each histological type patient were generated and the log-rank test was used to determine if the prognostic nomogram can effectively distinguish among patients in various risk stratifications.

Splicing correlation network construction

It has been reported that SF plays a vital role in regulating AS events [30,31]. Therefore, to clarify the potential AS-SF regulatory network, a correlation network between SFs and prognostic AS events was constructed. First, data on the expression of 71 SF genes were downloaded from TCGA. Then, the correlation between SFs and AS event was determined by Spearman correlation analysis. A correlation with $p<0.05$ and $r>0.6$ was considered to be statistically significant. Cytoscape (version 3.7.2) was used to visualize the correlation network.

Functional enrichment

In our study, we used Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways to explore the prognostic utility of AS events. Because the overlapping AS events play an essential role in cancer progression, corresponding genes of overlapping AS events between OS- and DFS-related AS events were selected to perform the functional enrichment. Terms with a p-value <0.05, a minimum count of 3, and an enrichment factor >1.5 were collected and grouped into clusters based on their membership similarities. The enrichment analyses were performed in Metascape (*http://metascape.org*).

Evaluation of correlation between tumor microenvironment score and immune cell proportion

To better understand the correlation between different molecular subtypes and immune features, the unsupervised consensus approach implemented with the Consensus Cluster Plus package was used to classify the sarcoma cohort based on the overlapping prognostic AS events. Because the overlapping AS events are the most highly conserved, we suggested they are the most likely to be associated with prognosis. The CIBERSORT package was used to determine the proportions of 22 immune cell types, and only patients with CIBERSORT $P<0.05$ were considered eligible for further analysis [32]. In addition, the Estimation of STromal and Immune cells in Malignant Tumors using Expression data (ESTIMATE) algorithm, a method that uses gene expression signatures to infer the fraction of stromal and immune cells in tumor samples, was performed in R software to identify the immune score, stromal score, and estimate score by using the estimate package [33]. Then, one-way ANOVA was performed to compare the difference of immune cells proportion and tumor microenvironment score between 3 clusters, and a heat map was generated to show the differences in features among the 3 clusters.

Results

Overview of AS events in sarcoma cohort

One hundred ninety-five sarcoma patients who met the criterion were included in our research. Overall, 40 182 AS events and 18 996 corresponding genes were detected in the sarcoma cohort, including 15 311 Exon Skip (ES), 8287 AT, 7837 AP, 3196 AA, 2815 AD, 2572 RI, and 164 ME (Figure 1A, 1B). In Figure 1A, most of the AS events were from 1 gene, which could have several types of AS events. For example, 1 gene might contain up to 6 AS types, such as ES, AP, AT, AA, AD, and RI (Figure 1A).

Identification of prognostic AS events

To identify the OS-AS events and DFS-AS events in sarcoma patients, univariate Cox analysis was performed. A total of 1945 AS events with 1404 parent genes and 1831 AS events with 1351 genes were identified as OS- and DFS-related AS events, respectively. Two UpSet plots were created to visualize the interaction among all types of AS events, and 2 bar plots were generated to show the number of AS events and corresponding genes (Figure 1C–1F). Meanwhile, 763 overlapping AS events

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

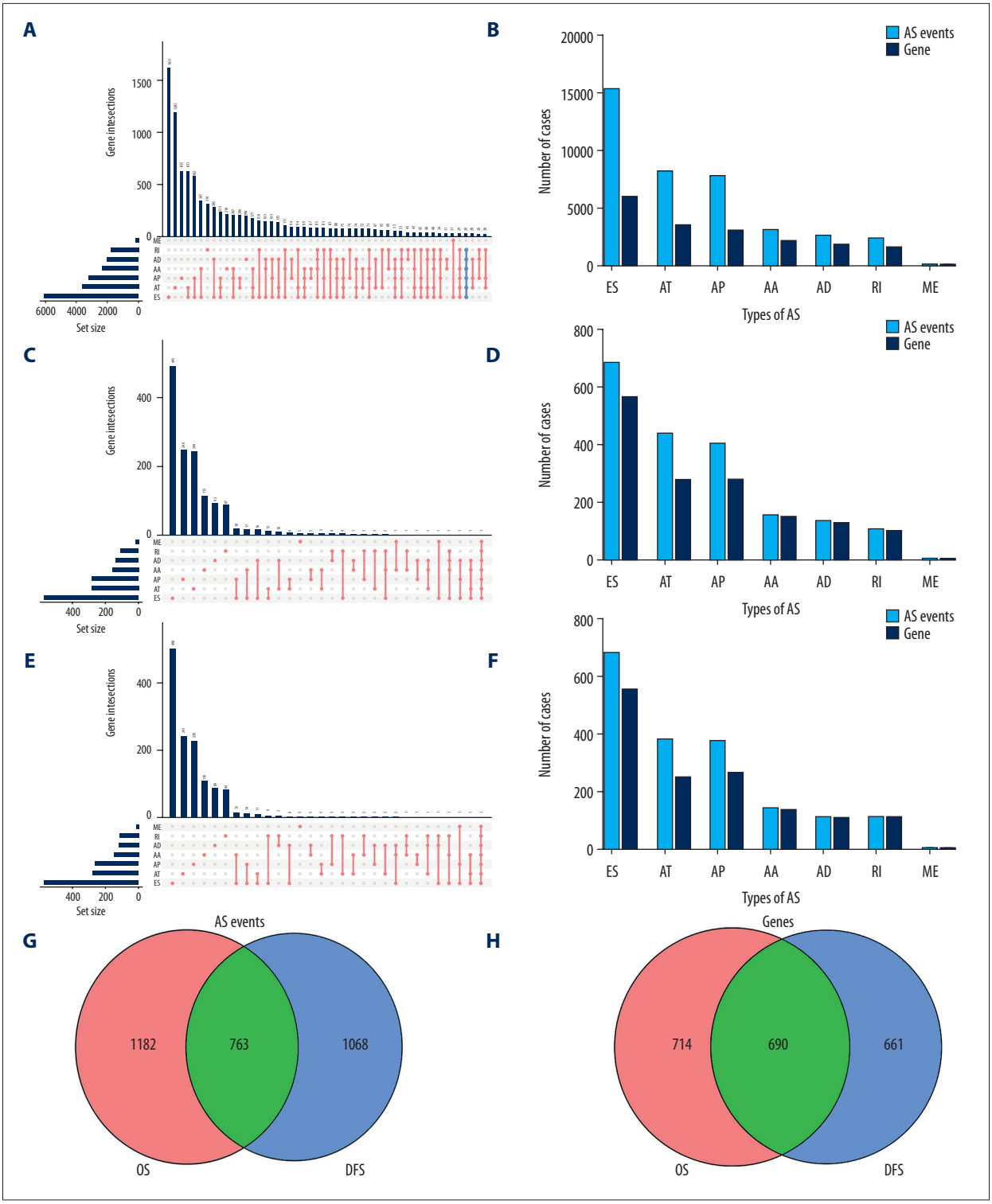


Figure 1. Overview of AS events and prognostic AS events in SARC. (**A**) UpSet plot of interactions between 7 types of AS events in SARC. (**B**) Number of AS events and related genes in SARC. (**C**) UpSet plot of interactions between 7 types of OS-AS events in SARC. (**D**)The number of OS-AS events and related genes in SARC. (**E**) UpSet plot of interactions between 7 types of DFS-AS events in SARC. (**F**) Number of DFS-AS events and related genes in SARC. (**G, H**) Venn diagram of OS-AS events, DFS-AS events, overlapping AS-events, and overlapping genes. AS – alternative splicing; SARC – sarcoma; OS – overall survival; DFS – disease-free survival.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

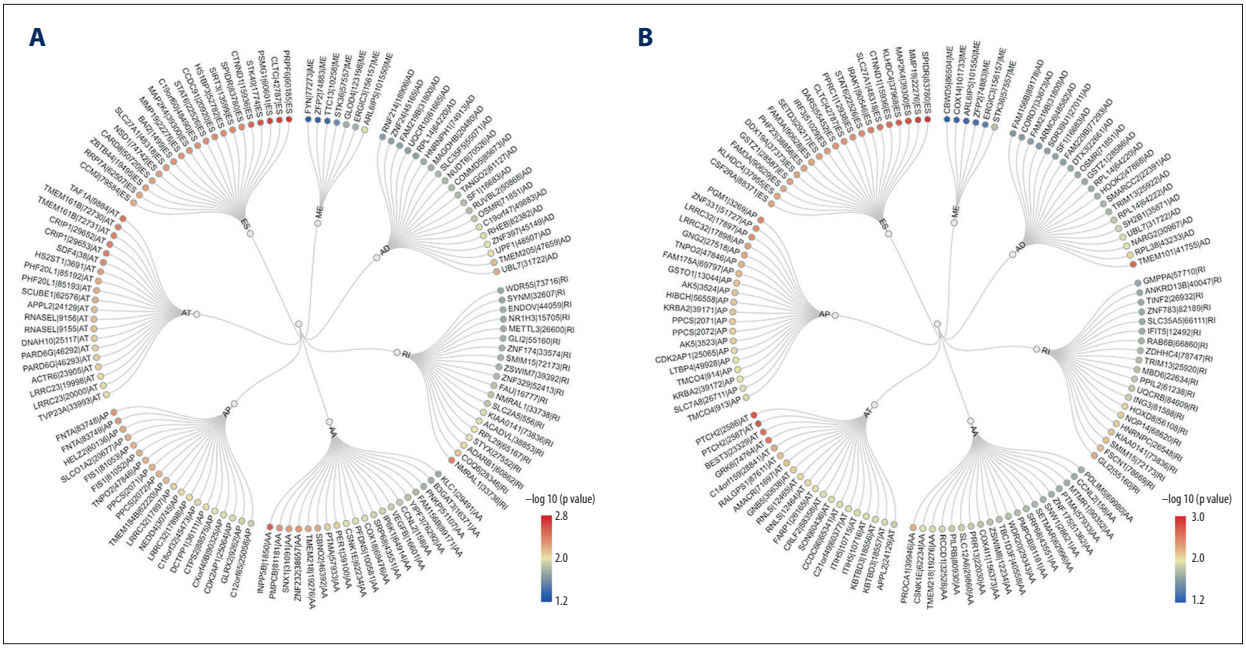


Figure 2. Prognostic AS events in the SARC cohort. (**A**) The top 20 significant OS-AS events for 7 splicing types (except ME). (**B**) The top 20 significant DFS-AS events for 7 splicing types (except ME). SARC – sarcoma; OS – overall survival; ME – mutually exclusive exon; DFS – disease-free survival.

