

# Comprehensive analysis of alternative splicing and transcriptome related to the prognosis of cervical squamous cell cancer

**Chen Li**

Medical College of Nantong University

**Fang Yan**

Sun Yat-sen University Cancer Center

**Gang Chen**

Sun Yat-sen University Cancer Center

**Qiang Li**

Sun Yat-sen University Cancer Center

**Hua Xian**

Affiliated Hospital of Nantong University

**Jingdun Xie** (✉ [xiejd6@mail.sysu.edu.cn](mailto:xiejd6@mail.sysu.edu.cn))

Sun Yat-sen University Cancer Center <https://orcid.org/0000-0002-3820-8435>

---

## Primary research

**Keywords:** Alternative splicing, Splicing factor, Pathway, Cervical squamous cell cancer, Spliceseq database

**Posted Date:** May 15th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-28212/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** Cervical squamous cell cancer (CESC) is a common malignancy tumor with high incidence and mortality in women globally. Increasing studies have indicated that there was an indivisible association between alternative mRNA splicing (AS) and multiple types of cancer. However, comprehensive analysis of alternative mRNA splicing events is scarce in CESC.

**Methods:** In this study, alternative mRNA splicing events data and clinical information of 216 CESC patients were downloaded from TCGA SpliceSeq database and TCGA website. We used identified survival-associated splicing events (SASEs) to construct prognostic signatures. Kaplan-Meier survival analysis and receiver operating characteristic (ROC) curve were performed to evaluate the clinical value of prognostic signatures. A nomogram was carried out to quantitatively predict individuals' survival probability. Regulatory network between splicing factors (SFs) and SASE was analyzed to explore the upstream regulators of a certain SASE. Additionally, we explored potential downstream pathways of a certain SASE.

**Results:** A total of 41776 alternative splicing events in 9961 genes were collected. After sorting out SASEs, multivariable regression analysis was used to acquire 70 SASEs that could be independent prognostic factors for overall survival in CESC. Kaplan-Meier survival analyses showed that the difference was statistically significant ( $P \leq 0.01$ ). The area under ROC curve (AUC) for all AS pattern was 0.932. A nomogram was constructed with a concordance index of 0.82 to predict individuals' survival probability. EIF3A positively regulated SEC23A-27346-AP with highest correlation coefficient ( $P \leq 0.001$ ,  $R=0.69$ ). Besides, the most significant SASEs and KEGG pathways were SEC23A-27346-AP and adherens junction pathway ( $P=0.02$ ,  $R=0.65$ ).

**Conclusions:** Prognostic signatures of survival-associated splicing events were independent prognostic factors for overall survival among CESC patients. A nomogram could quantitatively predict individuals' survival probability. Moreover, we proposed that splicing factor EIF3A positively regulated SEC23A-27346-AP, and adherens junction pathway may be the downstream pathway of SEC23A-27346-AP. The aforementioned signaling pathway could play a crucial role in the development of CESC.

## Background

Cervical squamous cell cancer (CESC) is the fourth most frequently detected malignancy and the fourth leading cause of cancer-associated death in women globally[1]. On the basis of the recent global cancer statistics, there were 570,000 newly diagnosed CESC cases and 311,000 deaths in 2018 worldwide[2]. The incidence and mortality of CESC in developing countries still drastically continue to rise, and CESC tends to occur in younger women[3]. Clinically, surgical resection, chemotherapy and radiotherapy, alone or combined, sometimes may be effective when treating early-stage, focal, sensitive CESC[4, 5]. However, for a considerable proportion of CESC patients with unsatisfying treatment effect, we starve for a novel therapeutic strategy. In addition, clinical prediction for survival time depends heavily on histopathologic

type, TNM stage and a fraction of tumor markers, while aforementioned indicators fail to predict survival time individually[6]. Along with the development of scientific research, we discovered a growing number of biomarker or molecular mechanisms underlying tumorigenesis and progression. These molecular markers gradually guide the nosological classification, precisely evaluate the prognostic outcome, and even become an effective therapeutic target[7–9].

Alternative mRNA splicing (AS) is a ubiquitous biological process that can remarkably increase protein species diversity under the circumstance of a relatively finite number of genes[10]. The exact mechanism is that AS enables precursor mRNA to generate different types of mature mRNA by multiple combinations of introns and exons. There are seven types of alternative splicing pattern, that is, mutually exclusive exons (ME), retained intron (RI), alternate donor site (AD), alternate acceptor site (AA), alternate promotor (AP), alternate terminator (AT) and exon skip (ES)[11]. Under the physiological condition, AS plays a vitally important regulating effect in the process of development, tissue identity, differentiation, cell-to-cell communication, cell senescence and apoptosis[12–14]. AS is accurately regulated by extensive splicing factors which participate in the precise selection of splicing sites and subsequent splicing events[15]. Undoubtedly, aberrant splicing factors or abnormal expression of splicing factors directly leads to changes of splicing events expression, thus causing abnormal cell growth[16]. There is an increasing number of studies that AS also makes a critical difference in the development and progression of cancer[17]. The main mechanism may be the involvement of AS throughout the whole stage of cancer, including tumorigenesis, abnormal proliferation, progression, angiogenesis, metastasis, and immune escape[18–20]. A recent study hinted that there was an indivisible association between AS and cancer prognosis[21, 22]. Therefore, it may be a credible prognostic biomarker and effective therapeutic target. Previous studies demonstrated that AS event-associated signatures may act as prognostic biomarkers in some cancers, including stomach adenocarcinoma, colorectal cancer, prostate adenocarcinoma, sarcoma, bladder urothelial carcinoma and so on[23–27]. However, comprehensive profiles of alternative splicing events are still scarce in cervical squamous cell cancer.

In this study, we constructed a splicing event-associated prognostic signature that could predict overall survival in cervical squamous cell carcinoma, and integrated multiple risk factors, including splicing event-associated prognostic signature and clinical characteristics, to predict the individualized survival probability. Moreover, we investigated the potential interaction network between splicing events and splicing factors, and downstream regulatory network of alternative splicing events in CESC.

## Methods

### Data acquirement of alternative splicing events

The data of alternative splicing events for CESC were downloaded from the TCGA SpliceSeq database[28]. Clinical characteristics of 308 CESC patients were obtained by searching the TCGA website. 261 patients whose survival time was more than 90 days met the inclusion criteria. After we

matched the patients with their TCGA SpliceSeq database items according to their sample ID, a total of 216 patients were included in our study.

## Identification and construction of the prognostic signature

Univariable cox regression analysis was carried out to identify survival-associated splicing events (SASEs). We used these candidate SASEs to construct prognostic signatures that could predict overall survival time. In order to avoid prognostic signatures overfitting and build an optimal prognostic model, the least absolute shrinkage and selection operator (LASSO) regression analysis was applied to screen out splicing events whose absolute value of coefficients were greater than a predetermined value by using the R package “glmnet”[29]. After excluding SASEs with zero coefficients in the LASSO regression analysis, we calculated risk score of each patient for overall survival prediction by using the formula:

$$\text{risk score} = \sum_i^n PSI_i * \beta_i$$

PSI<sub>i</sub>, that is, percent-spliced-in, is a ratio that indicates the efficiency of splicing of sequences of interest into transcripts, and can be used to undertake an intuitive quantitative comparison of splicing events[30].  $\beta_i$  signifies the coefficients of SASEs in the LASSO regression model. CESC patients were divided into low-risk and high-risk groups by using the median risk score as a cutoff value, respectively.

