



A Novel Prognostic Index Based on Alternative Splicing in Papillary Renal Cell Carcinoma

Zhipeng Wu^{1†}, Jinhui Liu^{2†}, Rui Sun^{2†}, Dongming Chen¹, Kai Wang¹, Changchun Cao^{1*} and Xianlin Xu^{1*}

¹ Department of Urology, Sir Run Run Hospital, Nanjing Medical University, Nanjing, China, ² Department of Gynecology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China

OPEN ACCESS

Edited by:

Julio R. Banga,
Spanish National Research Council
(CSIC), Spain

Reviewed by:

Zheng Xia,
Oregon Health & Science University,
United States
Yuriy L. Orlov,
First Moscow State Medical
University, Russia

*Correspondence:

Xianlin Xu
xianlinxu@njmu.edu.cn
Changchun Cao
caochangchun@njmu.edu.cn

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Bioinformatics and
Computational Biology,
a section of the journal
Frontiers in Genetics

Received: 25 August 2019

Accepted: 05 December 2019

Published: 29 January 2020

Citation:

Wu Z, Liu J, Sun R, Chen D, Wang K,
Cao C and Xu X (2020) A Novel
Prognostic Index Based on
Alternative Splicing in Papillary
Renal Cell Carcinoma.
Front. Genet. 10:1333.
doi: 10.3389/fgene.2019.01333

Background: Papillary renal cell carcinoma (pRCC) is a heterogeneous multifocal or isolated tumor with an invasive phenotype. Previous studies presented that alternative splicing, as a crucial posttranscriptional regulator in gene expression, is associated with tumorigenesis. However, the association between alternative splicing and pRCC has not been clarified

Methods: The RNA sequencing data and clinical information were downloaded from The Cancer Genome Atlas database and mRNA splicing profiles from TCGASpliceSeq. The percent spliced in data of alternative splicing merged with survival information was firstly calculated by univariate Cox regression analysis to screen for survival-associated alternative splicing events, and survival-associated alternative splicing events were then analyzed by Gene Ontology categories using Kyoto Encyclopedia of Genes and Genomes. Meanwhile, the least absolute shrinkage and selection operator Cox analysis and multivariate Cox analysis were performed to calculate the prognostic index for each alternative splicing type. In addition, clinical factors were introduced to assess the performance of prognostic index.

Results: A total of 4,084 candidate survival-associated alternative splicing events in 2,558 genes were screened out. Patients were divided into the low-risk group and the high-risk group based on the median prognostic index value. The Kaplan-Meier survival analysis ($p < 0.05$) and receiver operating characteristics curves ($AUC > 0.9$) indicated that prognostic index was effective and stable for predicting the prognosis of pRCC patients. Furthermore, a regulatory network was constructed incorporating alternative splicing events and survival-associated splicing factors.

Conclusion: Our study provides new insights into the mechanism of alternative splicing events in tumorigenesis and their clinical potential for pRCC.

Keywords: alternative splicing, prognostic index, papillary renal cell carcinoma, splicing factor, The Cancer Genome Atlas

INTRODUCTION

Papillary renal cell carcinoma (pRCC), which accounts for up to 15% of renal cell carcinoma, is the second most common histological subtype of kidney cancer (Jonasch et al., 2014; Malouf et al., 2016). pRCC emerges as either indolent localized tumor or aggressive metastatic cancer (Delahunt and Eble, 1997), but the biological basis for this difference remains unidentified. Vascular endothelial growth factor (VEGF) pathway has been proven to be involved in metastatic pRCC (Armstrong et al., 2016), but we still speculate that multiple mechanisms lie behind these pRCC with diverse presentations. Thus, we designed this systematic and comprehensive analysis to drill into the oncogenic mechanism of pRCC.