with 690 genes between OS- and DFS-related AS events were also identified and visualized (Figure 1G, 1H). Furthermore, 2 circular dendrograms were developed to depict the top 20 (except ME) significant OS-related AS events (Figure 2A) and DFS-related AS events among 7 types (Figure 2B). The results showed that the distribution of 7 types of OS-related AS events (Figure 1D) was similar to the overall AS events (Figure 1B), and the distribution of 7 types of DFS-related AS events (Figure 1F) was also similar to the overall AS events (Figure 1B).

Construction of an AS events prognostic model for the sarcoma cohort

Based on the OS- and DFS-related AS events, the LASSO analysis was used to avoid overfitting, and 17 AS events and 11 AS events were determined as significant prognostic AS events for OS and DFS, respectively (Figure 3A–3D). Finally, 9 AS events were determined as the independent OS-related AS events and 8 for DFS-related AS events (Tables 1, 2). The risk scores for each patient were calculated by multivariate Cox analysis based on the independent predictors, and all patients were stratified into low-, middle-, and high-risk groups by use of X-tile software (Supplementary Figure 1). The results of the survival analysis indicated that the risk score was significantly associated with the prognosis of sarcoma patients, including OS and DFS. Supplementary Figure 2 contains the survival status distributions and heat map showing the PSI of independent AS events of sarcoma patients, which also showed that the AS event is an ideal predictor for both survival and recurrence in sarcoma patients.

Development of nomograms based on the AS events and clinicopathologic data

Nomograms are a novel type of predictive model and are convenient for use in clinical practice. Hence, a nomogram based on AS events and clinicopathologic data was developed in the present study. First, the cutoffs of age were determined using X-tile software (Supplementary Figure 1). Then, univariate Cox analysis showed that the risk classification based on the AS events, age, disease multifocal indicator, and surgical margin resection status were identified as OS-related predictors, and the same prognostic predictors were also identified in DFS-related predictors (Table 3). Subsequently, the risk classification based on AS events was identified as independent OS- and DFS-related predictors, and the disease multifocal indicator was confirmed as an independent predictor of OS, while surgical margin resection status was confirmed as an independent predictor of DFS (Table 4). Then, based on the independent prognostic predictors, 2 nomograms were established (Figures 4A, 5A). The ROC curves of 3-, 5-, and 7-year survival of the nomograms in predicting the OS and DFS in sarcoma patients were generated and the time-AUC was calculated (Figures 4B, 5B). The AUC of the nomogram ranged from 0.807 to 0.894, which were higher than the single predictors in all situations. The calibration curve of 3-, 5-, and 7-year survival showed good agreement between predictive and actual outcome (Figures 4C, 5C). DCA analysis was also performed in the present study, and the results showed great clinical utility in all situations (Figures 4D, 5D).

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

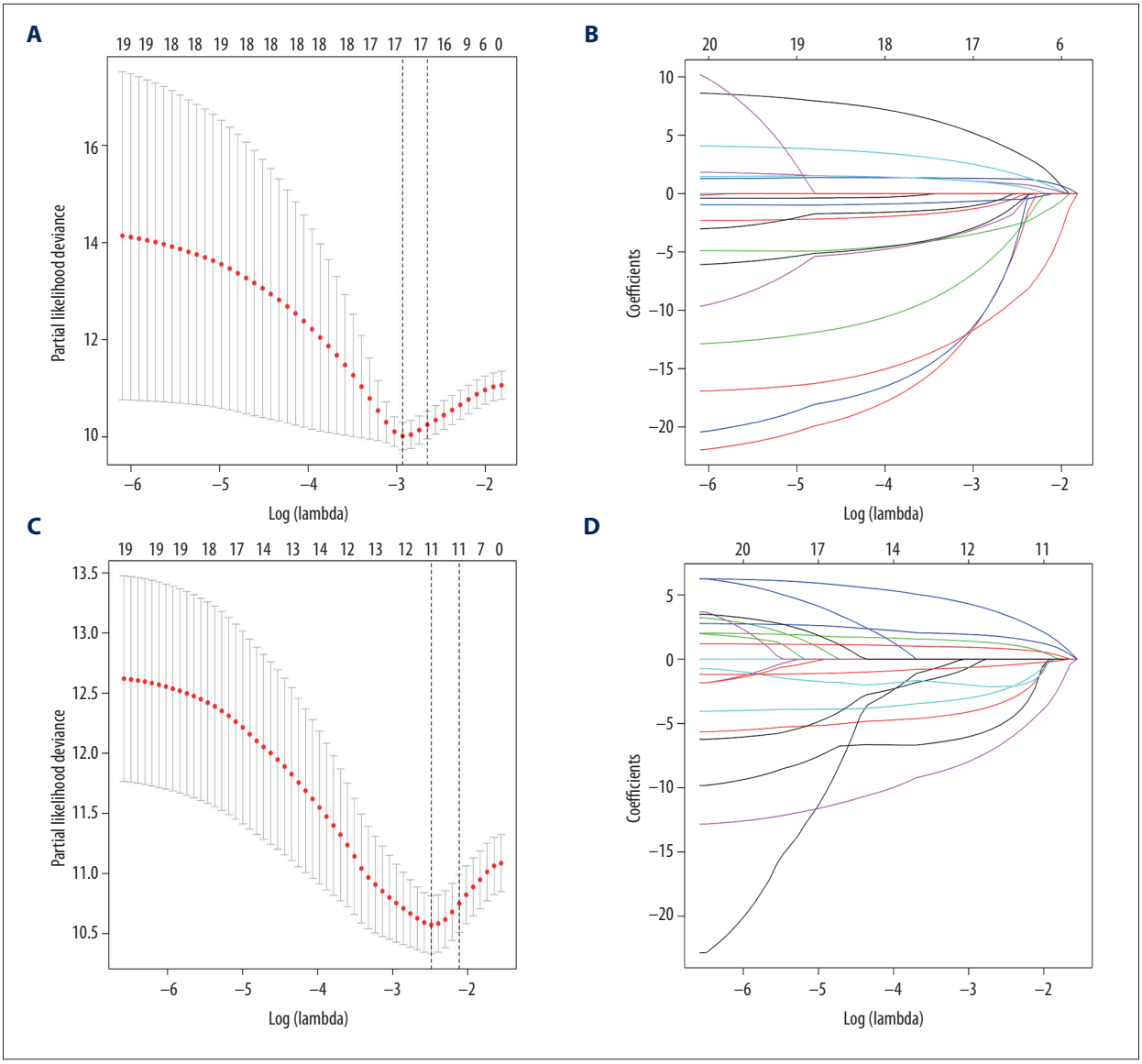


Figure 3. LASSO regression to select the most significant OS-AS events (**A, B**) and DFS-AS events (**C, D**), respectively. LASSO – least absolute shrinkage and selection operator; AS – alternative splicing; OS – overall survival; DFS – disease-free survival.

Subgroup analysis for the nomograms

We assessed the predictive ability of the nomogram in 4 different histological sarcomas. The ROC curves in 8 subgroups showed good discrimination, with all AUC>0.700 (Figure 6A–6D, 6I–6L). In addition, in 8 subgroups, all survival curves suggested that the prognosis of high-risk patients is significantly worse than that of low-risk patients (all $p<0.05$) (Figure 6E–6H, 6M–6P).

Prognostic AS events correlation network of splicing factors

The splicing-regulatory network between OS-AS events and SFs is shown in Figure 7A and the corresponding correlation coefficient is shown in Figure 7B. In addition, the network between DFS-AS events and SFs is shown in Figure 7C and the corresponding correlation coefficient is shown in Figure 7D. We found that 1 SF can regulate different AS events and even cause different effects, and some AS events can be regulated by 2 SFs. For example, as a SF, MBNL1 can regulate 18 AS events, including 12 positive regulations and 6 negative regulations. Moreover, as a DFS-related AS event, TMEM189-59769-AP can be regulated by MBNL1 and PTBP2. In addition, in the present cohort, we also found that the majority of prognostic AS events with poor prognosis were positively regulated by SFs. In contrast, the majority of AS events with favorable prognosis were negatively regulated by SFs.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

Table 1. Multivariate COX analysis of overall survival-related AS events.