## Assessment of the clinical value of risk scores for prognostic signature

Kaplan-Meier (K-M) survival analysis and receiver operating characteristic (ROC) curve were performed to assess the prognostic value of risk scores, with the area under ROC curve (AUC) indicating the predictive efficacy of prognostic signature construction. Multivariable cox proportional hazards regression analysis was conducted to validate whether risk scores were independent prognostic factors by using the “forestplot” package for R software.

## Construction of nomogram

We constructed a nomogram to quantitatively figure out individuals' survival probability by incorporating multiple risk factors, including age, stage T, N, M and prognostic signatures of SASEs. According to risk contribution to survival probability, each risk factor was set as a specific point. To assess the consistency between actual and predicted survival time, calibration curves were performed for 3-year and 5-year survival. Besides, concordance index, that is C-index, was calculated to evaluate whether the nomogram model had excellent performance for predicting survival time[31].

## Construction of the potential correlation network

The data of splicing factors was extracted from the SpliceAid2 database, including 404 splicing factors[32]. Pearson correlation analysis was carried out to discover the relationship between PSI value of SASEs and expression level of SFs. Correlation coefficient greater than 0.400 and P value less than

0.05 were considered to be statistically significant. Cytoscape 3.6.0 was used to build the potential correlation network[33].

## **Correlation of SASEs and KEGG pathways**

We performed Gene Set Variation Analysis (GSVA) to screen out the differential Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways in CESC[34]. Univariable cox regression analysis was utilized to identify the prognostic KEGG pathways related to overall survival. Pearson correlation analysis was carried out to compute the correlation coefficient between SASEs and prognostic KEGG pathways.

## **Statistical analysis**

All statistical analyses were performed using R software version 3.4.1. Univariable and multivariable cox proportional hazards regression method were used to verify the association between alternative mRNA splicing events/prognostic signatures and overall survival. The least absolute shrinkage and selection operator (LASSO) regression analysis was applied to avoid prognostic signature overfitting. The receiver operating characteristic (ROC) curve was applied to assess the sensitivity and specificity of prognostic prediction of risk score. Pearson correlation analysis was performed to calculate the correlation coefficients between SFs/KEGG pathways and SASEs. A two-tailed  $P < 0.05$  was considered to be statistically significant.

## **Results**

### **Identification of survival-associated alternative mRNA splicing events in CESC**

The data of clinical characteristics and AS events of 216 CESC patients were analyzed. In total, 41776 AS events in 9961 genes were detected, including 209 MEs in 202 genes, 2723 RIs in 1800 genes, 3017 ADs in 2106 genes, 3424 AAs in 2398 genes, 8066 APs in 3258 genes, 8395 ATs in 3664 genes and 15942 ESs in 6278 genes. The intersection between AS events related genes and alternative splicing event patterns in 216 CESC patients are displayed. ES was the most common AS pattern and ME was the least common AS pattern in CESC, respectively. Besides, more than half of genes comprised ES events. In order to screen out survival associated alternative splicing events, univariable cox regression analysis was performed. Identified SASEs are illustrated in Upset plot. The most significant top 10 or 20 SASEs in each splicing pattern was selected.

### **Construction of prognostic signature for CESC**

In order to construct an optimal prognostic signature, we integrated these most significant candidate SASEs into LASSO cox regression analysis, and calculated regression coefficients. After excluding AS events which could make prognostic signature overfitting, we obtained 106 AS events significantly correlated with survival. Multivariable cox regression analysis was carried out to acquire 70 SASEs that could be an independent prognostic predictor for overall survival in CESC. Then, we separately calculated

risk score of each patient for overall survival prediction in each AS pattern as well as in all AS pattern. CESC patients were divided into low-risk and high-risk groups by using the median risk score as a cutoff value. The median risk score was set as 0.951 in AA pattern, 0.850 in AD pattern, 0.842 in AP pattern, 1.026 in AT pattern, 0.944 in ES pattern, 0.957 in ME pattern, 0.937 in RI pattern, and 0.809 in all pattern.

## Assessment of the clinical value of prognostic signature

To validate the predictive efficiency of prognostic signature, K-M survival analysis and ROC curve were conducted. K-M survival analyses showed that the difference was statistically significant between low-risk and high-risk groups for overall survival ( $P \leq 0.01$ ) (Fig. 1A-H). As shown in Fig. 1I-P, prognostic signatures of each AS pattern and all AS pattern had considerable predictive efficacy in differentiating better or worse outcomes for CESC patients. Prognostic signatures for AA (AUC = 0.951) and RI (AUC = 0.895) had optimal performance for predicting overall survival, and the area under ROC curve (AUC) for all AS pattern was 0.932. The correlation between risk score and overall survival was displayed in scatter plot and risk plot for 216 CESC samples. The red and green dots presented high-risk and low-risk group in risk plot. While in the scatterplot, these dots indicated survival status and overall survival of CESC patients, divided by median risk score of each AS pattern or all AS pattern. Expression level of SASEs in each AS pattern was visualized in the heatmap, with the red and green stripe indicating high and low PSI values, and the red and blue bars presenting low-risk and high-risk groups (Fig. 2A-H). Next, univariable and multivariable cox proportional hazards regression analyses were used to confirm whether prognostic signatures were independent prognostic factors in each AS pattern and all AS pattern. After integrating age, grade and cancer status, multivariate cox regression analyses confirmed that prognostic signatures for risk score were exactly independent prognostic factors for CESC patients ( $P \leq 0.01$ ). The aforementioned results were shown in forest map (Fig. 3A-H).

## Nomogram

As we all known, age and TNM stage are independent prognostic factors in CESC. In order to obtain individuals' survival probability, we performed a comprehensive analysis of multiple risk factors. We constructed a nomogram to quantitatively predict the 3-year and 5-year survival by integrating age, stage T, N, M and prognostic signature of SASEs. According to risk contribution to survival probability, each risk factor was set as a specific point. As shown in Fig. 4A, age less than 50 years old was assigned 68.75 points, stage T1-T2 was assigned 0 points, stage N1 was assigned 7.50 points, stage M0 was assigned 0 points, high-risk score was assigned 71.25 points and so on. Calibration curves showed that there was excellent consistency between actual and predicted survival time (Fig. 4B-C). The concordance index (C-index) was 0.82, indicating that the model had an excellent performance for predicting overall survival in CESC. For example, a 45-year old patient with high-risk score was in stage\_T1-T2, stage\_Nx, stage\_M0 of cervical squamous cell cancer, she would get 147.5 points. Her 3-year survival rate was 65%, and 5-year survival rate was 50%.

