

# Prognostic Significance of Alternative Splicing Genes in Cervical Squamous Cell Carcinoma and Endocervical Adenocarcinoma

Xiaoyu Wang\*

Weichun Tang\*

Yilin Lu

Jun You

Yun Han 

Yanli Zheng

Department of Obstetrics and  
Gynecology, Nantong First People's  
Hospital, Nantong, Jiangsu, 226001,  
People's Republic of China

\*These authors contributed equally to  
this work

**Background:** Alternative splicing (AS) acts on many tumors and its relationship with cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC) needs to be researched.

**Methods:** RNA sequencing data and clinical information of CESC cohorts were obtained from the Cancer Genome Atlas (TCGA) and SpliceSeq was used to analyze the splicing profile of mRNA in CESC. UpSetR displayed the intersections among AS events and univariate analysis chose survival-associated AS and splicing factor (SF) genes. Functional analysis was operated on Enrichr, STRING database and MCODE analysis were used to evaluate protein-protein interaction (PPI) information. LASSO and multivariate analysis constructed prognostic model and risk analysis of tumor infiltrating immune cells was also conducted.

**Results:** A total of 402 AS-generated genes were found to be associated with CESC prognosis. Functional analysis showed that Golgi to lysosome transport was enriched. PPI network suggested that UBA52 was most functional. Dendritic cells activated, dendritic cells resting, macrophages M0, mast cells resting, T cells CD4 memory activated and T cells CD8 were most correlative with the risk score.

**Conclusion:** SFs and AS events can directly or indirectly affect the prognosis of CESC patients and this study identified SNRPA and CELF2 as two CESC-engaged SFs.

**Keywords:** cervical squamous cell carcinoma and endocervical adenocarcinoma, CESC, alternative splicing, AS, prognosis, TCGA

## Introduction

In the population of developing countries, cervical cancer incidence ranks third among all tumors and first among gynecological tumors.<sup>1</sup> The rising incidence of cervical cancer is related to the following factors: early sexual activities, group sex, infrequent use of condoms, multiple pregnancies with Chlamydia infection, and immunosuppression with HIV.<sup>2</sup> Effective diagnosis in the early stage can significantly reduce the incidence of cervical cancer.<sup>3</sup> It has been shown that timely treatment of dysplasia can effectively reduce the incidence and mortality of cervical cancer.<sup>4</sup> However, the molecular mechanism of cervical cancer has not been fully understood. If lesions can be detected early on the micro level, the prevention and treatment of cervical cancer will become more efficient.

Alternative splicing (AS) allows selective removal of introns and ligation of exons, two processes during which multiple isoforms of mRNA are produced and

Correspondence: Yanli Zheng  
Department of Obstetrics and  
Gynecology, Nantong First People's  
Hospital, Nantong, Jiangsu, 226001,  
People's Republic of China  
Tel/Fax +86-513-85051875  
Email gaoshan1189@sina.com

post-transcriptional proteome diversified.<sup>5,6</sup> AS is currently categorized into 7 transition types: Alternate Acceptor site (AA); Alternate Donor site (AD); Alternate Promoter (AP); Alternate Terminator (AT); Exon Skip (ES); Retained Intron (RI); and Mutually Exclusive Exons (ME). A growing number of studies have shown that tumorigenesis was associated with the abnormal expression of genes and proteins which were induced by AS.<sup>7–12</sup> Referring to different locations and effect on the usage of a splice site, cis-regulatory sequences could be classified into exonic splicing enhancers, exonic splicing silencers, intronic splicing enhancers and intronic splicing silencers, which determined their affinities to cognate splicing factors (SFs).<sup>13</sup> Differences in gene expression levels of splicing regulatory factors have been observed in many cancers, and these proteins often affect the splicing patterns of many genes that function in certain cancer-specific biological pathways, including cell cycle progression, cellular proliferation, migration, and RNA processing.<sup>14,15</sup>

In the current study, integrated bioinformatics analysis of the AS and gene expression profiles was conducted to identify the splicing events that have prognostic value in CESC. The Univariate Cox proportional hazards regression, LASSO Cox regression and multivariate Cox proportional hazards regression models were performed to screen for the potential prognostic splicing events and genes, uncovering a number of survival-associated AS events. Regulatory network of between AS events and SFs was constructed to elucidate the underlying mechanisms. However, the molecular mechanism of AS in cervical cancer has not been elucidated. This study aimed to screen AS-generated genes related to CESC and provide new candidate molecules for targeted therapy.

## Materials and Methods

### Process of AS Data Curation

TCGA data portal (<https://portal.gdc.cancer.gov/>) provided the RNA sequencing data and clinical information of CESC cohorts. mRNA splicing profiles in CESC was analyzed with the aid of SpliceSeq, a java program that explicitly quantifies RNA-Seq reads and identifies its possible functional changes as a consequence of AS in the context of transcript splice graphs. We downloaded the percent spliced in (PSI) value of seven 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). PSI (range: 0–1) is a common value for scoring AS events.

### Survival-Associated AS Events in CESC

A total of 306 CESC tissues and were analyzed in this study and the overall survival (OS) of the 306 CESC patients were at least 90 days. We used UpSetR (version 1.3.3)<sup>16</sup> to present the intersections between seven types of AS. Interactive sets among the seven types of AS events were visualized by UpSet plot. Patients were divided into the low-risk group and the high-risk group based on the median levels of risk score.

Univariate Cox regression analysis was conducted to identify survival-associated AS events ( $P < 0.05$ ). Gene Ontology (GO) function and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways involving the two lists of genes were determined using an online tool Enrichr database (<http://amp.pharm.mssm.edu/Enrichr/>).<sup>17</sup> The biological functions of the overlap genes were interpreted.  $P < 0.05$  was considered statistically significant. Since PPI (protein–protein interaction) analysis help identify the hub genes with core functions, we utilized STRING (Search Tool for the Retrieval of Interacting Genes Database) (<http://www.string-db.org/>) to evaluate PPI of hub modules.<sup>18</sup> To visualize the findings, we exported the results from STRING database to Cytoscape software.<sup>19</sup> The topological analysis method Degree and the centrality analysis methods Closeness and Betweenness were used separately to identify the hub nodes in the PPI network. Among the top 10 hub nodes identified by each method, only genes with consistent high Degree, Closeness, and Betweenness values (larger than the median value) were selected as hub genes. MCODE plug-in was employed for searching clustered sub-networks. Default parameters were set: node score cut-off = 0.2, k-core = 2, degree cut-off = 2, and max. depth = 100. Enrichr was conducted for functional analysis of hub clusters

### Prognostic Models Featured by AS Events for CESC

The result of univariate Cox regression analysis in identifying survival-associated AS events was also shown by Volcano plot using “ggplot2” in R. Seven types of AS events were revealed by Bubble chart respectively. Each chart contained the top 20 significant survival-associated AS events of corresponding types. Following univariate Cox, we used the least absolute shrinkage and selection operator

(LASSO) Cox analysis. LASSO method is used to calculate the regression of high-dimensional predictors.<sup>20</sup> In Cox regression survival analysis, LASSO is ideal for high-dimensional data. Then, we used multivariate Cox regression to construct prognostic feature for each AS type. The prognostic model was constructed based on a linear integration of the expression level multiplied regression model ( $\beta$ ) using the following formula: Risk score =  $\beta_{\text{gene1}} \times \text{expr}_{\text{gene1}} + \beta_{\text{gene2}} \times \text{expr}_{\text{gene2}} + \dots + \beta_{\text{genen}} \times \text{expr}_{\text{genen}}$ . Patients were then divided into two groups based on the median levels of risk score. This prognostic model and patient survival information were merged. Kaplan–Meier survival curves were used to compare the prognostic ability of prediction models. Area under the curve (AUC) value for the ROC curves of each prognostic model was calculated by survival ROC package in R. Besides, univariate and multivariate Cox regression analysis were performed to compare the hazard ratios (HRs) of prognostic models and important clinical features for CESC. Finally, Chi-square tests were used to compare the tumor status, age, grade, clinical stage and histological type of the two risk groups. Risk analysis related to tumor-infiltrating immune cells was also performed.

## Correlation Network of SFs and Survival-Associated AS Events

The data of SFs was obtained from SpliceAid 2.<sup>21</sup> The expression profiles of SF genes were curated from the TCGA dataset and further coped with univariate Cox analysis. We selected axes between the expression value of SFs and PSI values of prognosis-related AS events to construct the SF-AS regulatory network according to the following conditions:  $P < 0.05$ , Pearson correlation coefficient  $> 0.5$ . Then, we built the correlation plots via Cytoscape version 3.6.1. Univariate and multivariate Cox regression analyses and LASSO Cox regression analysis were further performed for SF genes.

## Results

### Association Between AS Events and Cervical Cancer Survival

In total, 41,426 AS events were detected from 19,696 genes of 256 CESC patients and classified into seven AS types (Table 1). The events were mainly enriched in ES, followed by AT and AP. In detail, we detected 15,942 ESs in 6271 genes, 8395 ATs in 3663 genes, 8066 APs in 3256 genes, 3424 AAs in 2397 genes, 3017 ADs in 2106 genes, 2373 RIs