Symbol	AS type	HR	Lower.95	Upper.95	P value
CLTC	ES	1.07E-07	5.61E-12	0.002039	0.001413
PSMG1	ES	0.001137	2.26E-05	0.057229	0.000697
INPP5B	AA	81.747094	16.300191	409.969893	8.67E-08
CTNND1	ES	0.000118	2.85E-07	0.048717	0.003252
TAF1A	AT	10073.277430	92.247074	1099990.640000	0.000118
TMEM161B	AT	0.006279	0.000411	0.095853	0.000266
NMRAL1	RI	5.839665	1.239303	27.516829	0.025665
SIRT3	ES	2.74E-09	1.15E-18	6.539018	0.073531
HS1BP3	ES	8.42E-08	2.03E-11	0.000349	0.000126
CCDC91	ES	4.748142	1.195326	18.860833	0.026864
C19orf60	ES	0.001407	1.08E-06	1.825446	0.072588

AS – alternative splicing; ES – exon skip; RI – retained intron; AT – alternate terminator; AA – alternate acceptor.

Table 2. Multivariate COX analysis of disease-free survival-related AS events.

Symbol	AS type	HR	Lower.95	Upper.95	P value
SPIDR	ES	0.000367	1.37E-06	0.09814	0.005537
MMP19	ES	0.003436	2.96E-05	0.398999	0.019351
PTCH2	AT	18.55505	1.876145	183.5092	0.012485
BEST3	AT	8.490379	1.244015	57.94666	0.029052
SLC27A1	ES	0.015654	0.001242	0.197346	0.001304
IRAK1	ES	1.99E-05	4.86E-08	0.008167	0.000421
ZNF331	AP	3.224539	1.09384	9.505643	0.03379
PPRC1	ES	418.7191	7.698737	22773.3	0.003066

AS – alternative splicing; ES – exon skip; AP – alternate promoter; AT – alternate terminator.

Enrichment analysis of overlapping AS events

To explore the potential functions and pathways of prognostic AS events, GO and KEGG analyses were performed. As shown in Figure 8A, the GO analysis showed that the parent genes of overlapping AS events were enriched in many tumor-related features, such as mitochondrial matrix, NK-kappa B signaling, DNA repair, and response to hypobaric hypoxia. In the KEGG analysis, we also found many tumor-related pathways, such as Toll-like receptor (TLR) signaling pathway, Pathways in cancer, MAPK signaling pathway, p53 signaling pathway, Transcriptional misregulation in cancer, and Basal cell carcinoma (Figure 8B). Interestingly, some immune-related features were discovered, including NK-kappa B signaling and the Toll-like receptor signaling pathway, which suggests a significant association between AS events and immune functions.

AS-based clusters were significantly associated with prognosis, tumor microenvironment score, and immune cell proportion

Based on the overlapping prognostic AS events between OS-related AS events and DFS-related AS events, 3 clusters of samples were identified: C1 (n=86, 44.1%), C2 (n=56, 28.7%), and C3 (n=53, 27.2%) (Figure 9A–9C). The results of survival analysis showed that clusters were significantly associated with different survival patterns, including OS and DFS (Figure 9D, 9E). In addition, 131 patients (C1=64, C2=29, and C3=38, Supplementary Table 1) with complete immune cell data were further studied to assess the correlation between immune cell features and the 3 clusters. Interestingly, the results showed that the 10/22 type of immune cells was significantly different between the 3 clusters (Figure 9F). Furthermore, the distinctions of tumor microenvironment among the 3 clusters were also studied, and the results showed that C2 was associated with lower immune

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

Table 3. Univariate Cox analysis of clinicopathologic factors for sarcoma patients.

	Overall survival				Disease-free survival			
	HR	95.0% CI		P	HR	95.0% CI		P
Risk classification based on AS events								
Low-risk				0.000				0.000
Mid-risk	4.620	2.345	9.103	0.000	3.688	2.175	6.251	0.000
High-risk	26.918	12.932	56.027	0.000	14.003	7.559	25.943	0.000
Age								
Low-age				0.008				0.075
Mid-age	0.989	0.508	1.925	0.975	1.094	0.664	1.802	0.725
High-age	2.371	1.269	4.430	0.007	1.659	1.027	2.679	0.039
Sex (Female)	1.080	0.637	1.830	0.776	1.214	0.824	1.787	0.326
Neoplasm histologic type name								
Dedifferentiated liposarcoma				0.350				0.194
Leiomyosarcoma	0.617	0.331	1.153	0.130	0.665	0.417	1.059	0.086
Undifferentiated pleomorphic sarcoma	0.703	0.315	1.566	0.388	0.716	0.396	1.294	0.268
Other	0.510	0.213	1.224	0.132	0.531	0.277	1.017	0.056
Disease multifocal indicator	2.628	1.469	4.700	0.001	2.119	1.310	3.427	0.002
Surgical margin resection status (rother)	2.638	1.552	4.484	0.000	2.216	1.497	3.280	0.000
Tumor tissue site								
Chest				0.943				0.668
Gynecological	2.460	0.302	20.038	0.400	2.259	0.499	10.215	0.290
Head and neck	0.000	0.000	2.876E+250	0.972	3.171	0.443	22.695	0.251
Lower abdominal/pelvic	1.086	0.098	12.012	0.946	1.637	0.317	8.463	0.556
Lower extremity	1.958	0.260	14.740	0.514	1.852	0.442	7.759	0.399
Retroperitoneum/upper abdominal	2.220	0.301	16.388	0.434	2.130	0.516	8.789	0.296
Superficial trunk	1.799	0.163	19.909	0.632	0.878	0.123	6.265	0.897
Upper extremity	1.015	0.063	16.298	0.992	3.339	0.671	16.615	0.141

AS – alternative splicing; HR – hazard ratio.

Table 4. Multivariate Cox analysis of clinicopathologic factors for sarcoma patients.

	Overall survival				Disease-free survival			
	HR	95.0% CI		P	HR	95.0% CI		P
Risk classification based on AS events								
Low-risk				0.000				0.000
Mid-risk	5.192	2.599	10.371	0.000	3.537	2.084	6.006	0.000
High-risk	32.126	15.140	68.172	0.000	12.045	6.439	22.532	0.000
Disease multifocal indicator	3.559	1.967	6.439	0.000				
Surgical margin resection status (rother)					1.665	1.114	2.489	0.013

AS – alternative splicing; HR – hazard ratio.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