## SASEs and SFs correlation network in CESC

To explore the upstream regulators of SASEs, we carried out the Pearson correlation analysis to discover the relationship between SASEs and SFs. The correlation network between 116 SASEs and 15 SFs were illustrated in Fig. 5. The purple arrows signified SFs, red and green circles represented high and low hazard ratio (HR) of SASEs by univariable cox regression analysis, as well as red and green lines indicated positive and negative regulation between SASEs and SFs. Our results showed that EIF3A had a positive regulation of SEC23A-27346-AP with highest correlation coefficient.

## Correlation of SASEs and KEGG pathways

To explore the downstream pathways of SASEs, we screened out 185 differential KEGG pathways in CESC by using Gene Set Variation Analysis (GSVA). Univariable cox regression analysis was applied to identify 16 prognostic KEGG pathways related to overall survival. Then, the correlation coefficient of SASEs and prognostic KEGG pathways was calculated by using Pearson correlation analysis. The correlation between SASEs and prognostic KEGG pathways was illustrated in heatmap (Fig. 6). Red and blue dots represented positive and negative correlation between SASEs and prognostic KEGG pathways. The most significant SASEs and prognostic KEGG pathways were SEC23A-27346-AP and KEGG\_ADHERENS\_JUNCTION ( $P = 0.02$ ,  $R = 0.65$ ). Combining the correlation network between SASEs and SFs, EIF3A was the most significant SF positively related to SEC23A-27346-AP ( $P < 0.001$ ,  $R = 0.69$ ). As a consequence, the most significant SF, SASEs, and downstream pathway were EIF3A, SEC23A-27346-AP and adherens junction pathway, respectively.

## Discussions

CESC is one of the frequently detected malignancies and pose a health threat to women all over the world[1–3]. Clinical stage, pathologic type, response to therapy and long-term prognosis can vary considerably among CESC patients. It is important for CESC patients to obtain survival time individually. Research suggested that biomarkers, such as alternative mRNA splicing, might be an effective prognostic predictor[7]. AS is a vital biological process that can expand protein species diversity[10]. There is an increasing number of studies that AS plays a crucial role in the oncogenesis and development of cancer. Previous studies reported that AS may accurately evaluate prognostic outcomes in some cancers, including glioblastoma multiforme, colorectal cancer, prostate adenocarcinoma, bladder urothelial carcinoma, gastric cancer and so on[24–26, 35]. However, the relevant research of cervical squamous cell cancer is scarce.

In the present study, we constructed AS-associated prognostic signatures. AS-associated prognostic signatures were exactly independent prognostic factors in CESC. And prognostic signatures had high AUC values, proving prognostic signatures had reliable predictive efficacy. The findings highlight the clinical significance of AS-associated prognostic signatures in CESC. After integrating age, stage T, N, M and AS-associated prognostic signatures, we carried out a comprehensive analysis, that is nomogram, to quantitatively predict individuals' survival probability. Compared with histopathological type or TNM stage, it had a more excellent performance in predicting survival time in CESC. These findings suggest

that AS-associated prognostic signatures may eventually act as efficient prognostic biomarkers to assess survival time in CESC. This is the first and the most comprehensive analysis to construct a prognostic prediction model, which incorporate age, TNM stage and AS-associated prognostic signatures, in cervical squamous cell cancer.

Alternative splicing events are thought to be accurately regulated by hundreds of splicing factors, which specifically combine with the splicing sites and participate in subsequent splicing events. There is no doubt that aberrant splicing factors or abnormal expression of splicing factors directly leads to changes of splicing events pattern and splicing events expression[15, 16]. Some studies reported that there was correlation between splicing factors and poor prognosis in chronic lymphocytic leukemia, lung adenocarcinoma, pancreatic cancer, breast cancer and melanoma[36–38]. In our study, splicing factors were collected from the SpliceAid2 database, and correlation network between SASEs and SFs was explored. The most significant SASEs and SFs were SEC23A-27346-AP and EIF3A with the maximum correlation coefficient. EIF3A had a strongly positive regulation with SEC23A-27346-AP. Our findings revealed that EIF3A might be the upstream regulator of SEC23A-27346-AP. The in-depth correlation analysis of regulation network between SESAs and SFs provides a new insight into the mechanism of AS involved in the unfavorable prognosis of cervical squamous cell cancer.

Alternative splicing events not only can be regulated by splicing factors, but also can respond to multiple signaling pathways. As we all know, weakening or lacking tight cell-cell adherens junction has been considered to be a fatal hallmark of cancer[39, 40]. In normal physiological conditions, cell-cell adherens junction is the foundation of tissues architecture[41]. However, tumor cells tend to loss of adherens junctions, and display local dissemination and distant metastasis especially in advanced cancers. Cadherins and catenins are the core proteins of cell-cell adherens junction. Low or loss of expression of cadherins and catenins inevitably lead to tumor progression and unfavorable prognosis[42]. In the current study, we selected prognostic KEGG pathways in CESC by Gene Set Variation Analysis (GSVA). The correlation analysis between SASEs and prognostic KEGG pathways was carried out. The most significant SASEs and prognostic KEGG pathways were SEC23A-27346-AP and adherens junction pathway with the maximum correlation coefficient. Therefore, we supposed that adherens junction pathway might be one of the downstream pathways of SEC23A-27346-AP in CESC.

SEC23A, as a core component of coat protein complex  $\sigma$  vesicles, can not only transport proteins but also cause human diseases. Aberrant SEC23A or abnormal expression of SEC23A was reported to be relevant to the development and progression of human cancer[43]. A previous study implicated that RNA binding motif 5 (RBM5) participate in regulating mRNA splicing of SEC23A[44]. However, our results indicated that SEC23A-27346-AP was associated with poor prognosis in CESC patients, and splicing factor EIF3A positively regulated SEC23A-27346-AP. Above all, SEC23A regulated by EIF3A may facilitate the development of cervical squamous cell cancer via adherens junction pathway.

There are some limitations in this study. Firstly, the number of CESC patients in TCGA database was relatively small. Clinical information of 308 CESC patients was extracted from the TCGA website. After

excluding the patients who didn't accord with the inclusive criteria, a total of 216 CESC patients were enrolled in our study. Secondly, we performed a nomogram to more precisely obtain individuals' survival probability by integrating multiple risk factors. However, we failed to take into consideration other risk factors, such as pathological grade, therapeutic strategies and so on. Moreover, the mechanism of alternative splicing events in the development of CESC was unclear. We merely carried out correlation analysis to explore upstream regulators and downstream pathways of alternative splicing events. A number of known upstream regulators or downstream pathways of SEC23A were displayed, but other more important SFs or downstream pathways are needed to be identified and verified. In the follow-up study, the molecular mechanisms of alternative splicing events in the occurrence and development of cervical squamous cell cancer need further investigations to verify.

## Conclusions

We constructed a splicing event-associated prognostic signature that could predict overall survival in cervical squamous cell carcinoma, and integrated multiple risk factors, including splicing event-associated prognostic signature and clinical characteristics, to predict the individualized survival probability. Moreover, the potential interaction network between splicing events and splicing factors as well as downstream regulatory mechanisms of AS in CESC were provided. Our results suggested that SEC23A positively regulated by EIF3A may contribute to cervical squamous cell cancer via adherens junction pathway. This study may offer new direction for subsequent experimental research in cervical squamous cell cancer.

