Alternative Splicing Events and Subtype Analysis of Esophageal Cancer

Yuanyuan Zheng

Zhengzhou University First Affiliated Hospital

Zhibo Shen

Zhengzhou University First Affiliated Hospital

Zhirui Fan

Zhengzhou University First Affiliated Hospital

Wenbin Wang

Zhengzhou University First Affiliated Hospital

Qishun Geng

Zhengzhou University First Affiliated Hospital

Quancheng Kan

Zhengzhou University First Affiliated Hospital

Wenhua Xue

Zhengzhou University First Affiliated Hospital

Jie Zhao (✉ jiezhaoz2016@163.com)

Professor of The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, Henan, China

---

Research article

Keywords: alternative splicing, esophageal cancer, prognosis, carcinogenesis, biomarker

Posted Date: October 6th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-80935/v1>

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

[Read Full License](#)

Abstract

Aim: Alternative splicing (AS) has been widely demonstrated in the occurrence and progression of many cancers. Nevertheless, the involvement of cancer-associated splicing in the development of esophageal carcinoma (ESCA) is still ambiguous.

Method: RNA-Seq data and the corresponding clinical information of the ESCA cohort was downloaded from The Cancer Genome Atlas database. The splicing percentage value was calculated using a Java application called SpliceSeq, and differently expressed AS (DEAS) events and their splicing network were further analyzed using bioinformatics methods. Kaplan–Meier, Cox regression, and unsupervised cluster analyses were used to assess the association between AS events and clinical characteristics of ESCA patients.

Results: A total of 50,342 AS events were identified, of which 3,988 were DEAS events; 46 of these were associated with overall survival (OS) of ESCA patients, and the 5-year OS rate was 0.941. By constructing a network of AS events with survival-related splicing factors and variable-shear events associated with prognosis, the regulatory relationship was further predicted. Four clusters with different survival patterns were revealed using unsupervised cluster analysis.

Conclusion: ESCA-associated AS events and splicing networks are of great value in deciphering the underlying mechanisms of AS in ESCA and providing clues for therapeutic goals for further validation.

Introduction

According to the 2018 global cancer epidemiology report, esophageal cancer (ESCA) currently has the 7th highest incidence and is the 6th leading cause of mortality, while the ESCA mortality in developing countries tends to rise [1]. Although the efficiency of advanced diagnosis and multidisciplinary therapy for ESCA have demonstrated a notable improvement recently, data from the last decade showed that the 5-year overall survival (OS) rates in the US, China, and Europe are 22.0%, 20.9%, and 12.6%, respectively, and 70% of patients have missed the opportunity of undergoing radical surgery on first diagnosis because of their late disease stage [2–5]. Related studies have shown that immunotherapy significantly prolong the survival of advanced gastric or gastro-oesophageal junction cancer patients (8.2 months: 7.1 months, hazard ratio (HR) = 0.78, 95% confidence interval (CI): 0.63–0.96, $P = 0.0095$), but in all patients with ESCA, the effective rate is only 13.1% (41/314) [6]. Similarly, a phase III trial of epidermal growth factor receptor inhibitors have revealed that there is no difference in OS between the two treatment groups (median gefitinib 3.73 months, CI 3.23–4.50 95%, placebo group 3.67 months, CI 2.97–4.37 95%; hazard ratio [HR] 0.90, 95% CI 0.74–1.09, $p = 0.29$). OS [7]. Accurate medicines and new biomarkers relying on genomic data provide new methods for the prevention, diagnosis, and therapy of ESCA, which is an advancing area of research.

Developments in diversified clinical databases and high-throughput genomic technologies has made the exploration of cancer pathogenesis at the molecular level easier. In recent years, the emergence of

comprehensive studies on the genomics of ESCA including whole-exome sequencing, gene mutation analysis, DNA methylation profiling, and deregulated pathways has contributed to a deeper understanding of ESCA [8]. Based on this situation, we explored the possible influencing factors of ESCA at the RNA level to identify more effective therapeutic targets, which are valuable to predict treatment response and prognosis.