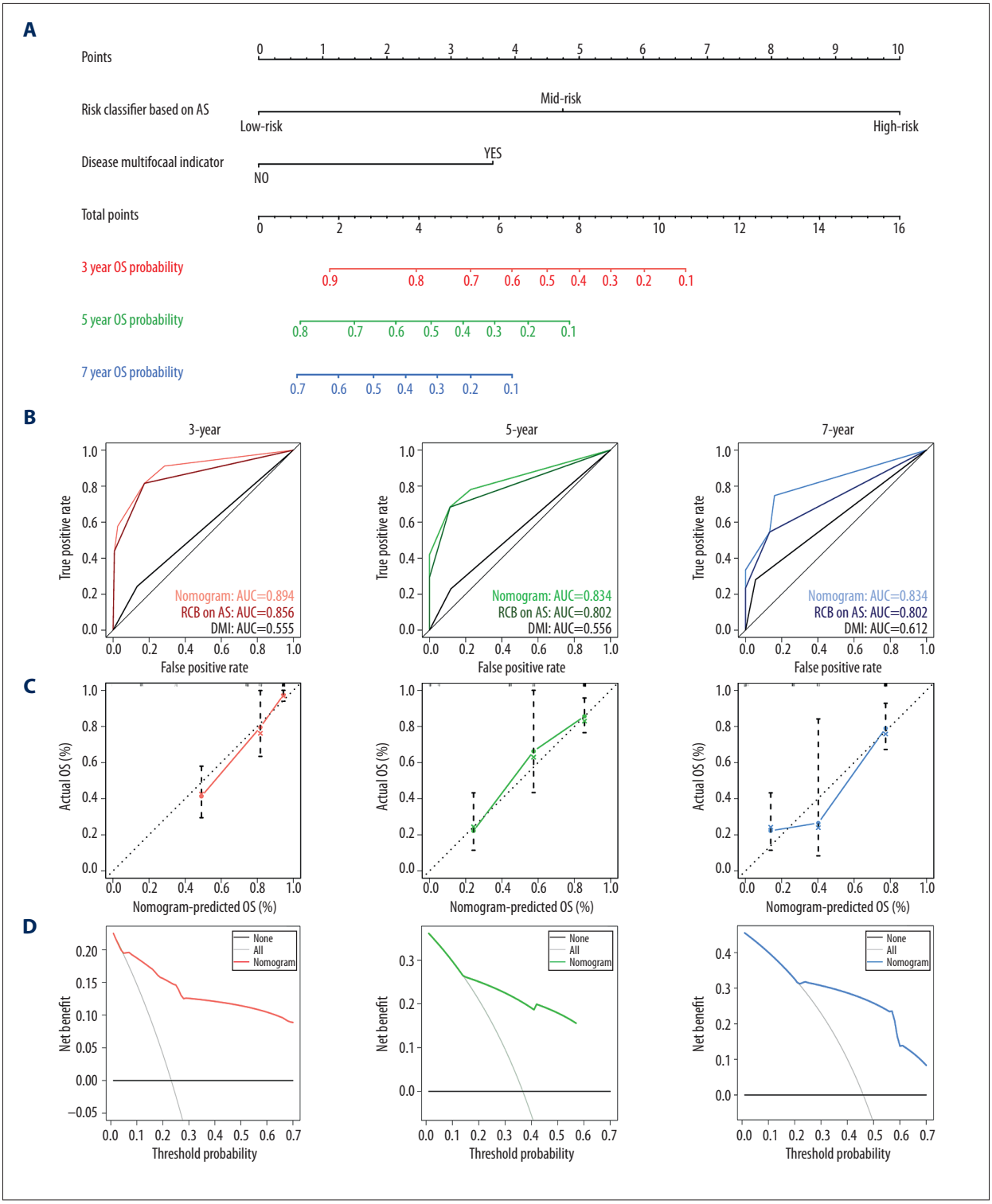


Figure 4. The nomograms and corresponding results showed the prognostic value of AS events and clinicopathologic data. (**A**) Nomogram for predicting OS in the SARC cohort. (**B**) Time-dependent ROC curves for 3- (red), 5- (green), and 7- (blue) year OS. (**C**) Calibration plot of the AS-clinicopathologic nomogram in terms of the agreement between nomogram-predicted and observed 3- (red), 5- (green), and 7- (blue) year OS in the SARC cohort. (**D**) Decision curve analysis of the AS-clinicopathologic nomogram for 3- (red), 5- (green), and 7- (blue) year risk in the SARC cohort. AS – alternative splicing; OS – overall survival; SARC – sarcoma; ROC – receiver operating characteristic.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

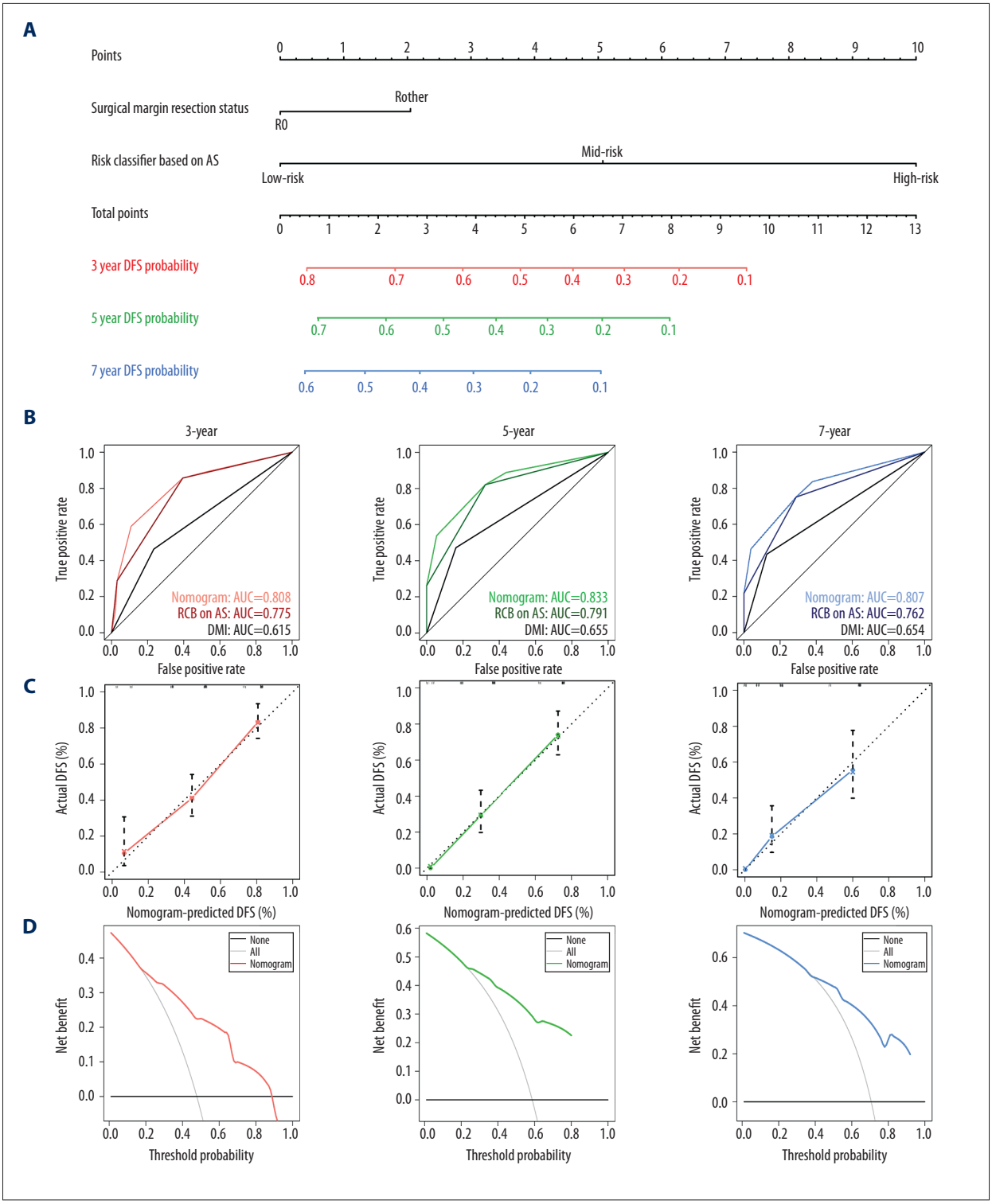


Figure 5. The nomograms and corresponding results showed the prognostic value of AS events and clinicopathologic data. (**A**) Nomogram for predicting DFS in the SARC cohort. (**B**) Time-dependent ROC curves for 3- (red), 5- (green), and 7- (blue) year DFS. (**C**) Calibration plot of the AS-clinicopathologic nomogram in terms of the agreement between nomogram-predicted and observed 3- (red), 5- (green), and 7- (blue) year DFS in the SARC cohort. (**D**) Decision curve analysis of the AS-clinicopathologic nomogram for 3- (red), 5- (green), and 7- (blue) year risk in the SARC cohort. AS – alternative splicing; DFS – disease-free survival; SARC – sarcoma; ROC – receiver operating characteristic.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

