

RESEARCH

Open Access



Systematic profiling of alternative splicing signature reveals prognostic predictor for cervical cancer

Yue-Xin Hu^{1†}, Ming-Jun Zheng^{2†}, Wen-Chao Zhang¹, Xiao Li¹, Rui Gou¹, Xin Nie¹, Qing Liu¹, Ying-Ying Hao¹, Juan-Juan Liu¹ and Bei Lin^{1*}

Abstract

Aim: Cervical cancer is a common malignant carcinoma of the gynecological tract with high morbidity and mortality. Therefore, it is crucial to elucidate the pathogenesis, prevention, diagnosis and prognosis of cervical cancer by searching for the involved key genes.

Method: In this study, the alternative splicing (AS) events of 253 patients with cervical cancer were analyzed, and 41,766 AS events were detected in 9961 genes. Univariate analysis was performed to screen prognostic AS events. Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was used to identify the pathways in which these AS events were involved.

Results: We found that exon skip (ES) is the main AS event in patients with cervical cancer. There was pronounced consistency between the genes involved in overall survival and those involved in recurrence. At the same time, we found that a gene may exhibit several different types of AS events, and these different AS events may be related to prognosis. Four characteristic genes, HSPA14, SDHAF2, CAMKK2 and TM9SF1, that can be used as prognostic markers for cervical cancer were selected. Conclusion: The importance of AS events in the development of cervical cancer and prediction of prognosis was revealed by a large amount of data at the whole genome level, which may provide a potential target for cervical cancer treatment. We also provide a new method for exploring the pathogenesis of cervical cancer to determine clinical treatment and prognosis more accurately.

Keywords: Alternative splicing, TCGA, Whole genome level, Cervical cancer, Prognosis

Background

Cervical cancer is one of the most common malignant gynecological tumors, with an incidence rate second to only breast cancer and the third highest mortality rate following breast and lung cancers [1]. Approximately 470,000 people suffer from this disease worldwide, with 233,000 patients dying from it each year [2]. In developing countries, the incidence and mortality of cervical cancer are on the rise. China has approximately 150,000 new cases of cervical cancer each year, and the incidence

tends occur in younger people [1, 2]. Although the relationship between persistent high-risk human papillomavirus (HPV) infection and cervical cancer has been confirmed, only 1% to 3% of the population with HPV infection develop cervical cancer. HPV infection is a necessary factor for the occurrence of cervical cancer and plays an important role in the process of infection and cancer. The three-level prevention strategy for cervical cancer has effectively reduced morbidity and mortality, yet it remains a serious threat to women's health. Therefore, it is crucial to elucidate the pathogenesis, diagnosis, treatment and prognosis of cervical cancer by searching for the involved key genes and to determine their roles in cervical cancer transformation and progression. At the

*Correspondence: linbei88@hotmail.com

[†]Yue-Xin Hu and Ming-Jun Zheng contributed equally to this work

¹ Department of Obstetrics and Gynecology, Shengjing Hospital Affiliated to China Medical University, Shenyang, Liaoning 110004, China
Full list of author information is available at the end of the article



same time, these genes also provide potential targets for drug action for developing targeted therapeutics.

Protein diversity plays a significant role in the regulation and functional complexity of eukaryotic cellular processes, and alternative splicing (AS) events greatly increase the phenotypes of eukaryotic genes and the diversity of protein function [3]. Selective pre-mRNA splicing is a universal mechanism to produce mRNA isomers using a limited set of genes. Splicing complexes recognize intron–exon junctions by identifying loci with characteristic markers on a mRNA molecule, then perform intron excision and finally connect the adjacent exons. In other words, AS is a posttranscriptional process that can produce different mRNAs from a single gene. Splicing patterns are closely related to the occurrence of human diseases [4, 5]. The influencing factors of AS events include CIS action elements, splice site consistency, conservative sequences and trans-action factors, any change in which can lead to splicing abnormalities.

The Cancer Genome Atlas (TCGA) project, a public database that was completed by the National Cancer Institute (NCI) and the National Human Genome Research Institute (NHGRI) [6], aims to map the genome variants of all human cancers by using large-scale genome sequencing techniques. With the advantage of high-throughput RNA sequencing, TCGA has been a great source for the study of AS events [7]. SpliceSeq is a Java program that can provide information on each exon and splice site. Ryan et al. expanded the calculation of SpliceSeq and obtained potential events for 33 types of cancer-specific AS events [8]. Our study systematically analyses all AS events in cervical cancer for the first time using the TCGA database and describes several newly discovered selective splicing events related to survival. It also explores the genes associated with cervical cancer and provides a new way to explore the pathogenesis of cervical cancer, which will guide clinical treatment more accurately and help determine patient prognoses.

This study included 253 cervical cancer cases that were shared in the TCGASpliceSeq database and RNA-seq expression spectra, from which 19,754 genes with expression values were obtained. There are seven types of AS events for cervical cancer, including 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 data distributions of all encoded genes from the samples regarding the seven different types of AS events were analyzed, and ES events were found to be the main type of AS event in cervical cancer. We then used an R package [9] for univariate analysis, and the results showed that most ES events were not associated with prognosis, but approximately 10% of AT

events were significantly correlated with recurrence in cervical cancer. By constructing a network of genes that have significant roles in AS associated with prognosis, we found that most genes have protein interactions. KEGG enrichment analysis revealed the selective splicing pathways associated with prognosis. To observe whether the relationship between gene expression and prognosis in AS events significantly correlated with prognosis, univariate analysis of each gene was performed using TCGA RNA-seq expression profile data. Using Pearson correlation analysis, four recurrence-related genes were screened from 46 genes associated with overall survival: HSPA14, SDHAF2, CAMKK2 and TM9SF1. These four characteristic genes were used to construct a multivariate survival model, and the results showed that the four genes may be used as prognostic markers of cervical cancer. This discovery may provide a new idea for the early diagnosis and treatment of cervical cancer.

Methods

Data download and preprocessing

RNA-seq variable splicing event data for cervical cancer were obtained from the TCGASpliceSeq database and included 256 samples, three of which were normal samples. The RNA-seq expression profiles of cervical cancer were obtained from the TCGA database and included 309 samples, three of which were normal samples. We downloaded all clinical follow-up information provided in TCGA for a total of 307 samples with clinical follow-up information. The RNA-seq expression profile data set in fragments per kilobase per million fragments mapped (FPKM) was downloaded and converted into trusted platform module (TPM) data. At the same time, the IDs were converted using the GRCh38.p2 version of GENCODE's genome file to obtain each encoded protein's gene, and 253 samples of TCGASpliceSeq and RNA-seq were selected for further study.

Screening for prognostic AS events in cervical cancer

TCGASpliceSeq is a database based on TCGA RNA-seq data that uses SpliceSeq tools to analyze mRNA splicing patterns with seven types of selective splicing events: ES, ME, RI, AP, AT, AD and AA. We analyzed the distributions of all encoded genes in each of the seven different types of AS events in cervical cancer. Different AS events in genes led to diversity in outcomes, and changes in gene expression affected survival time. The *Survival.R* package was used for the univariate analysis of different AS events. A gene with significant $p < 0.05$ was selected as an AS event with different prognoses.

Prognostic selective splice event gene type analysis

AS events affect gene translation and protein diversity, so we selected AS events related to prognosis and analyzed the distributions of genes for different types of AS events related to prognosis. To determine whether AS events can be used as prognostic factors, we selected the most significant genes using multivariate regression model analysis and observed their effects on prognosis.