## Abbreviations

AA: Alternate acceptor site; AD: Alternate donor site; AP: Alternate promoter; AS: Alternative mRNA splicing; AT: Alternate terminator; AUC: Area under ROC curve; CESC: Cervical squamous cell cancer; CI: Confidence interval; Coef; Coefficient; ES: Exon skip; GSVA: Gene set variation analysis; HR: Hazard ratio; KEGG: Kyoto Encyclopedia of Genes and Genomes; LASSO: Least absolute shrinkage and selection operator; ME: Mutually exclusive exons; RI: Retained intron; ROC: Receiver operating characteristic; SASEs: Survival associated splicing events; SEC23: Secretory 23; SF: Splicing factor; TCGA: The cancer genome atlas;

## Declarations

### Acknowledgements

We thank the TCGA website and TCGA SpliceSeq database for providing platforms and valuable data sets.

### Authors' Contributions

CL, FY, HX and JDX conceived and designed the study. GC and QL provided the material and patient data of the study. CL and FY analyzed the data. FY wrote the manuscript. HX and JDX revised the manuscript. All authors read and approved the final manuscript

## **Funding**

This work was supported by the National Natural Science Foundation of China (grant # 81870878) and Guang Dong Natural Science Foundation (2018A030313092).

## **Availability of data and materials**

The data were available from the TCGA database.

## **Ethics approval and consent to participate**

The data in this study were obtained mainly from the TCGA databases, which are open access.

## **Consent for publication**

Not applicable.

## **Competing Interests**

The authors have declared that they have no competing interests.

## **References**

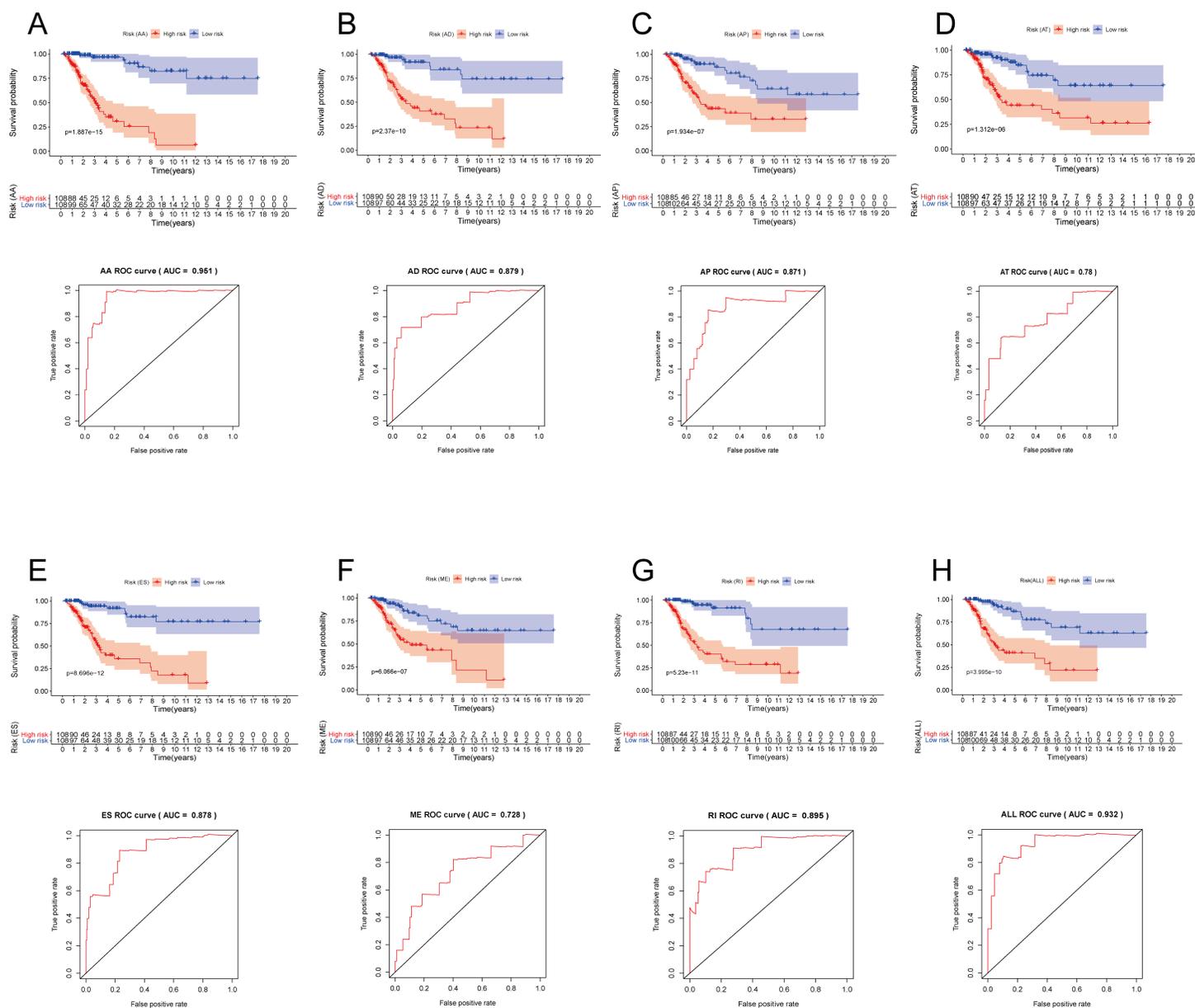
1. Fader AN. Surgery in Cervical Cancer. *N Engl J Med*. 2018;379(20):1955–7.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Cancer J Clin*. 2018;68(6):394–424.
3. Vu M, Yu J, Awolude OA, Chuang L. Cervical cancer worldwide. *Curr Probl Cancer*. 2018;42(5):457–65.
4. Naga Ch P, Gurram L, Chopra S, Mahantshetty U. The management of locally advanced cervical cancer. *Current opinion in oncology*. 2018;30(5):323–9.
5. Wang C, Yang C, Wang W, Xia B, Li K, Sun F, Hou Y. A Prognostic Nomogram for Cervical Cancer after Surgery from SEER Database. *J Cancer*. 2018;9(21):3923–8.
6. Park KJ, Braschi-Amirfarzan M, DiPiro PJ, Giardino AA, Jagannathan JP, Howard SA, Shinagare AB, Krajewski KM. Multimodality imaging of locally recurrent and metastatic cervical cancer: emphasis on histology, prognosis, and management. *Abdominal radiology (New York)*. 2016;41(12):2496–508.
7. Gao C, Zhou C, Zhuang J, Liu L, Liu C, Li H, Liu G, Wei J, Sun C. MicroRNA expression in cervical cancer: Novel diagnostic and prognostic biomarkers. *Journal of cellular biochemistry*.