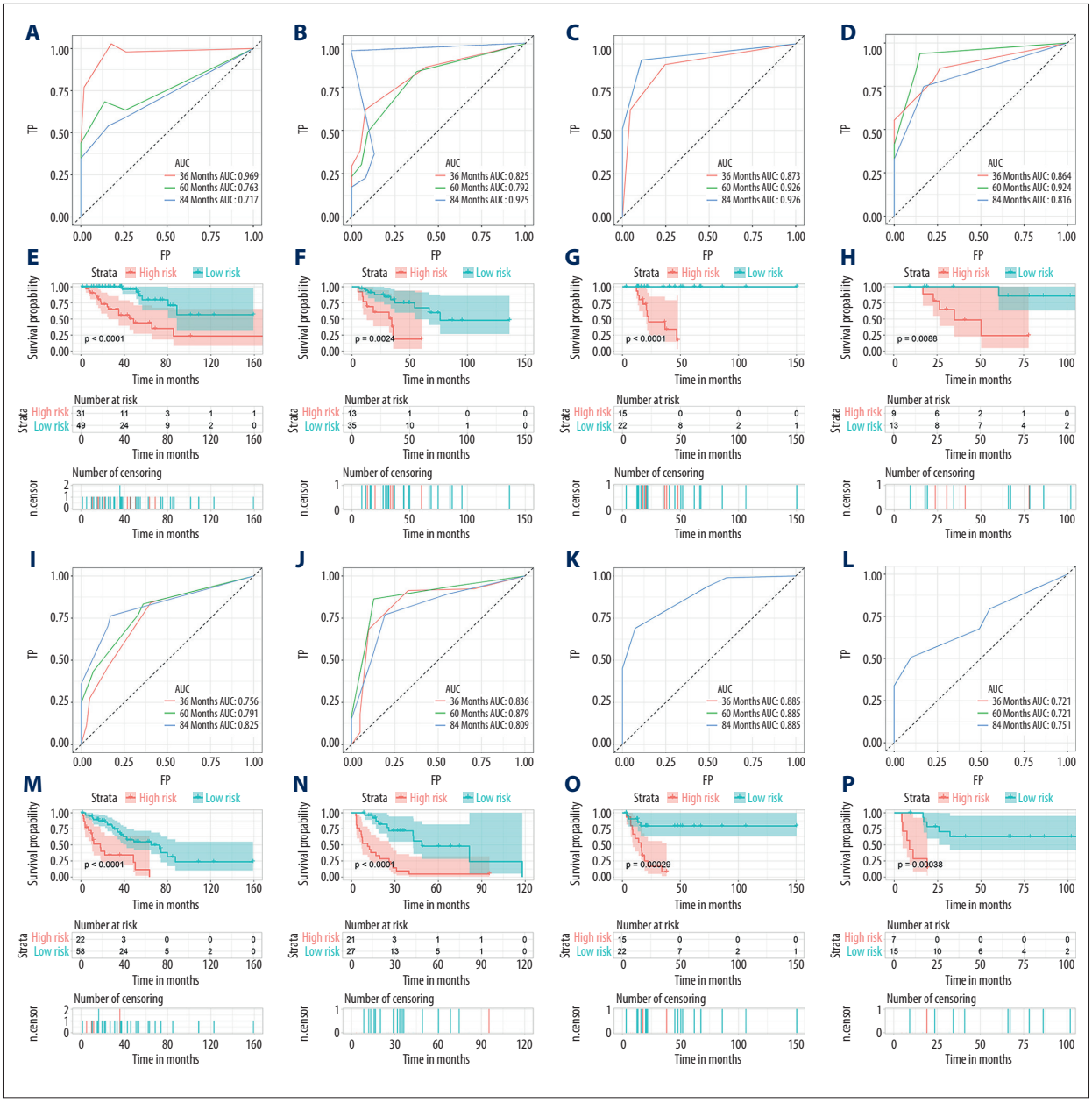


Figure 6. ROC curves and survival curves in 8 subgroups. ROC curves of OS nomogram in LMS (**A**), DLP (**B**), UPS (**C**), and MFS (**D**). Survival curves of OS in LMS (**E**), DLP (**F**), UPS (**G**), and MFS (**H**). ROC curves of DFS nomogram in LMS (**I**), DLP (**J**), UPS (**K**), and MFS (**L**). Survival curves of DFS in LMS (**M**), DLP (**N**), UPS (**O**), and MFS (**P**). LMS – leiomyosarcoma; DLP – dedifferentiated liposarcoma; UPS – undifferentiated pleomorphic sarcoma; MFS – myxofibrosarcoma; ROC – receiver operating characteristic; OS – overall survival; DFS – disease-free survival.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

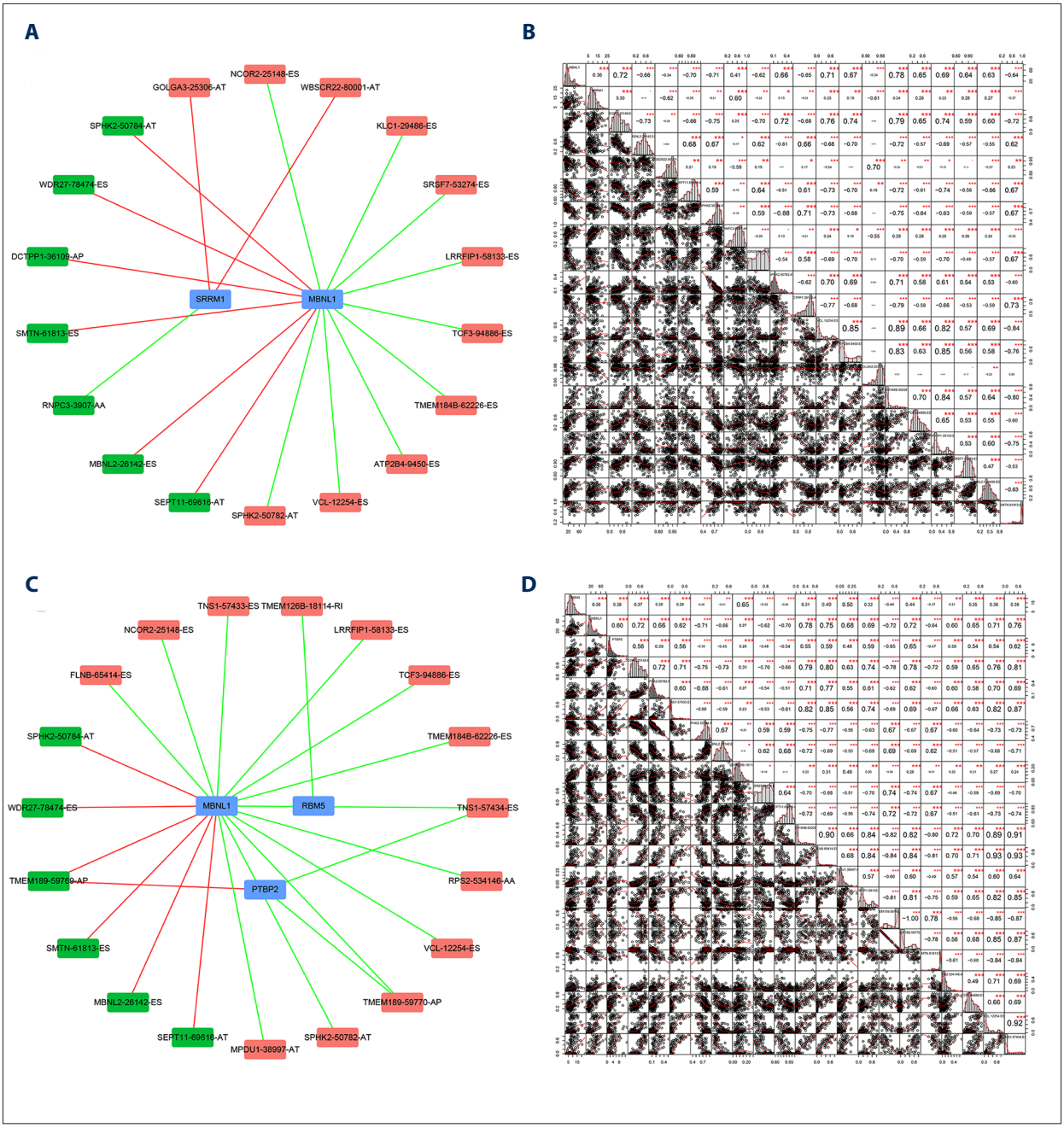


Figure 7. Correlation network between OS-AS events (**A, B**) and DFS-AS events (**C, D**) and SFs in SARC. The majority of prognostic AS events with poor prognosis (green dots) were positively (red lines) correlated with the expression of SFs (blue dots), while the majority of AS events with good prognosis (red dots) were negatively (green lines) correlated with the expression of SFs. AS – alternative splicing; OS – overall survival; DFS – disease-free survival; SF – splicing factor; SARC – sarcoma.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

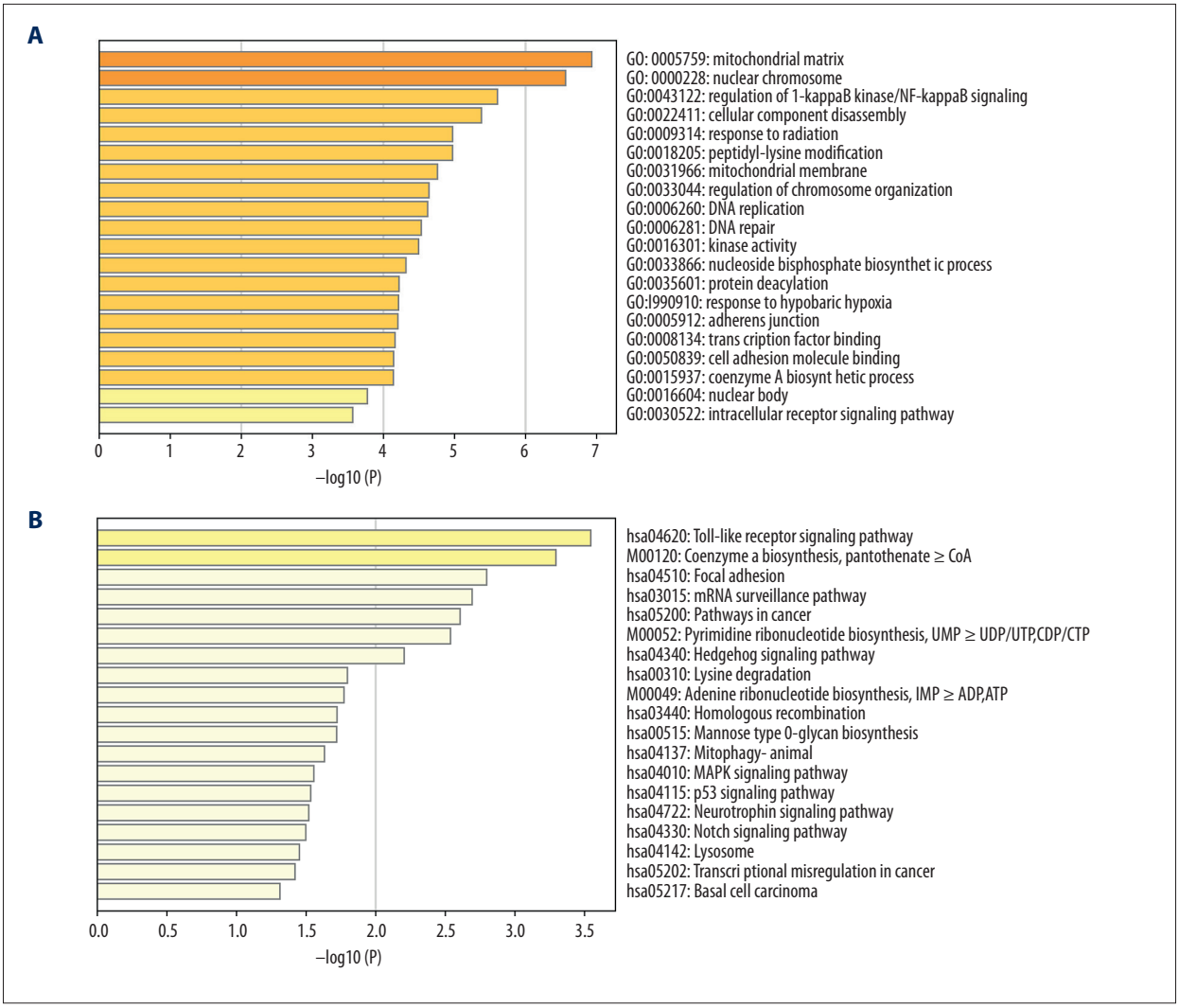


Figure 8. Enrichment of parent genes from overlapping prognostic AS events in SARC. (**A**) GO analysis of parent genes from overlapping AS events. (**B**) KEGG pathway analysis of parent genes from overlapping AS events. AS – alternative splicing; SARC – sarcoma; GO – Gene Ontology; OS – overall survival; DFS – disease-free survival; KEGG – Kyoto Encyclopedia of Genes and Genomes.

scores (Figure 9G), stromal scores (Figure 9H), and estimate scores (Figure 9I) compared with C1 and C3.