Many precursor mRNAs are processed to produce only one mature mRNA, which is translated into a corresponding polypeptide, and some can be spliced into mRNAs with different structures. This phenomenon is called alternative splicing (AS). AS increases the use of a limited number of genes and is one of the mechanisms for increasing the diversity of biological proteins in multicellular eukaryotes [9]. AS plays a crucial role in biological processes, molecular functions, signal pathways, and cellular components. The disruption of AS likely results in abnormal cell differentiation even in the occurrence of cancer and other diseases [10]. AS occurs differently in different tissues; approximately 30% of AS is tissue-specific on account of the exon of genes are expressed in difference tissues, and changes in AS will occur with the adaption of cancer progression [11, 12]. Increasing evidence has shown that widespread splicing disorders result in the mutation of targeted genes and promote the progression of tumors [13, 14]. Therefore, a study of AS will identify potential biomarkers for cancer therapy.

Materials And Methods

ESCA cell lines culture

Human esophageal epithelial cell lines Het-1A and the ESCA cell lines (TE-1, KYSE-150, and EC-109) were all purchased from Shanghai cell Resource Center Academy of Sciences. Cells were cultured in complete medium containing 90% DMEM (4mM glutamine and 1% penicillin streptomycin) and 10% fetal bovine serum, at 37°C in a CO₂ cell culture incubator. Cells were passaged at intervals of 2 to 3 days, and logarithmic growth phase cells were selected for the experiment.

RNA isolation, reverse transcription, and quantitative RT-PCR

Total RNA was extracted using Trizol reagent (American life Technologies), detected the absorbance of RNA at 260 nm-280nm with spectrophotometer to obtain concentration value. Then, the total RNA extracted was reverse transcribed into first-strand cDNA using a reverse transcription kit from Takara Corporation of Japan. The PCR primer were synthesized by Shanghai Biological Engineering Co., Ltd, and more details could be found in Table 1. Using GAPTH mRNA as an internal reference and the 2- $\Delta\Delta CT$ method was employed to calculate the relative expression levels of the molecules. Reaction conditions: 95°C pre-denatured for 30s, 95°C 5s, 60°C 34s, 95°C 15s, 60°C 60s, a total of 40 cycles.

AS events from The Cancer Genome Atlas RNA sequences

RNA sequence expression data are available at The Cancer Genome Atlas (TCGA) database. We obtained data on the AS events related to ESCA and analyzed it using SpliceSeq, which is a java program that provides a comprehensive view of AS patterns and highlight its biological consequences [15]. AS events are sorted into seven patterns: 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). The percent splicing in (PSI) value, ranging from 0 to 1 to summarize the rate of splicing a specific exon into the transcript population of a gene. The score of PSI indicated the AS events for specific exons without the need to know the comprehensive synthetic of full-length transcripts [16].

Analysis of differential variable splicing events

On analysis of the AS spectrum from ESCA and adjacent tissue samples, the AS events were required to fulfil the condition that a t-test yielded a P-value of <0.05 and fold-change (FC) >1.5 or FC $<2/3$.

Survival analysis

Download the data of patients with complete clinical data and OS greater than 30 months from TCGA database. To explore the effect of AS on the OS of patients with ESCA, we divided the patients into two groups according to the median PSI and generated a univariate Cox regression model with seven patterns of AS events and OS. ClusterProfiler analysis was used to conduct Geneontology (GO) enrichment analysis to identify the effects of survival-related AS events on biological processes, cellular components, and molecular functions. Multivariable Cox regression was conducted to analyze the difference between the seven patterns of AS events and OS, and then eliminated the genes that had no relevance with survival and confirmed the prognostic predictor. In addition, we plotted the accuracy of the receiver-operating characteristic (ROC) curves to compare the prognostic models for each type of AS.

Splicing-related network construction

Through database filtering, a list of 67 human splicing factors was created [17]. Expression of the splicing factor gene in the mRNA splicing pathway was derived from grade 3 mRNA-seq data in TCGA. Survival analysis was performed to identify survival-related shear factors, and the correlation between survival-related splicing factor gene expression and survival-related AS PSI values was analyzed using Spearman test. $P<0.05$ was defined as significantly correlated. The final interaction network of variable shear events and shear factors was constructed using Cytoscape (3.6.0).

