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

Xiaoyu Wang
Nantong first people's hospital
Weichun Tang
Nantong first people's hospital
Yilin Lu
Nantong first people's hospital
Jun You
Nantong first people's hospital
Yun Han
Nantong first people's hospital
zheng yanli (gaoshan1189@sina.com)
Nantong first people's hospital, Nantong

Research article

Keywords: cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), Alternative splicing (AS), prognosis, TCGA

Posted Date: December 8th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-121117/v1

Abstract

AS acts on many tumors and its relationship with CESC needs to be researched. TCGA provided Data. SpliceSeq analyzed Splicing profiles. UpSetR displayed the intersections among AS events. Univariate analysis chose survival-associated AS and SF genes. Functional analysis was operated on Enrichr. STRING evaluated PPI information. MCODE searched clustered sub-networks. LASSO and multivariate analysis constructed prognostic model. Risk analysis of tumor infiltrating immune cells was also conducted. 402 AS-generated genes were found to be associated with CESC prognosis. Functional analysis showed that Golgi to lysosome transport was enriched. PPI 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 riskscore. This study identified SNRPA and CELF2 as two CESC-engaged SFs.

Introduction

In the population of developing countries, Cervical cancer incidence ranks third among all tumors and first among gynecological tumors[1]. 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[2]. Effective diagnosis in the early stage can significantly reduce the incidence of cervical cancer [3]. It has been shown that timely treatment of dysplasia can effectively reduce the incidence and mortality of cervical cancer [4]. 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 post-transcriptional proteome diversified [5, 6]. 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. [7–12]. 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)[13]. 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[14, 15].

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 value (range 0–1) is a commonly used for scoring AS events.

Survival-associated AS events in CESC

A total of 306 CESC tissues were analyzed in this study and the overall survival (OS) of the 306 CESC patients were at least 90 days. We used UpSetR (version1.3.3) [16] 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/)[17]. The biological functions of the overlap genes were interpreted. $P < 0.05$ was considered statistically significant. We utilized STRING (Search Tool for the Retrieval of Interacting Genes Database) (http://www.string-db.org/) to evaluate PPI (protein-protein interaction) of hub modules[18]. To visualize the findings, we exported the results from STRING database to Cytoscape software [19]. 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 [20]. 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 (β) using the following formula: Risk score = $\beta_{gene1} \times expr_{gene1} + \beta_{gene2} \times expr_{gene2} + \cdots + \beta_{genen} \times expr_{genen}$. 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 ratio (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[21] .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, 41426 AS events were detected from 19696 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 15942 ESs in 6271 genes, 8395 ATs in 3663 genes, 8066 APs in 3256 genes, 3424 AAs in 2397 genes, 3017 ADs in2106 genes, 2373 RIs in 1801 genes, and 209 MEs in 202 genes (Fig. 1A). On average, one gene might possess up to six types of AS. 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 (Fig. 1B) were screened out.

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

Table 2
3708 survival-associated AS events in seven AS types in 3355 genes.

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

Biological terms enriched in 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 (FIGS1C). 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)**. 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-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 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 were displayed in Fig. 6A. Among them, the MEs have no statistical significance. The 20 most significant prognosis-related AS events were also shown **(**Fig. 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 **(**Fig. 3**)**. We also performed the receiver operating characteristic curve (ROC) analysis. As shown in Fig. 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 patients was displayed in **Figure S6-7**, Fig. 5. To assess whether these model was an independent predictor of CESC, univariate and multivariate Cox regression analyses were performed, including clinical

factors and risk scores. The results showed that this prognostic model showed moderate and independent prognostic power for all seven AS events (Fig. 6–7).

Table 3
Seven prognostic models based on 7 types of AS events.

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
Type	**Formula**	**AUC**
AA	-ENTPD6|58863|AA*9.39-GATAD2A|48635|AA*13.54-ENAH|9992|AA*4.58-S100PBP|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	FCF1|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-ERBB2IP|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	FCF1|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

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 |Cor|>0.5, we constructed the SF-AS regulatory network with 43 significant SFs **(**Fig. 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 **(**Fig. 8B-C**)**. Finally,Multivariate Cox proportional hazards regression analysis screened out three genes (SNRPA, CELF2 and HSPA8) **(**Fig. 8D**)**. Risk score=-0.033* SNRPA-0.34*CELF2 + 0.003*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 **(**Fig. 9A**)**. In ROC curve analysis, risk score of 0.685 reaped the highest specificity and sensitivity in predicting the 5-year survival **(**Fig. 9B**)**.

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 **(**Fig. 9C-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 riskscore and tumor immune microenvironment was shown in (Fig. 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 riskscore.

Discussion

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

This is the first to investigate the performance of AS events in CESC using 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 RM et al. have found that lysosome drives pancreatic cancer metabolism [24]. Li G et al. have found that glycosaminoglycan can mark lung cancer [25]. 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 [26]. Tang D et al. found TNF-Alpha promoted the invasion and metastasis via NF-Kappa B pathway in oral squamous cell carcinoma [27]. Pacifico F et al. found that f NF-kappa B pathway also acted in human thyroid carcinomas [28].

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 [29]. Zhou Q et al. have found that UBA52 regulates colorectal cancer tumorigenesis [30]. 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 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 W et al. have found

that C1QTNF1 affects the progression of human hepatocellular carcinoma [31]. Xu ES et al. have found that ECE-1 overexpression in worsens tumor differentiation and clinical outcome of head and neck cancer and colon cancer[32]. SLC39A6 functions in hepatocellular carcinoma and Esophageal carcinoma [33, 34]. Yun M et al. have found that ACSS1 decides the survival of hepatocellular carcinoma [35]. Tada Y et al. have demonstrated that ENTPD6 monitors cisplatin resistance in testis and testicular cancer [36]. Rosin G et al. have found that DYX1C1 can indicate the poor survival in breast cancer[37]. 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 S et al. found RNA binding protein CELF2 might be a putative tumor suppressor gene in colon cancer [38]. CELF2 can inhibit the development of glioma and non-small cell lung carcinoma[39, 40]. TC2N can accelerate lung cancer progression by suppressing p53 signaling pathway [41]. 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[42]. Kurumi H 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 [43]. GALNT4 is implicated in clear cell renal cell carcinoma and colon cancer [44, 45]. 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.

Research on Prognostic alternative splicing signature by Lin Wang et al. reveals the landscape of immune infiltration in Pancreatic Cancer[46]. In fact, Immunotherapy is currently widely used in tumor treatment options[47]. 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 riskscore. 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[48]. Dendritic cells have also been shown to be closely linked to cervical cancer treatment[49, 50]. Macrophages M0 was shown to be associated with the risk of cervical cancer in our previous study[51]. 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[52–55]. 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.

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.

Declarations

DATA AVAILABILITY 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/).

FUNDING

This research was supported by Maternal and Child Health Research Project of Jiangsu Provincial Health Commission (F201836).

CONFLICT OF INTEREST STATEMENT

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 Morbidity and mortality weekly report* 2014, **63**(44):1004-1009.
2. Gustafsson L, Ponten J, Zack M, Adami HO: **International incidence rates of invasive cervical cancer after introduction of cytological screening**. *Cancer causes & control : CCC* 1997, **8**(5):755-763.
3. Scarinci IC, Garcia FA, Kobetz E, Partridge EE, Brandt HM, Bell MC, Dignan M, Ma GX, Daye JL, Castle PE: **Cervical cancer prevention: new tools and old barriers**. *Cancer* 2010, **116**(11):2531-2542.
4. Soutter WP, de Barros Lopes A, Fletcher A, Monaghan JM, Duncan ID, Paraskevaidis E, Kitchener HC: **Invasive cervical cancer after conservative therapy for cervical intraepithelial neoplasia**. *Lancet (London, England)* 1997, **349**(9057):978-980.
5. Tang JY, Lee JC, Hou MF, Wang CL, Chen CC, Huang HW, Chang HW: **Alternative splicing for diseases, cancers, drugs, and databases**. *TheScientificWorldJournal* 2013, **2013**:703568.
6. Bowler E, Porazinski S, Uzor S, Thibault P, Durand M, Lapointe E, Rouschop KMA, Hancock J, Wilson I, Ladomery M: **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.
7. Liang Y, Song J, He D, Xia Y, Wu Y, Yin X, Liu J: **Systematic analysis of survival-associated alternative splicing signatures uncovers prognostic predictors for head and neck cancer**. *Journal of cellular physiology* 2019.
8. Xu J, Qiu Y: **Role of androgen receptor splice variants in prostate cancer metastasis**. *Asian journal of urology* 2016, **3**(4):177-184.

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**. *Journal of cellular physiology* 2019.
10. Lin P, He RQ, Huang ZG, Zhang R, Wu HY, Shi L, Li XJ, Li Q, Chen G, Yang H *et al*: **Role of global aberrant alternative splicing events in papillary thyroid cancer prognosis**. *Aging* 2019, **11**(7):2082-2097.
11. Mao S, Li Y, Lu Z, Che Y, Sun S, Huang J, Lei Y, Wang X, Liu C, Zheng S *et al*: **Survival-associated alternative splicing signatures in esophageal carcinoma**. *Carcinogenesis* 2019, **40**(1):121-130.
12. Zhu J, Chen Z, Yong L: **Systematic profiling of alternative splicing signature reveals prognostic predictor for ovarian cancer**. *Gynecologic oncology* 2018, **148**(2):368-374.
13. Kornblihtt AR, Schor IE, Allo M, Dujardin G, Petrillo E, Munoz MJ: **Alternative splicing: a pivotal step between eukaryotic transcription and translation**. *Nature reviews Molecular cell biology* 2013, **14**(3):153-165.
14. Anczukow O, Krainer AR: **Splicing-factor alterations in cancers**. *RNA (New York, NY)* 2016, **22**(9):1285-1301.
15. Paschalis A, Sharp A, Welti JC, Neeb A, Raj GV, Luo J, Plymate SR, de Bono JS: **Alternative splicing in prostate cancer**. *Nature reviews Clinical oncology* 2018, **15**(11):663-675.
16. Shen B, Yuan Y, Zhang Y, Yu S, Peng W, Huang X, Feng J: **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.
17. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A *et al*: **Enrichr: a comprehensive gene set enrichment analysis web server 2016 update**. *Nucleic acids research* 2016, **44**(W1):W90-97.
18. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP *et al*: **STRING v10: protein-protein interaction networks, integrated over the tree of life**. *Nucleic acids research* 2015, **43**(Database issue):D447-452.
19. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: **Cytoscape: a software environment for integrated models of biomolecular interaction networks**. *Genome research* 2003, **13**(11):2498-2504.
20. Tibshirani R: **The lasso method for variable selection in the Cox model**. *Statistics in medicine* 1997, **16**(4):385-395.
21. Piva F, Giulietti M, Burini AB, Principato G: **SpliceAid 2: a database of human splicing factors expression data and RNA target motifs**. *Human mutation* 2012, **33**(1):81-85.
22. Climente-Gonzalez H, Porta-Pardo E, Godzik A, Eyras E: **The Functional Impact of Alternative Splicing in Cancer**. *Cell reports* 2017, **20**(9):2215-2226.
23. Li Y, Sun N, Lu Z, Sun S, Huang J, Chen Z, He J: **Prognostic alternative mRNA splicing signature in non-small cell lung cancer**. *Cancer letters* 2017, **393**:40-51.