- 2018;119(8):7080–90.
8. Aalijahan H, Ghorbian S. Long non-coding RNAs and cervical cancer. *Exp Mol Pathol*. 2019;106:7–16.
  9. Liu S, Meng F, Ding J, Ji H, Lin M, Zhu J, Ma R. **High TRIM44 expression as a valuable biomarker for diagnosis and prognosis in cervical cancer.** *Bioscience reports* 2019, 39(3).
  10. Ule J, Blencowe BJ. Alternative Splicing Regulatory Networks: Functions, Mechanisms, and Evolution. *Molecular cell*. 2019;76(2):329–45.
  11. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li Yi, Kosmicki JA, Arbelaez J, Cui W, Schwartz GB, et al. Predicting Splicing from Primary Sequence with Deep Learning. *Cell*. 2019;176(3):535–48.e524.
  12. Baralle FE, Giudice J. Alternative splicing as a regulator of development and tissue identity. *Nature reviews Molecular cell biology*. 2017;18(7):437–51.
  13. Furlanis E, Scheiffle P. Regulation of Neuronal Differentiation, Function, and Plasticity by Alternative Splicing. *Annu Rev Cell Dev Biol*. 2018;34:451–69.
  14. Lin JC, Tsao MF, Lin YJ. **Differential Impacts of Alternative Splicing Networks on Apoptosis.** *International journal of molecular sciences* 2016, 17(12).
  15. Lee Y, Rio DC. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annual review of biochemistry*. 2015;84:291–323.
  16. Yoshida K, Ogawa S. Splicing factor mutations and cancer. *Wiley interdisciplinary reviews RNA*. 2014;5(4):445–59.
  17. Climente-Gonzalez H, Porta-Pardo E, Godzik A, Eyraes E. The Functional Impact of Alternative Splicing in Cancer. *Cell reports*. 2017;20(9):2215–26.
  18. Bowler E, Oltean S. **Alternative Splicing in Angiogenesis.** *International journal of molecular sciences* 2019, 20(9).
  19. Marzese DM, Manughian-Peter AO, Orozco JIJ, Hoon DSB. Alternative splicing and cancer metastasis: prognostic and therapeutic applications. *Clin Exp Metastasis*. 2018;35(5–6):393–402.
  20. Oltean S, Bates DO. Hallmarks of alternative splicing in cancer. *Oncogene*. 2014;33(46):5311–8.
  21. Xiping Z, Qingshan W, Shuai Z, Hongjian Y, Xiaowen D. A summary of relationships between alternative splicing and breast cancer. *Oncotarget*. 2017;8(31):51986–93.
  22. Cao ZX, Xiao GA, Zhang W, Ji J, Ye C, Liu D, Tian QQ, Prof YS. Comprehensive investigation of alternative splicing and development of a prognostic risk score for prostate cancer based on six-gene signatures. *J Cancer*. 2019;10(22):5585–96.
  23. Shi Y, Chen Z, Gao J, Wu S, Gao H, Feng G. Transcriptome-wide analysis of alternative mRNA splicing signature in the diagnosis and prognosis of stomach adenocarcinoma. *Oncol Rep*. 2018;40(4):2014–22.
  24. Liu J, Li H, Shen S, Sun L, Yuan Y, Xing C. Alternative splicing events implicated in carcinogenesis and prognosis of colorectal cancer. *J Cancer*. 2018;9(10):1754–64.

25. Huang ZG, He RQ, Mo ZN. Prognostic value and potential function of splicing events in prostate adenocarcinoma. *Int J Oncol.* 2018;53(6):2473–87.
26. Hong W, Zhang W, Guan R, Liang Y, Hu S, Ji Y, Liu M, Lu H, Yu M, Ma L. Genome-wide profiling of prognosis-related alternative splicing signatures in sarcoma. *Annals of translational medicine.* 2019;7(20):557.
27. He RQ, Zhou XG, Yi QY, Deng CW, Gao JM, Chen G, Wang QY. Prognostic Signature of Alternative Splicing Events in Bladder Urothelial Carcinoma Based on Spliceseq Data from 317 Cases. *Cellular physiology biochemistry: international journal of experimental cellular physiology biochemistry pharmacology.* 2018;48(3):1355–68.
28. Ryan M, Wong WC, Brown R, Akbani R, Su X, Broom B, Melott J, Weinstein J. TCGASpliceSeq a compendium of alternative mRNA splicing in cancer. *Nucleic acids research.* 2016;44(D1):D1018–22.
29. Shen Y, Zhu H, Yang Y, Shen Y. A sparse loudspeaker array for surround sound reproduction using the least absolute shrinkage and selection operator algorithm. *J Acoust Soc Am.* 2019;145(5):E1430.
30. Schafer S, Miao K, Benson CC, Heinig M, Cook SA, Hubner N: **Alternative Splicing Signatures in RNA-seq Data: Percent Spliced in (PSI).** *Current protocols in human genetics* 2015, **87**:11.16.11–11.16.14.
31. Park SY. Nomogram: An analogue tool to deliver digital knowledge. *J Thorac Cardiovasc Surg.* 2018;155(4):1793.
32. 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):81–5.
33. Su G, Morris JH, Demchak B, Bader GD. Biological network exploration with Cytoscape 3. *Current protocols in bioinformatics.* 2014;47:8.13.11–24.
34. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research.* 2000;28(1):27–30.
35. Huang R, Li Z, Li C, Wang G, Yan P, Peng L, Wang J, Zhu X, Hu P, Zhang J, et al. Germ Cell-Specific Gene 1-Like Protein Regulated by Splicing Factor CUGBP Elav-Like Family Member 5 and Primary Bile Acid Biosynthesis are Prognostic in Glioblastoma Multiforme. *Frontiers in genetics.* 2019;10:1380.
36. Liao W, Jordaan G, Nham P, Phan RT, Pelegrini M, Sharma S. Gene expression and splicing alterations analyzed by high throughput RNA sequencing of chronic lymphocytic leukemia specimens. *BMC Cancer.* 2015;15:714.
37. Jia R, Ajiro M, Yu L, McCoy P Jr, Zheng ZM. Oncogenic splicing factor SRSF3 regulates ILF3 alternative splicing to promote cancer cell proliferation and transformation. *RNA.* 2019;25(5):630–44.
38. Seiler M, Peng S, Agrawal AA, Palacino J, Teng T, Zhu P, Smith PG, Buonamici S, Yu L. Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types. *Cell reports.* 2018;23(1):282–96.e284.
39. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011;144(5):646–74.