Identifying the effects of AS on the prognosis of patients with different ESCA subtypes

In consideration of the difference in AS events at the individual level, the association between ESCA subtypes based on Consensus Cluster Plus and OS were analyzed using an unsupervised consensus approach.

Results

AS events in ESCA

Overall, 50342 AS events of 10765 genes were identified in patients with ESCA: 4144 AA events in 2870 genes, 2589 AD events in 2463 genes, 10033 AP events in 4046 genes, 8848 AT events in 3690 genes, 20842 ES events in 7173 genes, 245 MT events in 237 genes, and 3038 RI events in 2001 genes (Figure 1). In short, the number of AS events was greater in ESCA tissue than in adjacent tissue, but the types of AS events were roughly the same. Among them, ES was the highest in number, accounting for one-third of the AS events, followed by AP, AT, AA, AD, RI, and ME in succession (Figure 1A).

Differently expressed AS in ESCA

There were 3988 differently expressed AS (DEAS) events in 2818 genes of patients with ESCA (2758 up-regulated AS events and 1230 down-regulated AS events, Figure 1B). Interestingly, the number of AS events and involved genes were nonmatched, implying that one gene underwent more than one type of splicing event. Some genes had up to four variable types on DEAS. The UpSet plot (Figure 1C) was used to visualize the overlapping different AS events among genes. In the differential expression analysis, 5269 differentially expressed genes were identified in ESCA (2711 up-regulated differentially expressed genes and 2558 down-regulated differentially expressed genes).

Survival-associated AS events in ESCA

To illustrate the relationship between the OS in patients with ESCA-related AS events, we assessed the prognostic value of AS events in patients with ESCA using univariate Cox regression analysis in patients with intact clinical survival information. We detected a total of 217 survival-related AS events ($P < 0.05$) among all AS events that occurred in ESCA, and plotted a forest map based on six types of AS with the top 15 most important survival-related AS events (Figure 2A-F). The AS events included AA, AD, AP, AT, ES, and RI, except ME, in which there was only one splicing event associated with survival (ZFP2, HR = 0.56, 95% CI: 0.35–0.91). Obviously, most of these survival-related AS events were adverse prognostic factors. For instance, RI events showed a negative association with the survival of patients with ESCA (HR > 1). The forest maps revealed the genes participating in ESCA carcinogenesis. The EIF4B in AA of splicing events predicted poor survival of patients with ESCA, and other studies have experimentally validated that the EIF4B result in poor prognosis of tumor patients, which can increase the credibility of our data [18, 19]. RPS21 and MYL6B in RI were associated with poor prognosis of patients with ESCA; they have been validated as oncoproteins in prostate tumor and hepatocellular carcinoma, respectively [20, 21]. To

further explore the underlying molecular mechanism among the survival-associated events, we conducted GO functional enrichment analysis on 206 genes involved in AS events in ESCA. The GO categories showed that the genes in survival-associated AS events were significantly enriched in different biological processes, such as translational initiation, mRNA catabolic processes, and cellular responses to external stimuli, among others (Figure 2G).

Prognostic predictors for patients with ESCA

Considering the interaction between AS events and the prognosis of patients with ESCA, we attempted to verify if the six types of AS events could serve as the predictive element of ESCA by choosing the highly correlated AS events as the candidates to analyze the predictive factors in ESCA using multivariate Cox analysis ($P < 0.05$). There was one respective prognostic factor in AA and RI, three respective prognostic factors in AD, 22 respective prognostic factors in AP, 13 respective prognostic factors in ES, and no respective prognostic factor in ME. According to our data, each splicing type performed well in predicting favorable or unfavorable consequences of the prognosis of patients with ESCA (all adjusted $P < 0.05$, Figure 3A-F). To better confirm the validity of AS events in forecasting the prognosis of patients with ESCA, the ROC was applied in the prognostic models. In ESCA, the AS events including AA, AD, AP, AT, ES, and RI all demonstrated areas under the curve (AUCs) of the ROC curves of greater than 0.6, indicating powerful efficiency in distinguishing the prognosis of patients, and ES showed the best performance in predicting the survival of ESCA patients (Figure 3H). Furthermore, candidate AS events (a total of 46 AS events) were combined to build the final predictive model, which showed the best manifestation in predicting survival of ESCA patients (ROC=0.941) (Figure 3H, G). Meanwhile, the AUCs of the ROC curves of 3-, 5-, and 7-year survival was 0.953, performing remarkably in prognosis assessment for ESCA.