Discussion

In recent years, with the development of high-throughput sequencing techniques and computer technology for bioinformation, more and more research has focused on the profiling of AS events for tumors. However, the significance of AS events in sarcoma remains unclear, especially in tumor recurrence. In our study, several prognostic AS events were identified, and the predictive models based on the AS event and clinicopathologic data showed excellent performance. To the best of our knowledge, this is the first published study to determine the relationship between AS events and immune cell infiltration and tumor microenvironment, and our results may increase understanding of the role of AS events in tumors.

Our research focused on the prognostic significance of AS events in sarcoma patients. Although the prognostic ability of AS events have been widely confirmed in other cancers [14,19–23,34–36], the association of AS events in sarcoma patients remains unclear. In our study, 1945 AS events were determined as OS-related AS events and 1831 AS events were determined as DFS-related AS events. To further investigate the predictive ability of AS events in sarcoma patients, 2 nomograms were developed based on the independent prognostic AS events and clinicopathologic data, which could serve as a useful tool for management of sarcoma patients.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

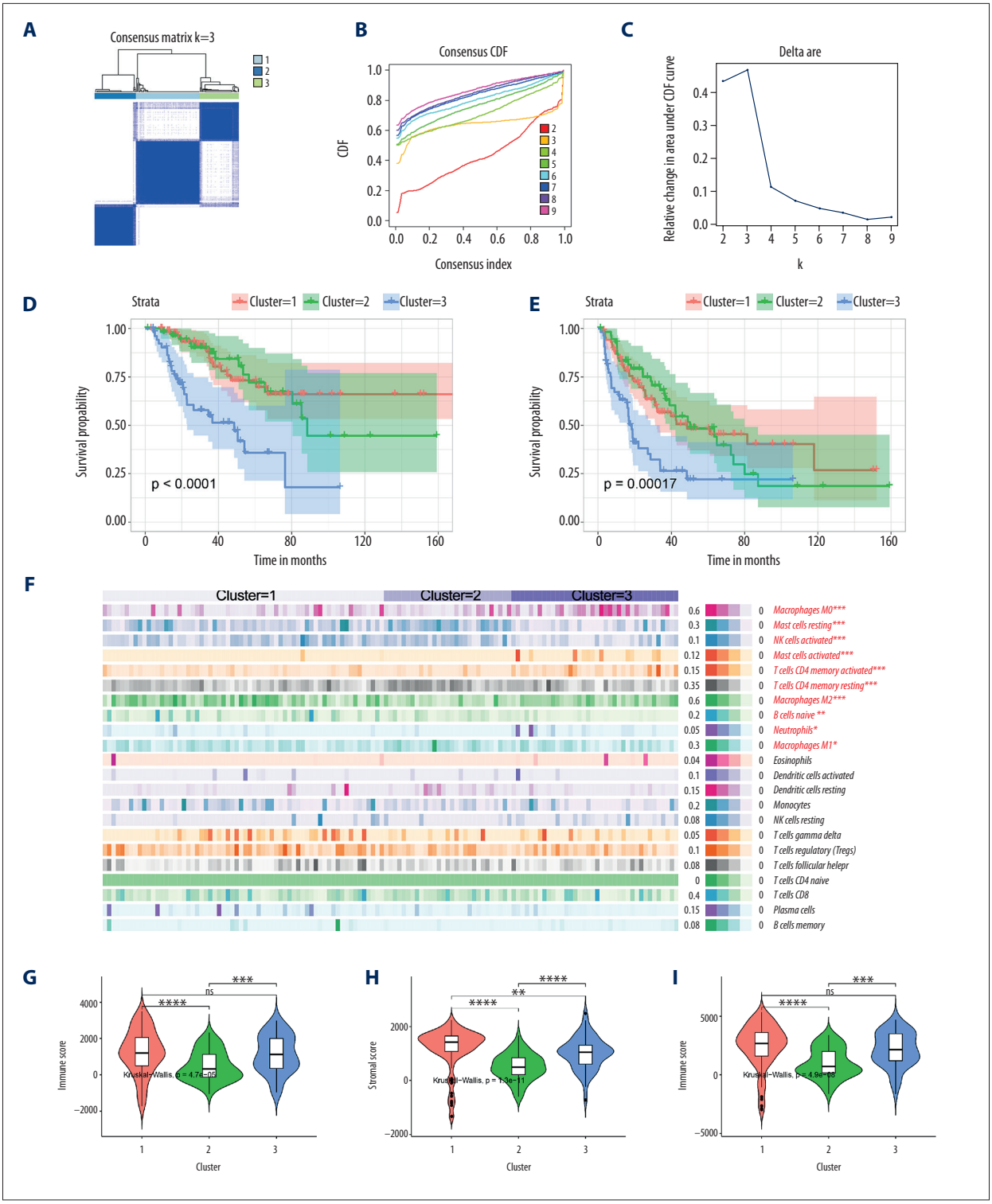


Figure 9. AS-based clusters significantly associated with prognosis, immune microenvironment scores, and immune cells. (**A–C**) Unsupervised clustering analysis based on the overlapping prognostic AS events identified 3 clusters. (**D**) Kaplan-Meier survival curve showing the OS probability over time for 3 clusters. (**E**) Kaplan-Meier survival curve showing the DFS probability over time for 3 clusters. (**F**) Heat map showing the distribution of 22 immune cells between 3 clusters. (**G–I**) The immune score, stromal score, and estimate score between AS-based clusters. AS – alternative splicing; OS – overall survival; DFS – disease-free survival.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

X-tile software was used to stratify the sarcoma patients into 3 subgroups according to the optimal cutoff risk scores. Most previous studies have used a common cutoff value to divide patients into high- and low-risk groups, without considering survival data. As a new bioinformatics tool, X-tile can identify the optimal cutoff of tumor biomarker expression and has been validated in many previous studies [27]. In addition, many predictive models for sarcoma patients have been constructed based on clinicopathologic data, lncRNA, plasmacytoma variant translocation 1, and other predictors [3–5], but the discriminative ability of the previous models was low, with AUC or C-statistic less than 0.800 [3–5,37]. In the present study, the time-dependent AUC at 3, 5, and 7 years of the 2 nomograms were higher than 0.800, which means that both nomograms have good discrimination in predicting the outcome of sarcoma patients [37].

In the present study, enrichment analysis of overlapping genes indicated that many pathways are involved in tumor progression. Interestingly, both in the GO and KEGG analysis, several pathways were found to play an important role in tumor progression. For example, the MAPK signaling pathway has been confirmed was involved in some tumors [38–40]. More importantly, some immune-related pathways have been discovered. For instance, TLRs are key proteins in innate immunity, stimulation of which generates a series of antitumor effects through intermediary immune cells [41]. Dendritic cells (DCs), a kind of immune cell, express all TLRs and exert effects on T and B lymphocytes to connect innate and adaptive immune responses [42]. DCs also play important roles in tumorigenesis, progression, and immunotherapy [43–45]. NF-κB was firstly found by Ranjan Sen and David Baltimore in 1986 in the context of the expression of a gene encoding immunoglobulin-j light chain in B lymphocytes [46], which can regulate the transcription of genes to promote immune cell development [47]. Hence, it is possible that NF-κB can generate immune responses in the development of sarcoma.

Our study is also the first to report that the AS clusters presented distinct immune features in sarcoma. Recently, Li et al. [48] reported an association between AS events and immune features in head and neck squamous cell carcinoma. Hence, we studied the association between clusters based on the overlapping prognostic AS events and clinical data, tumor microenvironment, and immune cell fractions. To the best of our knowledge, the present study is the first to perform a comprehensive analysis of an AS-based cluster of sarcoma patients. As in the study by Li et al., tumor microenvironment scores were significantly different between clusters [48]. More importantly, among the 22 types of immune cells, 10 of them were determined to be different in the 3 AS clusters. Based on the novel phenomenon presented in the present research and previous studies, we hypothesized that the interactions between the immune response and AS events play an important role in regulating tumor progression [48–51]. However, further research is needed to better understand their correlations. Generally, these findings enrich our knowledge of the molecular classification of sarcoma and provide a large number of biomarker candidates and potential targets for the treatment of sarcoma patients.