Gene interactions and functional analysis in different types of prognostic AS events

To observe the association between significant genes in different types of AS events and prognosis, we mapped the genes to the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database using a score greater than 0.4 to obtain the interaction relationship of these genes and used Cytoscape [10] for visualization. Then, we used the *clusterprofiler.R* package to conduct KEGG enrichment analysis on the AS genes significantly correlated with prognosis in each type of AS event and observed the pathways enriched by these genes.

Relationship between gene expression and prognosis in AS events

To determine the relationship between gene expression and prognosis in AS events that significantly correlated with prognosis, we used TCGA RNA-seq expression profile data to perform univariate analysis survival of each gene separately.

Characteristic gene selection and construction of a prognostic model

Genes with Pearson correlation coefficients greater than 0.2 or less than -0.2 were selected as prognostic feature genes to construct prognostic models suitable for cervical cancer. The classification of prognosis by these

prognostic characteristic genes were observed at the level of AS events and expression profiles.

Results

ES is the main selective splicing event in patients with cervical cancer

We counted all AS events in 253 cancer samples, as shown in Additional file 1: Table S1; the seven AS events are shown in Fig. 1a. We detected 41,766 AS events in 9961 genes, of which the seven types of AS events were distributed as shown in Fig. 1b. A single gene might display several types of mRNA AS events, and we found that nearly one-third of all AS events were ES events.

Pronounced consistency between the genes involved in overall survival and the genes involved in recurrence

To observe the relationship between AS events and prognosis, we integrated clinical follow-up data as shown in Additional file 2: Table S2. A total of 41,766 AS events were analyzed by univariate survival analysis to observe the relationship between these AS events and the prognosis of patients with cervical cancer. By selecting $p < 0.05$, we obtained a total of 3306 AS events that were significantly related to survival and 2077 genes. There were 2174 AS events significantly associated with the recurrence of cervical cancer, including 1443 genes (Additional file 3: Table S3). There were 524 intersections of AS events that were significantly related to overall survival and recurrence (Fig. 2a) that contain a total of 589 overlapping genes (Fig. 2b). This shows that there is a pronounced consistency between the genes involved in overall survival and the genes associated with recurrence. The statistics of the AS events significantly related to overall survival are shown in Fig. 2, which shows that the largest proportion of ES events are reduced in relation to recurrence. In contrast, AP and AT events increased,

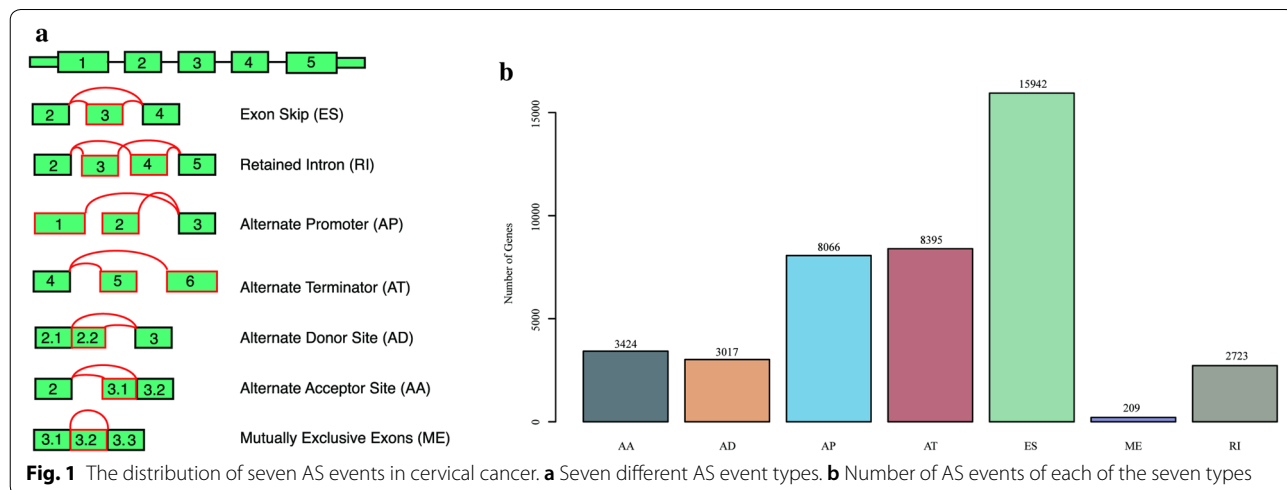
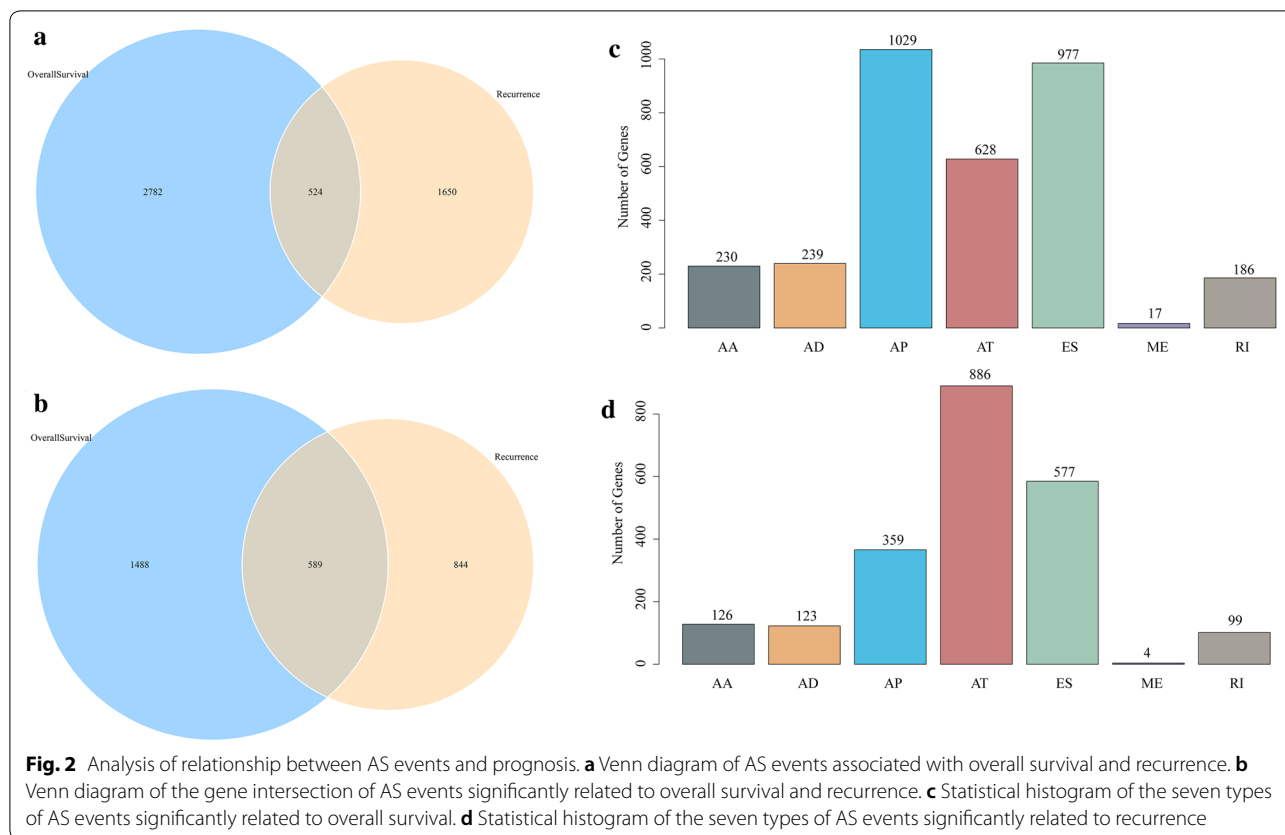


Fig. 1 The distribution of seven AS events in cervical cancer. **a** Seven different AS event types. **b** Number of AS events of each of the seven types



and among the AS events associated with recurrence, AT events occurred the most (Fig. 2d), indicating that most ES events were not associated with prognosis, while approximately 10% of AT events were significantly associated with recurrence of cervical cancer.