High-throughput sequencing has revolutionized human genomics and the research in this field. The current number of human genes is still controversial. Up to now, people's statistics on the number of genes are constantly changing (Pertea and Salzberg, 2010; Pertea et al., 2018). The GENCODE (Frankish et al., 2019) genome maintained by EBI currently counts 19,965 protein-coding genes, 17,910 long noncoding RNA genes, and 7,576 small noncoding genes in human (<https://www.encodegenes.org/human/stats.html>). The database RefSeq (O'Leary et al., 2016), managed by the National Center for Biotechnology Information, lists 20,203 protein-coding genes and 17,871 noncoding genes. Regardless of the specific number of genes, given the limited number of human genes, alternative splicing (AS) serves as a key mechanism producing myriads of proteins (Tang et al., 2013; Bowler et al., 2018). AS is regulated by spliceosome, a large and highly dynamic protein complex constructed by nearly 200 protein components and five small nuclear ribonucleic acids (Agrawal et al., 2018). Dysregulation of

splicing factors (SFs) can distort mRNA splicing programs, which could result in cancer development and progression (Grosso et al., 2008). Studies have also shown that aberrant AS events during transcription, which are tissue-specific and stage-specific, can evoke tumorigenesis (Wang et al., 2008; Kahles et al., 2018; Wan et al., 2019).

In this study, to clarify the AS events and its clinical implications in pRCC, AS events and complete clinical information from the TCGA database were analyzed. A prognostic model was formed to predict the prognosis of pRCC according to the survival information. Meanwhile, a regulatory network was constructed to evaluate the correlation between AS events and SFs, and identify several key factors which might exert important functions in occurrence and development of pRCC.

MATERIALS AND METHODS

Data Collection of AS Events

RNA sequencing data (level 3) and clinical information of The Cancer Genome Atlas (TCGA) KIRP cohorts were obtained from the TCGA data portal (<https://portal.gdc.cancer.gov/>). Analysis of mRNA splicing profiles in pRCC was conducted with the aid of SpliceSeq (Ryan et al., 2016), java that explicitly quantifies RNA-Seq reads and identifies its possible functional changes as a consequence of AS in the context of transcript splice graphs. AS events were divided into seven types 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) (Figure 1). Meanwhile, we downloaded the Percent Spliced In (PSI) value

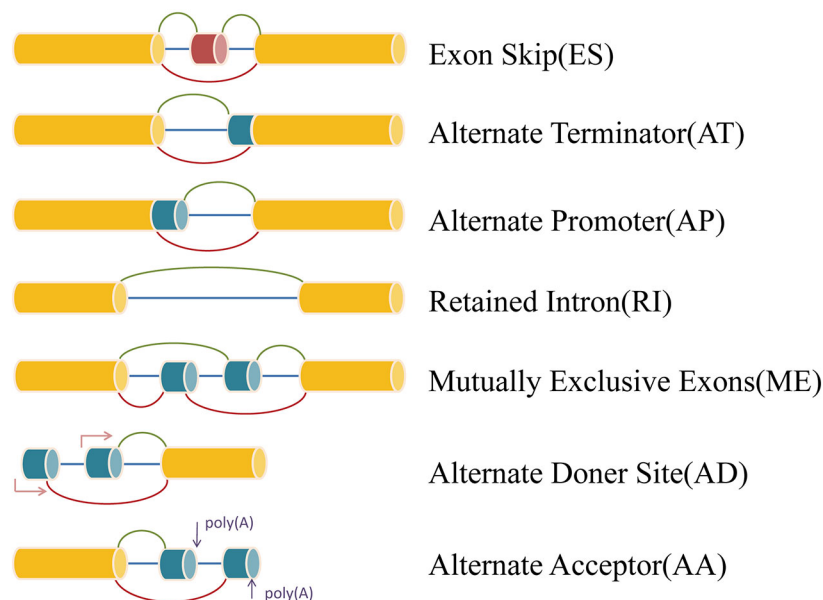


FIGURE 1 | Representative model of seven types of alternative splicing.

(>75%) for pRCC patients. The PSI value, ranging from zero to one, was used in quantifying AS events.