Network of survival-associated AS events and splicing factors

To determine the specific splicing factors in the AS events associated with the survival of patients with ESCA, we conducted survival analysis based on gene expression of the splicing factors. Five splicing factors (hnRNP J, hnRNP A3, hnRNP G, FMRP, and Fox-2) showed significant correlation with OS in patients with ESCA. The representative splicing factors and survival-associated AS events were shown in Figure 4C. Additionally, Spearman test was also used to study the correlation between the PSI of AS events and splicing factors associated with survival. The correlation network indicated that a total of 27 survival associated AS events were highly correlated with five splicing factors (green points), with 16 positive (red points) and 11 negative (green points) AS events (Figure 4C). Splicing factors such as Fox-2 and hnRNA G were involved in 17 and 13 AS events, respectively. Among these survival-associated splicing factors, some were associated with poor prognosis whereas others were linked with favorable prognosis (Figure 4A, B). Splicing genes (*IAH1*, *NSUN4*, *SERAC1*, and *TRIM4*) were linked with up to four splicing factors. Among them, *IAH1* was a gene that related to hnRNA G (positive correlation, Figure 4D)

and Fox-2 (negative correlation, Figure 4E) and its high expression can cause poor prognosis of patients with ESCA (Figure 4F).

Expression of splicing factors related to prognosis in ESCA cell lines

Next, we examined the expression levels of the splicing factors Fox-2 and hnRNA G in the ESCA cell lines and normal esophageal epithelial cell. qRT-PCR results showed that the expression of FOX-2 in ESCA cell lines KYSE-150, EC-109, and TE-1 were significantly lower than that of normal esophageal epithelial cell line HET-1A (Figure 5A), while the expression of HNRNP in ESCA cell lines KYSE-150, EC-109, and TE-1 were significantly higher than that of normal esophageal epithelial cell line HET-1A (Figure 5B). The consistency between the results of RT-PCR and the above survival analysis demonstrated that our results were reliable.

Molecular subtype clustering

According to 46 survival-related AS events, we identified different AS types by performing unsupervised analysis on all samples. Here, we introduced the functional heatmap, which revealed the hidden trends driven by different AS types and reduced the manual labor in discovery and comparison of different patterns [22]. Using the Elbow method to determine the optimal number of clusters, and based on the distribution of the consensus values ranging from 0 (white, no samples aggregation) to 1 (blue, sample always aggregation), we finally determined four sets of samples: C1 (n=36, 20.7%), C2 (n=73, 42.0%), C3 (n=46, 26.3%), and C4 (n=19, 11.0%) (Figure 6A-B). Kaplan–Meier analysis was then used to explore the correlation between these clusters and the survival rate of patients with ESCA, indicating that the clusters were significantly associated with the outcome of survival (Figure 6C). Therefore, we can acquire molecular subtype clusters associated with prognosis through a small number of AS events.

Discussion

More than 95% of precursor messenger RNA (pre-mRNA) are processed to multiple mRNAs through AS events [23]. Growing evidence has demonstrated that AS plays a crucial role not only in physiological processes and cell development programs, but also in the differentiation of cells, leading to the development of tumors and other diseases [12, 24]. AS events related to aberrant pre-mRNA precursors, including changes in splicing types as well as mutations in splicing factors and regulatory signals, has been widely shown in the oncogenesis of tumors, along with increased invasiveness, drug resistance, and other factors of tumor progression [25]. Increasing evidence in the past several years has indicated that AS events promise to recapitulate cancer epigenetics [26, 27]. For example, the key regulator of AS events in lung cancer, RNA-binding protein QKI, is significantly associated with poor prognosis when it is down-regulated [28]. Similarly, changes in AS events in lung cancer will also affect the transcripts of *VEGFA*, *MACF1*, *APP*, and *NUMB* genes, thereby promoting the process of tumorigenesis [29]. Mutations of SF3B1-encoding proteins involved in RNA splicing may be a driving factor and novel therapeutic target in breast cancers [30]. Similarly, breast tumors adopt AS events to remove deleterious germline *BRCA1*