A gene may have multiple AS events of different types

Because AS events affect gene translation and subsequent protein diversity, we selected AS events related to prognosis to analyze the distributions of the genes involved in these events. As shown in Fig. 3a, b, a gene can be seen in several different types of AS events, and each of these AS events may be related to prognosis.

AS events can be used as a new prognostic classification for cervical cancer

To observe whether an AS event can be used as an independent prognostic factor for overall survival, the multivariate Cox regression model was applied to the seven types of survival-related AS events. According to the results of the univariate Cox regression model, the first 15 genes with the most significant AS event p value per type were selected. Multivariate Cox analysis models were carried out according to the percent spliced in (PSI) value of the AS events of these 15 genes.

The results are presented in Fig. 4a–g, which shows that the seven predictors built with different types of AS events have considerable power in distinguishing favorable or poor outcomes for cervical cancer patients. As shown in Fig. 4h, the type of AS event can classify patient overall survival; the performance of the AP and AT models are optimal for overall survival (area under the curve (AUC) = 0.89), followed by the ES model (AUC = 0.86).

Then, we used multivariate Cox regression models to determine whether the seven types of AS events could be used as independent prognostic factors for predicting recurrence in patients with cervical cancer. As presented in Fig. 4i–o, the seven predictors built with different types of AS events showed considerable power in distinguishing favorable or poor outcomes for cervical cancer patients.

Regarding AS events associated with recurrence (Fig. 4p), the AT, AA, AD, AP and RI models performed optimally, with AUCs close to 1, suggesting that AS events could be used as a new method of prognostic classification.

Forest map analysis

The most significant top 15 genes in each type of AS event were selected for forest map analysis (Fig. 5).