24. Perera RM, Stoykova S, Nicolay BN, Ross KN, Fitamant J, Boukhali M, Lengrand J, Deshpande V, Selig MK, Ferrone CR *et al*: **Transcriptional control of autophagy-lysosome function drives pancreatic cancer metabolism**. *Nature* 2015, **524**(7565):361-365.
25. Li G, Li L, Joo EJ, Son JW, Kim YJ, Kang JK, Lee KB, Zhang F, Linhardt RJ: **Glycosaminoglycans and glycolipids as potential biomarkers in lung cancer**. *Glycoconjugate journal* 2017, **34**(5):661-669.
26. DiDonato JA, Mercurio F, Karin M: **NF-kappaB and the link between inflammation and cancer**. *Immunological reviews* 2012, **246**(1):379-400.
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**. *Medical science monitor basic research* 2017, **23**:141-149.
28. Pacifico F, Mauro C, Barone C, Crescenzi E, Mellone S, Monaco M, Chiappetta G, Terrazzano G, Liguoro D, Vito P *et al*: **Oncogenic and anti-apoptotic activity of NF-kappa B in human thyroid carcinomas**. *The Journal of biological chemistry* 2004, **279**(52):54610-54619.
29. Guo G, Gui Y, Gao S, Tang A, Hu X, Huang Y, Jia W, Li Z, He M, Sun L *et al*: **Frequent mutations of genes encoding ubiquitin-mediated proteolysis pathway components in clear cell renal cell carcinoma**. *Nature genetics* 2011, **44**(1):17-19.
30. Zhou Q, Hou Z, Zuo S, Zhou X, Feng Y, Sun Y, Yuan X: **LUCAT1 promotes colorectal cancer tumorigenesis by targeting the ribosomal protein L40-MDM2-p53 pathway through binding with UBA52**. *Cancer science* 2019, **110**(4):1194-1207.
31. Han W, Yu G, Meng X, Hong H, Zheng L, Wu X, Zhang D, Yan B, Ma Y, Li X *et al*: **Potential of C1QTNF1-AS1 regulation in human hepatocellular carcinoma**. *Molecular and cellular biochemistry* 2019.
32. Leon J, Casado J, Jimenez Ruiz SM, Zurita MS, Gonzalez-Puga C, Rejon JD, Gila A, Munoz de Rueda P, Pavon EJ, Reiter RJ *et al*: **Melatonin reduces endothelin-1 expression and secretion in colon cancer cells through the inactivation of FoxO-1 and NF-kappabeta**. *Journal of pineal research* 2014, **56**(4):415-426.
33. Cheng X, Wei L, Huang X, Zheng J, Shao M, Feng T, Li J, Han Y, Tan W, Tan W *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.
34. Lian J, Jing Y, Dong Q, Huan L, Chen D, Bao C, Wang Q, Zhao F, Li J, Yao M *et al*: **miR-192, a prognostic indicator, targets the SLC39A6/SNAIL pathway to reduce tumor metastasis in human hepatocellular carcinoma**. *Oncotarget* 2016, **7**(3):2672-2683.
35. Yun M, Bang SH, Kim JW, Park JY, Kim KS, Lee JD: **The importance of acetyl coenzyme A synthetase for 11C-acetate uptake and cell survival in hepatocellular carcinoma**. *J Nucl Med* 2009, **50**(8):1222-1228.
36. Tada Y, Yokomizo A, Shiota M, Song Y, Kashiwagi E, Kuroiwa K, Oda Y, Naito S: **Ectonucleoside triphosphate diphosphohydrolase 6 expression in testis and testicular cancer and its implication in cisplatin resistance**. *Oncology reports* 2011, **26**(1):161-167.

37. Rosin G, Hannelius U, Lindstrom L, Hall P, Bergh J, Hartman J, Kere J: **The dyslexia candidate gene DYX1C1 is a potential marker of poor survival in breast cancer**. *BMC cancer* 2012, **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, Yin S, Yang S, Nie W, Wang M, Zhou L, Li T, Li X *et al*: **CELF2 suppresses non-small cell lung carcinoma growth by inhibiting the PREX2-PTEN interaction**. *Carcinogenesis* 2019.
40. Fan B, Jiao BH, Fan FS, Lu SK, Song J, Guo CY, Yang JK, Yang L: **Downregulation of miR-95-3p inhibits proliferation, and invasion promoting apoptosis of glioma cells by targeting CELF2**. *International journal of oncology* 2015, **47**(3):1025-1033.
41. Hao XL, Han F, Zhang N, Chen HQ, Jiang X, Yin L, Liu WB, Wang DD, Chen JP, Cui ZH *et al*: **TC2N, a novel oncogene, accelerates tumor progression by suppressing p53 signaling pathway in lung cancer**. *Cell death and differentiation* 2019, **26**(7):1235-1250.
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 proliferation* 2018, **51**(5):e12484.
43. Kurumi H, Kanda T, Kawaguchi K, Yashima K, Koda H, Ogihara K, Matsushima K, Nakao K, Saito H, Fujiwara Y *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.
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.
45. Liu Y, Liu W, Xu L, Liu H, Zhang W, Zhu Y, Xu J, Gu J: **GALNT4 predicts clinical outcome in patients with clear cell renal cell carcinoma**. *J Urol* 2014, **192**(5):1534-1541.
46. 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**. *Journal of Cancer* 2020, **11**(22):6530-6544.
47. Yang Y: **Cancer immunotherapy: harnessing the immune system to battle cancer**. *The Journal of clinical investigation* 2015, **125**(9):3335-3337.
48. Wculek SK, Cueto FJ, Mujal AM, Melero I, Krummel MF, Sancho D: **Dendritic cells in cancer immunology and immunotherapy**. *Nature reviews Immunology* 2020, **20**(1):7-24.
49. Walch-Rückheim B, Ströder R, Theobald L, Pahne-Zeppenfeld J, Hegde S, Kim YJ, Bohle RM, Juhasz-Böss I, Solomayer EF, Smola S: **Cervical Cancer-Instructed Stromal Fibroblasts Enhance IL23 Expression in Dendritic Cells to Support Expansion of Th17 Cells**. *Cancer research* 2019, **79**(7):1573-1586.
50. 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**. *International journal of biological macromolecules* 2018, **113**:1182-1187.
51. Liu J, Yang J, Gao F, Li S, Nie S, Meng H, Sun R, Wan Y, Jiang Y, Ma X *et al*: **A microRNA-Messenger RNA Regulatory Network and Its Prognostic Value in Cervical Cancer**. *DNA and cell biology* 2020,

39(7):1328-1346.

52. van der Burg SH, Piersma SJ, de Jong A, van der Hulst JM, Kwappenberg KM, van den Hende M, Welters MJ, Van Rood JJ, Fleuren GJ, Melief CJ *et al*: **Association of cervical cancer with the presence of CD4+ regulatory T cells specific for human papillomavirus antigens**. *Proceedings of the National Academy of Sciences of the United States of America* 2007, **104**(29):12087-12092.
53. Garcia-Chagollan M, Jave-Suarez LF, Haramati J, Bueno-Topete MR, Aguilar-Lemarroy A, Estrada-Chavez C, Bastidas-Ramirez BE, Pereira-Suarez AL, Del Toro-Arreola S: **An approach to the immunophenotypic features of circulating CD4NKG2D T cells in invasive cervical carcinoma**. *Journal of biomedical science* 2015, **22**:91.
54. Pilch H, Hoehn H, Schmidt M, Steiner E, Tanner B, Seufert R, Maeurer M: **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**. *Zentralblatt fur Gynakologie* 2002, **124**(8-9):406-412.
55. Komdeur FL, Prins TM, van de Wall S, Plat A, Wisman GBA, Hollema H, Daemen T, Church DN, de Bruyn M, Nijman HW: **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.

Figures

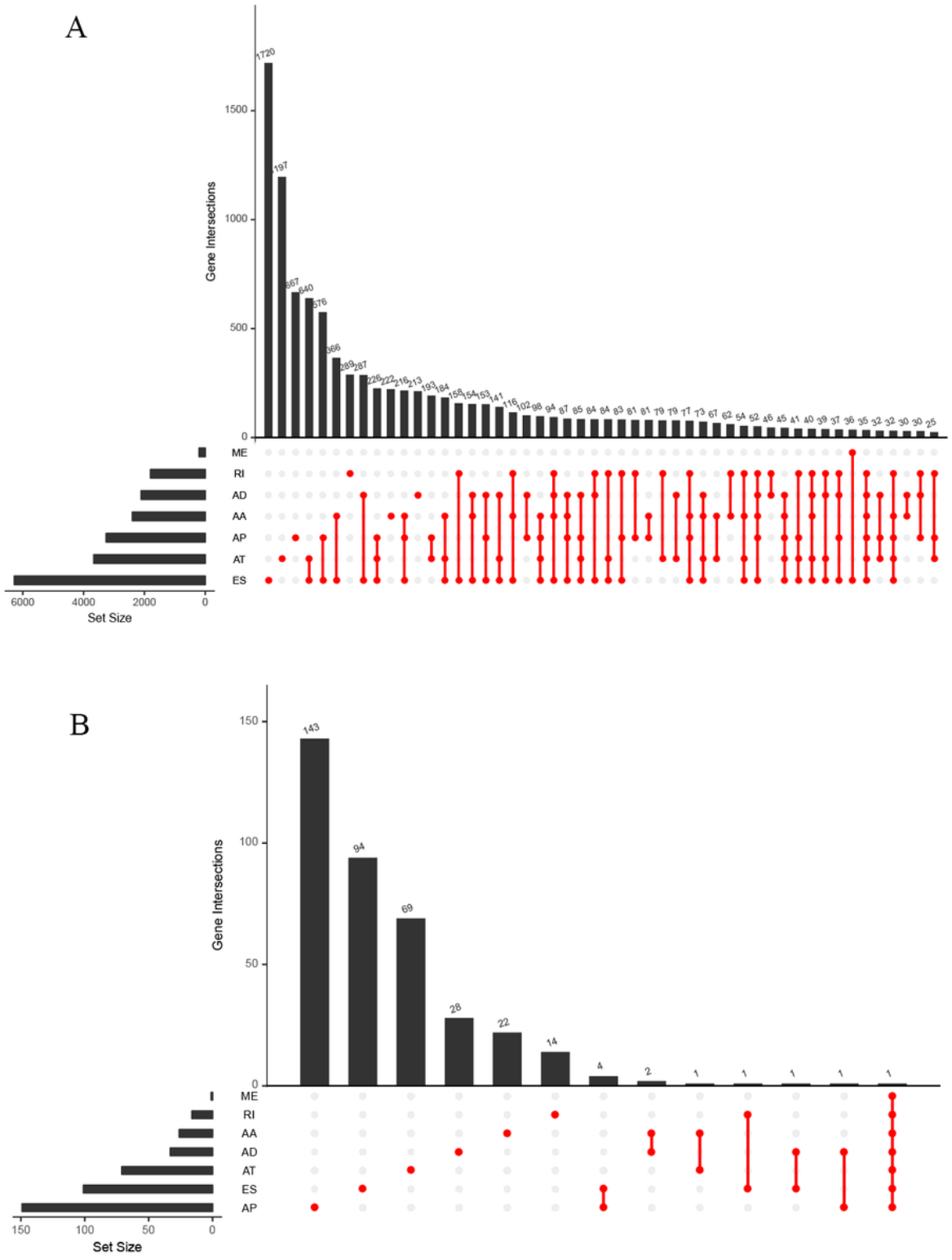

Figure 1

(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 gens in each AS type, respectively.