mutations by removing exon 11 to contribute to retaining activity and drug resistance [31]. Accordingly, mutations in splicing factors and regulatory pathways in AS events can cause abnormal splicing types, leading to multiple outcomes, including the development, invasiveness, substitution, and transformation of tumors. Therefore, the basic functional study of AS events and mechanisms of tumorigenesis and development can facilitate the discovery of novel biomarkers.

In this study, we identified AS events and profiles of regulatory splicing factors and established the interrelation between them in ESCA through an analysis of TCGA program. A total of 50342 AS events of 10765 genes were detected in samples of tumors with seven splicing types (AA, AD, AP, AT, ES, ME, and RI), of which ES was the most common splicing type. It seems that AS event is an ordinary process in ESCA, and that most patterns of splicing are active. There are 3988 DEAS events between ESCA and non-tumor tissues: 2758 upregulated AS events and 1230 downregulated AS events involving 2818 genes. Then, we found that 217 AS events among those DEAS events were significantly associated with the survival rate of patients with ESCA. Most of these AS events showed a critical influence on tumor biology, such as the top 15 splicing events of RI in DEAS events all lead to the poor survival of patients with ESCA. This is consistent with some findings regarding AS events and tumor prognosis [32, 33]. In addition, GO functional enrichment and KEGG pathway analysis provided insights into the enrichment of many DEAS events in many biological adhesions, cell organizations, and many other essential biological processes. Several genes (*EIF4B*, *RPS21*, *MYL6B*) have been demonstrated to make ESCA patients have poor survival, and further investigation concerning their potential impact on ESCA is still required.

Moreover, we established a prognostic model for each splicing type using multivariate Cox regression analysis, and each type of AS event performed reasonably well in showing the positive or negative prognosis. The six survival-associated types (AA, AD, AP, AT, ES, and RI) of AS events showed different AUCs of ROC curves, and ES showed the maximum efficiency in predicting the survival of ESCA patients with the best AUC value: 0.793. The integrated predictive model with six types together showed a high correlation with survival. Further specific functional experiments to determine how alternative splicing modulates tumor prognosis are needed.

Given that certain genes and splicing factors may exhibit extensive “spliceosomal mutations”, leading to cancer-specific mis-splicing, we therefore focused on the correlation network of splicing factors and survival-associated AS events [34]. Five splicing factors (hnRNP A3, hnRNP J, hnRNP G, FMRP, Fox-2) with four genes (*IAH1*, *NSUN4*, *SERAC1*, and *TRIM4*) showed a strong correlation with prognosis. To this end, we verified the expression of splicing factors related to prognosis in esophageal cells, and the results showed that splicing factors positively correlated with prognosis were highly expressed in ESCA cell lines, while the expression of splicing factors negatively correlated with prognosis in ESCA cell lines were significantly lower than that of normal esophageal epithelial cells. Consistent with the findings of other investigations, the splicing factors hnRNP A3 [35], FMRP [36], and Fox-2 [37] may be related to the development of tumors. *IAH1* (isoamyl acetate-hydrolyzing esterase 1 homolog) regulates the expression of genes related to lipid metabolism, and the accumulation of metabolic acids is a hallmark of all primary and metastatic tumors [38]. Increased expression of *NSUN4* and *SERAC1* has been described in breast

cancer [39, 40]. Differential expression of *TRIM4* makes cells sensitive to H₂O₂-induced death, which is common in tumor cell lines [41]. These splicing factors and genes can serve as prognostic biomarkers and expound a novel etiology of ESCA in the future.