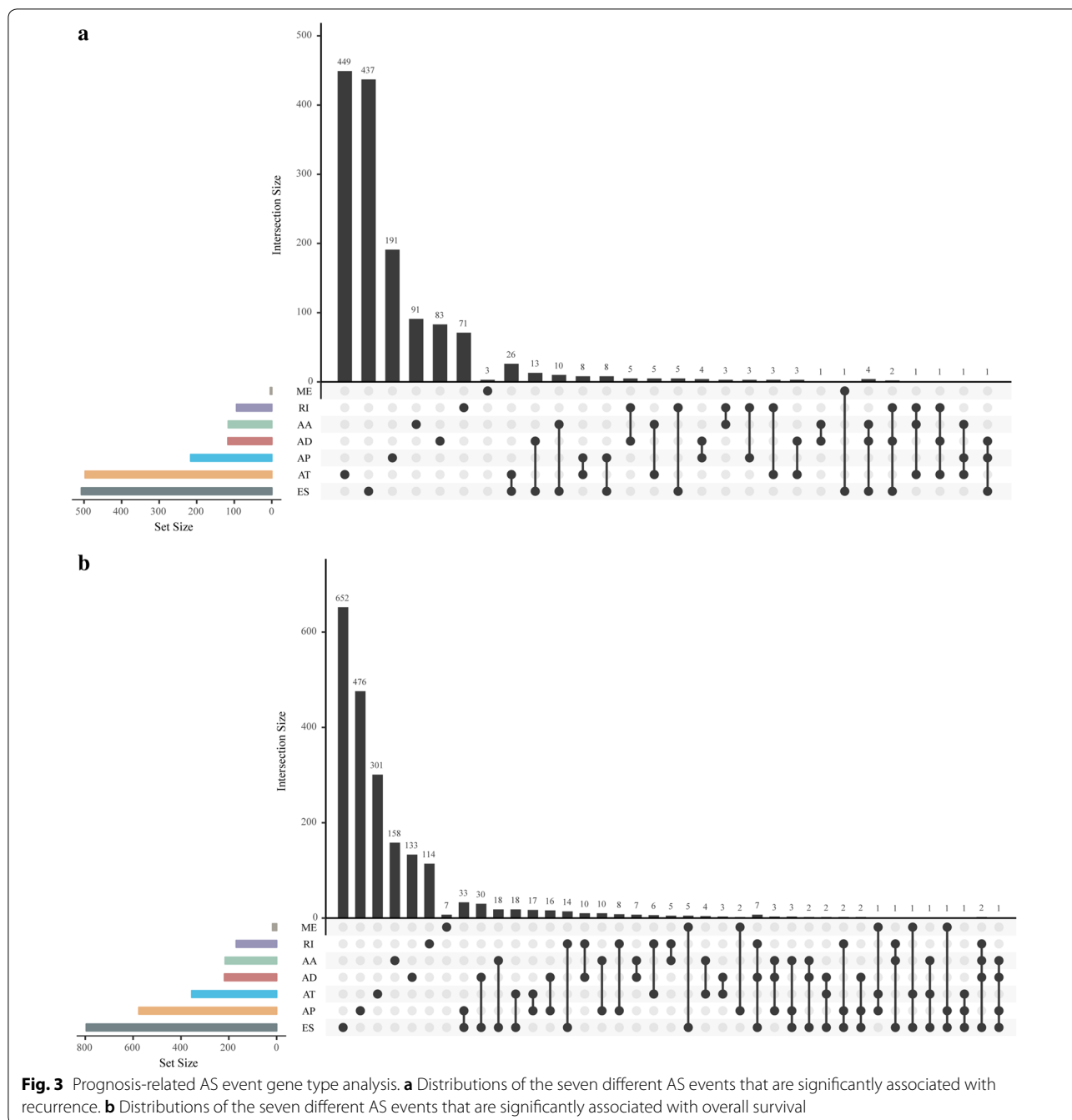


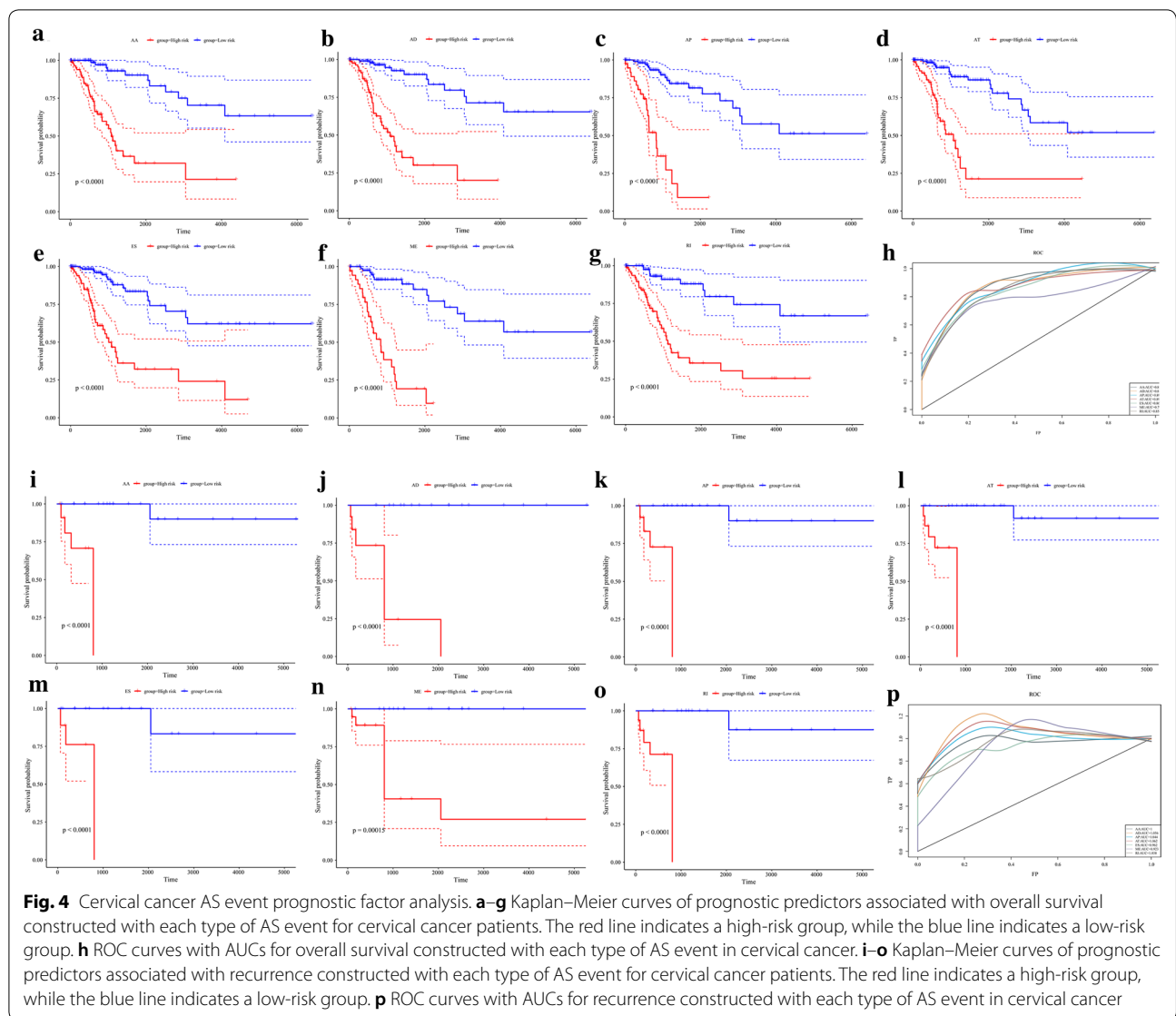
Fig. 3 Prognosis-related AS event gene type analysis. **a** Distributions of the seven different AS events that are significantly associated with recurrence. **b** Distributions of the seven different AS events that are significantly associated with overall survival

Most of the survival-associated AA events were favorable for prognosis (HR < 1) (Fig. 5a, h), which is consistent with recurrence risk, and the AP events were almost identical in recurrence and overall survival.

Gene interaction networks and functional analysis of the different types of AS events

As mentioned previously, one gene might have more than one AS event. To observe the relationship between different types of genes in AS events that are

significantly related to prognosis, we mapped the genes to the STRING database to obtain the interaction of these genes using a score greater than 0.4. Cytoscape was used to visualize the results shown in Fig. 6a, b. AD and RI show most of the interactions, and most of the genes in AS events related to prognosis have protein interactions. To observe the function of genes in different types of AS events that were significantly associated with prognosis, KEGG enrichment analysis was performed for each type of gene significantly



related to prognosis. The results are shown in Fig. 6c, d. We found that the enrichment of these genes is related to multiple disease pathways, suggesting that the genes are involved in many biological functions.

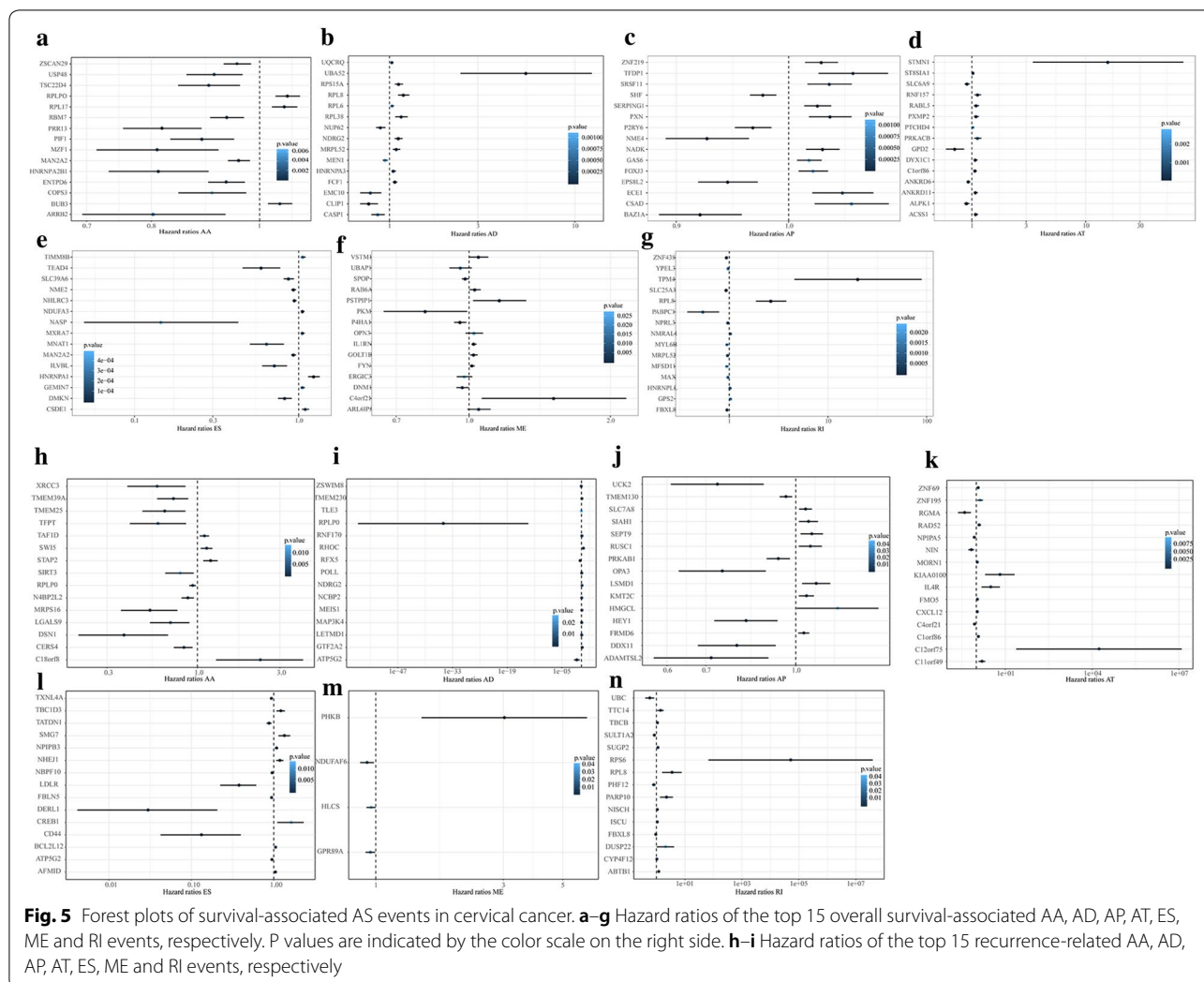
Relationship between gene expression profile and prognosis in AS events

To observe the relationship between gene expression and prognosis in AS events, we used TCGA RNA-seq expression profile data to analyze the survival rate related to each gene and found that 241 AS genes were associated with overall survival. The expression of 159 AS genes was significantly correlated with recurrence. Furthermore, we analyzed the correlation between the 241 genes associated with overall survival and the occurrence of AS events and found that 115 genes were significantly correlated with AS (Pearson $p < 0.05$). Similarly, 71 of the

159 genes associated with recurrence were significantly associated with AS events, indicating that the AS events of most genes were significantly associated with their expression.