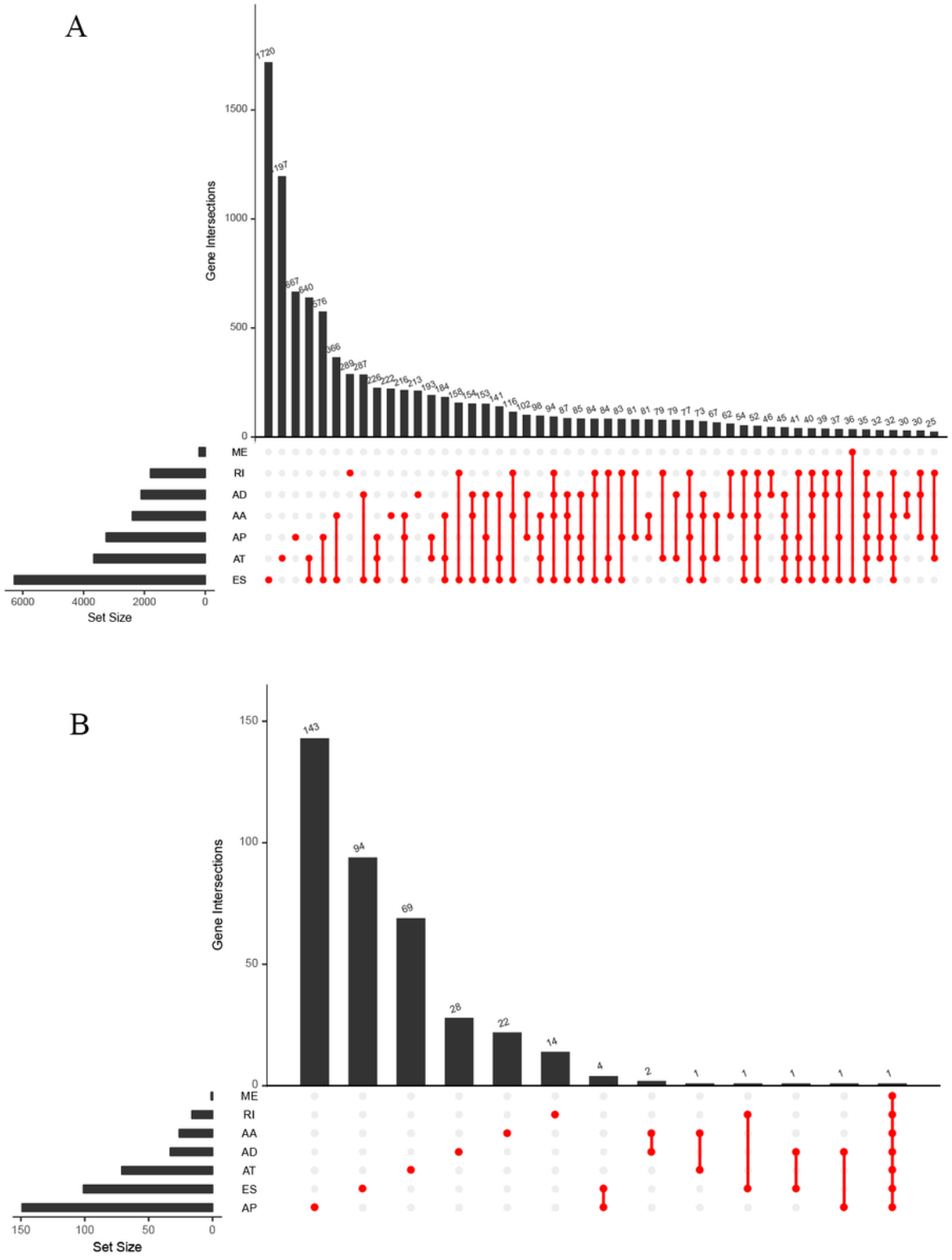

Figure 1

(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 gens in each AS type, respectively.

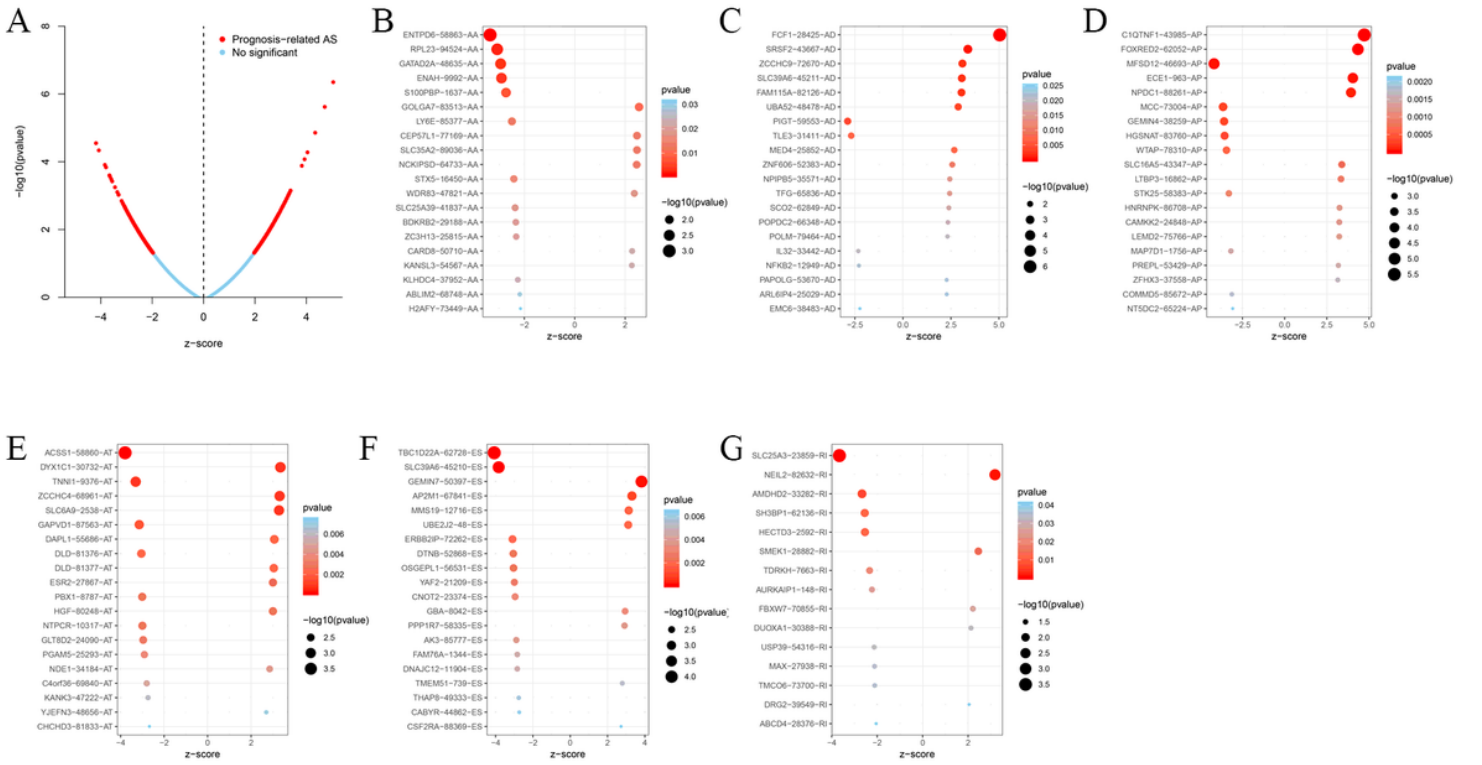

Figure 2

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.

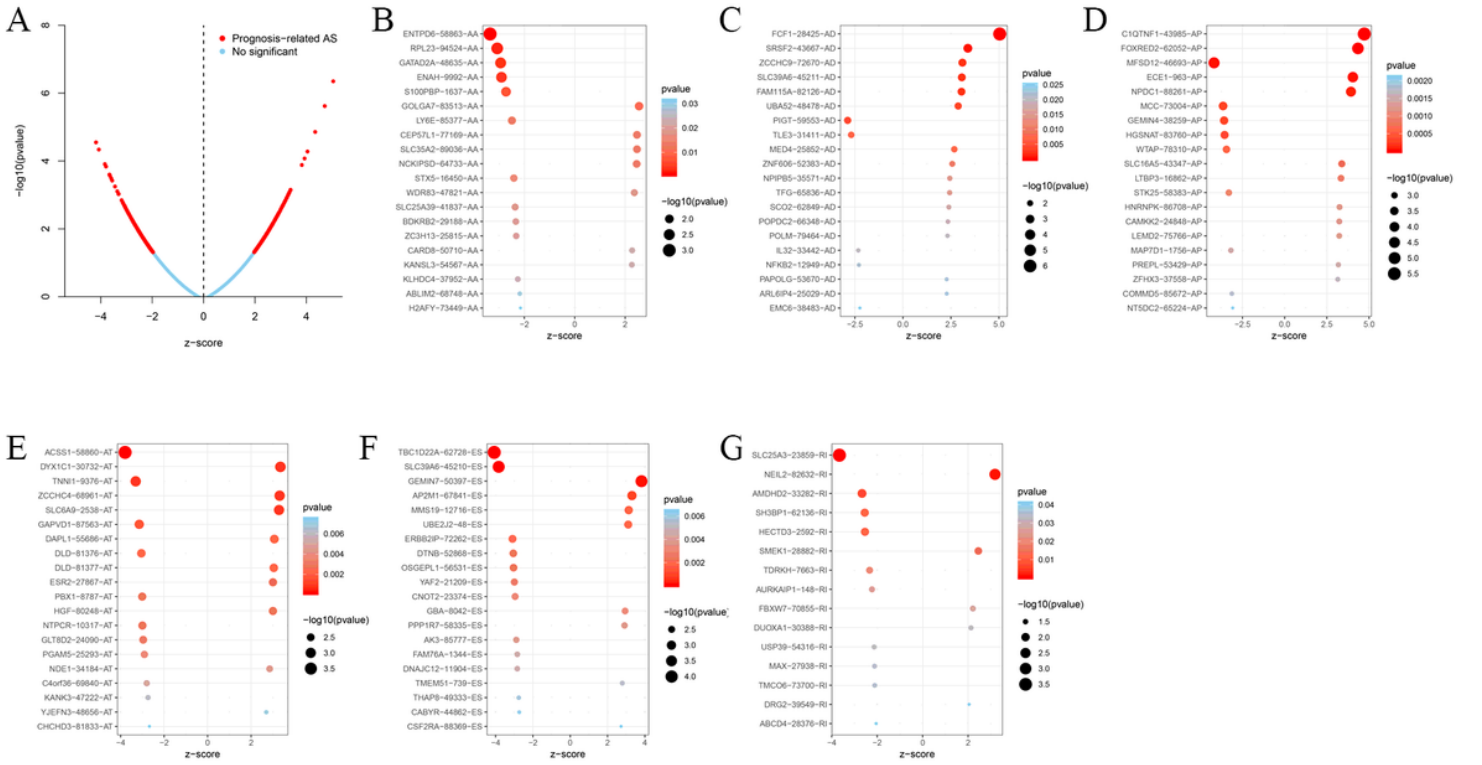

Figure 2

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.

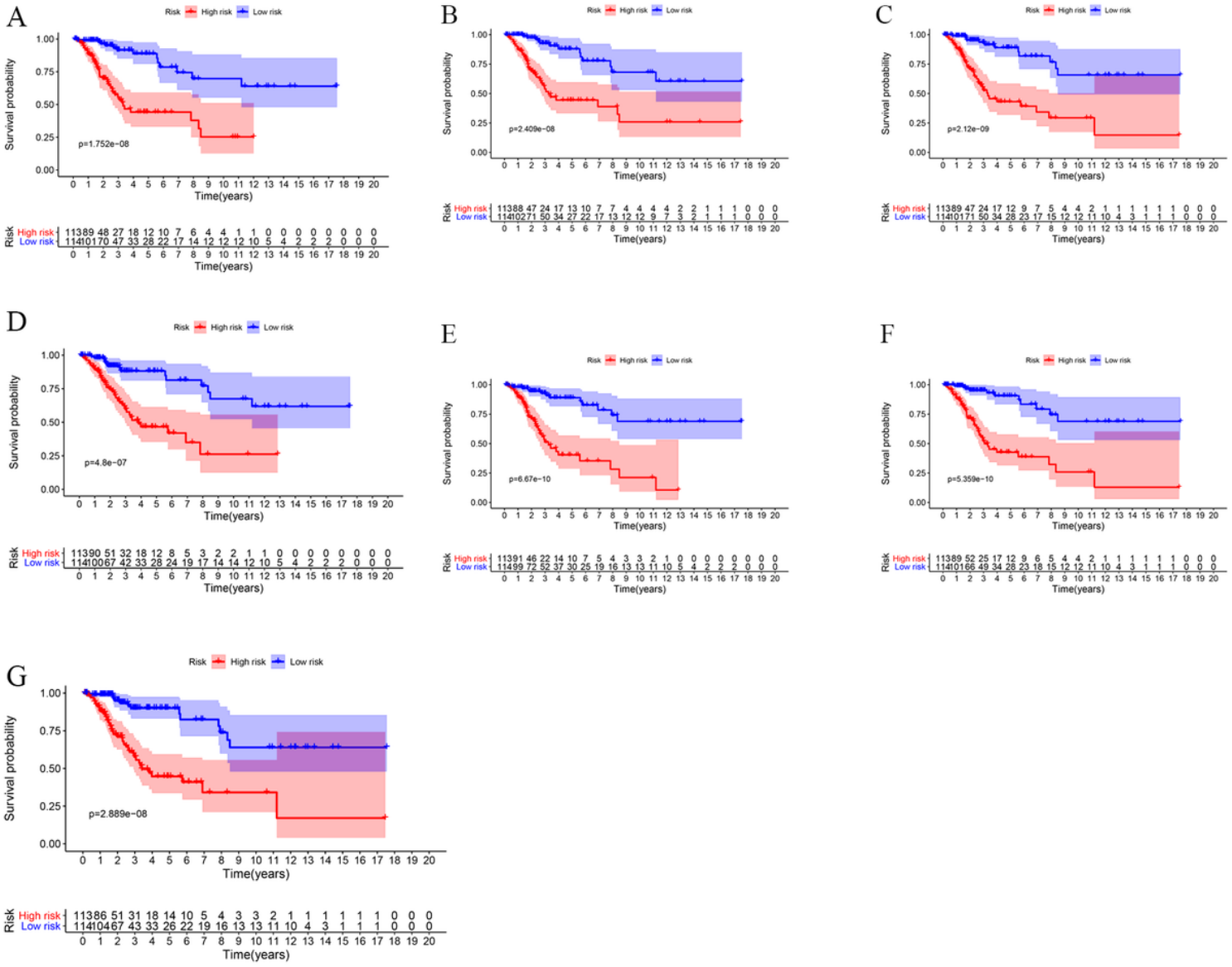

Figure 3

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.