Conclusion

In summary, this was a comprehensive and up-to-date profile of AS events between ESCA and its corresponding non-tumor tissues, uncovering the inter-event correlations in splicing factors and prognostic signatures in ESCA. The interaction network of splicing factors and prognostic AS events highlights the value of AS events and tumorigenesis of ESCA at the genome level. This analysis of survival-associated splicing factors and tumor-specific AS events also point out that new underlying clinical biomarkers, which still need to be validated in future mechanistic research and clinical trials.

Declarations

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships.

Author Contributions

JZ, WHX, conceived the designed experiments. YYZ, WBW and QSG conducted data analysis, and critical discussions of the results. QCK, ZBS and LFL provided material support and study supervision. All authors contributed to the writing and editing of the manuscript and approved the final draft of the manuscript.

Funding

This study was supported by grants from National Key R&D Program of China (Grant No. 2017YFC0909900), the National Natural Science Foundation of China (Grant No. 31670895, 71673254), and Science and Technology Project of Henan Provincial Department of Education (Grant No. 18A320044), Henan Province Medical Science and Technology Research Project Joint Construction Project (Grant No. LHGJ20190055).

Acknowledgments

The authors would like to acknowledge the contributions of specific colleagues, institutions, or agencies that aided the efforts of the authors.

Data Availability Statement

All datasets generated for this study are included in the manuscript/supplementary files.

References

- [1]. F, B., et al., Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 2018. 68(6): p. 394-424.
- [2]. B, N., et al., Trends in esophageal cancer survival in United States adults from 1973 to 2009: A SEER database analysis. *Journal of gastroenterology and hepatology*, 2016. 31(6): p. 1141-6.
- [3]. H, Z., et al., Cancer survival in China, 2003-2005: a population-based study. *International journal of cancer*, 2015. 136(8): p. 1921-30.
- [4]. LA, A., et al., Survival for oesophageal, stomach and small intestine cancers in Europe 1999-2007: Results from EURO CARE-5. *European journal of cancer (Oxford, England : 1990)*, 2015. 51(15): p. 2144-2157.
- [5]. Le Bras GF, et al., Esophageal cancer: The latest on chemoprevention and state of the art therapies. *Pharmacological research*, 2016. 113(null): p. 236-244.
- [6]. K, S., et al., Pembrolizumab versus paclitaxel for previously treated, advanced gastric or gastro-oesophageal junction cancer (KEYNOTE-061): a randomised, open-label, controlled, phase 3 trial. *Lancet (London, England)*, 2018. 392(10142): p. 123-133.
- [7]. SJ, D., et al., Gefitinib for oesophageal cancer progressing after chemotherapy (COG): a phase 3, multicentre, double-blind, placebo-controlled randomised trial. *The Lancet. Oncology*, 2014. 15(8): p. 894-904.
- [8]. FR, T., et al., Molecular landscape of esophageal cancer: implications for early detection and personalized therapy. *Annals of the New York Academy of Sciences*, 2018. 1434(1): p. 342-359.
- [9]. B, R., et al., Review: Alternative Splicing (AS) of Genes As An Approach for Generating Protein Complexity. *Current genomics*, 2013. 14(3): p. 182-94.
- [10]. ER, G., S. BE and A.I.H.I. Alternative Splicing Is A Major Cellular Mechanism In Metazoans For Generating Proteomic Diversity. A Large Proportion Of Protein-coding Genes In Multicellular Organisms Undergo Alternative Splicing, *Genomics of alternative splicing: evolution, development and pathophysiology. Human genetics*, 2014. 133(6): p. 679-87.
- [11]. Q, X., et al., Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. *Nucleic acids research*, 2002. 30(17): p. 3754-66.