Characteristic gene selection and construction of a prognostic model

We chose to study those genes with Pearson correlation coefficients greater than 0.2 or less than -0.2 for gene expression and AS events. There were 46 genes related to the overall survival of patients with cervical cancer and 82 related to recurrence. Four genes were associated with both overall survival and recurrence: HSPA14, SDHAF2, CAMKK2 and TM9SF1 (correlations with transcriptome levels are shown in Fig. 7). The graph shows that all of these genes have negative correlations between AS and transcription levels.



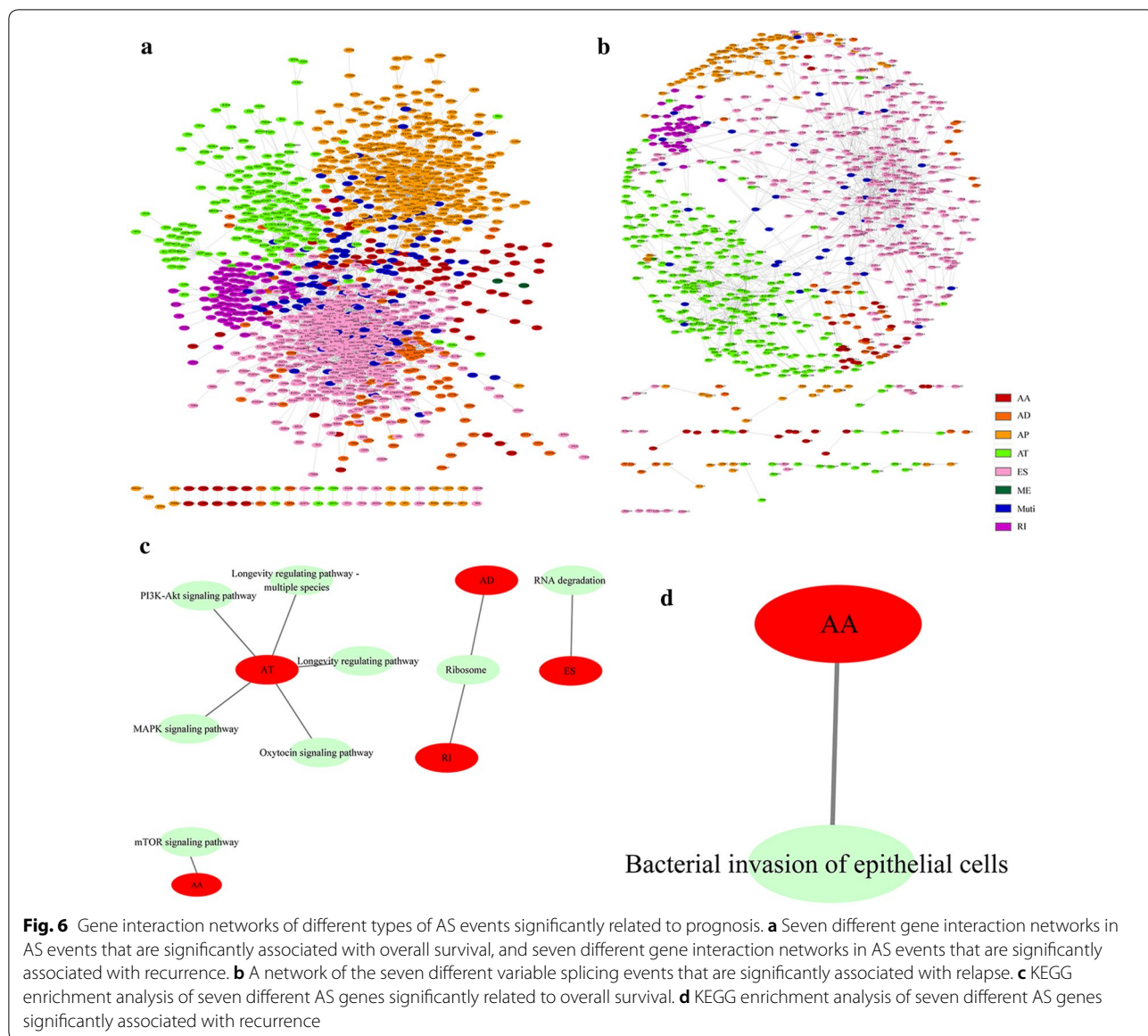
The four characteristic genes were selected for the construction of a multivariate survival model to classify prognosis regarding AS events and gene expression profiles. The map shows that these four genes have a high prognostic classification effect in both datasets (Fig. 8b, d, f, h), and the AUC value for 5 years is greater than 0.7, with high accuracy (Fig. 8a, c, e, g). These results suggest that these four genes can be used as prognostic markers of cervical cancer. The flow chart of the whole data analysis is shown in Fig. 9.

Discussion

Variable mRNA splicing is prevalent in eukaryotic cells, and Modrek et al. found that 35–59% of genes in the human genome display AS events [10]. Because AS requires the participation of many factors, an abnormality of any one factor will change the original splicing mode, thus precipitating abnormal or even opposing protein functions, leading to disease occurrence. There

is a direct relationship between the occurrence of some hereditary diseases [11] and AS events.

Current studies have confirmed that the abnormal selective splicing of pre-mRNAs is involved in the occurrence and development of cancer. The splicing products of some genes are specifically overexpressed in some tumors, and the isomer RON corresponding to the deletion of exon 11 of the RON gene has strong activity. The abnormal high expression of RON in breast and colon cancers can enhance the ability of cell migration and then affect embryogenesis, histogenesis and tumor metastasis [12]. The splicing product of the CD44 gene CD44v8-10 was significantly upregulated in gastric cancer. Its overexpression directly leads to the enhancement of the metastasis potential of tumor cells [13, 14]. KAI1 is a tumor suppressor gene that inhibits tumor metastasis. The isomer KAI1-SP, which lacks exon 7, is highly expressed in ovarian metastases. Further studies have found that this isomer has a carcinogenic effect. It can also improve



the invasive ability of tumor cells [15]. These specifically expressed AS products may become a new target for the treatment of tumors in the future. Although many cancer-specific mRNA isomers have been identified, the understanding of AS events and their functional pathways is still far from being elucidated. The rapid development of high-throughput sequencing and bioinformatics makes it possible to comprehensively explore the AS events of cervical cancer.