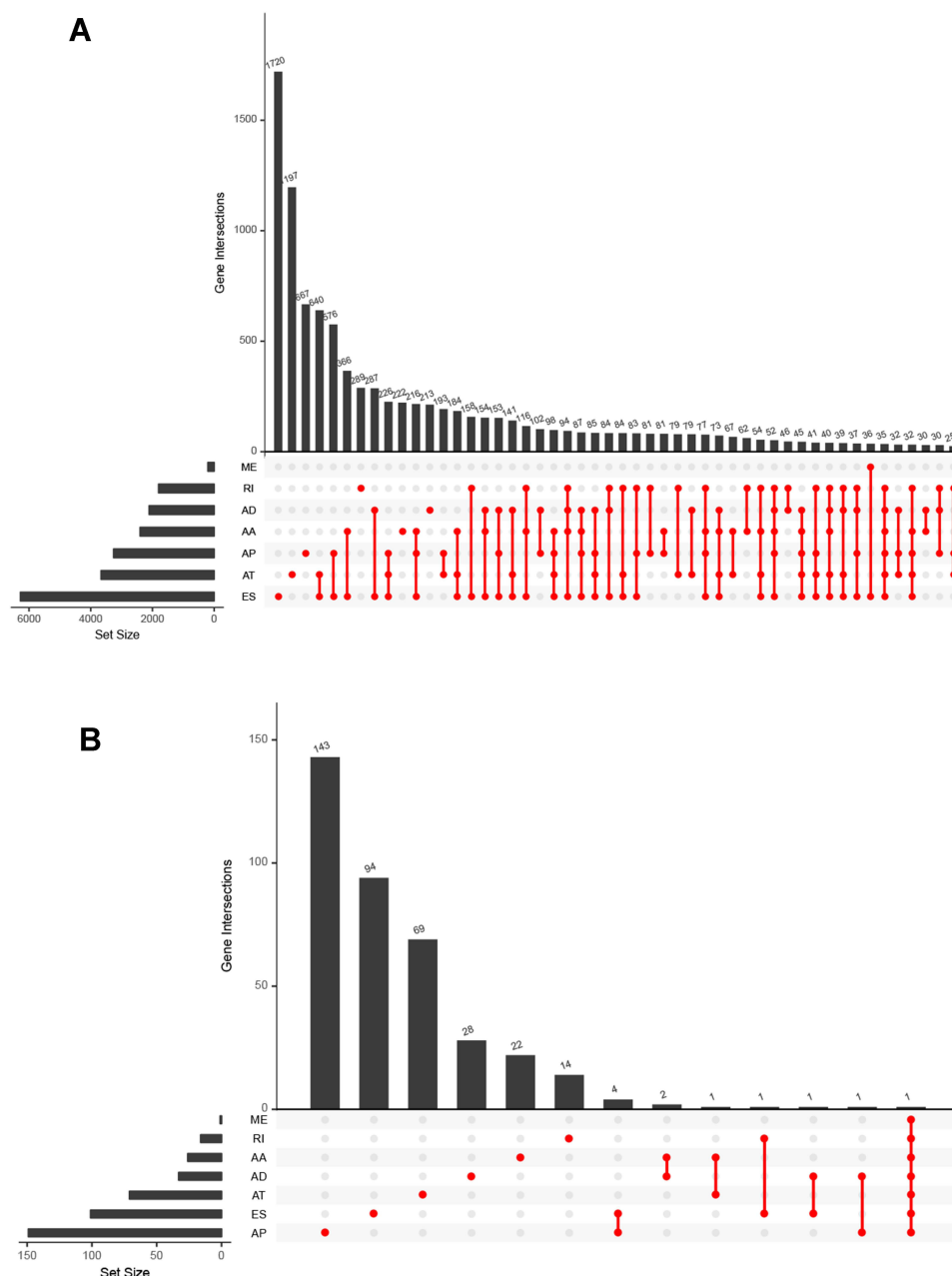
**Table 1** AS Events Were Detected and Classified into Seven as Types

Type	AS Events	Genes
AA	3424	2397
AD	3017	2106
AP	8066	3256
AT	8395	3663
ES	15942	6271
ME	209	202
RI	2373	1801
Total	41426	19696

in 1801 genes, and 209 MEs in 202 genes (Figure 1A). Overall, the results reveal that one gene might possess up an average of 2.5 AS events. To explore the relationship between AS events and OS of CESC patients, all AS events mentioned above were subjected to the univariate Cox regression analysis, and 3708 survival-associated AS events in seven AS types in 3355 genes (Table 2) which we identified as significant AS events (Figure 1B) were screened out.

### Enrichment Analysis of AS-Related Genes in Cervical Cancer

Since we identified 3708 survival-associated AS events in seven AS types in 3355 genes, to explore possible pathways for these AS events, the parent genes of survival-associated AS events were subjected to functional enrichment analyses. GO analysis was performed to investigate the biological functions and pathways enriched in these genes as well as the interaction network among them. In the aspect of biological processes, the genes were mostly enriched in Golgi to lysosome transport, negative regulation of mitotic nuclear division and negative regulation of nuclear division (Figure S1A). In the aspect of cellular component, the genes were mostly enriched in lysosome, mitochondrion and lysosomal lumen (Figure S1B). In the aspect of molecular functions, the genes were mostly enriched in death domain binding, nucleoside-diphosphatase activity and GTPase activator activity (Figure S1C). KEGG analysis showed that the genes were mostly enriched in glycosaminoglycan degradation, glyoxylate and dicarboxylate metabolism and NF-kappa B signaling pathway (Figure S1D). Molecular Complex Detection (MCODE) was used to screen out the modules in the protein-to-protein network. PPI network propped by these genes revealed that the AS events were engaged in nine MCODE components (Figure S2A).



**Figure 1** (A) UpSet plots of AS events in CESC (B) UpSet plots of top significant AS events in CESC. The horizontal axis and vertical axes represent number of genes in the interacting sets of one or multiple AS types and number of genes in each AS type, respectively.

Among these components, the top related cluster was composed of 12 nodes and 36 edges (Figure S3A), the second of 9 nodes and 17 edges and the third of 4 nodes and 6 edges (Figure S3B and C). The top related cluster was transferred to functional analysis. In the aspect of biological process, the genes were mostly enriched in regulation of oxidative stress-induced neuron intrinsic apoptotic signaling pathway, regulation of osteoclast development and regulation of DNA endoreduplication (Figure S4A). In the aspect of cellular component, the

genes were mostly enriched in AP-2 adaptor complex clathrin coat of endocytic vesicle and endolysosome membrane (Figure S4B). In the aspect of molecular function, the genes were mostly enriched in endocytic adaptor activity, clathrin adaptor activity and nitric-oxide synthase binding (Figure S4C). KEGG analysis showed that the genes were mostly enriched in ubiquitin mediated proteolysis, endocytosis and endocrine and other factor-regulated calcium reabsorption (Figure S4D). The top 10 hub genes by PPI network were screened according to their degree



**Table 2** 3708 Survival-Associated as Events in Seven as Types in 3355 Genes

Type	AS Events	Genes
AA	319	310
AD	289	279
AP	698	653
AT	658	620
ES	1471	1237
ME	14	14
RI	259	242
Total	3708	3355

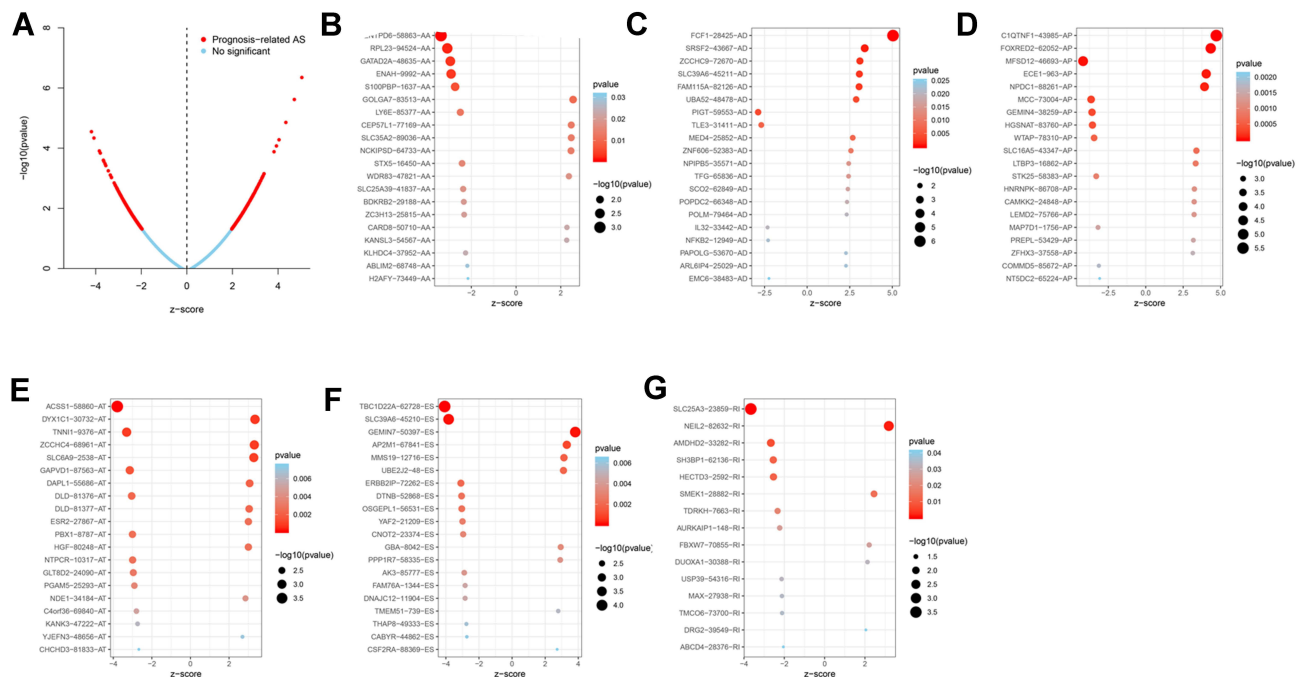
values, which were significantly associated with AS events (Figure S2B). Among them, UBA52 connecting 38 nodes showed the most prominence. The distributions of survival-associated AS events are displayed in Figure 2A. Among them, the MEs have no statistical significance. The 20 most significant prognosis-related AS events were also shown (Figure 2B–G).

## Prognostic Models Constructed with Seven Types of AS Events for CESC

To investigate the relationship between each AS event and the prognosis of CESC patients, we constructed proportional hazards regression analysis on genes related to AS events. With a cutoff of  $p < 0.05$ , 402 genes with possible

prognostic significance were screened out, including 33 of AD, 155 of AP, 103 of ES, 71 of AT, 15 of RI, and 25 of AA (Table S1).

These prognosis-related AS-event genes were further analyzed with the least absolute shrinkage and selection operator (LASSO) Cox regression algorithm (Figure S5). Then, multivariate Cox proportional hazards regression analysis was further performed to build the risk signature. Using the seven types of AS events, we constructed seven prognostic models. The risk scores were calculated (Table 3). Based on the risk score, CESC patients were divided into low-risk group and high-risk group based on the median levels of risk score. In each type of model, Kaplan–Meier survival analysis indicated that low-risk patients had significantly better overall survival than high-risk patients (Figure 3). We also performed the receiver operating characteristic curve (ROC) analysis. As shown in Figure 4, PI-AA demonstrated the highest capacity of estimating the prognosis of CESC patients (AUC 0.92), followed by PI-RI (AUC 0.871). The risk score, survival status and the expression of 7 types of AS-event genes in prognostic model for each patient was displayed in Figures S6 and 7, Figure 5. To assess whether these models were the independent predictor of CESC, univariate and multivariate Cox regression analyses were performed, including clinical factors and risk scores. The results showed that



**Figure 2** Top 20 most significant AS events in CESC. (A) The red dots represent AS events that are significantly correlated with patient survival. The blue dots represent AS events without correlation. The top 20 AS events correlated with clinical outcome based on acceptor sites (B) AA (C) AD (D) AP (E) AT (F) ES (G) RI.