We established and evaluated 2 predictive models based on AS events and clinicopathologic data. Although the nomograms had good predictive ability, the present study had some limitations. For instance, the predictive signature was developed based on a public database without external validation or experimental validation. In addition, this study was bioinformatics research based on the TCGA database, and further research is needed to explore the mechanism of the effects of AS events in the prognosis of sarcoma.

Conclusions

The present study showed the prognostic value of AS events in sarcoma patients. The cluster for sarcoma patients based on AS events was established and revealed the intrinsic relevance of molecular alterations and immune features.

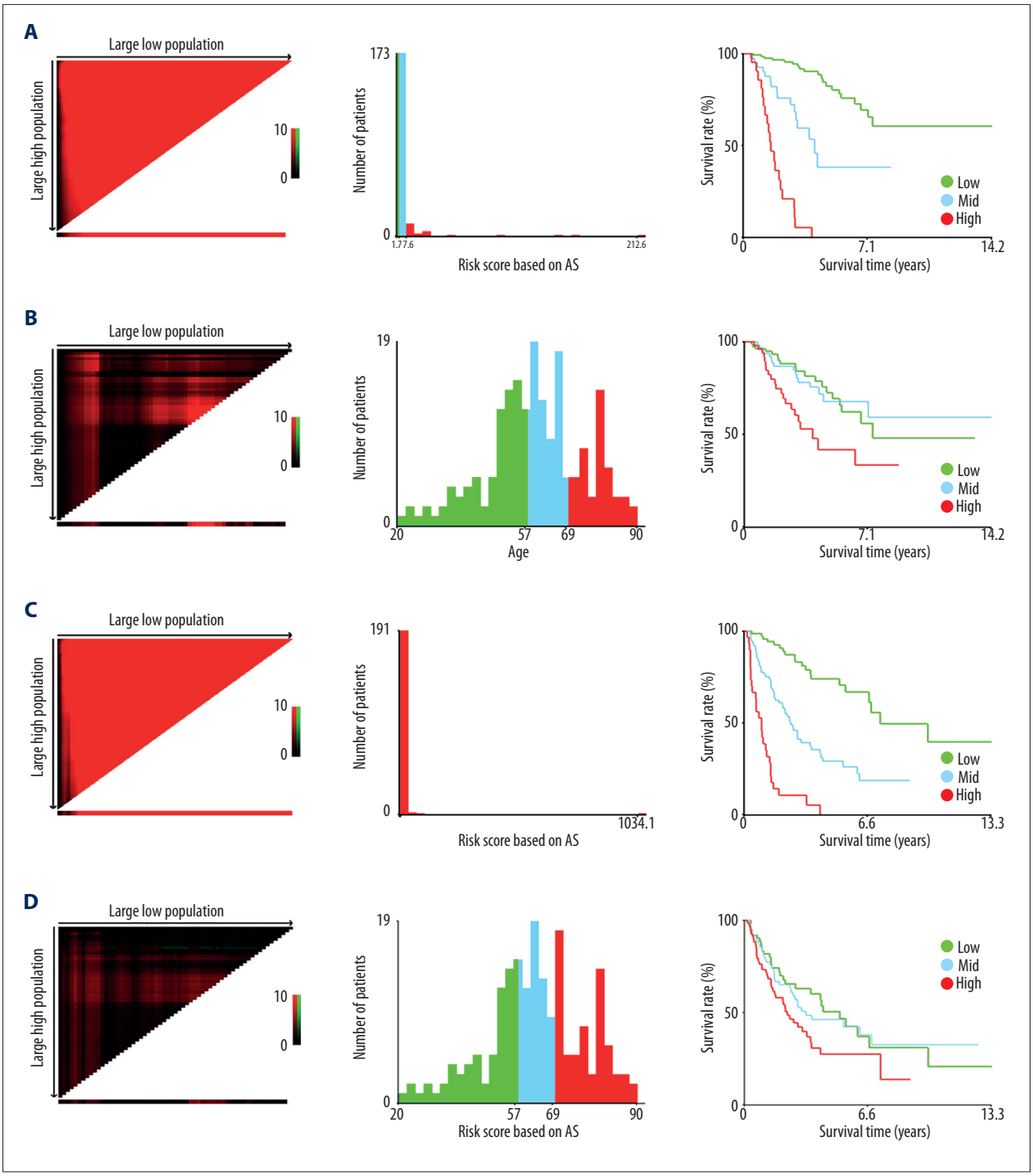
Conflicts of interest

None.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

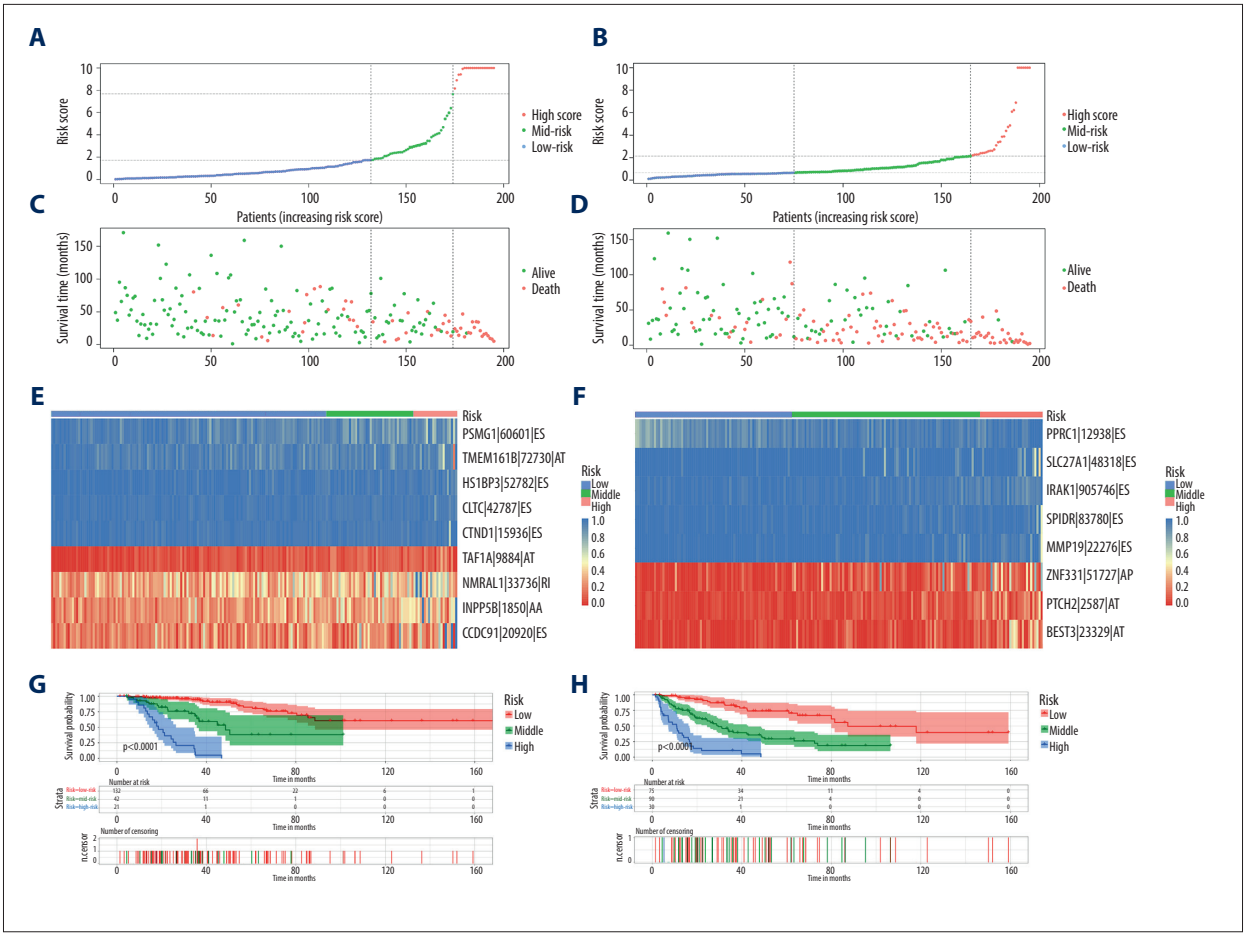
Supplementary Data

Supplementary/raw Table 1 available from the corresponding author on request.



Supplementary Figure 1. X-tile was used to determine the optimal cutoff value of the risk score and age in sarcoma patients. (**A**) The X-tile to identify the optimal cutoff value of risk score of OS. (**B**) The X-tile to identify the optimal cutoff value of the age of OS. (**C**) The X-tile to identify the optimal cutoff value of the risk score of DFS. (**D**) The X-tile to determine the optimal cutoff value of the age of DFS. OS – overall survival; DFS – disease-free survival.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]



Supplementary Figure 2. Establishment of the prognostic model for OS (**A, C, E**) and DFS (**B, D, F**) based on independent AS events. (**A, B**) Risk curve of each sample reordered by risk score. (**C, D**) Scatter plot showing the survival status of sarcoma patients. (**E, F**) Heat map of the expression level of 9 independent OS-AS events (**E**) and DFS-AS events (**F**) filtered by multivariate Cox regression. OS – overall survival; DFS – disease-free survival; AS – alternative splicing.