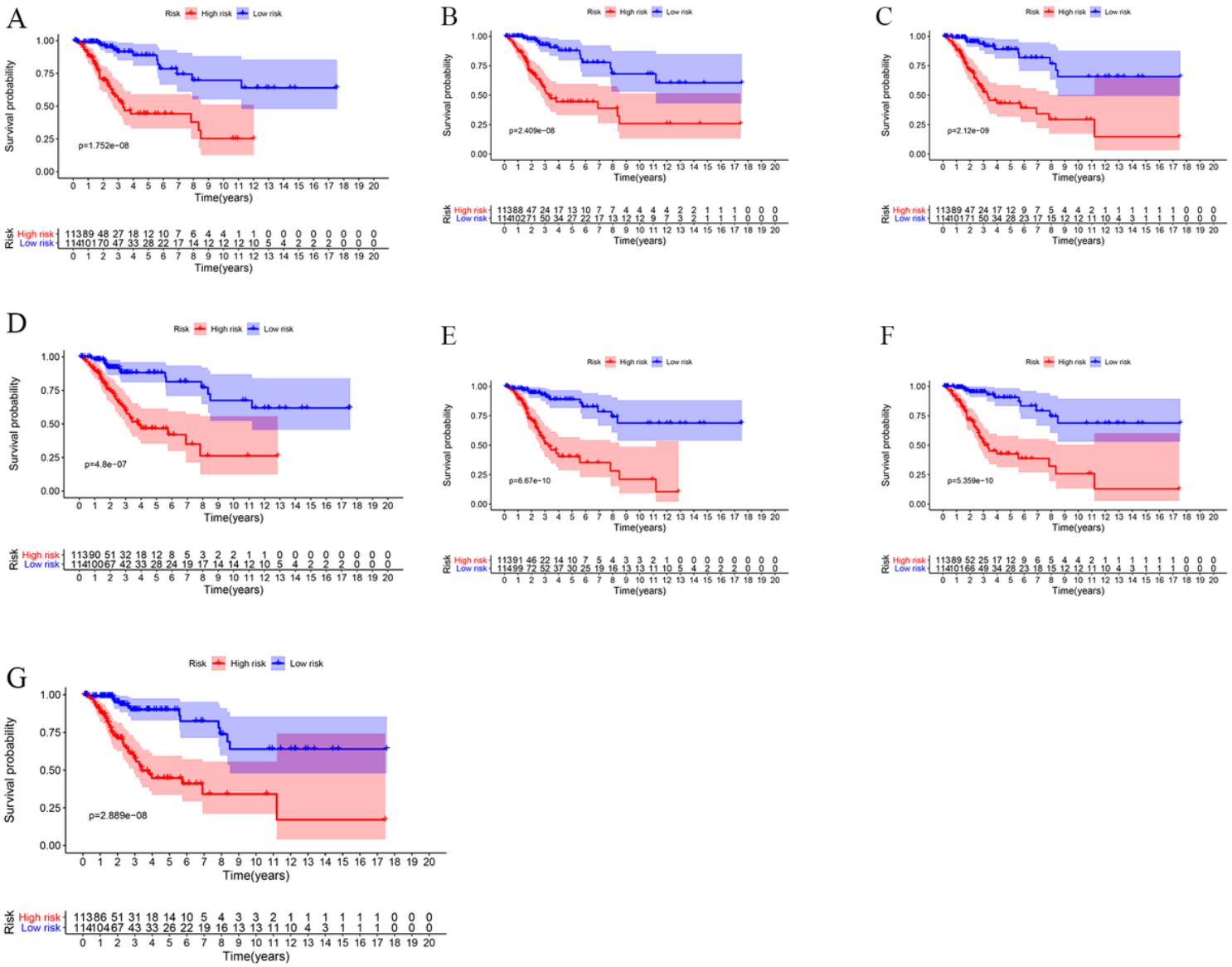

Figure 3

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.