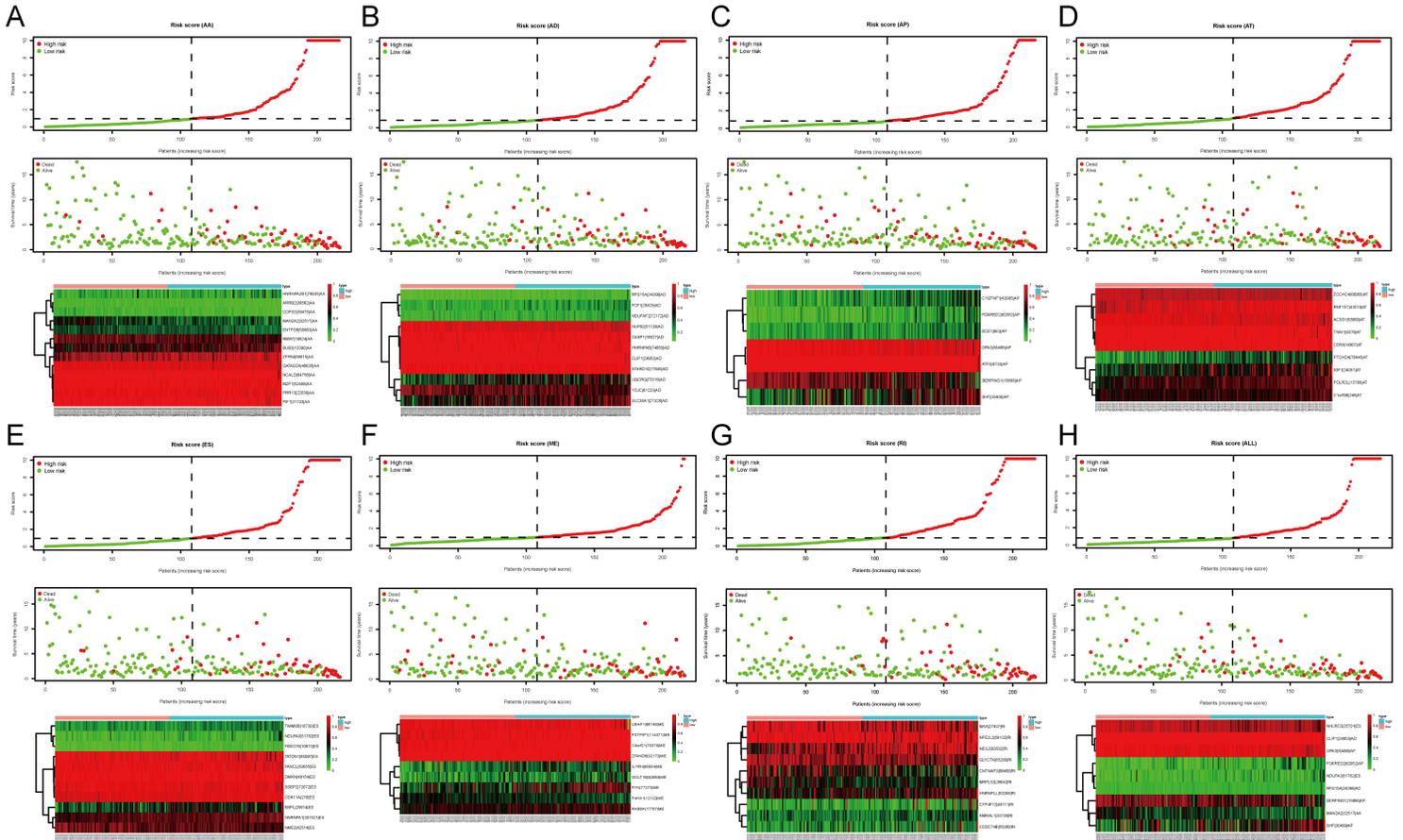
40. Gloushankova NA, Rubtsova SN, Zhitnyak IY. Cadherin-mediated cell-cell interactions in normal and cancer cells. *Tissue barriers*. 2017;5(3):e1356900.
41. Mege RM, Ishiyama N. **Integration of Cadherin Adhesion and Cytoskeleton at Adherens Junctions.** *Cold Spring Harbor perspectives in biology* 2017, 9(5).
42. Wong SHM, Fang CM, Chuah LH, Leong CO, Ngai SC: **E-cadherin: Its dysregulation in carcinogenesis and clinical implications.** *Critical reviews in oncology/hematology* 2018, **121**:11–22.
43. Jing J, Wang B, Liu P. The Functional Role of SEC23 in Vesicle Transportation, Autophagy and Cancer. *Int J Biol Sci*. 2019;15(11):2419–26.
44. Jackson TC, Kotermanski SE, Kochanek PM. Whole-transcriptome microarray analysis reveals regulation of Rab4 by RBM5 in neurons. *Neuroscience*. 2017;361:93–107.