In this study, we screened prognosis-related AS events and discussed the relationship between AS events and the prognosis of cervical cancer for the first time using the TCGA database, which provided a new idea for exploring the pathogenesis of cervical cancer to guide clinical treatment and judge prognosis more accurately. In this study,

the characteristics of AS events in 253 patients with cervical cancer were analyzed. A total of 41,766 AS events were detected in 9961 genes, of which ES events were the main type, accounting for 1/3 of the total AS events. AT events were the second most frequently occurring, followed by AP events. Subsequently, we performed a univariate survival analysis of 41,766 selective splicing events and found that most of the ES events were not associated with prognosis, but approximately 10% of AT events were significantly associated with recurrence of cervical cancer. At the same time, we found that multiple types of AS events may occur in a gene, and these different AS events may be related to prognosis. The seven types of AS events had high AUC values. For overall survival, the AP model performed the best, followed by the

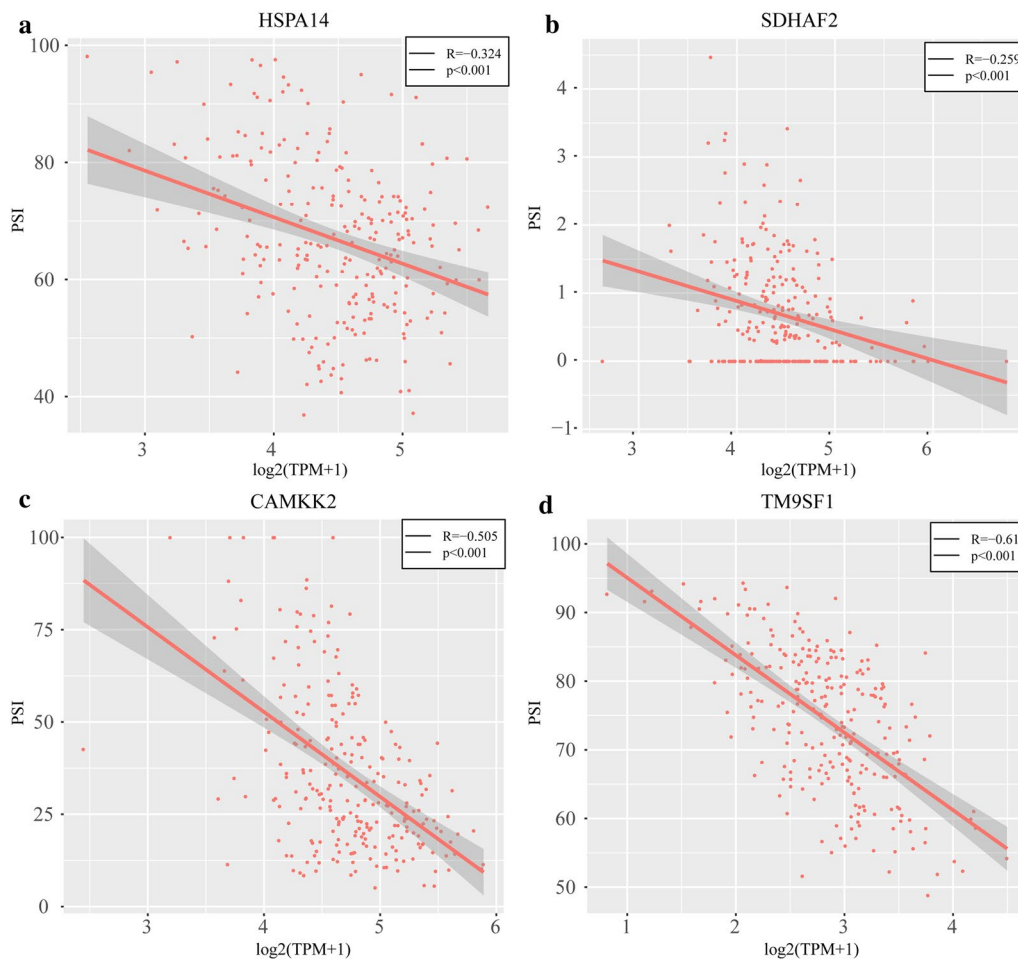
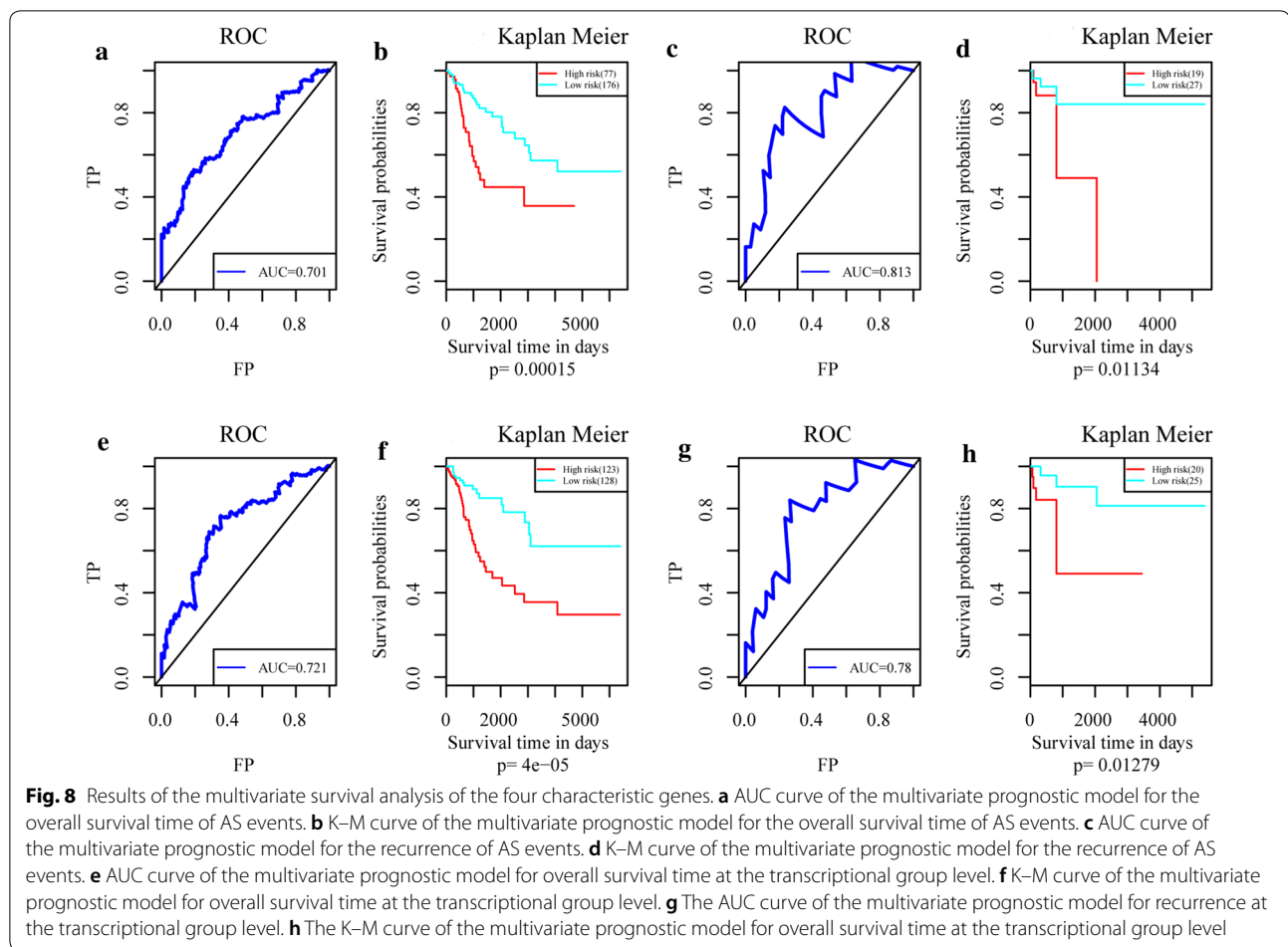