**Table 3** Seven Prognostic Models Based on 7 Types of as Events

Type	Formula	AUC
AA	-ENTPD6 58863 AA*9.39-GATAD2A 48635 AA*13.54-ENAH 9992 AA*4.58-S100BPB 1637 AA*6.26-LY6E 85377 AA*2.75+SLC35A2 89036 AA*8.23+NCKIPSD 64733 AA*2.08-STX5 16450 AA*24.42-SLC25A39 41837 AA*4.14-BDKRB2 29188 AA*1.61+KANSL3 54567 AA*4.64-KLHDC4 37952 AA*4.60-H2AFY 73449 AA*14.92	0.92
AD	FCFI 28425 AD*3.04+ZCCHC9 72670 AD*5.08+FAM115A 82126 AD*2.93-PIGT 59553 AD*2.94-TLE3 31411 AD*3.72+NPIPB5 35571 AD*4.39+TFG 65836 AD*3.56+POPDC2 66348 AD*2.60+POLM 79464 AD*6.22-NFKB2 12949 AD*4.29+PAPOLG 53670 AD*3.18	0.792
AP	C1QTNF1 43985 AP*2.11+FOXRED2 62052 AP*2.54-STK25 58383 AP*2.23+HNRNPK 86708 AP*8.33+PREPL 53429 AP*1.93-COMMD5 85672 AP*6.51-NT5DC2 65224 AP*1.89	0.807
AT	DYX1C1 30732 AT*10.05-TNNI1 9376 AT*10.37+ZCCHC4 68961 AT*5.06-GAPVD1 87563 AT*40.73+DAPL1 55686 AT*1.07+HGF 80248 AT*1.71-GLT8D2 24090 AT*16.48-PGAM5 25293 AT*16.41-C4orf36 69840 AT*5.67+YJEFN3 48656 AT*3.40-CHCHD3 81833 AT*2.82	0.826
ES	-TBC1D22A 62728 ES*27.75-SLC39A6 45210 ES*15.41+GEMIN7 50397 ES*5.26+AP2M1 67841 ES*5.23-ERBB2 72262 ES*1.89-YAF2 21209 ES*12.61+GBA 8042 ES*38.51-FAM76A 1344 ES*1.63-DNAJC12 11904 ES*2.30+TMEM51 739 ES*1.79-THAP8 49333 ES*5.57	0.849
RI	-SLC25A3 23859 RI*3.66+NEIL2 82632 RI*3.33-AMDHD2 33282 RI*3.69-SH3BP1 62136 RI*18.95+SMEK1 28882 RI*2.94-TDRKH 7663 RI*2.27+DUOXA1 30388 RI*6.54-MAX 27938 RI*2.34-TMCO6 73700 RI*3.24+DRG2 39549 RI*2.17-ABCD4 28376 RI*2.76	0.871
All	FCFI 28425 AD*4.32+C1QTNF1 43985 AP*2.10-TBC1D22A 62728 ES*17.75+ECE1 963 AP*3.33-SLC39A6 45210 ES*20+GEMIN7 50397 ES*4.63-ACSS1 58860 AT*4.59-ENTPD6 58863 AA*5.89+DYX1C1 30732 AT*10.38	0.865

this prognostic model showed moderate and independent prognostic power for all seven AS events (Figures 6 and 7).

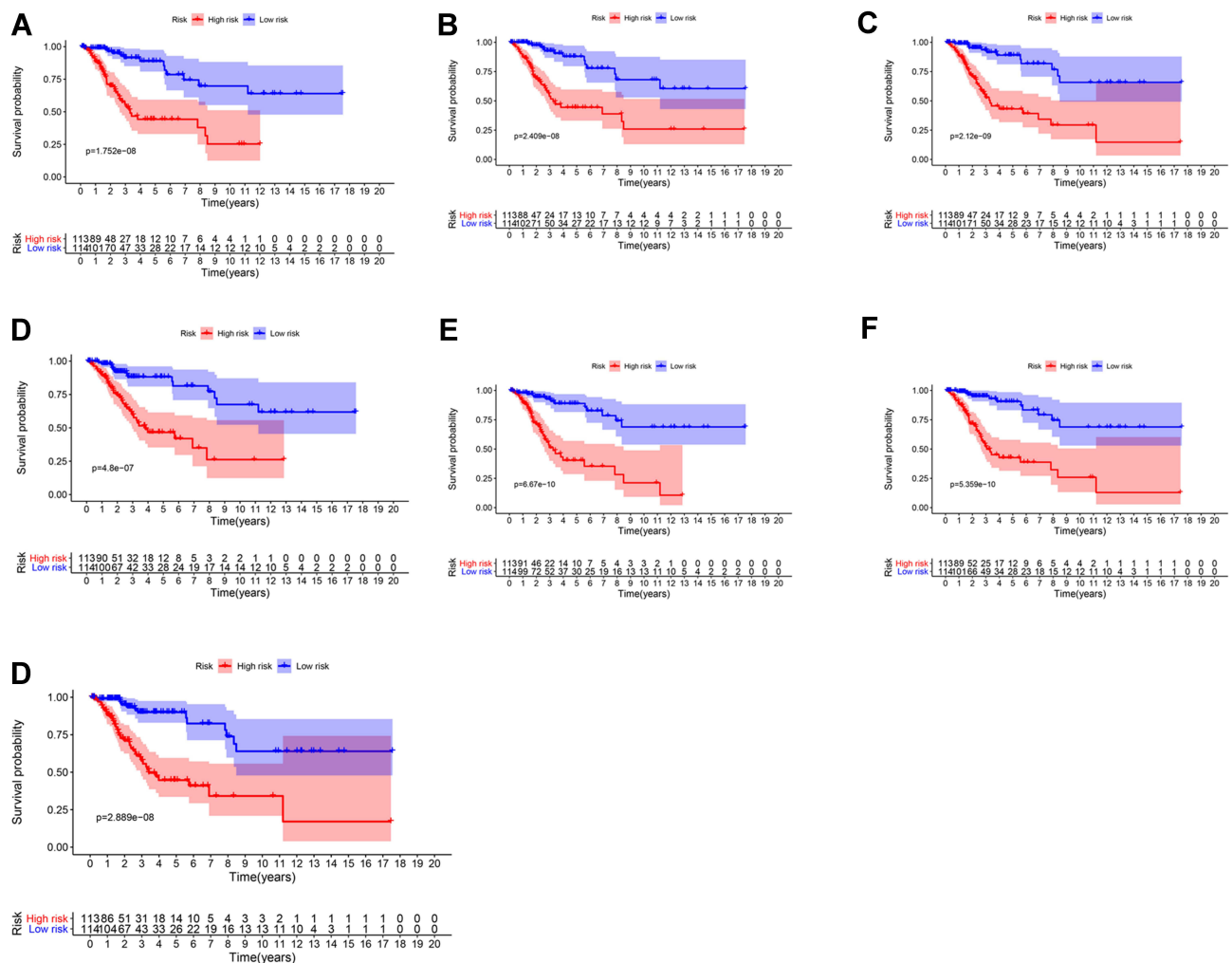
## Association Between AS Events' Risk Scores and Clinical Features

We analyzed the association between AS event risk score and the clinical features of CESC (Figure S8). Interestingly, we found that all the AS genes had significant correlation with tumor status in both the high- and low-risk groups. Among them, AT-related genes were associated with tumor status and grade, RI-related genes with tumor status and stage.

## Association Between Splicing Factors and AS Events

SFs play an important part in the occurrence and development of AS events via changing exon selection and splicing site. Therefore, it is necessary to uncover the correlation between SFs and AS events. We got 385 SFs from SpliceAid 2, and downloaded the expression profiles of these 385 genes from the TCGA database. Then, we used the Pearson's correlation matrices to calculate the relationship between the SFs and AS-events.

With a cutoff of  $p < 0.05$  and  $|\text{Cor}| > 0.5$ , we constructed the SF-AS regulatory network with 43 significant SFs (Figure 8A) (Table S3). The relationship between 43 SFs and CESC prognosis was determined with univariate Cox proportional hazard regression analyses. With a cutoff of  $p < 0.05$ , seven genes (SNRPA, PQBP1, CCDC12, SNRPF, CELF2, HSPA8 and TXNL4A) were screened out (Table S2). Then, seven genes were further submitted to LASSO Cox regression algorithm. The optimal values of the penalty parameter lambda were determined through 10-times cross-validations. The simplest (smallest parameter) model of prognostic genes signature within one standard error of the best lambda value was screened out (Figure 8B and C). Finally, multivariate Cox proportional hazards regression analysis screened out three genes (SNRPA, CELF2 and HSPA8) (Figure 8D). Risk score =  $-0.033 * \text{SNRPA} - 0.34 * \text{CELF2} + 0.003 * \text{HSPA8}$ . Based on the median levels of risk score, CESC patients were divided into low-risk and high-risk groups. Kaplan–Meier survival analysis indicated that low-risk group had better overall survival than high-risk group (Figure 9A). In ROC curve analysis, risk score of 0.685 reaped the highest specificity and sensitivity in predicting the 5-year survival (Figure 9B).