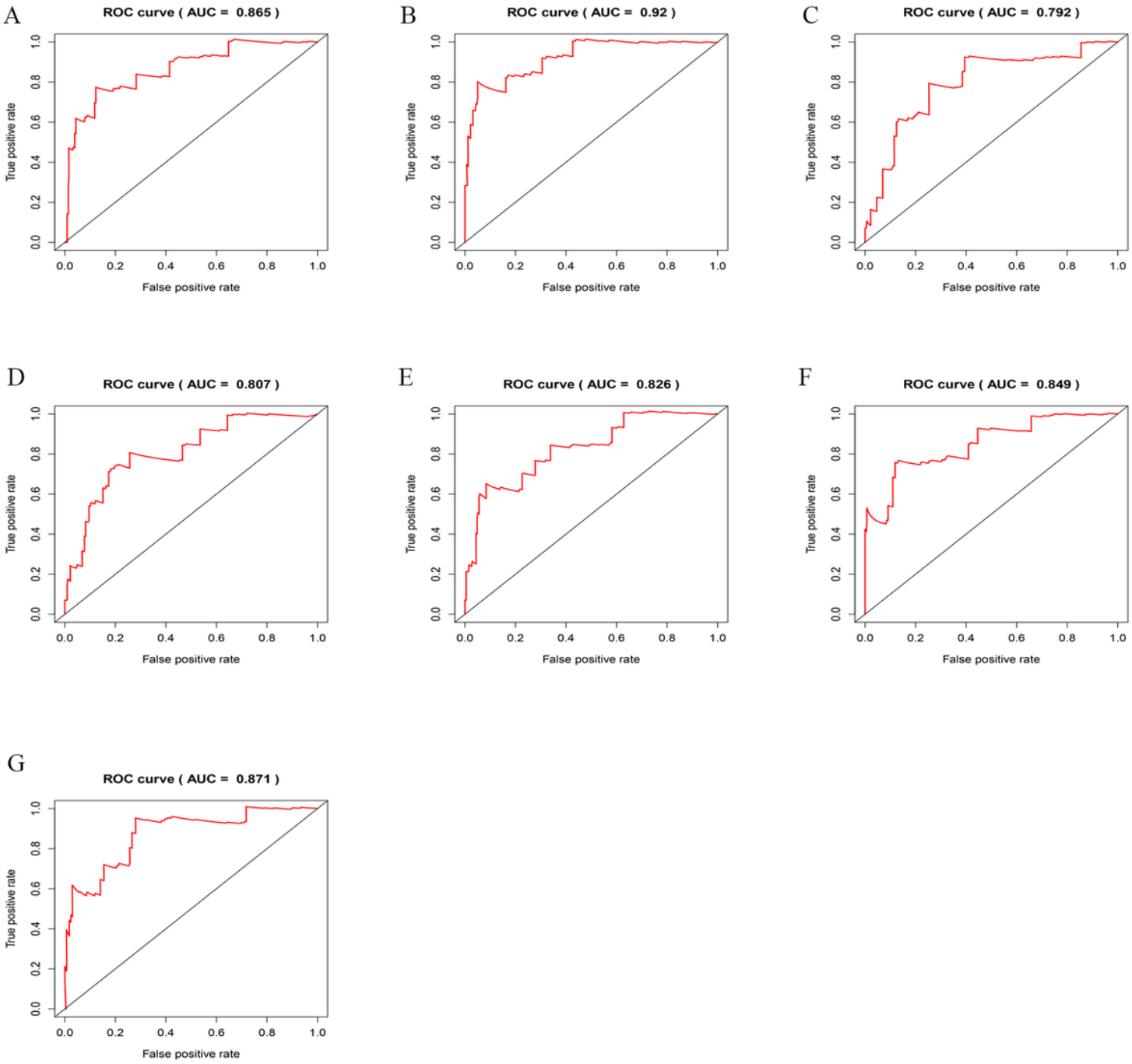

Figure 4

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

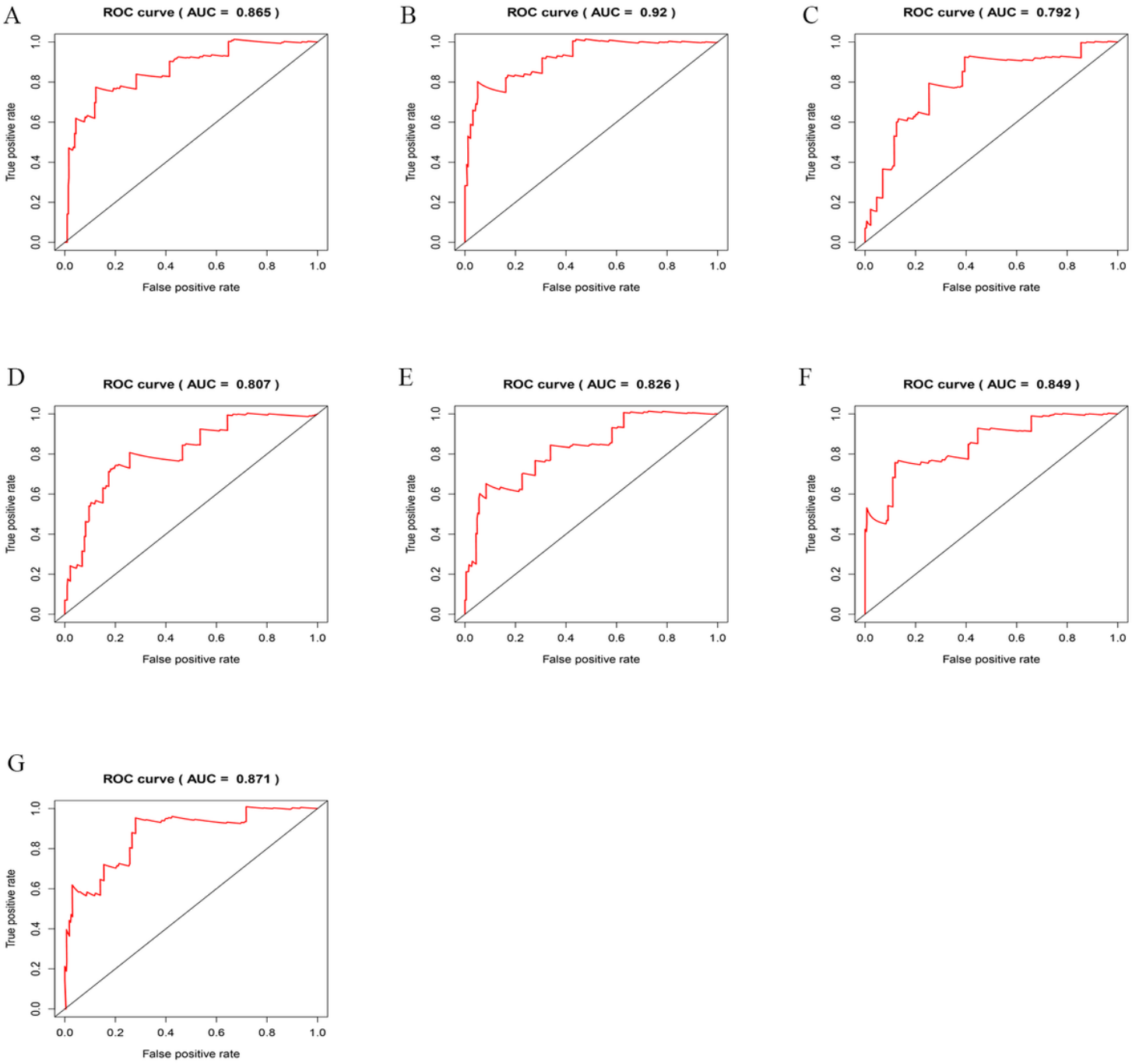

Figure 4

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

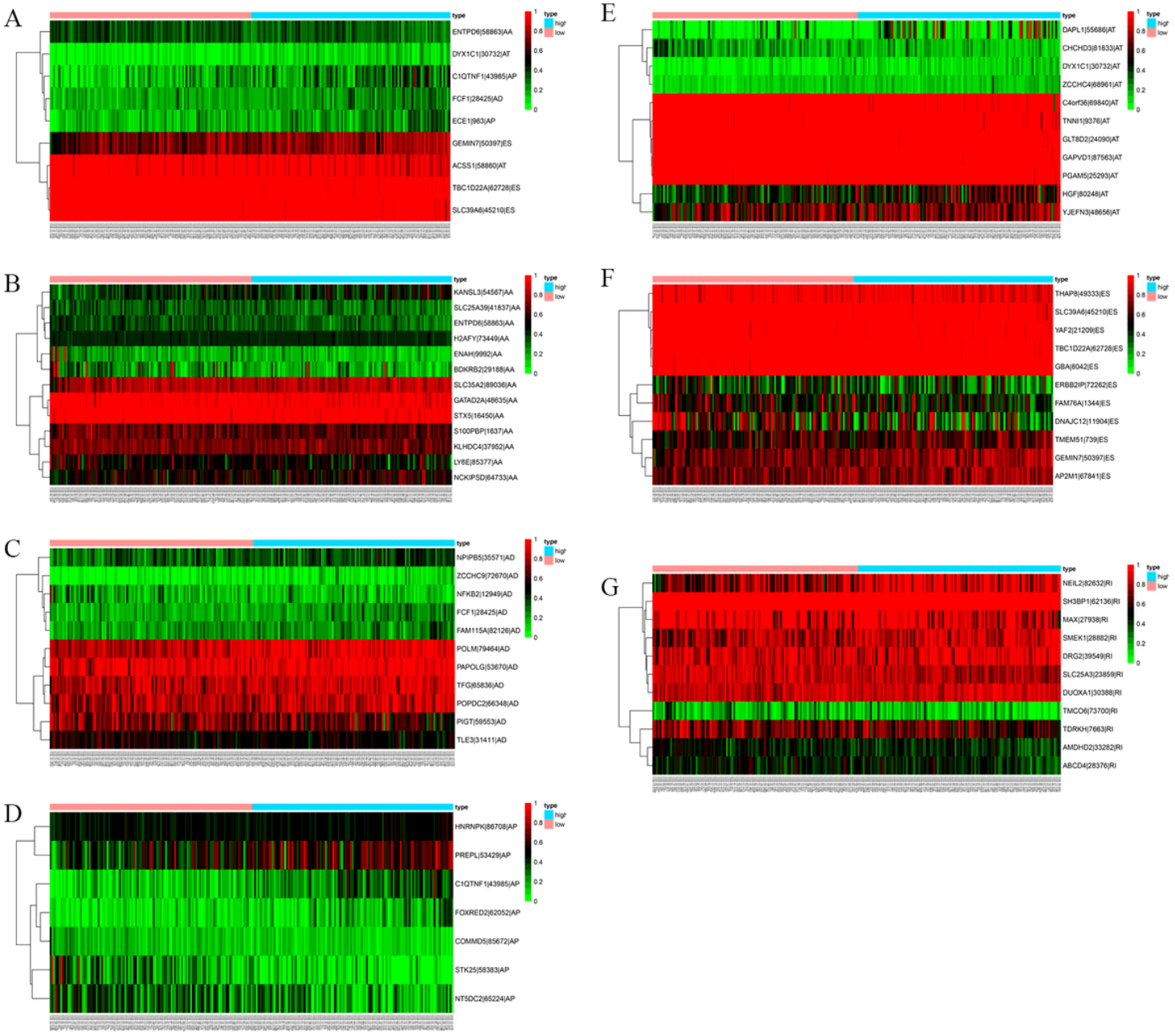

Figure 5

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.

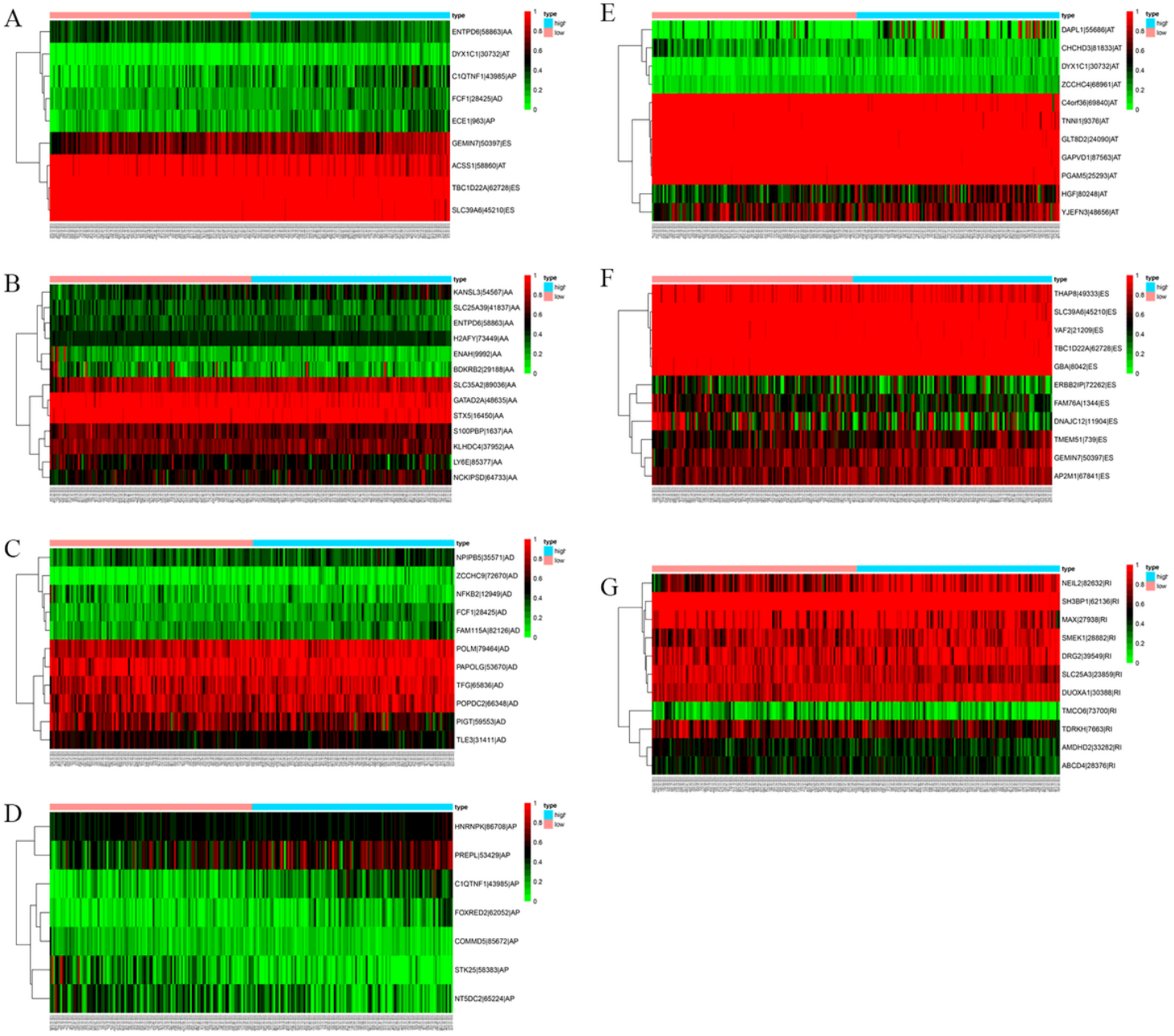

Figure 5

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.

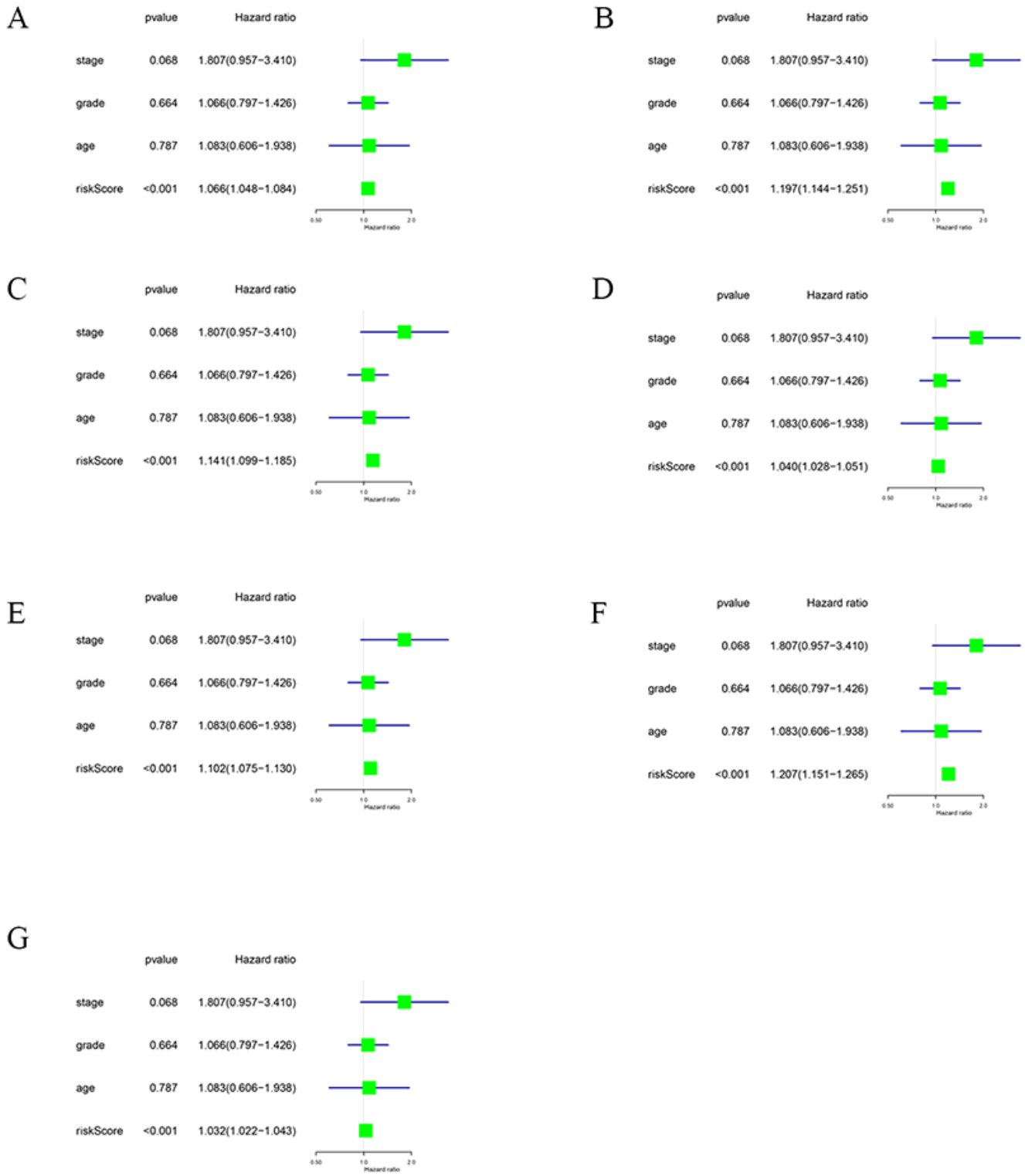

Figure 6

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.