References:

1. Cormier JN, Pollock RE: Soft tissue sarcomas. Cancer J Clin, 2004; 54(2): 94–109
2. Helman LJ, Meltzer P: Mechanisms of sarcoma development. Nat Rev Cancer, 2003; 3(9): 685–94
3. Callegaro D, Miceli R, Bonvalot S et al: Development and external validation of two nomograms to predict overall survival and occurrence of distant metastases in adults after surgical resection of localised soft-tissue sarcomas of the extremities: A retrospective analysis. Lancet Oncol, 2016; 17(5): 671–80
4. He R-q, Wei Q-j, Tang R-x et al: Prediction of clinical outcome and survival in soft-tissue sarcoma using a ten-lncRNA signature. Oncotarget, 2017; 8(46): 80336
5. Liu J, Li R, Liao X et al: Comprehensive investigation of the clinical significance and molecular mechanisms of plasmacytoma variant translocation 1 in sarcoma using genome-wide RNA sequencing data. J Cancer, 2019; 10(20): 4961–77
6. Han Q, Li C, Cao Y et al: CBX2 is a functional target of miRNA let-7a and acts as a tumor promoter in osteosarcoma. Cancer Med, 2019; 8(8): 3981–91
7. Hu Q, Zhou S, Hu X et al: Systematic screening identifies a 2-gene signature as a high-potential prognostic marker of undifferentiated pleomorphic sarcoma/myxofibrosarcoma. J Cell Mol Med, 2019; 24(1): 1010–21
8. Wang ET, Sandberg R, Luo S et al: Alternative isoform regulation in human tissue transcriptomes. Nature, 2008; 456(7221): 470–76
9. Nilsen TW, Graveley BR: Expansion of the eukaryotic proteome by alternative splicing. Nature, 2010; 463(7280): 457–63
10. Fu X-D, Ares M Jr.: Context-dependent control of alternative splicing by RNA-binding proteins. Nat Rev Genet, 2014; 15(10): 689–701
11. Ge Y, Porse BT: The functional consequences of intron retention: Alternative splicing coupled to NMD as a regulator of gene expression. Bioessays, 2014; 36(3): 236–43
12. Lee Y, Rio DC: Mechanisms and regulation of alternative pre-mRNA splicing. Annu Rev Biochem, 2015; 84: 291–323
13. Paschalis A, Sharp A, Welti JC et al: Alternative splicing in prostate cancer. Nat Rev Clin Oncol, 2018; 15(11): 663–75
14. Li S, Hu Z, Zhao Y et al: Transcriptome-wide analysis reveals the landscape of aberrant alternative splicing events in liver cancer. Hepatology, 2019; 69(1): 359–75
15. Marzese DM, Manughian-Peter AO, Orozco JI, Hoon DS: Alternative splicing and cancer metastasis: Prognostic and therapeutic applications. Clin Exp Metastasis, 2018; 35(5–6): 393–402
16. Siegfried Z, Karni R: The role of alternative splicing in cancer drug resistance. Curr Opin Genet Dev, 2018; 48: 16–21

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

17. Pal S, Gupta R, Davuluri RV: Alternative transcription and alternative splicing in cancer. Pharmacol Ther, 2012; 136(3): 283–94
18. Venables JP, Klinck R, Koh C et al: Cancer-associated regulation of alternative splicing. Nat Struct Mol Biol, 2009; 16(6): 670–76
19. Zhang S, Wu X, Diao P et al: Identification of a prognostic alternative splicing signature in oral squamous cell carcinoma. J Cell Physiol, 2019; 235(5): 4804–13
20. Zhang D, Duan Y, Cun J: Identification of prognostic alternative splicing signature in breast carcinoma. Front Genet, 2019; 10: 278
21. Zuo Y, Zhang L, Tang W, Tang W: Identification of prognosis-related alternative splicing events in kidney renal clear cell carcinoma. J Cell Mol Med, 2019; 23(11): 7762–72
22. Xiong Y, Deng Y, Wang K et al: Profiles of alternative splicing in colorectal cancer and their clinical significance: A study based on large-scale sequencing data. EBioMedicine, 2018; 36: 183–95
23. Xie Z-C, Wu H-Y, Ma F-C et al: Prognostic alternative splicing signatures and underlying regulatory network in esophageal carcinoma. Am J Transl Res, 2019; 11(7): 4010–28
24. Ryan MC, Cleland J, Kim R et al: SpliceSeq: A resource for analysis and visualization of RNA-Seq data on alternative splicing and its functional impacts. Bioinformatics, 2012; 28(18): 2385–87
25. Lex A, Gehlenborg N, Strobelt H et al: UpSet: Visualization of intersecting sets. IEEE Trans Vis Comput Graph, 2014; 20(12): 1983–92
26. Tibshirani R: Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological), 1996; 58(1): 267–88
27. Camp RL, Dolled-Filhart M, Rimm DL: X-tile: A new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. Clin Cancer Res, 2004; 10(21): 7252–59
28. Heagerty PJ, Lumley T, Pepe MS: Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics, 2000; 56(2): 337–44
29. Vickers AJ, Elkin EB: Decision curve analysis: A novel method for evaluating prediction models. Med Decis Making, 2006; 26(6): 565–74
30. Siam A, Baker M, Amit L et al: Regulation of alternative splicing by p300-mediated acetylation of splicing factors. RNA, 2019; 25(7): 813–24
31. Singh RK, Kolonin AM, Fiorotto ML, Cooper TA: Rbfox-splicing factors maintain skeletal muscle mass by regulating calpain3 and proteostasis. Cell Rep, 2018; 24(1): 197–208
32. Newman AM, Liu CL, Green MR et al: Robust enumeration of cell subsets from tissue expression profiles. Nat Methods, 2015; 12(5): 453–57
33. Yoshihara K, Shahmoradgoli M, Martínez E et al: Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun, 2013; 4: 2612
34. Liang Y, Song J, He D et al: Systematic analysis of survival-associated alternative splicing signatures uncovers prognostic predictors for head and neck cancer. J Cell Physiol, 2019; 234(9): 15836–46
35. Huang X, Liu J, Mo X et al: Systematic profiling of alternative splicing events and splicing factors in left-and right-sided colon cancer. Aging (Albany NY), 2019; 11(19): 8270–93
36. Chen X, Zhao C, Guo B et al: Systematic profiling of alternative mRNA splicing signature for predicting glioblastoma prognosis. Front Oncol, 2019; 9: 928
37. Tape TG: Interpretation of diagnostic tests. Ann Intern Med, 2001; 135(1): 72–72
38. Shen X, Shen P, Yang Q et al: Knockdown of long non-coding RNA PCAT-1 inhibits myeloma cell growth and drug resistance via p38 and JNK MAPK pathways. J Cancer, 2019; 10(26): 6502–10
39. Choi JH, Hwang YP, Jin SW et al: Suppression of PMA-induced human fibrosarcoma HT-1080 invasion and metastasis by kahweol via inhibiting Akt/JNK1/2/p38 MAPK signal pathway and NF-κB dependent transcriptional activities. Food Chem Toxicol, 2019; 125: 1–9
40. Wu R, Li D, Tang Q et al: A novel peptide from vespa ducalis induces apoptosis in osteosarcoma cells by activating the p38 MAPK and JNK signaling pathways. Biol Pharm Bull, 2018; 41(4): 458–64
41. Lowe EL, Crother TR, Rabizadeh S et al: Toll-like receptor 2 signaling protects mice from tumor development in a mouse model of colitis-induced cancer. PLoS One, 2010; 5(9): e13027
42. Diamond MS, Kinder M, Matsuhita H et al: Type I interferon is selectively required by dendritic cells for immune rejection of tumors. J Exp Med, 2011; 208(10): 1989–2003
43. Moghadan SO, Nowroozi MR: Toll-like receptors: The role in bladder cancer development, progression and immunotherapy. Scand J Immunol, 2019; 90(6): e12818
44. Braunstein MJ, Kucharczyk J, Adams S: Targeting toll-like receptors for cancer therapy. Target Oncol, 2018; 13(5): 583–98
45. Messaritakis I, Stogiannitsi M, Koulouridi A et al: Evaluation of the detection of Toll-like receptors (TLRs) in cancer development and progression in patients with colorectal cancer. PLoS One, 2018; 13(6): e0197327
46. Ranjan S, Baltimore D: Multiple nuclear factors interact with the immunoglobulin enhancer sequences. Cell, 1986; 46(5): 705–16
47. Christian F, Smith El, Carmody RJ: The regulation of NF-κB subunits by phosphorylation. Cells, 2016; 5(1): 12
48. Li Z-X, Zheng Z-Q, Wei Z-H 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
49. Frankiw L, Mann M, Li G et al: Alternative splicing coupled with transcript degradation modulates OAS1g antiviral activity. RNA, 2019; 26(2): 126–36
50. Meng T, Huang R, Zeng Z et al: Identification of prognostic and metastatic alternative splicing signatures in kidney renal clear cell carcinoma. Front Bioeng Biotechnol, 2019; 7: 270
51. Huang X, Liu J, Mo X et al. Systematic profiling of alternative splicing events and splicing factors in left- and right-sided colon cancer. Aging, 2019; 11(19): 8270–93

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]