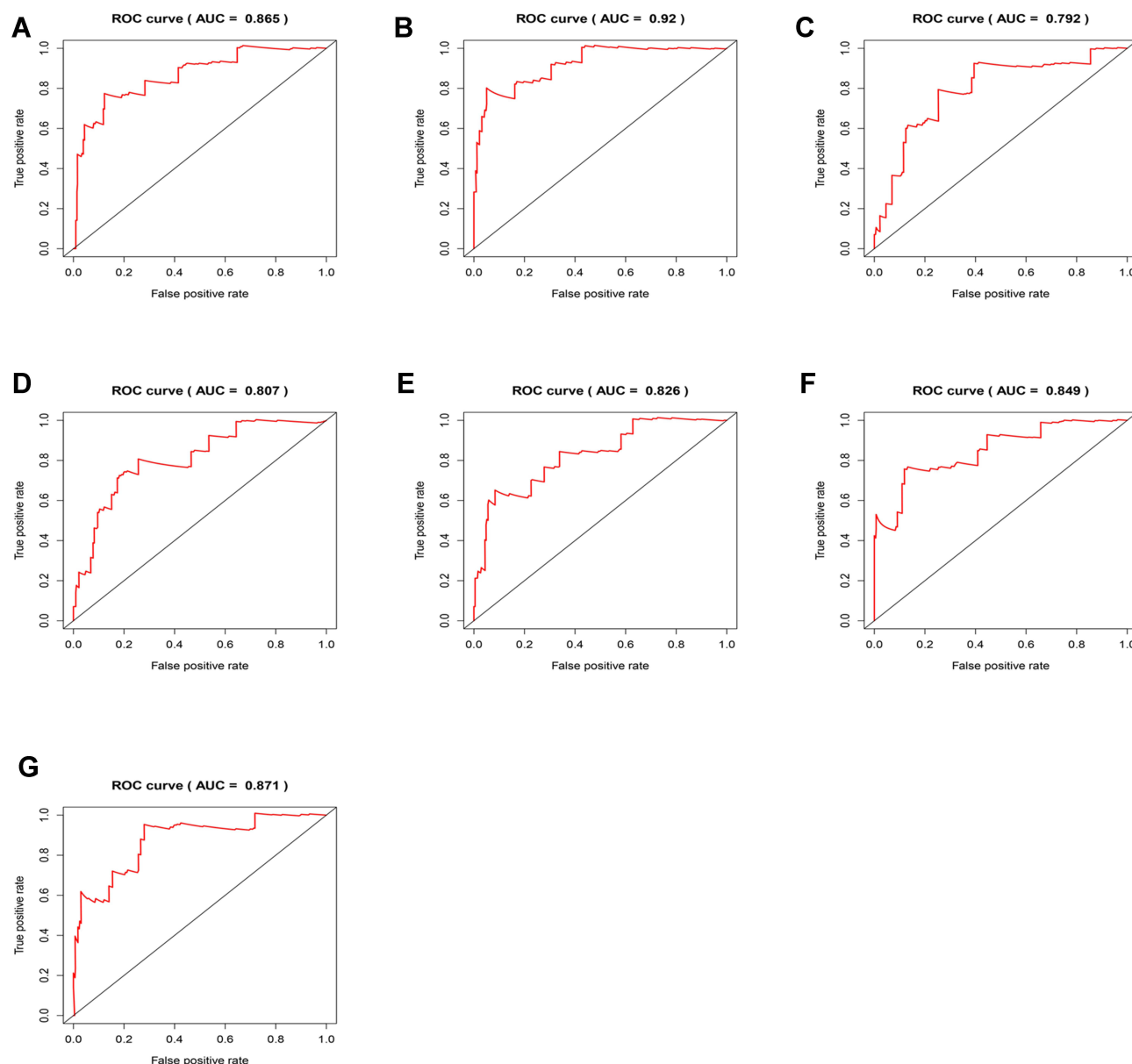
**Figure 3** Survival analysis based on seven types of AS-event genes in TCGA cohort. (A) ALL (B) AA (C) AD (D) AP (E) AT (F) ES (G) RI.

Meanwhile, we analyzed the association between the clinical parameters and the risk score of three genes. The univariate and multivariate Cox proportional hazards regression showed that only the risk score calculated based on three genes was an independent prognostic indicator of CESC (Figure 9C and D). Of the three genes, SNRPA and CELF2 were protective genes to CESC. Interestingly, the AS most closely related to CELF2 was TC2N, type AP. Our analysis showed that TC2N was negatively correlated with CELF2, and univariate Cox proportional hazard regression analysis showed that TC2N was a causative factor, and the two results were consistent. Similarly, CALCOCO2, PLS3 and FAM76A developed from ES, NGLY1 from AP, COX15 and PPOX from AT; all these AS-developed genes were positively correlated with SNRPA. The results of univariate Cox proportional hazard regression analysis showed that these AS-developed genes were protective factors.

PLEKHB2, PPP1CC and GALNT4 developed from AP, LYRM5 and NPIPA7 from AT, and these genes were negatively correlated with SNRPA. Univariate Cox proportional hazard regression analysis proved that they were all pathogenic factors.

## Association Between AS Event and Tumor Immune Microenvironment

In order to explore the role of AS-related genes in the tumor immune microenvironment, We divide the samples into high- and low-risk groups, combined with comprehensive analysis of immune cells. The correlation of risk score and tumor immune microenvironment was shown in (Figure 10A). Figure 10B–G indicated that dendritic cells activated, dendritic cells resting, Macrophages M0, Mast cells resting, T cells CD4 memory activated and T cells CD8 were most correlative with risk core.



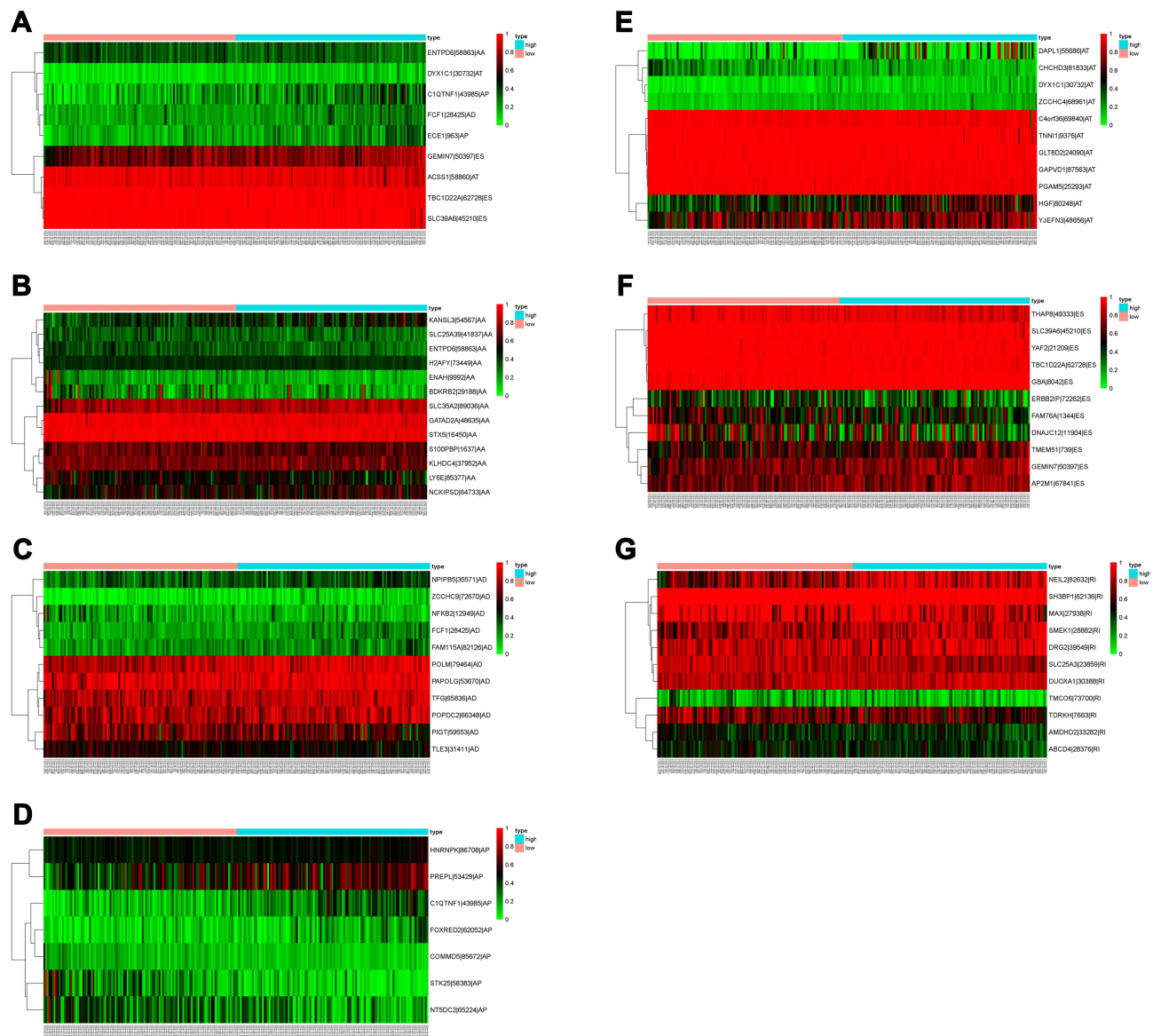
**Figure 4** ROC analysis based on seven types of AS-event genes. (A) ALL (B) AA (C) AD (D) AP (E) AT (F) ES (G) RI.

## Discussion

AS during mRNA editing diversifies the transcription of proteins. Studies have verified the regulatory role of variable scission in the development of tumors.<sup>12,22,23</sup> In the present study, we shifted our focus to cervical cancer.

In this paper, we firstly investigate the performance of AS events in CESC based on the TCGA data. We divided AS events into seven types, established CESC prognostic models using AS-related genes, and screened those with the strongest correlation with CESC. GO analysis showed that those genes were mostly enriched

in Golgi to lysosome transport, lysosome and death domain binding. KEGG analysis showed that those genes were mostly enriched in glycosaminoglycan degradation and NF-kappa B signaling pathway. Perera et al have found that lysosome drives pancreatic cancer metabolism.<sup>24</sup> Li et al have found that glycosaminoglycan can mark lung cancer.<sup>25</sup> Studies have not yet probed into mechanism of Golgi to lysosome transport and death domain binding in tumorigenesis. The NF-kappa B signaling pathway engages tumorigenesis with inflammation.<sup>26</sup> Tang et al found TNF-Alpha promoted the invasion and metastasis via NF-Kappa B pathway in



**Figure 5** Heatmap of the 7 types of AS-event genes expression profiles. (A) ALL (B) AA (C) AD (D) AP (E) AT (F) ES (G) RI.