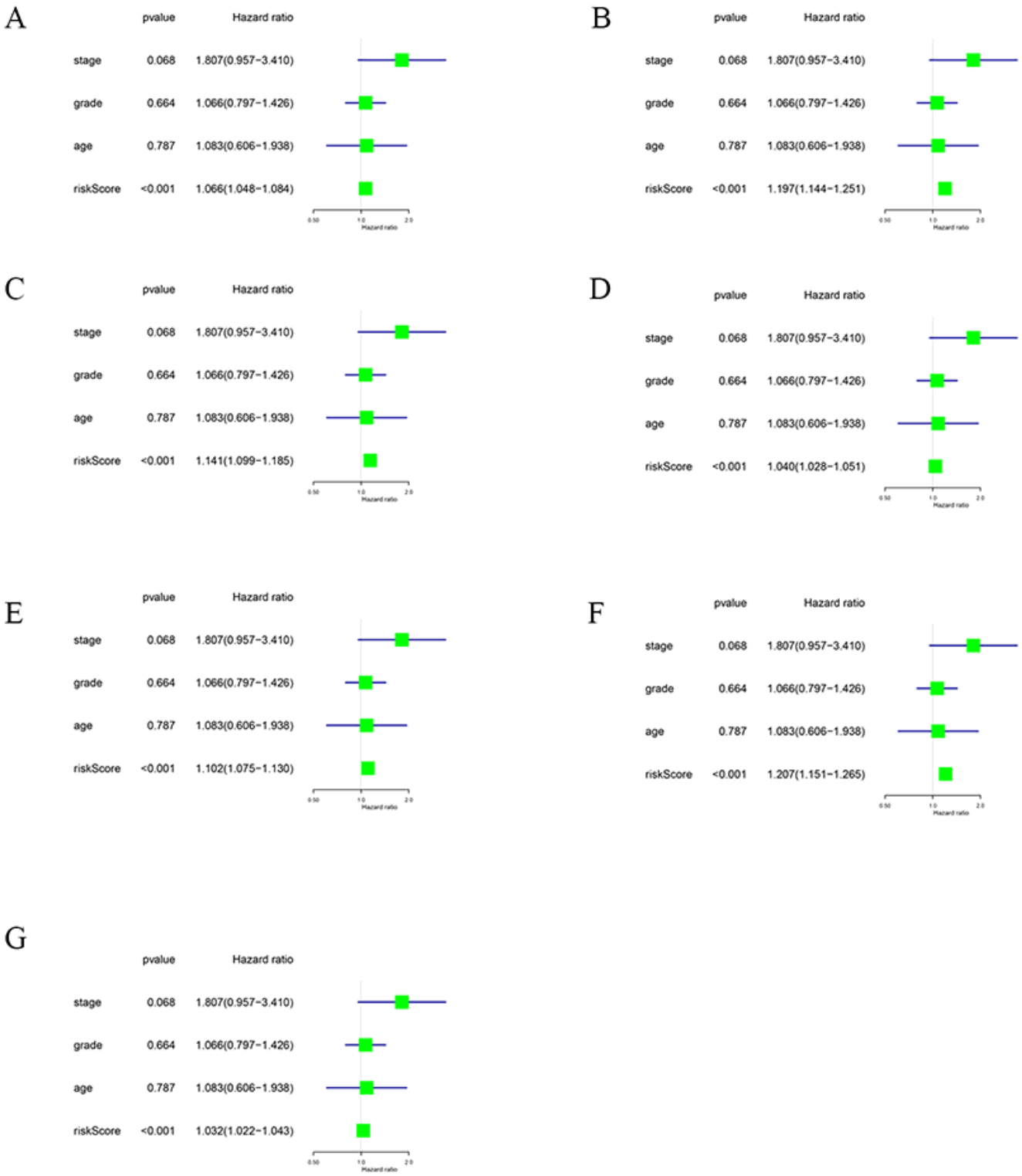

Figure 6

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.

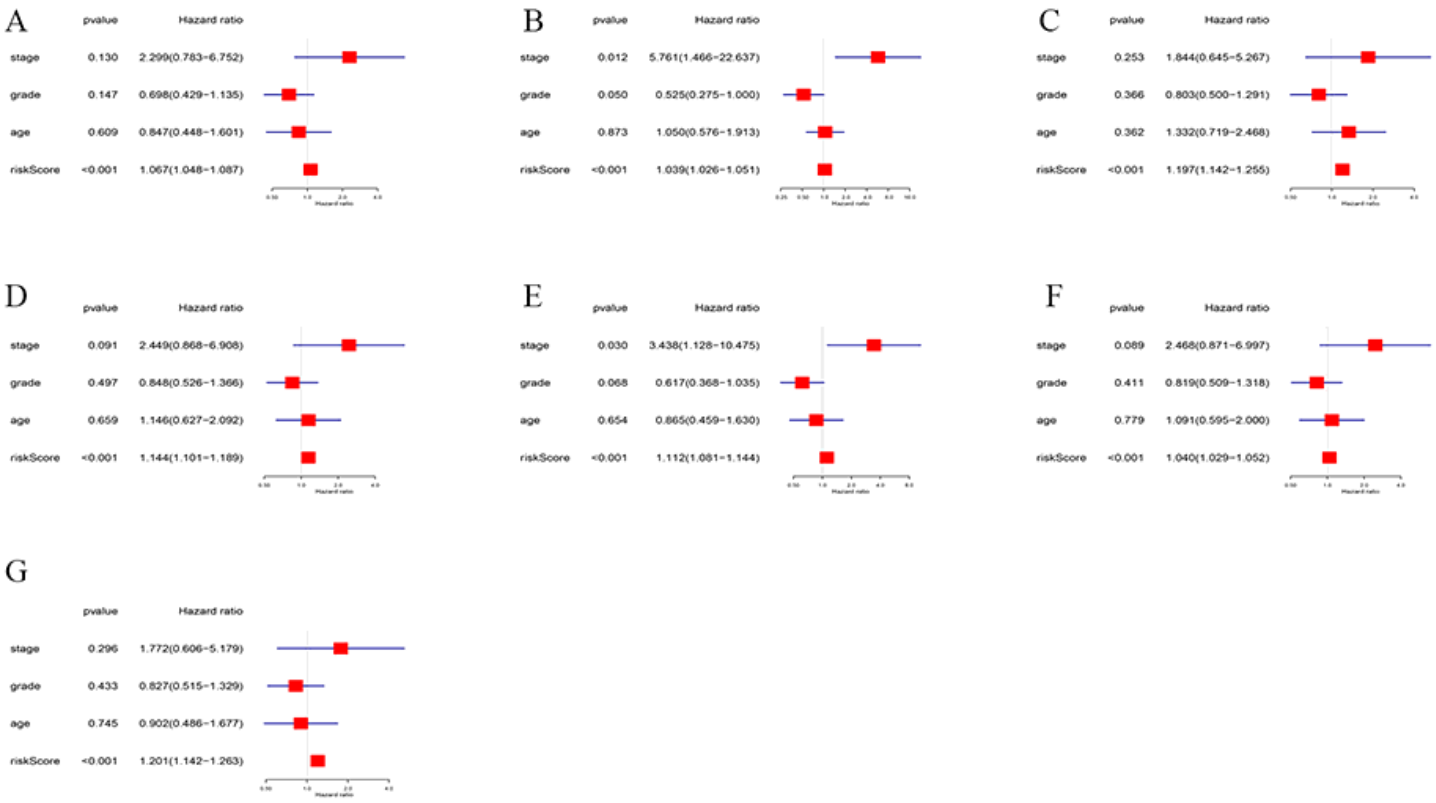

Figure 7

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.

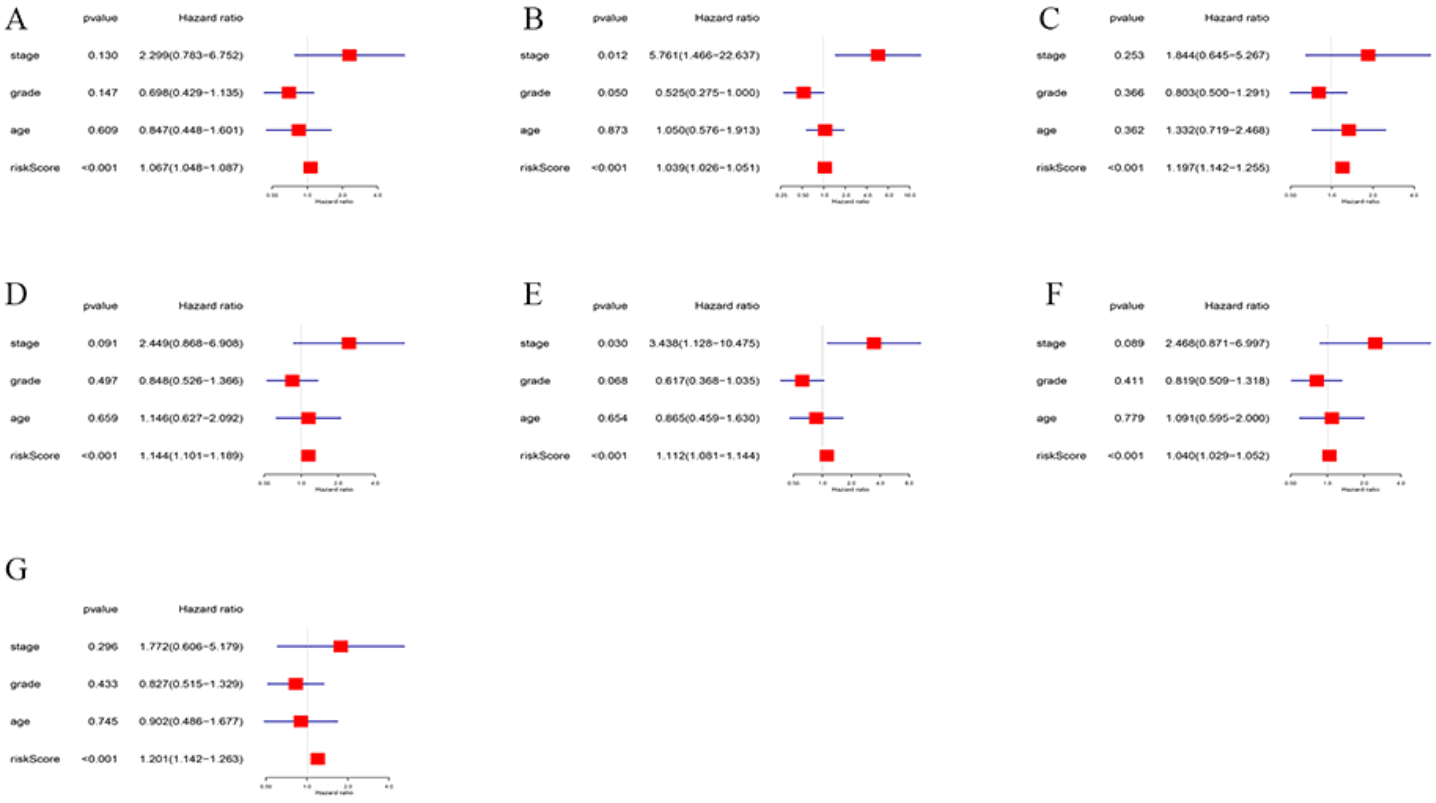

Figure 7

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.