## Figures



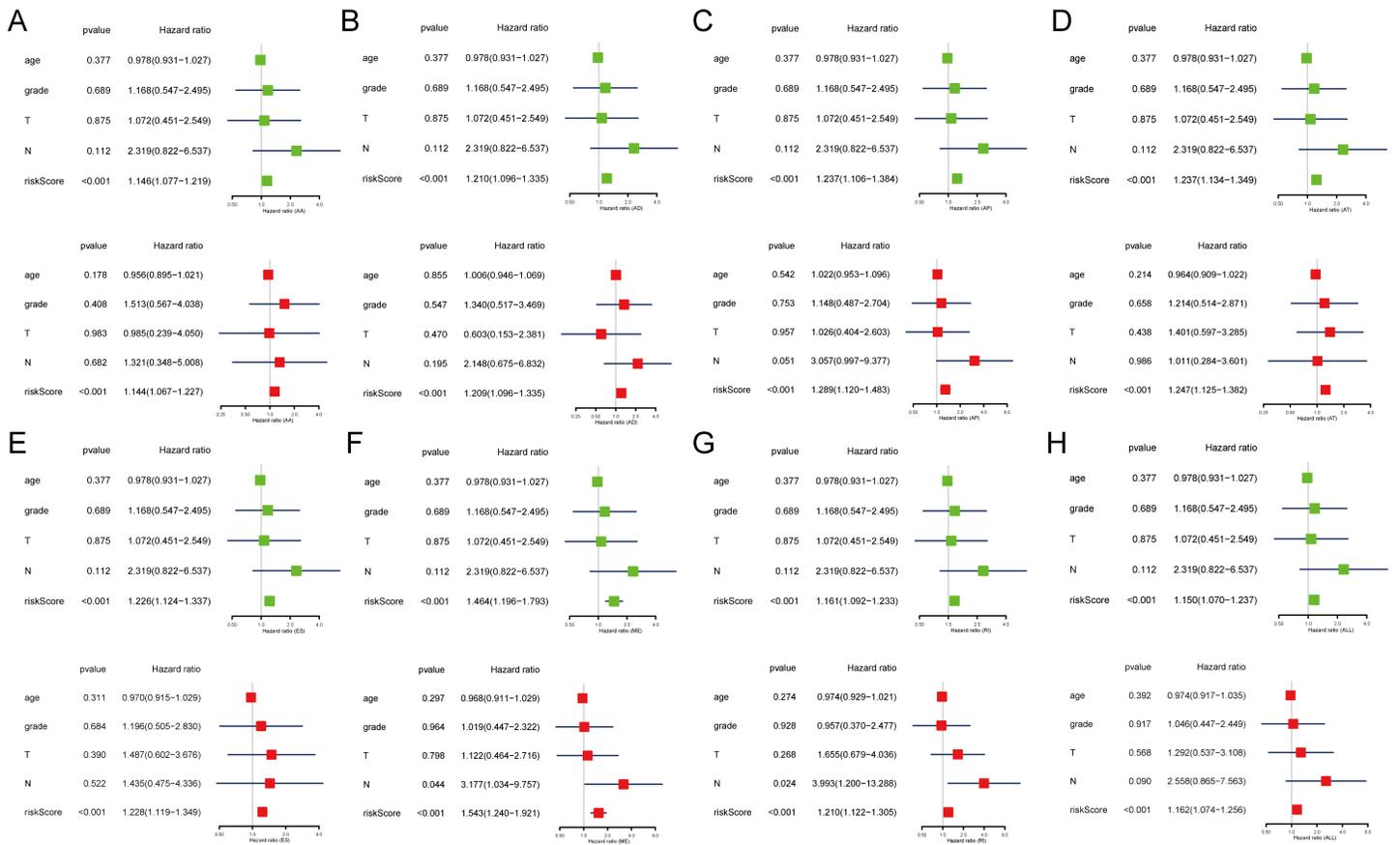
**Figure 1**

Kaplan-Meier survival curve and ROC curve. Kaplan-Meier survival curve of each AS pattern and all AS pattern in CESC patients, divided into low-risk and high-risk groups by using the median risk score as a cutoff value (A-H). ROC curve of each AS pattern and all AS pattern in CESC patients (I-P). AS, Alternative splicing; AA, Alternate acceptor site; AD, Alternate donor site; AP, Alternate promoter; AT, Alternate terminator; ES, Exon skip; ME, Mutually exclusive exons; RI, Retained intron.



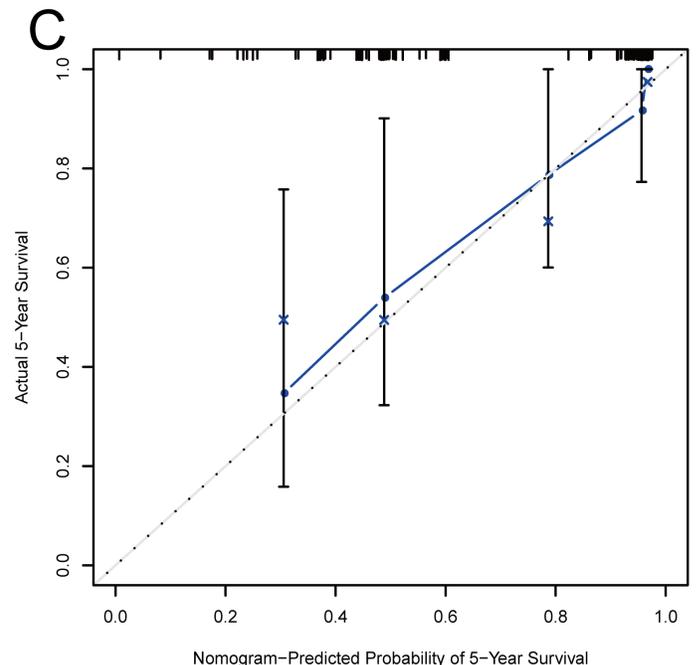
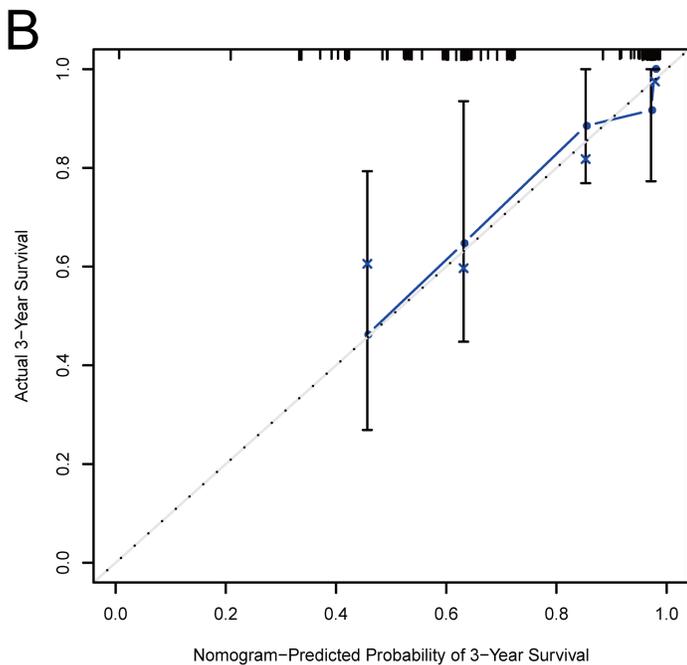
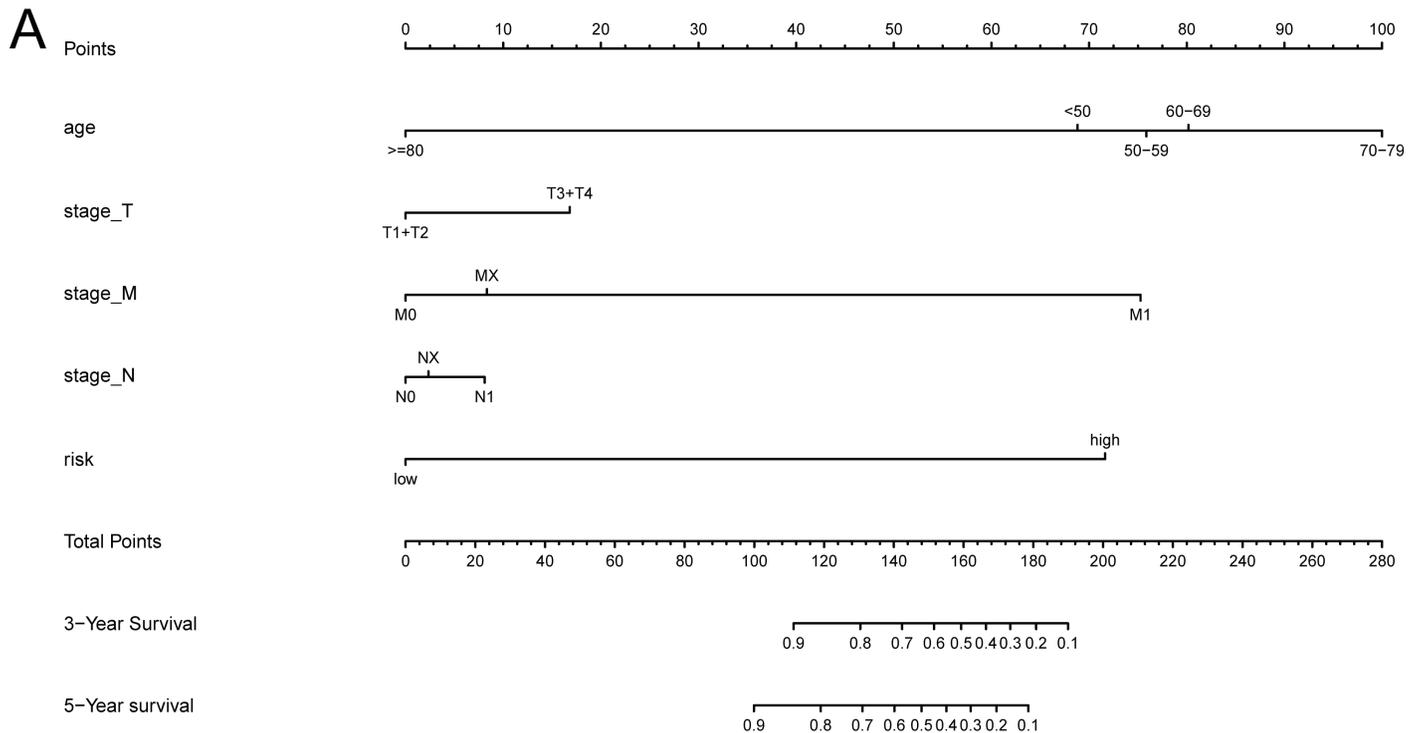
**Figure 2**