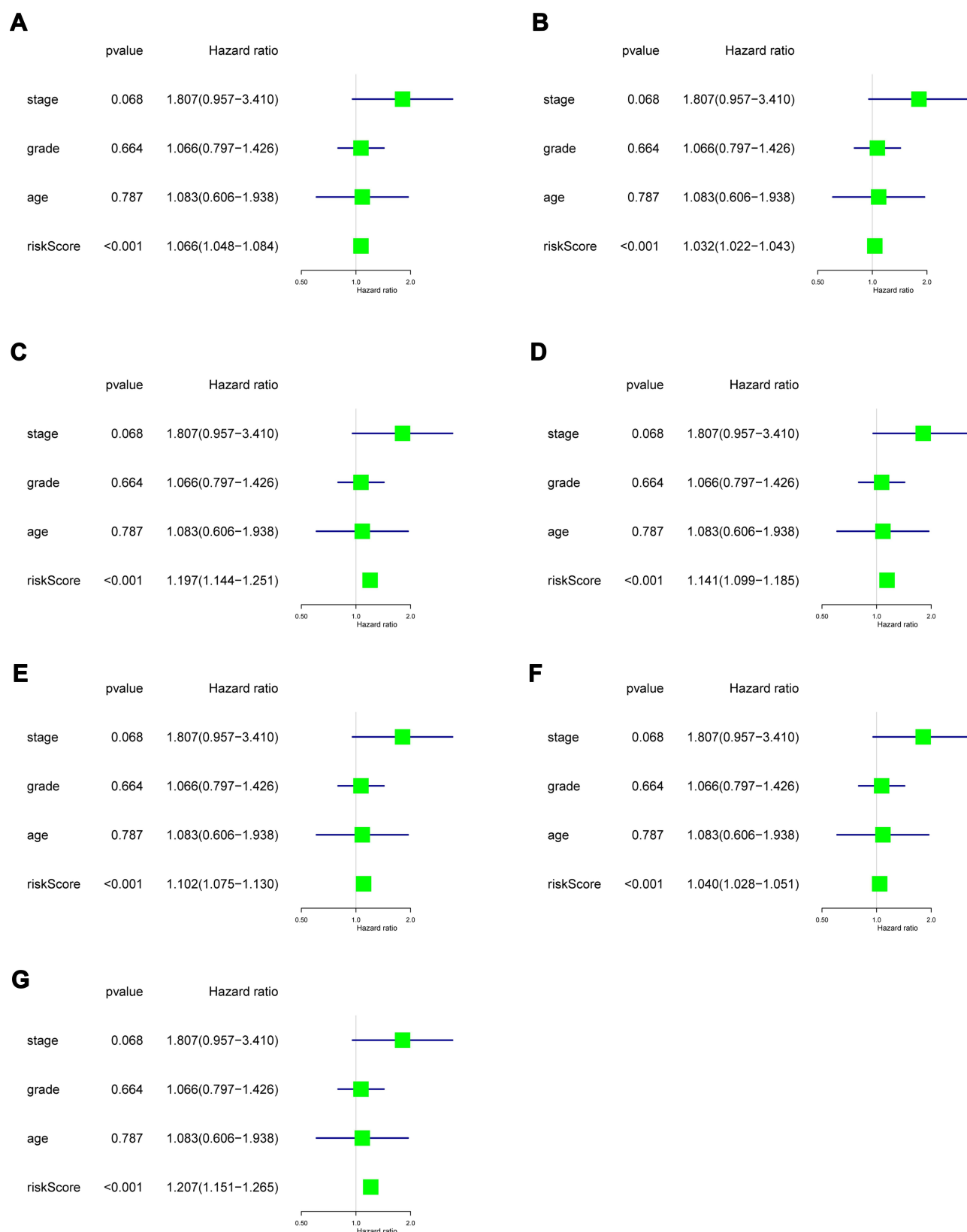
oral squamous cell carcinoma.<sup>27</sup> Pacifico et al found that NF-kappa B pathway also acted in human thyroid carcinomas.<sup>28</sup> AS events in our model are mainly active in proliferation, migration, apoptosis and tumor immune-related pathways, which may indicate that abnormal AS mainly affects tumor biological processes through these pathways.

We next analyzed the core modules in the PPI network. GO analysis showed that the core AS-related genes were mostly enriched in regulation of oxidative stress-induced neuron intrinsic apoptotic signaling pathway, AP-2 adaptor complex, and endocytic adaptor activity. KEGG analysis showed that these core genes

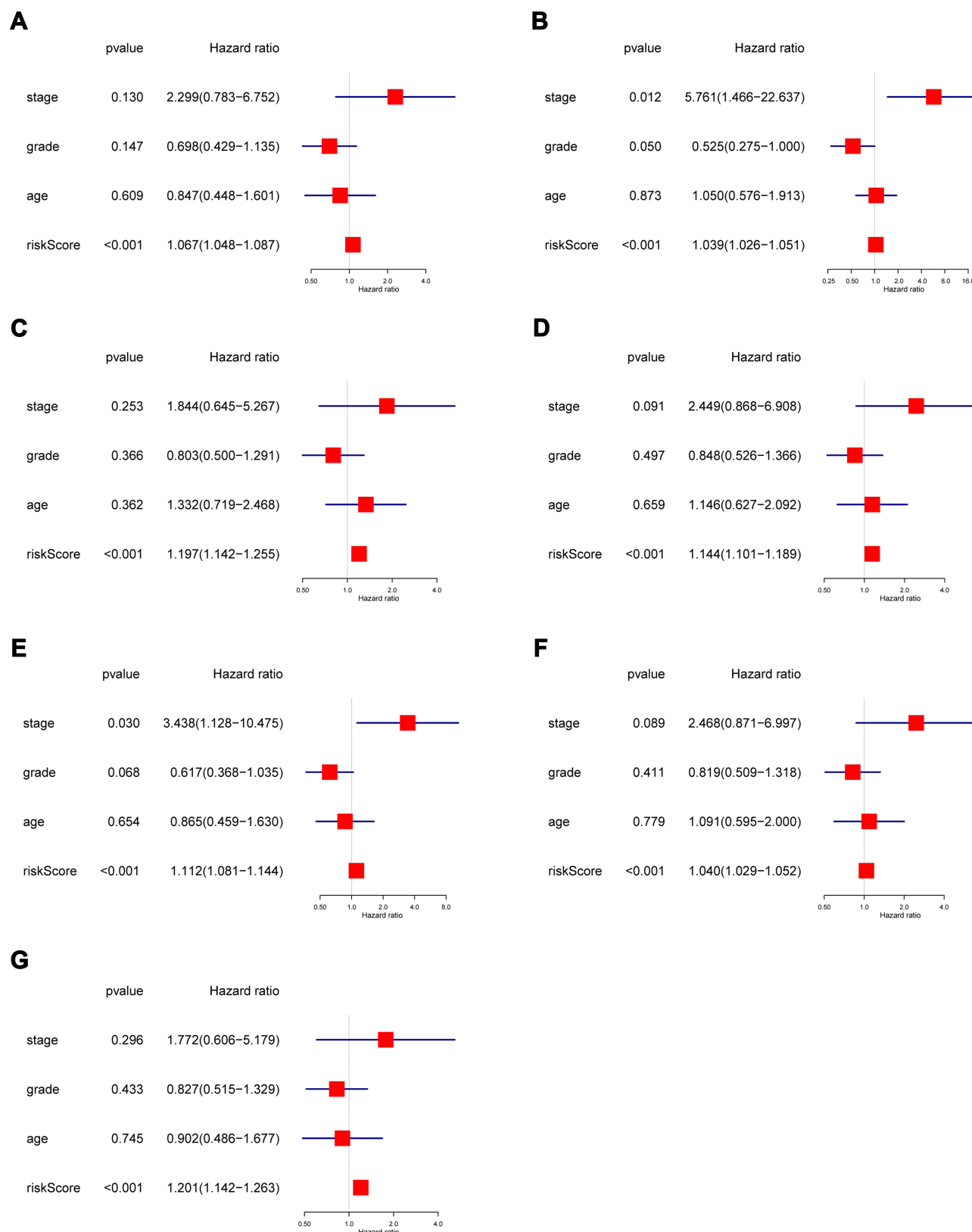
were mostly enriched in ubiquitin mediated proteolysis, endocytosis and endocrine and other factor-regulated calcium reabsorption. Among these AS-related genes, UBA52 showed the strongest intimacy with CESC. Instructive is that the biological terms we teased out were all related to tumor diseases. For example, ubiquitin-mediated proteolysis plays in the regulation of clear cell renal cell carcinoma.<sup>29</sup> Zhou et al have found that UBA52 regulates colorectal cancer tumorigenesis.<sup>30</sup> Our results demonstrate a range of genes and pathways pertaining to CESC.

We also found that each type of AS event could specify certain features of CESC patients. These AS