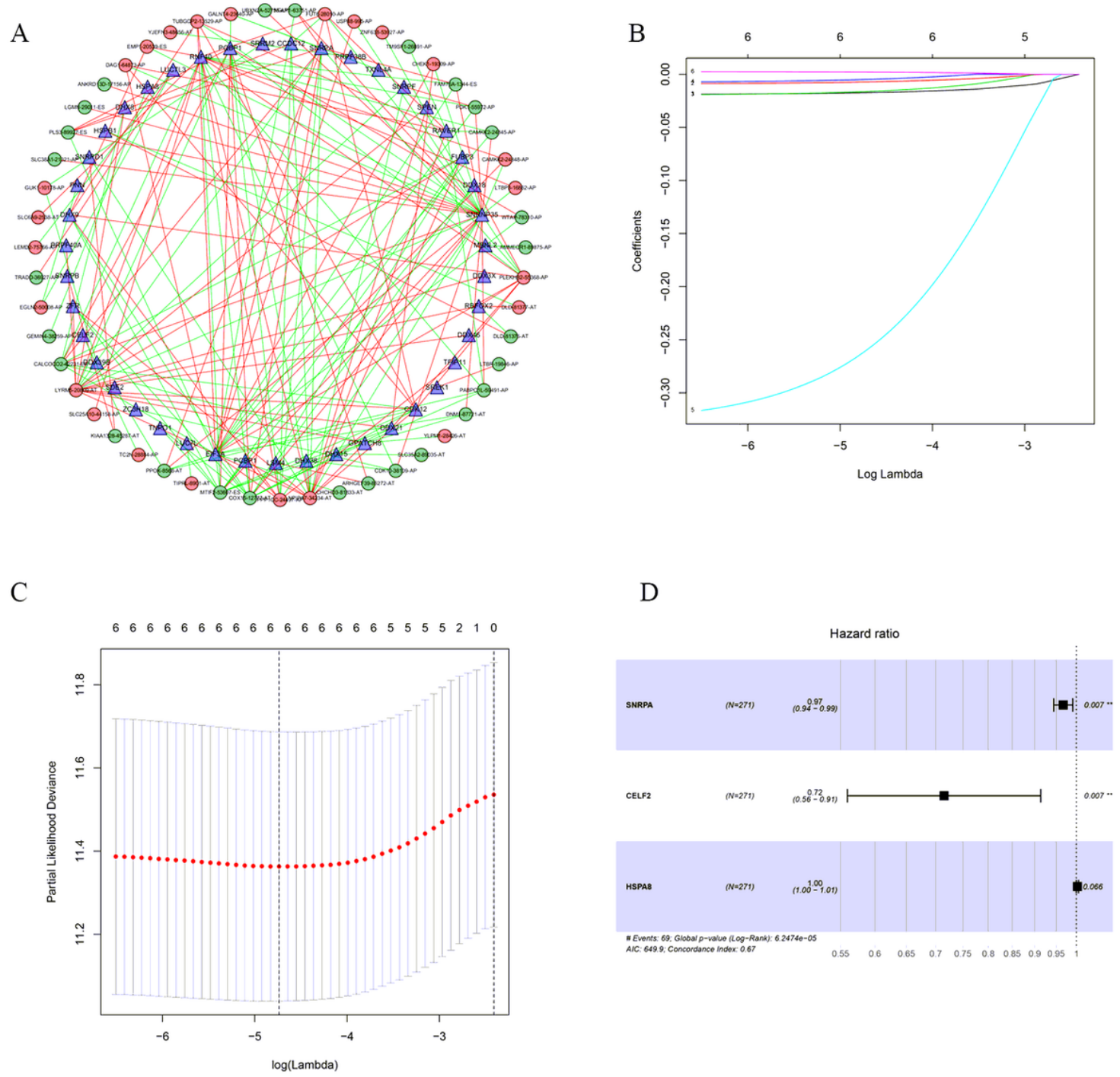

Figure 8

(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 represent AS events positively correlated with SF, and green dots represent AS events negatively correlated with SF. (B-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.

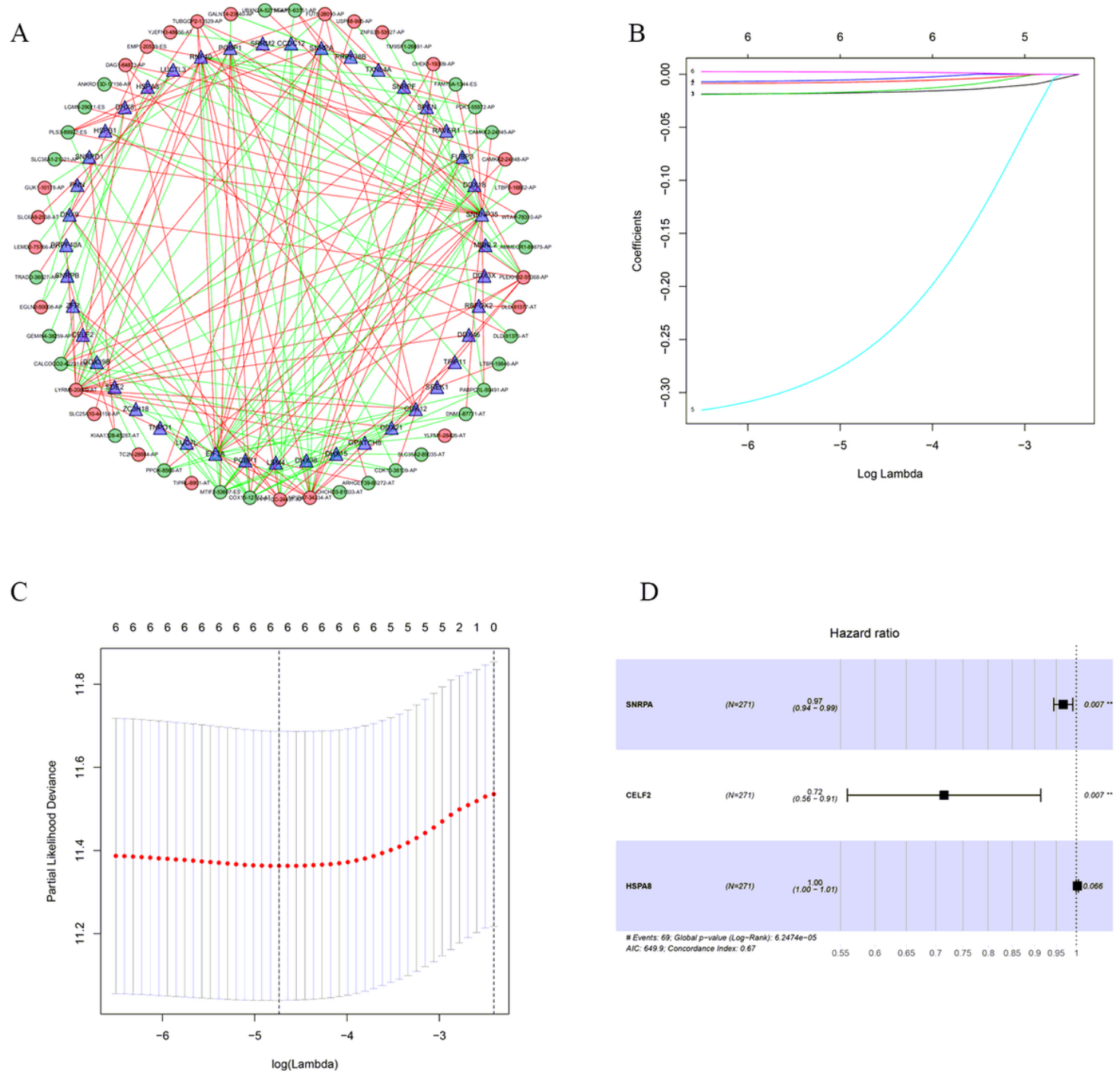

Figure 8

(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 represent AS events positively correlated with SF, and green dots represent AS events negatively correlated with SF. (B-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.

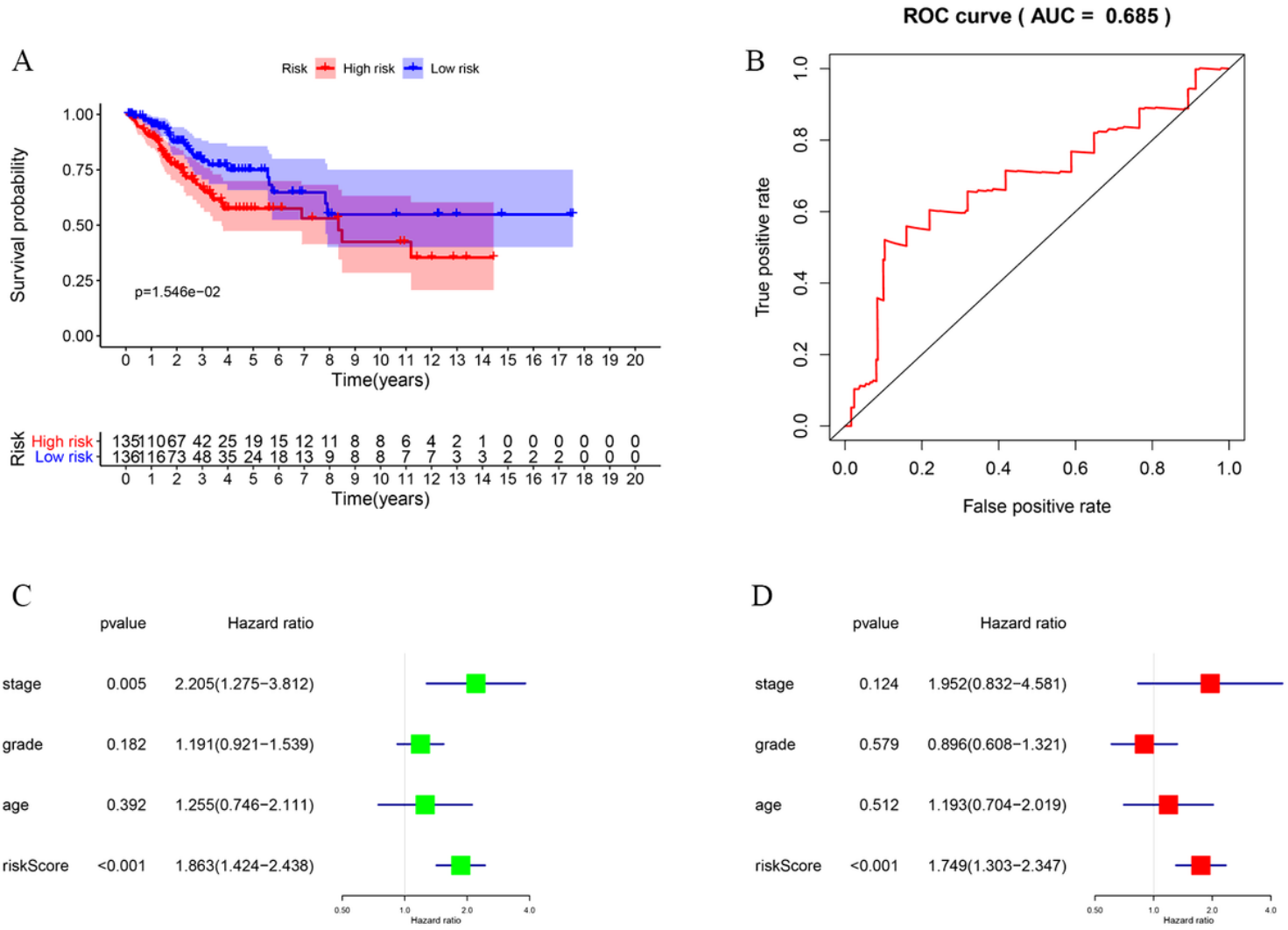

Figure 9

(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.