Risk plot and scatterplot. Risk plot and scatterplot of low-risk and high-risk groups for each CESC sample in AA, AD, AP, AT, ES, ME, RI and all AS pattern (A-H). The green and red dots in risk plot present low-risk and high-risk groups divided by median risk score. These dots in scatter plot indicate survival status and overall survival. Heatmap of survival associated splicing events expression level in each AS pattern and all AS pattern (A-H). Red and green stripe indicate high and low PSI values; Red and blue bars present low-risk and high-risk groups. AS, Alternative splicing; AA, Alternate acceptor site; AD, Alternate donor site; AP, Alternate promoter; AT, Alternate terminator; ES, Exon skip; ME, Mutually exclusive exons; RI, Retained intron.



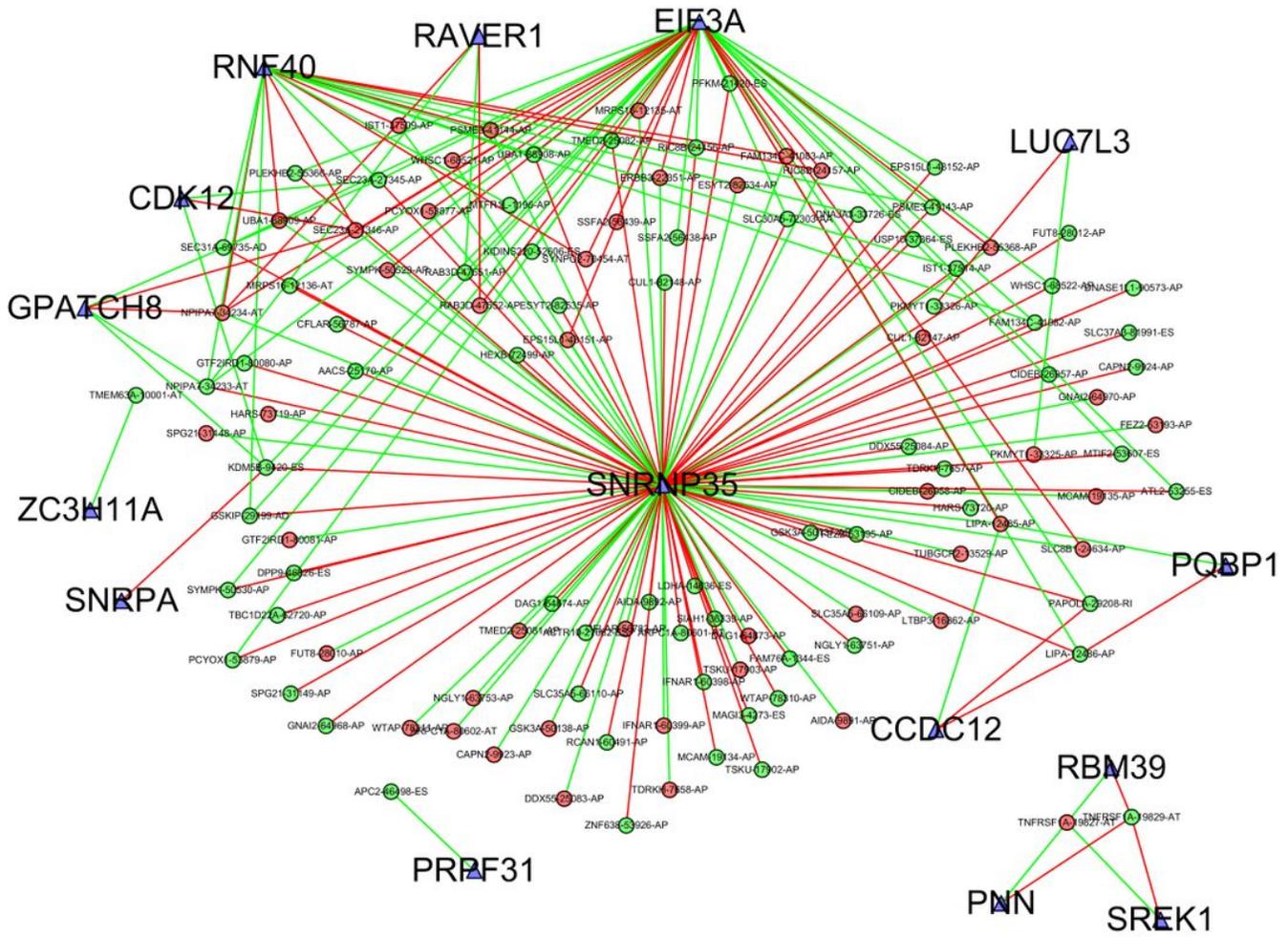
**Figure 3**

Forest plot. Forest plots of the HRs of age, grade, TNM stage and AS event associated prognostic signatures in AA, AD, AP, AT, ES, ME, RI and all AS pattern (A-H). Green represents univariate cox regression analysis, and red represents multivariate cox regression analysis. Dots represent the HR, and horizontal lines represent the 95% CIs. AS, Alternative splicing; AA, Alternate acceptor site; AD, Alternate donor site; AP, Alternate promoter; AT, Alternate terminator; ES, Exon skip; ME, Mutually exclusive exons; RI, Retained intron; HR, Hazard ratio; CI, Confidence interval.



**Figure 4**

Nomogram. Nomogram for predicting individuals' survival probability in CESC (A). Calibration curves for predicting 3-year (B) and 5-year (C) survival. The x axis represents nomogram predicted survival, and y axis represents actual survival. The blue line represents the predicted survival probability, and vertical lines represent the 95% CIs. CI, Confidence interval.



**Figure 5**

Correlation network between SASEs and SFs. Purple arrows represent SFs; Red and green circles represent high and low HR; Red and green lines represent positive and negative regulation. SASEs, Survival associated splicing events; SFs, Splicing factors; HR, Hazard ratio.



Figure 6

Heatmap of co-expression SASEs and KEGG pathways. Red and blue dots represent positive and negative correlation between SASEs and KEGG pathways. SASEs, Survival associated splicing events; KEGG, Kyoto Encyclopedia of Genes and Genomes.