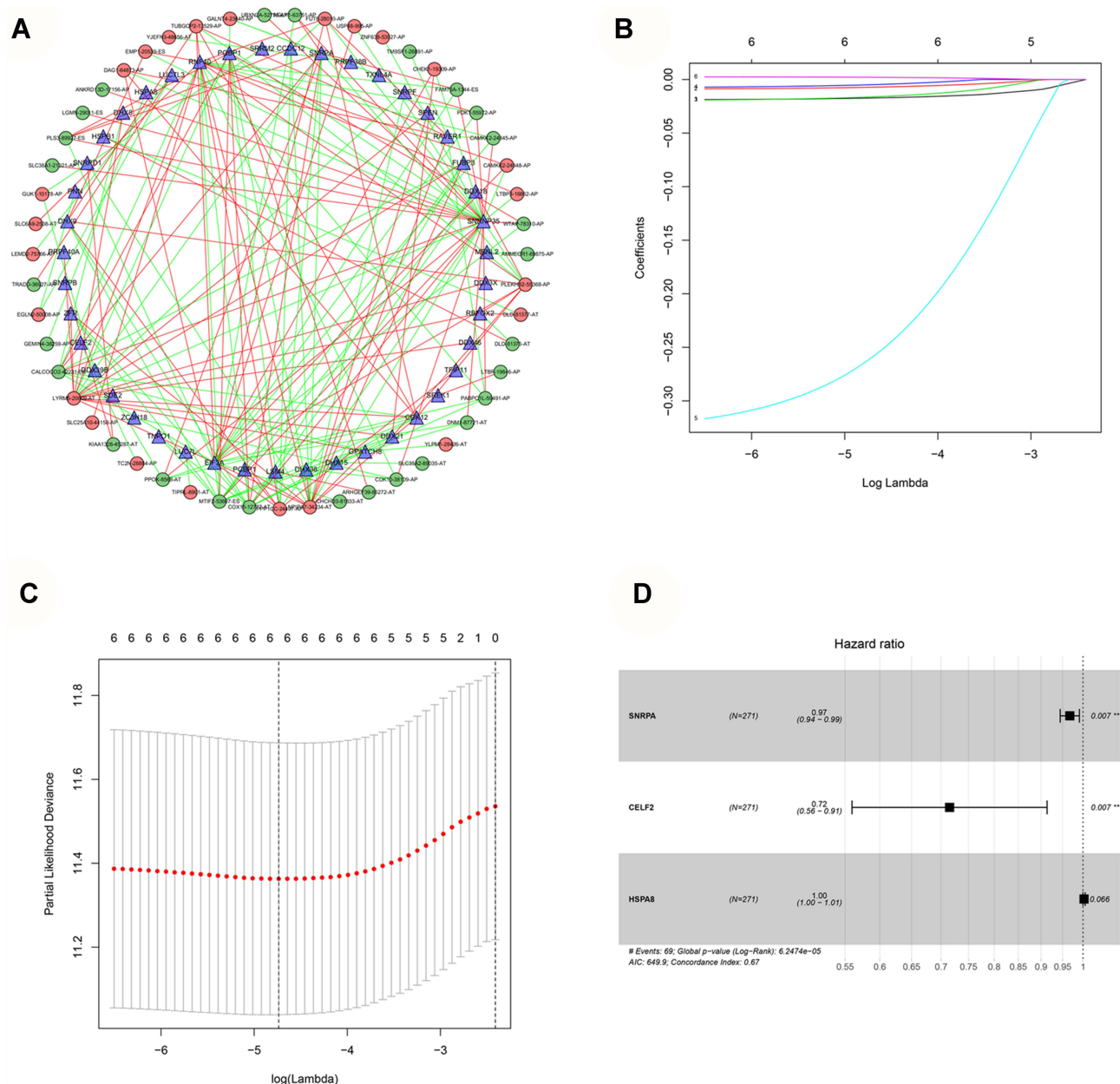




**Figure 6** Univariate Cox regression analysis was performed including clinical factors and risk score based on 7 types of AS-event genes. **(A)** ALL **(B)** AA **(C)** AD **(D)** AP **(E)** AT **(F)** ES **(G)** RI.



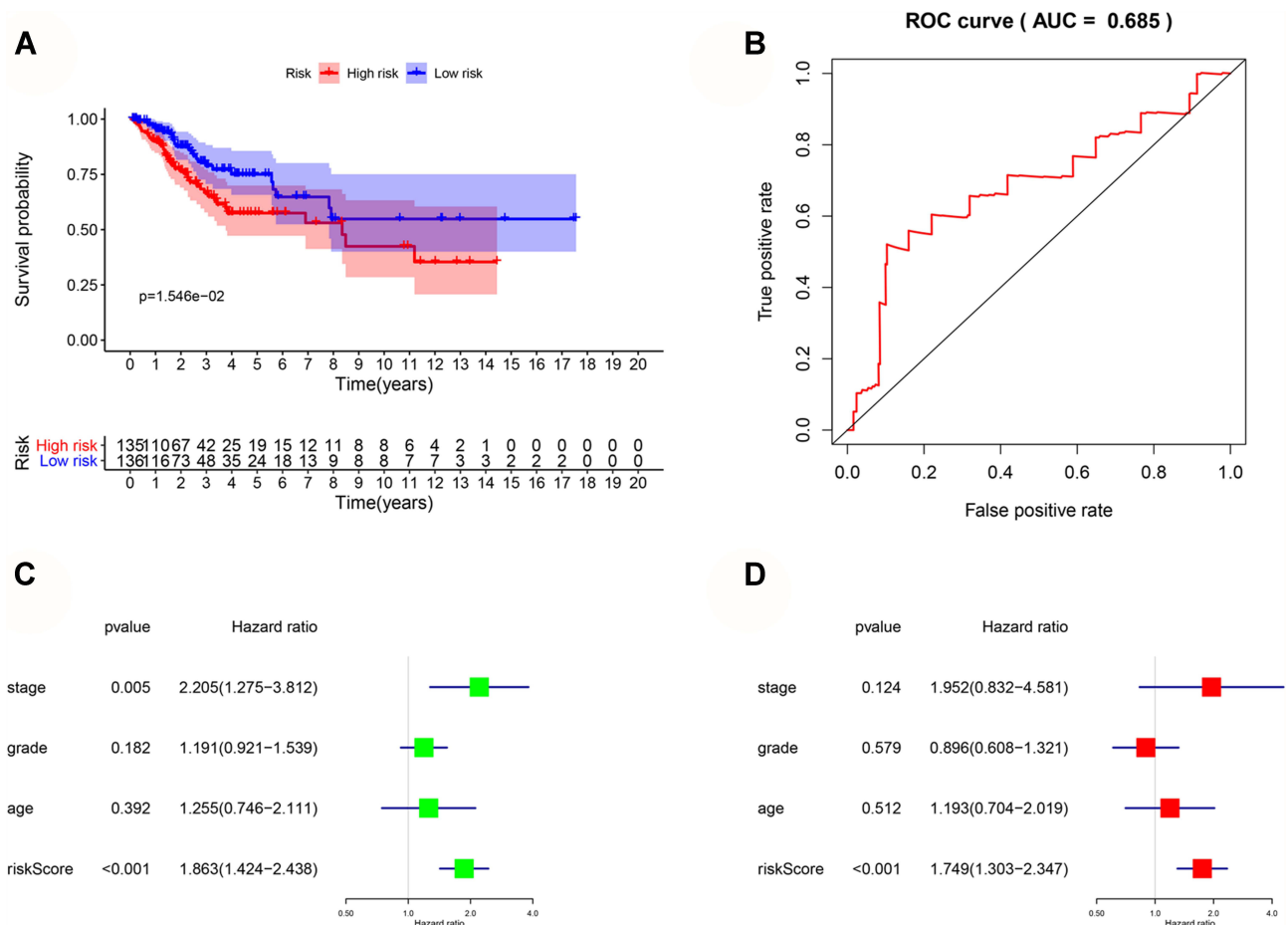
**Figure 7** Multivariate Cox regression analysis was performed including clinical factors and risk score based on 7 types of AS-event genes. (A) ALL (B) AA (C) AD (D) AP (E) AT (F) ES (G) RI.



**Figure 8** (A) Correlation network between significant survival-associated SFs and significant survival-associated AS events. The connecting line represents the association between the SF and the AS event. Triangles represent SF, dots represent AS events, red dots and red connecting lines represent AS events positively correlated with SF, green dots and green connecting lines represent AS events negatively correlated with SF. (B and C) Construction of prognostic signatures based on LASSO COX analysis target significant survival-associated SFs. (D) Multivariate Cox proportional hazards regression analysis based on the hub significant survival-associated SFs.

events, either combined or single, performed well in predicting overall survival time of CESC patients (AUC > 0.8). Kaplan–Meier survival curves also proved that all the AS-events-based models could efficiently predict the risks and survival outcome. Meanwhile, each prognostic model could serve as an independent prognostic factor. To be more specific, Prognostic models were established with nine genes

screened out from the model that incorporated all AS events, including FCF1|28425|AD, C1QTNF1|43985|AP, TBC1D22A|62728|ES, ECE1|963|AP, SLC39A6|45210|ES, GEMIN7|50397|ES, ACSS1|58860|AT, ENTPD6|58863|AA and DYX1C1|30732|AT. The prognostic efficiency of these prognostic models verified that the nine genes could be taken as independent prognostic factors for CESC. Han et al have found



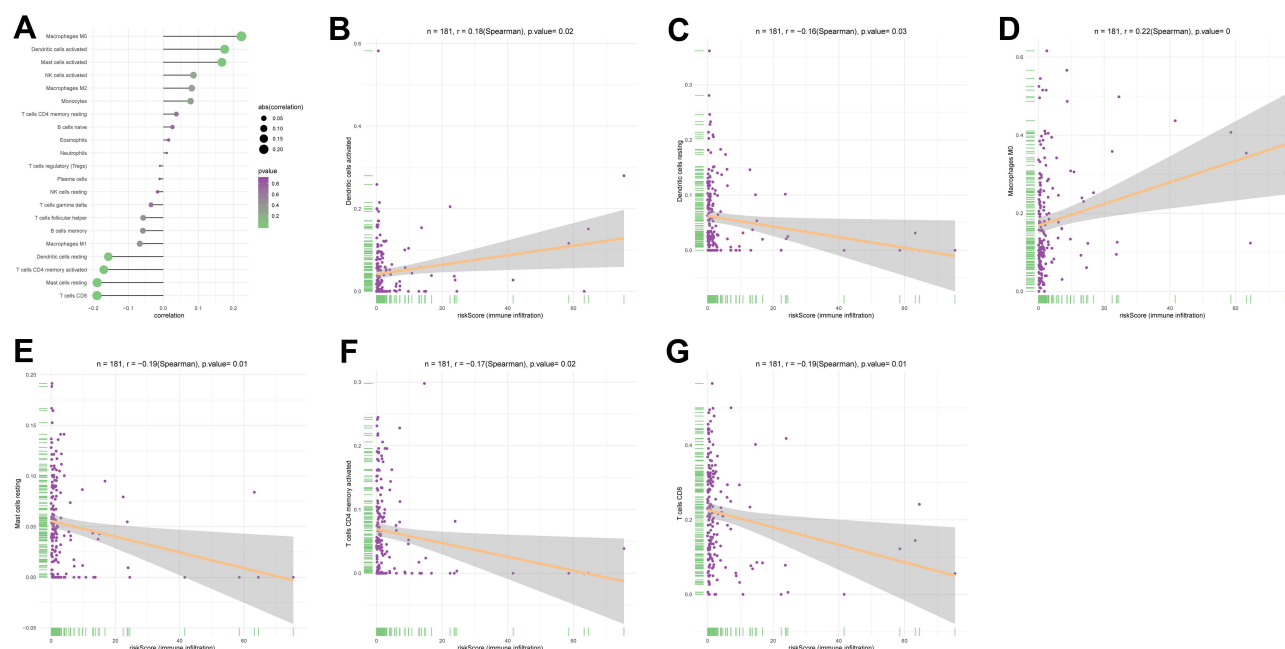
**Figure 9** (A) Survival analysis of hub significant survival-associated SFs. (B) ROC analysis of hub significant survival-associated SFs. (C) Univariate Cox regression analysis was performed including clinical factors and risk score based on hub significant survival-associated SFs. (D) Multivariate Cox regression analysis was performed including clinical factors and risk score based on hub significant survival-associated SFs.

that C1QTNF1 affects the progression of human hepatocellular carcinoma.<sup>31</sup> Xu et al have found that ECE-1 overexpression in worsens tumor differentiation and clinical outcome of head and neck cancer and colon cancer.<sup>32</sup> SLC39A6 functions in hepatocellular carcinoma and esophageal carcinoma.<sup>33,34</sup> Yun et al have found that ACSS1 decides the survival of hepatocellular carcinoma.<sup>35</sup> Tada et al have demonstrated that ENTPD6 monitors cisplatin resistance in testis and testicular cancer.<sup>36</sup> Rosin et al have found that DYX1C1 can indicate the poor survival in breast cancer.<sup>37</sup> How FCF1 and TBC1D22A act in CESC is waiting to be answered by future research. However, due to the limited database, we cannot use another cohort to verify the efficiency of the model.