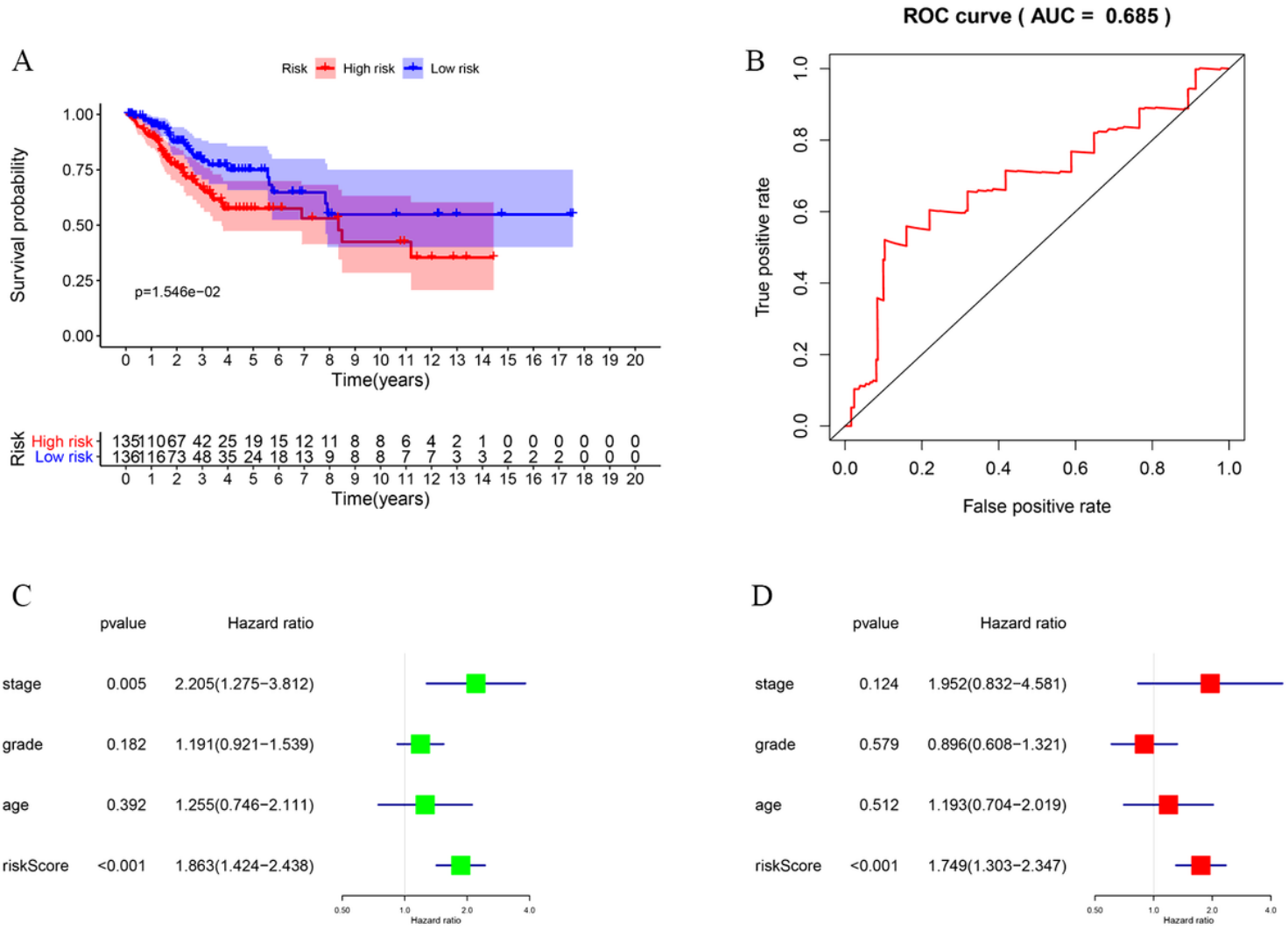

Figure 9

(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.

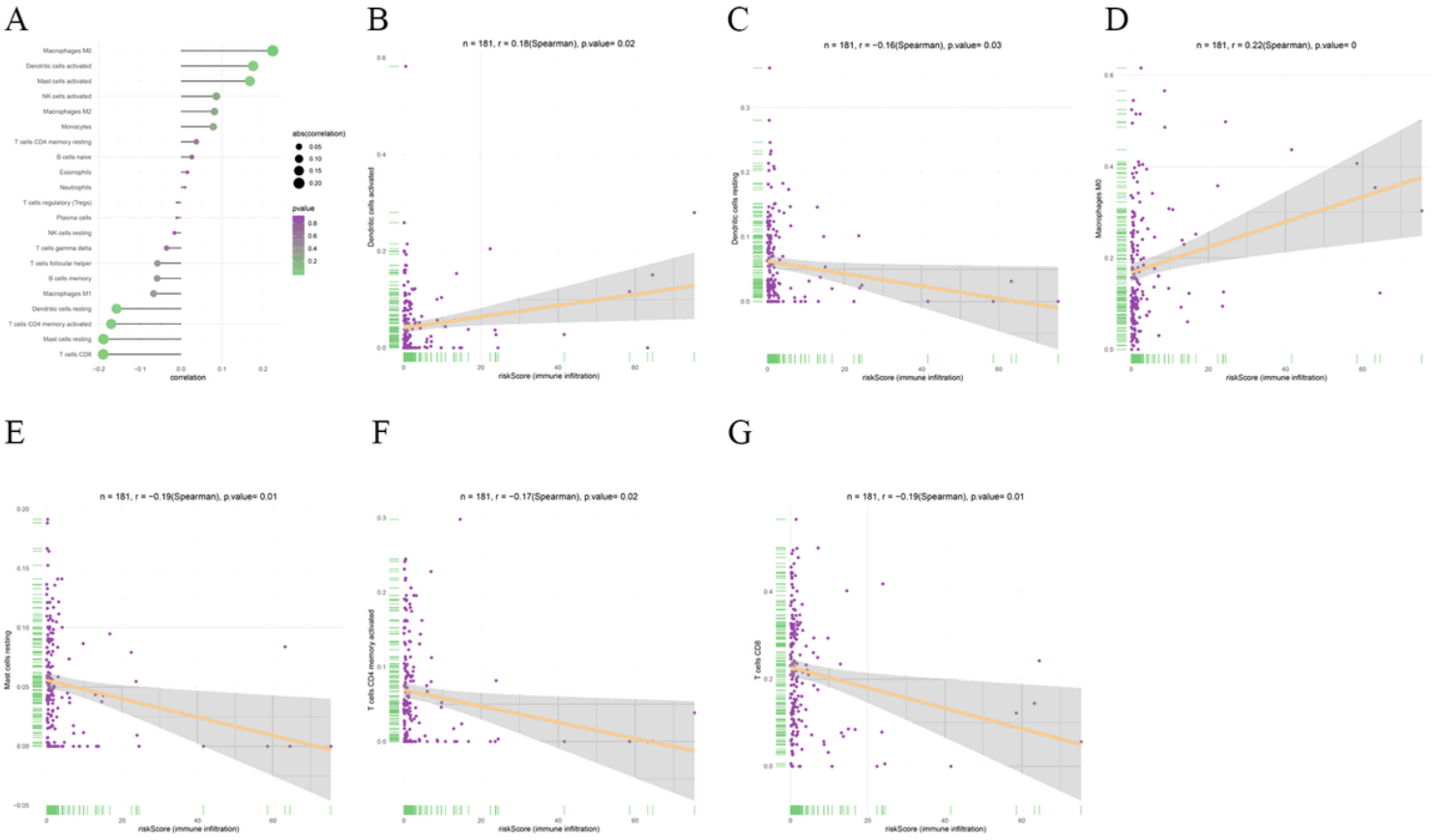

Figure 10

(A) The correlation of riskscore and tumor infiltrating immune cells. (B) The correlation of riskscore and Dendritic cells activated. (C) The correlation of riskscore and Dendritic cells resting. (D) The correlation of riskscore and Macrophages M0.(E) The correlation of riskscore and Mast cells resting. (F) The correlation of riskscore and T cells CD4 memory activated. (G) The correlation of riskscore and T cells CD8.

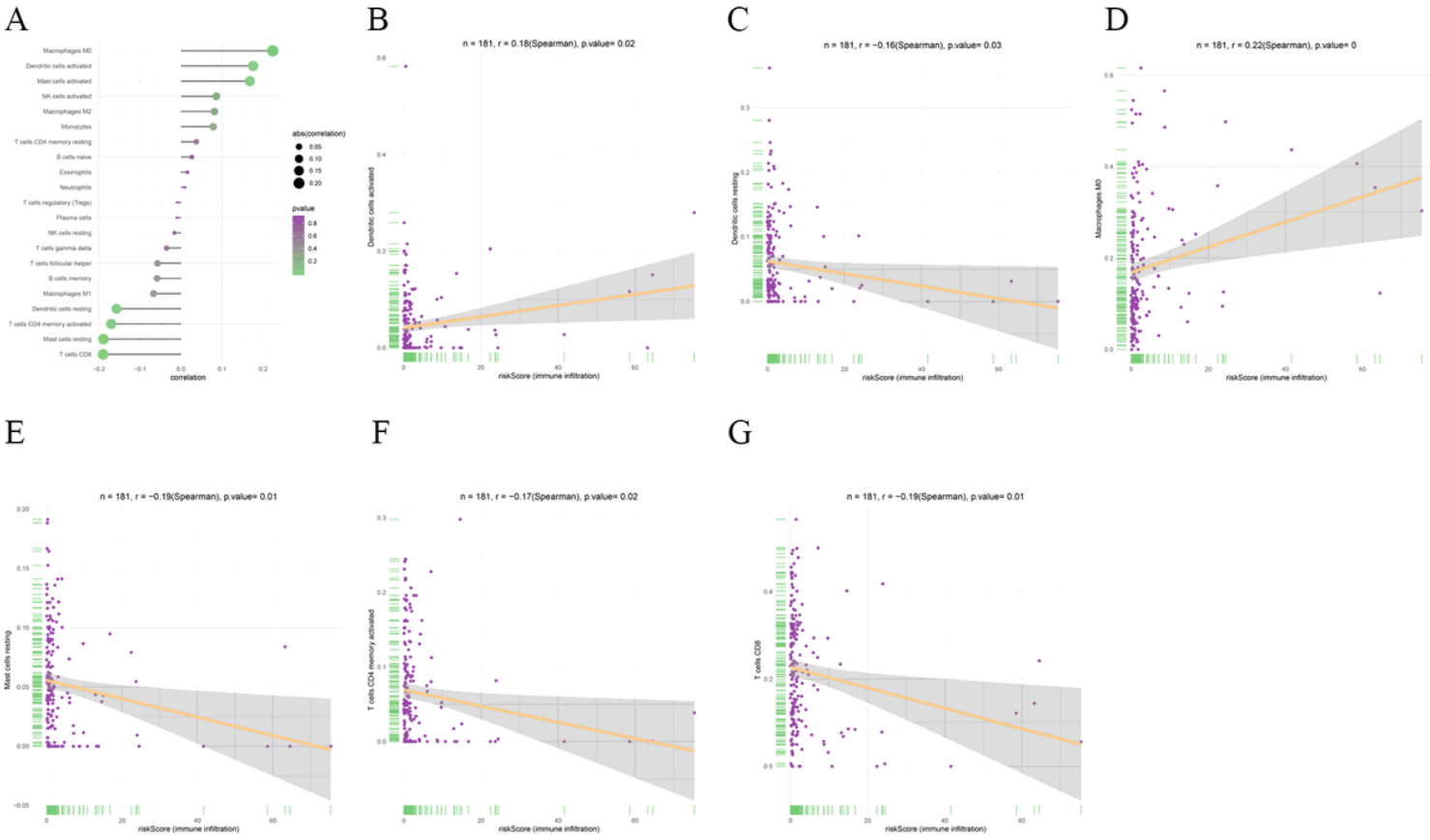

Figure 10

(A) The correlation of riskscore and tumor infiltrating immune cells. (B) The correlation of riskscore and Dendritic cells activated. (C) The correlation of riskscore and Dendritic cells resting. (D) The correlation of riskscore and Macrophages M0.(E) The correlation of riskscore and Mast cells resting. (F) The correlation of riskscore and T cells CD4 memory activated. (G) The correlation of riskscore and T cells CD8.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryfigurelegend.docx
- Supplementaryfigurelegend.docx
- tableS1.XLSX
- tableS1.XLSX
- tableS2.XLSX
- tableS2.XLSX
- tableS3.xlsx
- tableS3.xlsx