- [12]. Le KQ, et al., Alternative splicing as a biomarker and potential target for drug discovery. *Acta pharmacologica Sinica*, 2015. 36(10): p. 1212-8.
- [13]. A, S., et al., Aberrant RNA splicing in cancer; expression changes and driver mutations of splicing factor genes. *Oncogene*, 2016. 35(19): p. 2413-27.
- [14]. S, M., et al., Survival-associated alternative splicing signatures in esophageal carcinoma. *Carcinogenesis*, 2019. 40(1): p. 121-130.
- [15]. M, R., et al., TCGASpliceSeq a compendium of alternative mRNA splicing in cancer. *Nucleic acids research*, 2016. 44(null): p. D1018-22.
- [16]. S, S., et al., Alternative Splicing Signatures in RNA-seq Data: Percent Spliced in (PSI). *Current protocols in human genetics*, 2015. 87(undefined): p. 11.16.1-14.
- [17]. M, G., et al., SpliceAid-F: a database of human splicing factors and their RNA-binding sites. *Nucleic acids research*, 2013. 41(null): p. D125-31.
- [18]. K, R., et al., The over-expression of Pim-2 promote the tumorigenesis of prostatic carcinoma through phosphorylating eIF4B. *The Prostate*, 2013. 73(13): p. 1462-9.
- [19]. J, Y., et al., eIF4B phosphorylation by pim kinases plays a critical role in cellular transformation by Abl oncogenes. *Cancer research*, 2013. 73(15): p. 4898-908.
- [20]. Z, L., et al., Identification of candidate diagnostic and prognostic biomarkers for human prostate cancer: RPL22L1 and RPS21. *Medical oncology (Northwood, London, England)*, 2019. 36(6): p. 56.
- [21]. X, X., et al., MYL6B, a myosin light chain, promotes MDM2-mediated p53 degradation and drives HCC development. *Journal of experimental & clinical cancer research : CR*, 2018. 37(1): p. 28.
- [22]. JR, W., et al., Functional Heatmap: an automated and interactive pattern recognition tool to integrate time with multi-omics assays. *BMC bioinformatics*, 2019. 20(1): p. 81.
- [23]. TW, N., et al., Expansion of the eukaryotic proteome by alternative splicing. *Nature*, 2010. 463(7280): p. 457-63.
- [24]. ER, G., S. BE and A.I.H.I. Alternative Splicing Is A Major Cellular Mechanism In Metazoans For Generating Proteomic Diversity. A Large Proportion Of Protein-coding Genes In Multicellular Organisms Undergo Alternative Splicing, *Genomics of alternative splicing: evolution, development and pathophysiology. Human genetics*, 2014. 133(6): p. 679-87.
- [25]. B, S., E. E and A.W.A.M. The Functional Capacity Of Cells Is Defined By The Transcriptome. Many Recent Studies Have Identified Variations In The Transcriptome Of Tumors Due To Alternative Splicing Changes, *The role of alternative splicing in cancer. Transcription*, 2017. 8(2): p. 91-98.

- [26]. SP, N., et al., A saga of cancer epigenetics: linking epigenetics to alternative splicing. *The Biochemical journal*, 2017. 474(6): p. 885-896.
- [27]. C, A.F., et al., RNA splicing, cell signaling, and response to therapies. *Current opinion in oncology*, 2016. 28(1): p. 58-64.
- [28]. FY, Z., et al., The RNA-binding protein QKI suppresses cancer-associated aberrant splicing. *PLoS genetics*, 2014. 10(4): p. e1004289.
- [29]. CM, M., et al., Global profiling and molecular characterization of alternative splicing events misregulated in lung cancer. *Molecular and cellular biology*, 2011. 31(1): p. 138-50.
- [30]. SL, M., et al., SF3B1 mutations constitute a novel therapeutic target in breast cancer. *The Journal of pathology*, 2015. 235(4): p. 571-80.
- [31]. Y, W., et al., The BRCA1- Δ 11q Alternative Splice Isoform Bypasses Germline Mutations and Promotes Therapeutic Resistance to PARP Inhibition and Cisplatin. *Cancer research*, 2016. 76(9): p. 2778-90.
- [32]. J, L., et al., Alternative splicing events implicated in carcinogenesis and prognosis of colorectal cancer. *Journal of Cancer*, 2018. 9(10): p. 1754-1764.
- [33]. S, M., et al., Survival-associated alternative splicing signatures in esophageal carcinoma. *Carcinogenesis*, 2019. 40(1): p. 121-130.
- [34]. H, D., et al., RNA splicing factors as oncoproteins and tumour suppressors. *Nature reviews. Cancer*, 2016. 16(7): p. 413-30.
- [35]. G, B., et al., Deregulated expression of hnRNP A/B proteins in human non-small cell lung cancer: parallel assessment of protein and mRNA levels in paired tumour/non-tumour tissues. *BMC cancer*, 2010. 10(undefined): p. 434.
- [36]. YJ, Z., et al., Circular RNAs negatively regulate cancer stem cells by physically binding FMRP against CCAR1 complex in hepatocellular carcinoma. *Theranostics*, 2019. 9(12): p. 3526-3540.
- [37]. S, B., et al., The Epstein-Barr virus EBNA1 protein modulates the alternative splicing of cellular genes. *Virology journal*, 2019. 16(1): p. 29.
- [38]. M, K., et al., Detection of differentially expressed candidate genes for a fatty liver QTL on mouse chromosome 12. *BMC genetics*, 2016. 17(1): p. 73.
- [39]. SP, K., et al., Genome-Wide Meta-Analyses of Breast, Ovarian, and Prostate Cancer Association Studies Identify Multiple New Susceptibility Loci Shared by at Least Two Cancer Types. *Cancer discovery*, 2016. 6(9): p. 1052-67.