SFs are indispensable for AS events. We constructed an SF-AS regulatory network based on SFs. Cox regression analysis identified that AS was most

closely related to CELF2 and TC2N. Ramalingam et al found RNA binding protein CELF2 might be a putative tumor suppressor gene in colon cancer.<sup>38</sup> CELF2 can inhibit the development of glioma and non-small cell lung carcinoma.<sup>39,40</sup> TC2N can accelerate lung cancer progression by suppressing p53 signaling pathway.<sup>41</sup> Our research validated these opinions. As our new finding, the role of CELF2 in CESC needs new evidence.

We found that SNRPA interacted with CALCOCO2, PLS3, FAM76A, NGLY1, COX15, PPOX, PLEKHB2, PPP1CC, GALNT4, LYRM5 and NPIPA7. Among them, CALCOCO2, PLS3, FAM76A, NGLY1, COX15, PPOX and PLEKHB2 were all protective factors for their positive association with SNRPA. In contrast, PPP1CC, GALNT4, LYRM5 and NPIPA7 were pathogenic factors for their negative association with SNRPA. SNRPA functions in the progression of gastric cancer.<sup>42</sup>



**Figure 10** (A) The correlation of risk score and tumor infiltrating immune cells. (B) The correlation of risk score and dendritic cells activated. (C) The correlation of risk score and dendritic cells resting. (D) The correlation of risk score and Macrophages M0. (E) The correlation of risk score and Mast cells resting. (F) The correlation of risk score and T cells CD4 memory activated. (G) The correlation of risk score and T cells CD8.

Kurumi et al have found that PPOX can change the fluorescence intensity in 5-aminolevulinic acid-mediated laser-based photodynamic endoscopic diagnosis for early gastric cancer.<sup>43</sup> GALNT4 is implicated in clear cell renal cell carcinoma and colon cancer.<sup>44,45</sup> However, how these AS genes cooperate with SFs to regulate tumor development has not been fully studied, which is a task lying in our future research. In addition, we found that HSPA8 as SF negatively regulates the GEMIN4-38259-AP in CESC. HSPA8 protein integrates compensatory mechanisms to drive cell growth and is dysregulated in a variety of chronic stress diseases, including cancer.<sup>46</sup> In addition, HSPA8 could be used as candidate biomarkers and potential target for treatment in multiple tumors.<sup>47,48</sup> The results of this study showed that SNRPA, CELF2, HSPA8 affected the prognosis by affecting AS events. The relevant mechanisms and relationship involved in SNRPA, CELF2, HSPA8 have not been found yet, which can be further explore in the further.

Research on prognostic alternative splicing signature by Wang et al reveals the landscape of immune infiltration in pancreatic cancer.<sup>49</sup> In fact, Immunotherapy is currently widely used in tumor treatment options.<sup>50</sup> In our study, dendritic cells activated, dendritic cells resting, macrophages M0, Mast cells resting, T cells CD4

memory activated and T cells CD8 were found to be the most correlative with risk score. Dendritic cells are a special type of antigen presenting cells, which play a key role in the process of innate and adaptive immune response. The literature has proved that it is closely related to cancer immunology and immunotherapy.<sup>51</sup> Dendritic cells have also been shown to be closely linked to cervical cancer treatment.<sup>52,53</sup> Macrophages M0 was shown to be associated with the risk of cervical cancer in our previous study.<sup>54</sup> T cells CD4 and CD8 are the most common immune cells. There are not a few studies about their application in the treatment of cervical cancer.<sup>55–58</sup> All of the above support our analysis. The role of Mast cells has not been fully explored in cervical cancer. This will become a new research direction.

The study also has the following shortcomings. First, no other cohorts can be used to re-validate the prognostic model. In addition, more in vivo and in vitro functional experiments are needed to further explore the impact of dysregulated AS events and SFs on cancer development.

## Conclusion

This study screened out the AS events related to CESC and their most enriched functions and pathways. SNRPA



and CELF2 were two CESC-engaged SFs cooperating with a list of AS events. This scientific proposition may provide direct guidance for future biological experiments and contribute to the management and individualized treatment of CESC patients.

## Abbreviations

CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; SF, splicing factors; AS, Alternative splicing; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PPI, protein-protein interaction; STRING, Search Tool for the Retrieval of Interacting Genes Database; MCODE, Molecular Complex Detection; ROC, Receiver operating characteristic; AA, alternate acceptor; AD, alternate donor; AP, alternate promoter; AT, alternate terminator; ES, exon skip; ME, mutually exclusive exons; RI, retained intron.

## Data Sharing Statement

We confirm that my article contains a Data Availability Statement. We confirm that we have included a citation for available data in my references section. All the data in this study are available from TCGA database (<https://cancergenome.nih.gov/>).

## Ethics Approval and Consent to Participate

TCGA belong to public databases. The patients involved in the database have obtained ethical approval. Users can download relevant data for free for research and publish relevant articles. Our study is based on open-source data, so there are no ethical issues and other conflicts of interest. This article did not contain any studies with human participants or animals performed by the authors. The authors did not have access to extra informed consent, which did not distort scientific meaning.

## Funding

This research was supported by Maternal and Child Health Research Project of Jiangsu Provincial Health Commission (F201836), Nantong Municipal Health Commission Youth Project A (QA202004) and the Research project of Nantong health Committee project B (MB2020003).

## Disclosure

The authors have no conflicts of interests to declare.

## References

1. Benard VB, Thomas CC, King J, Massetti GM, Doria-Rose VP, Saraiya M. Vital signs: cervical cancer incidence, mortality, and screening - United States, 2007–2012. *MMWR Morb Mortal Wkly Rep*. 2014;63(44):1004–1009.
2. Gustafsson L, Ponten J, Zack M, Adami HO, Adami H-O. International incidence rates of invasive cervical cancer after introduction of cytological screening. *Cancer Causes Control*. 1997;8(5):755–763. doi:10.1023/A:1018435522475
3. Scarinci IC, Garcia FA, Kobetz E, et al. Cervical cancer prevention: new tools and old barriers. *Cancer*. 2010;116(11):2531–2542.
4. Soutter WP, de Barros Lopes A, Fletcher A, et al. Invasive cervical cancer after conservative therapy for cervical intraepithelial neoplasia. *Lancet (London, England)*. 1997;349(9057):978–980. doi:10.1016/S0140-6736(96)08295-5
5. Tang JY, Lee JC, Hou MF, et al. Alternative splicing for diseases, cancers, drugs, and databases. *TheScientificWorldJournal*. 2013;2013:703568. doi:10.1155/2013/703568
6. Bowler E, Porazinski S, Uzun S, et al. Hypoxia leads to significant changes in alternative splicing and elevated expression of CLK splice factor kinases in PC3 prostate cancer cells. *BMC Cancer*. 2018;18(1):355. doi:10.1186/s12885-018-4227-7
7. 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;234:15836–15846. doi:10.1002/jcp.28241
8. Xu J, Qiu Y. Role of androgen receptor splice variants in prostate cancer metastasis. *Asian J Urol*. 2016;3(4):177–184. doi:10.1016/j.ajur.2016.08.003
9. Song J, Liu YD, Su J, Yuan D, Sun F, Zhu J. Systematic analysis of alternative splicing signature unveils prognostic predictor for kidney renal clear cell carcinoma. *J Cell Physiol*. 2019;234:22753–22764. doi:10.1002/jcp.28840
10. Lin P, He RQ, Huang ZG, et al. Role of global aberrant alternative splicing events in papillary thyroid cancer prognosis. *Aging*. 2019;11(7):2082–2097. doi:10.18632/aging.101902
11. Mao S, Li Y, Lu Z, et al. Survival-associated alternative splicing signatures in esophageal carcinoma. *Carcinogenesis*. 2019;40(1):121–130. doi:10.1093/carcin/bgy123
12. Zhu J, Chen Z, Yong L. Systematic profiling of alternative splicing signature reveals prognostic predictor for ovarian cancer. *Gynecol Oncol*. 2018;148(2):368–374. doi:10.1016/j.ygyno.2017.11.028
13. Kornblihtt AR, Schor IE, Allo M, Dujardin G, Petrillo E, Munoz MJ. Alternative splicing: a pivotal step between eukaryotic transcription and translation. *Nat Rev Mol Cell Biol*. 2013;14(3):153–165. doi:10.1038/nrm3525
14. Anczukow O, Krainer AR. Splicing-factor alterations in cancers. *RNA (New York, NY)*. 2016;22(9):1285–1301. doi:10.1261/rna.057919.116
15. Paschalis A, Sharp A, Welti JC, et al. Alternative splicing in prostate cancer. *Nat Rev Clin Oncol*. 2018;15(11):663–675. doi:10.1038/s41571-018-0085-0
16. Shen B, Yuan Y, Zhang Y, et al. Long non-coding RNA FBXL19-AS1 plays oncogenic role in colorectal cancer by sponging miR-203. *Biochem Biophys Res Commun*. 2017;488(1):67–73. doi:10.1016/j.bbrc.2017.05.008
17. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*. 2016;44(W1):W90–97. doi:10.1093/nar/gkw377
18. Szklarczyk D, Franceschini A, Wyder S, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res*. 2015;43(Database issue):D447–452. doi:10.1093/nar/gku1003
19. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–2504. doi:10.1101/gr.1239303

20. Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med*. 1997;16(4):385–395. doi:10.1002/(SICI)1097-0258(19970228)16:4<385::AID-SIM380>3.0.CO;2-3
21. 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. doi:10.1002/humu.21609
22. Climente-Gonzalez H, Porta-Pardo E, Godzik A, Eyraes E. The functional impact of alternative splicing in cancer. *Cell Rep*. 2017;20(9):2215–2226. doi:10.1016/j.celrep.2017.08.012
23. 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. doi:10.1016/j.canlet.2017.02.016
24. Perera RM, Stoykova S, Nicolay BN, et al. Transcriptional control of autophagy-lysosome function drives pancreatic cancer metabolism. *Nature*. 2015;524(7565):361–365. doi:10.1038/nature14587
25. Li G, Li L, Joo EJ, et al. Glycosaminoglycans and glycolipids as potential biomarkers in lung cancer. *Glycoconj J*. 2017;34(5):661–669. doi:10.1007/s10719-017-9790-7
26. DiDonato JA, Mercurio F, Karin M. NF-kappaB and the link between inflammation and cancer. *Immunol Rev*. 2012;246(1):379–400. doi:10.1111/j.1600-065X.2012.01099.x
27. Tang D, Tao D, Fang Y, Deng C, Xu Q, Zhou J. TNF-alpha promotes invasion and metastasis via NF-kappa B pathway in oral squamous cell carcinoma. *Med Sci Monit Basic Res*. 2017;23:141–149. doi:10.12659/MSMBR.903910
28. Pacifico F, Mauro C, Barone C, et al. Oncogenic and anti-apoptotic activity of NF-kappa B in human thyroid carcinomas. *J Biol Chem*. 2004;279(52):54610–54619. doi:10.1074/jbc.M403492200
29. Guo G, Gui Y, Gao S, et al. Frequent mutations of genes encoding ubiquitin-mediated proteolysis pathway components in clear cell renal cell carcinoma. *Nat Genet*. 2011;44(1):17–19. doi:10.1038/ng.1014
30. Zhou Q, Hou Z, Zuo S, et al. LUCAT1 promotes colorectal cancer tumorigenesis by targeting the ribosomal protein L40-MDM2-p53 pathway through binding with UBA52. *Cancer Sci*. 2019;110(4):1194–1207. doi:10.1111/cas.13951
31. Han W, Yu G, Meng X, et al. Potential of C1QTNF1-AS1 regulation in human hepatocellular carcinoma. *Mol Cell Biochem*. 2019;460:37–51. doi:10.1007/s11010-019-03569-w
32. Leon J, Casado J, Jimenez Ruiz SM, et al. Melatonin reduces endothelin-1 expression and secretion in colon cancer cells through the inactivation of FoxO-1 and NF-kappabeta. *J Pineal Res*. 2014;56(4):415–426. doi:10.1111/jpi.12131
33. Cheng X, Wei L, Huang X, et al. Solute carrier family 39 member 6 gene promotes aggressiveness of esophageal carcinoma cells by increasing intracellular levels of zinc, activating phosphatidylinositol 3-kinase signaling, and up-regulating genes that regulate metastasis. *Gastroenterology*. 2017;152(8):1985–1997.e1912. doi:10.1053/j.gastro.2017.02.006
34. Lian J, Jing Y, Dong Q, et al. miR-192, a prognostic indicator, targets the SLC39A6/SLN pathway to reduce tumor metastasis in human hepatocellular carcinoma. *Oncotarget*. 2016;7(3):2672–2683. doi:10.18632/oncotarget.6603
35. Yun M, Bang SH, Kim JW, Park JY, Kim KS, Lee JD. The importance of acetyl coenzyme A synthetase for IIC-acetate uptake and cell survival in hepatocellular carcinoma. *J Nucl Med*. 2009;50(8):1222–1228. doi:10.2967/jnumed.109.062703
36. Tada Y, Yokomizo A, Shiota M, et al. Ectonucleoside triphosphate diphosphohydrolase 6 expression in testis and testicular cancer and its implication in cisplatin resistance. *Oncol Rep*. 2011;26(1):161–167.
37. Rosin G, Hannelius U, Lindstrom L, et al. The dyslexia candidate gene DYX1C1 is a potential marker of poor survival in breast cancer. *BMC Cancer*. 2012;12:79. doi:10.1186/1471-2407-12-79
38. Ramalingam S, Ramamoorthy P, Subramaniam D, Anant S. Reduced expression of RNA binding protein CELF2, a putative tumor suppressor gene in colon cancer. *Immuno-Gastroenterology*. 2012;1(1):27–33.
39. Yeung YT, Fan S, Lu B, et al. CELF2 suppresses non-small cell lung carcinoma growth by inhibiting the PREX2-PTEN interaction. *Carcinogenesis*. 2019;41:377–389.
40. Fan B, Jiao BH, Fan FS, et al. Downregulation of miR-95-3p inhibits proliferation, and invasion promoting apoptosis of glioma cells by targeting CELF2. *Int J Oncol*. 2015;47(3):1025–1033. doi:10.3892/ijo.2015.3080
41. Hao XL, Han F, Zhang N, et al. TC2N, a novel oncogene, accelerates tumor progression by suppressing p53 signaling pathway in lung cancer. *Cell Death Differ*. 2019;26(7):1235–1250. doi:10.1038/s41418-018-0202-8
42. Dou N, Yang D, Yu S, Wu B, Gao Y, Li Y. SNRPA enhances tumour cell growth in gastric cancer through modulating NGF expression. *Cell Prolif*. 2018;51(5):e12484. doi:10.1111/cpr.12484
43. Kurumi H, Kanda T, Kawaguchi K, et al. Protoporphyrinogen oxidase is involved in the fluorescence intensity of 5-aminolevulinic acid-mediated laser-based photodynamic endoscopic diagnosis for early gastric cancer. *Photodiagnosis Photodyn Ther*. 2018;22:79–85. doi:10.1016/j.pdpdt.2018.02.005
44. Qu JJ, Qu XY, Zhou DZ. miR4262 inhibits colon cancer cell proliferation via targeting of GALNT4. *Mol Med Rep*. 2017;16(4):3731–3736. doi:10.3892/mmr.2017.7057
45. Liu Y, Liu W, Xu L, et al. GALNT4 predicts clinical outcome in patients with clear cell renal cell carcinoma. *J Urol*. 2014;192(5):1534–1541. doi:10.1016/j.juro.2014.04.084
46. Robert G, Jacquelin A, Auberger P. Chaperone-mediated autophagy and its emerging role in hematological malignancies. *Cells*. 2019;8(10):1260. doi:10.3390/cells8101260
47. Li H, Han G, Li X, et al. MAPK-RAP1A signaling enriched in hepatocellular carcinoma is associated with favorable tumor-infiltrating immune cells and clinical prognosis. *Front Oncol*. 2021;11:649980. doi:10.3389/fonc.2021.649980
48. Liu Z, Zheng W, Liu Y, Zhou B, Zhang Y, Wang F. Targeting HSPA8 inhibits proliferation via downregulating BCR-ABL and enhances chemosensitivity in imatinib-resistant chronic myeloid leukemia cells. *Exp Cell Res*. 2021;405(2):112708. doi:10.1016/j.yexcr.2021.112708
49. Wang L, Bi J, Li X, Wei M, He M, Zhao L. Prognostic alternative splicing signature reveals the landscape of immune infiltration in pancreatic cancer. *J Cancer*. 2020;11(22):6530–6544. doi:10.7150/jca.47877
50. Yang Y. Cancer immunotherapy: harnessing the immune system to battle cancer. *J Clin Invest*. 2015;125(9):3335–3337. doi:10.1172/JCI83871
51. Wculek SK, Cueto FJ, Mujal AM, Melero I, Krummel MF, Sancho D. Dendritic cells in cancer immunology and immunotherapy. *Nat Rev Immunol*. 2020;20(1):7–24. doi:10.1038/s41577-019-0210-z
52. Walch-Rückheim B, Ströder R, Theobald L, et al. Cervical cancer-instructed stromal fibroblasts enhance IL23 expression in dendritic cells to support expansion of Th17 cells. *Cancer Res*. 2019;79(7):1573–1586. doi:10.1158/0008-5472.CAN-18-1913
53. Chen S, Lv M, Fang S, Ye W, Gao Y, Xu Y. Poly(I:C) enhanced anti-cervical cancer immunities induced by dendritic cells-derived exosomes. *Int J Biol Macromol*. 2018;113:1182–1187. doi:10.1016/j.ijbiomac.2018.02.034
54. Liu J, Yang J, Gao F, et al. A microRNA-messenger RNA regulatory network and its prognostic value in cervical cancer. *DNA Cell Biol*. 2020;39(7):1328–1346. doi:10.1089/dna.2020.5590
55. van der Burg SH, Piersma SJ, de Jong A, et al. Association of cervical cancer with the presence of CD4+ regulatory T cells specific for human papillomavirus antigens. *Proc Natl Acad Sci USA*. 2007;104(29):12087–12092. doi:10.1073/pnas.0704672104
56. Garcia-Chagollan M, Jave-Suarez LF, Haramati J, et al. An approach to the immunophenotypic features of circulating CD4<sup>+</sup>NKG2D<sup>+</sup> T cells in invasive cervical carcinoma. *J Biomed Sci*. 2015;22:91. doi:10.1186/s12929-015-0190-7

57. Pilch H, Hoehn H, Schmidt M, et al. CD8+CD45RA+CD27-CD28-T-cell subset in PBL of cervical cancer patients representing CD8+T-cells being able to recognize cervical cancer associated antigens provided by HPV 16 E7. *Zentralbl Gynakol.* 2002;124(8-9):406-412. doi:10.1055/s-2002-38130
58. Komdeur FL, Prins TM, van de Wall S, et al. CD103+ tumor-infiltrating lymphocytes are tumor-reactive intraepithelial CD8+ T cells associated with prognostic benefit and therapy response in cervical cancer. *Oncoimmunology.* 2017;6(9):e1338230. doi:10.1080/2162402X.2017.1338230

### International Journal of General Medicine

Dovepress

### Publish your work in this journal

The International Journal of General Medicine is an international, peer-reviewed open-access journal that focuses on general and internal medicine, pathogenesis, epidemiology, diagnosis, monitoring and treatment protocols. The journal is characterized by the rapid reporting of reviews, original research and clinical studies

across all disease areas. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-general-medicine-journal>