Fig. 7 Correlation between AS and transcription levels of the four characteristic genes. **a** Correlation between AS and transcription levels for HSPA14. **b** Correlation between AS and transcription levels for SDHAF2. **c** Correlation between AS and transcription levels for CAMKK2. **d** Correlation between AS and transcription levels for TM9SF1

ES model. Among the AS events significantly related to recurrence, the AT and AP models performed best, suggesting that AS events can be used as a new prognostic classification method. By constructing the interaction network of genes in AS events, we found that most of the genes in AS events related to prognosis had protein interactions. To explore the functions of the genes involved in AS events, a KEGG enrichment analysis of the genes was carried out. The results show that AT events are involved in a variety of signaling pathways, including the PI3 K-Akt signaling pathway, MAPK signaling pathway, and oxytocin signaling pathway; AD and RI events are related to ribosomal pathways; and ES events are related to RNA degradation. To observe the relationship between gene expression and prognosis in AS events, univariate survival analysis and Pearson correlation analysis were conducted, and the results showed that the AS events of most genes were significantly correlated with

their expression. Subsequently, we selected genes whose Pearson correlation coefficient between gene expression and AS events was greater than 0.2 or less than -0.2 that were related to both overall survival and recurrence. There were 4 genes in total: HSPA14, SDHAF2, CAMKK2 and TM9SF1. These four characteristic genes were used to construct a multivariate survival model, which had a significant prognostic classification effect on AS events and gene expression profiles; the AUC value for 5 years was more than 0.7, which indicated high accuracy. These results suggest that the four genes can be used as prognostic markers of cervical cancer.

HSPA14 gene translation can produce a heat shock protein, which is expressed in the lens of zebrafish, forms a mammalian ribosomal complex with MPP11, and acts as a Th1 adjuvant to activate dendritic cells [16–19]. The siRNA-mediated inhibition of HSPA4 or HSPA14 could reduce the migration, invasion and transformation



activity of NBS1 overexpression cells in vitro. The expression of NBS1 is closely related to the prognosis of patients with advanced head and neck squamous cell carcinoma (HNSCC) [20]. SDHAF2 gene translation can produce a protein that interacts with succinate dehydrogenase (SDH). It has been confirmed that the mutation of the SDHAF2 gene can lead to hereditary paraganglioma-pheochromocytoma type 2. The product of the CaMKK2 gene belongs to the serine/threonine protein kinase subfamily and the Ca²⁺/calmodulin-dependent protein kinase subfamily. Its main function is to regulate its downstream kinases CaMK1 and CaMK4 to further activate or inactivate related proteins through phosphorylation. In addition, it can also phosphorylate AMPK and participate in the occurrence and development of a variety of diseases [21]. CaMKK2 was significantly upregulated in hepatocellular carcinoma (HCC) and negatively correlated with the survival rate of HCC patients. TM9SF1 is an autophagy-induced gene that has been shown to be associated with Fuchs' endothelial dystrophy and corneal disease. In summary, most of the above four characteristic genes play an

important role in the process of cancer. Cervical cancer is a genetic disease with the participation of a variety of oncogenes and anticancer genes and the synergistic effect of multiple stages and pathways. The results showed that the proto-oncogene Bmi-1 was associated with clinical stage, tumor grade and lymphatic metastasis, and the Hk6 gene was overexpressed in cervical cancer [22, 23]. EGFR, NOTCHI, RHOA and other genes are associated with the lymph node metastasis of cervical cancer [24]. These studies are of great significance in revealing the role of RNA splicing in the occurrence of cervical cancer. An in-depth analysis of the RNA splicing model may indeed reveal new cancer drivers and provide insights into the mechanisms of these pathways. There are still several limitations in our research. First, the number of patients included in TCGA is limited. Second, the mechanisms of the four characteristic genes in cervical cancer are unclear. The abnormal proteins produced by the translation of the characteristic genes of AS events of cervical cancer are very important. In the follow-up study, we will further verify the expression of these characteristic genes and the molecular

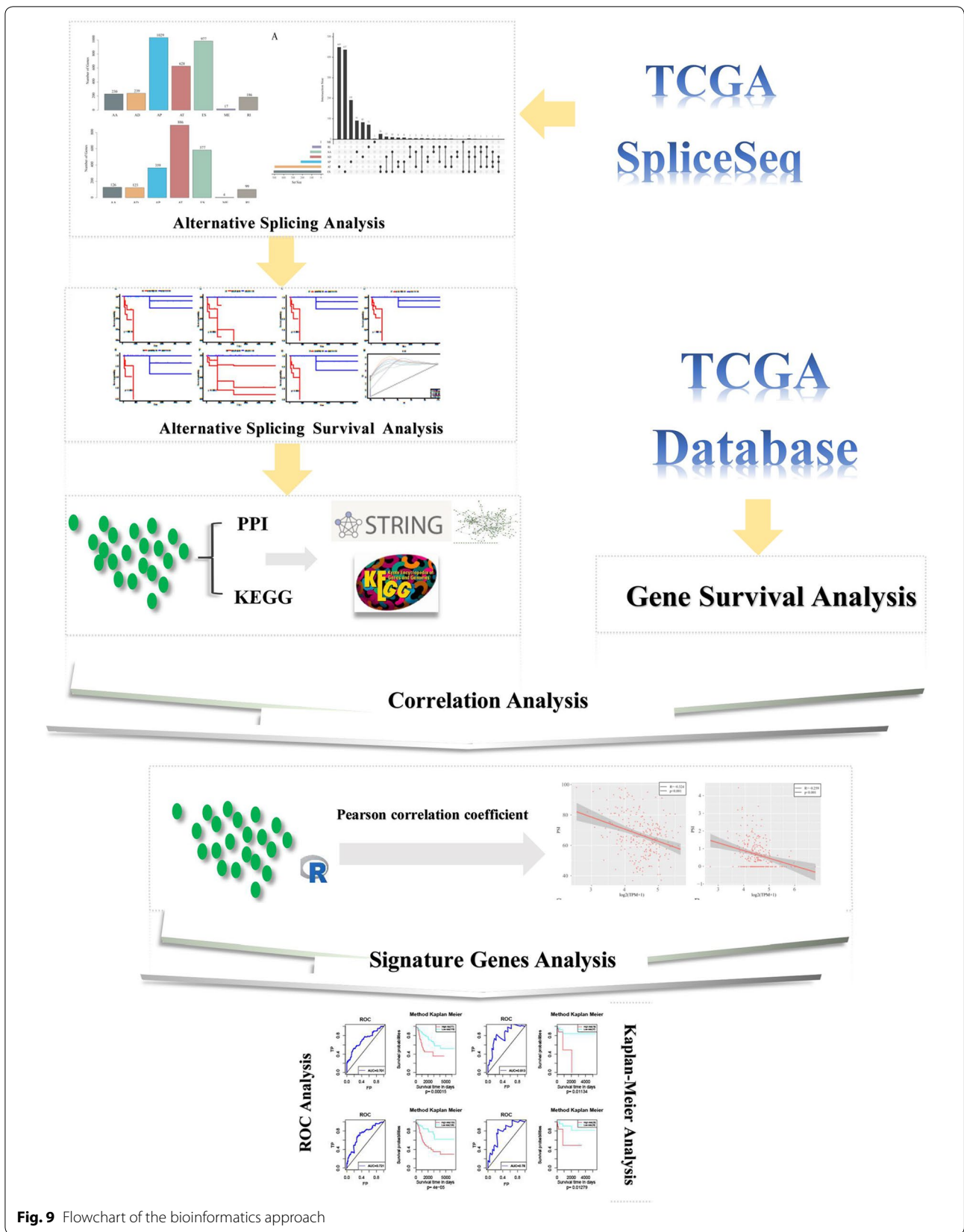


Fig. 9 Flowchart of the bioinformatics approach

mechanisms involved in the occurrence and development of cervical cancer.