[40]. L, Z., et al., Autoantibodies as potential biomarkers for breast cancer. Breast cancer research : BCR, 2008. 10(3): p. R40.

[41]. D, T., et al., TRIM4; a novel mitochondrial interacting RING E3 ligase, sensitizes the cells to hydrogen peroxide (H2O2) induced cell death. Free radical biology & medicine, 2015. 89(undefined): p. 1036-48.

Tables

Table 1 Primer sequences used for qRT-PCR

FOX2	F: TACAGTGACGGTTATGGCAGG
	R: CCTCGGTATAAACTCGCCACA
HNRNP G	F: CCATCAAGAGGCTATGGCGAT
	R: CCCTCGTGTAAGTGGAGCA
GAPDH	F: GACCACAGTCCATGCCATCAC
	R: GTCCACCACCCTGTTGCTGTA

Figures

Figure 1

(A) Each AS events type and its number of genes involved in ESCA; (B) Differential variable splicing event volcano map, the red nodes represent the up-regulated AS events and the blue nodes represent the down-regulated AS events; (C) The UpSet intersection diagram shows seven types of survival associated DEAS events.

Figure 2

(A-F) Survival-related top 15 AS events forest map in each AS type; (G) Functional enrichment analysis results of differentially spliced genes in ESCA from GO analysis.

Figure 3

(A-F) Kaplan–Meier curves of prognostic models built with AS events of AA, AD, AP, AT, ES and RI splice types for patients with ESCA. The blue line indicates the low-risk group, whereas the red line indicates the high-risk group. G Kaplan–Meier curves of prognostic models built with all type of AS events. (H) The ROC curves of predictive models for each splicing type. (I) The ROC curves of predictive models for all splicing type over 3years, 5 years, and 7years of survival.

Figure 4

(A, B) Kaplan-Meier curves for splicing factor hnRNP G and Fox-2 with high (red) and low (blue) expression group in patients with ESCA, respectively. (C) Splicing correlation network constructed by Cytoscape. Five survival-associated splicing factors (green dots) in AS events were up-regulated (red dots) or down-regulated (blue dots) by involved genes. (D) Negative correlations between Fox-2 expression and the AP PSI value of IAH1. (E) positive correlations between hnRNP G expression and the

AP PSI value of IAHI. (F) Kaplan-Meier curves for splicing gene IAHI with high (red) and low (blue) expression group in patients with ESCA.

Figure 5

The expression of hnRNP G in ESCA cell lines (TE-1, KYSE-150, EC-109) and normal esophageal epithelial cell line HET-1A was detected by qRT-PCR (A). The expression of FOX2 in ESCA cell lines (TE-1, KYSE-150, EC-109) and normal esophageal epithelial cell line HET-1A was detected by qRT-PCR (B). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ and **** $P < 0.0001$.

Figure 6

(A) Statistical analysis of elbows for different numbers of clusters ($k = 2$ to 8); (B) The consensus matrix heat maps, consensus values ranging from 0 (white, no samples aggregation) to 1 (blue, sample always aggregation); (C) Survival analysis in the identified four clusters.