Conclusions

In conclusion, these results highlight the value of AS events in the prognosis of cervical cancer and explore the potential regulatory mechanisms of cervical cancer. A large amount of data at the whole-genome level revealed the significance of AS events in the occurrence and development of cervical cancer, especially in the prediction of prognosis, which may provide a potential target for the treatment of cervical cancer.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12967-019-02140-x>.

Additional file 1: Table S1. Alternative splicing events in 253 cancer samples.

Additional file 2: Table S2. Integrated clinical follow-up data of cervical cancer patients.

Additional file 3: Table S3. Prognosis-related alternative splicing events of cervical cancer.

Abbreviations

KEGG: Kyoto Encyclopedia of Genes and Genomes; HPV: human papillomavirus; ES: exon skip; ME: mutually exclusive exons; RI: retained intron; AP: alternate promoter; AT: alternate terminator; AD: alternate donor site; AA: alternate acceptor site; HSPA14: Heat Shock Protein Family A (Hsp70) Member 14; SDHAF2: Succinate Dehydrogenase Complex Assembly Factor 2; CAMKK2: Calcium/Calmodulin-Dependent Protein Kinase 2; TM9SF1: Transmembrane 9 Superfamily Member 1.

Acknowledgements

We sincerely thank *American Journal Experts* for guidance with the English copyediting of our paper.

Authors' contributions

ZMJ performed the majority of the work described here. ZMJ and HYY wrote and edited the manuscript. ZWC translated the manuscript. LX, GR and HYY collected the data. NX, LQ and LJJ examined the data. LB supervised the project. All authors read and approved the final manuscript.

Funding

This work is supported by Grants from the Shengjing Freedom researchers' plan (201804). The funding body had no role in the design or conduct of the study.

Availability of data and materials

All data in the current study are based on public data available in The Cancer Genome Atlas (TCGA) datasets.

Ethics approval and consent to participate

Not applicable.

Consent for publication

All the authors have read and approved the paper and declare no potential conflicts of interest in the paper. If the paper is accepted, all the authors will observe the terms of the license to publish.

Competing interests

The authors declare that there are no competing interests.

Author details

¹ Department of Obstetrics and Gynecology, Shengjing Hospital Affiliated to China Medical University, Shenyang, Liaoning 110004, China. ² Department of Obstetrics and Gynecology, University Hospital, LMU Munich, Marchioninistr. 15, 81377 Munich, Germany.

Received: 14 February 2019 Accepted: 14 November 2019

Published online: 19 November 2019

References

- Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015. *CA Cancer J Clin*. 2016;66:115–32.
- Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin*. 2015;65:87–108.
- Pan Q, Shai O, Lee LJ, et al. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet*. 2008;40(12):1413–5.
- Gilbert W. Why genes in pieces? *Nature*. 1978;271:501.
- Buratti E, Baralle M, Baralle FE. Defective splicing, disease and therapy: searching for master checkpoints in exon definition. *Nucleic Acids Res*. 2006;34:3494–510.
- Deng M, Bragelmann J, Schultze JL, et al. Web-TCGA: an online platform for integrated analysis of molecular cancer data sets. *BMC Bioinform*. 2016;6(17):72.
- Tomczak K, Czerwinska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge [J]. *Contemp Oncol (Pozn)*. 2015; 19 (1A): A68-A77.
- Comprehensive molecular portraits of human breast tumours. *Nature*. 2012;490(7418):61–70.
- O'Quigley J, Moreau T. Cox's regression model: computing a goodness of fit statistic. *Comput Methods Prog Biomed*. 1986;22(3):253–6.
- Modrek B, Lee C. A genomic view of alternative splicing. *Nature Genetics*. 2002;30:13.
- Cáceres JF, Kornbliht AR. Alternative splicing: multiple control mechanisms and involvement in human disease. *Trends Genet*. 2002;18(4):186–93.
- Ghigna C, Giordano S, Shen H, et al. Cell motility is controlled by SF2/ASF through alternative splicing of the Ron proto oncogene. *Mol Cell*. 2005;20(6):881–90.
- Lau WM, Teng E, Chong HS, et al. CD44v8-10 is a cancer-specific marker for gastric cancer stem cells. *Cancer Res*. 2014;74(9):2630–41.
- Yae T, Tsuchihashi K, Ishimoto T, et al. Alternative splicing of CD44 mRNA by ESRP1 enhances lung colonization of metastatic cancer cell. *Nat Commun*. 2012;3:883.
- Upheber S, Karle A, Miller J, et al. Alternative splicing of KAI1 abrogates its tumor-suppressive effects on integrin $\alpha\beta 3$ -mediated ovarian cancer biology [J]. *Cell Signal*. 2015;27(3):652–62.
- Krone PH, Evans TG, Blechinger SR. Heat shock gene expression and function during zebrafish embryogenesis. *Sem Cell Dev Biol*. 2003;14:267–74.
- Otto H, Conz C, Maier P, Wolfle T, Suzuki CK, Jeno P, Rucknagel P, Stahl J, Rospert S. The chaperones MPP11 and Hsp70L1 form the mammalian ribosome-associated complex. *Proc Natl Acad Sci USA*. 2005;102:10064–9.
- Wu Y, Wan T, Zhou X, Wang B, Yang F, Li N, Chen G, Dai S, Liu S, Zhang M, Cao X. Hsp70-like protein 1 fusion protein enhances induction of carcinoembryonic antigen-specific CD8+ CTL response by dendritic cell vaccine. *Cancer Res*. 2005;65:4947–54.
- Wan T, Zhou X, Chen G, An H, Chen T, Zhang W, Liu S, Jiang Y, Yang F, Wu Y, Cao X. Novel heat shock protein Hsp70L1 activates dendritic cells and acts as a Th1 polarizing adjuvant. *Blood*. 2004;103:1747–54.
- Yang MH, Chiang WC, Chang SY, Chang SY, Chen PM, Teng SC, Wu KJ. Increased NBS1 expression is a marker of aggressive head and neck cancer and overexpression of NBS1 promotes transformation. *Clin Cancer Res*. 2006;12:507–15.
- Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*. 2004;116:281–97.
- Jiao Li, Ping Han, Jie Bai. The expression of Bmi-1 in cervical lesions and association with HR-HPV infection. *Modern Oncol*. 2015;23(09):1275–9.

23. Lu Z, Chen H, Zheng XM, et al. Expression and clinical significance of high risk human papillomavirus and invasive gene in cervical carcinoma. *Asian Pac J Trop Med*. 2017;10(2):187–91.
24. Fan SY, Li CX, Wang T, et al. Random Forests algorithm—based bioinformatic screening of functional genes involved in lymph metastasis of cervical cancer. *Chin J Biochem Pharm*. 2016;36(4):5–8.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