Identification of Survival-Associated AS Events

A total of 32 adjacent normal tissues and 289 pRCC tissues were collected from TCGA. The number of AS events and genes involved was showed by UpSet plot using “UpSetR” package in R (Conway et al., 2017). Univariate Cox regression analysis was used to screen out the candidate AS events ($P < 0.05$).

Functional Annotation

The parent genes of survival-associated AS events were subjected to functional enrichment analyses. Gene ontology (GO) term enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were performed using “clusterProfiler” package in R (Yu et al., 2012). A p-value and q-value both smaller than 0.05 in GO and KEGG was considered significant.

Survival Analysis

The result of univariate Cox regression analysis in identifying survival-associated AS events was shown by Volcano plot using “ggplot2” in R (Figure 4A). Seven types of AS events were revealed by Bubble chart respectively. Each chart contained the top 20 significant survival-associated AS events of corresponding types. LASSO method was then employed for the regression of high-dimensional predictors. LASSO Cox regression model was used to determine the ideal coefficient for each prognostic feature and to estimate the deviance likelihood via 1-standard error (SE) criteria. The coefficients and partial likelihood deviance were calculated with the “glmnet” package in R (Friedman et al., 2010).

Construction and Validation of a Prognostic Model

The result of LASSO Cox regression was then submitted to multivariate Cox analysis to evaluate the independent prognostic value of each gene and construct an independent prognosis model.

The risk score model for prediction based on survival-associated AS events were calculated by multiplying the PSI values of prognostic indicators and the regression coefficient calculated by the multivariate Cox regression analysis.

Risk Score based on events(patient)

$$= \sum_i \text{Coefficient}(\text{mRNA}_i) \times \text{PSI value}(\text{mRNA}_i)$$

The specific calculation formula was shown in Table 1. Patients were then divided into two groups based on the median levels of risk score. This prognostic model and patient survival information were merged. Kaplan-Meier survival curves were conducted to identify the prognostic ability of prediction models. Area under the curve (AUC) value for the ROC curves of each prognostic model was calculated by survivalROC package in R. Besides, univariate and multivariate analysis were performed

containing the risk score of prognostic models and important clinical features for pRCC. Patients with incomplete clinical information and less than 90 days of OS were excluded, and only 129 samples were deemed qualified. Finally, a nomogram was constructed using the “rms” package on R. The calibration of this nomogram was assessed by calibration curves.

Correlation Network of SF and Survival-Associated AS Events

The data of SFs was obtained from SpliceAid 2 (Piva et al., 2012). SF files and patient survival information were merged and calculated by univariate Cox regression analysis to get survival-associated SFs. Correlation network was constructed using the gene expression of SFs and PSI values of prognosis-related AS events with the conditions of P value less than 0.001 and Pearson correlation coefficient more than 0.7. The correlation network was plotted by Cytoscape (version 3.6.1).

The risk score model based on survival-associated SFs was the sum of each optimal prognostic mRNA expression level multiplying relative regression coefficient weight calculated from the multivariate Cox regression model.

Risk Score based on SFs(patient)

$$= \sum_i \text{Coefficient}(\text{mRNA}_i) \times \text{Expression}(\text{mRNA}_i)$$

RESULTS

Overview of AS Events in pRCC

We collected 41,673 AS events from 10,026 genes in 32 adjacent normal tissues and 289 pRCC tissues. The numbers of the genes showing seven types of AS events were plotted by UpSet plot (Figure 2A). Several genes only have one kind of AS event, ES was found in 1,699 genes (the largest number) and ME in 36 genes (the smallest number). The plot also showed that one gene might involve two or more AS events, leading to multiple transcripts from one gene. AS data was merged with clinical survival data and calculated by univariate Cox regression analysis. As a result, 4,084 AS events in 2,558 genes were deemed associated with the overall survival (OS) ($p < 0.05$). The result was also shown by UpSet plot (Figure 2B).

Functional and Pathway Enrichment Analysis

Based on survival-associated AS events, GO (Figure 3A), and KEGG (Figure 3B) were conducted by “clusterProfiler” package in R. The involved functions and pathways included “ciliary basal body-plasma membrane docking,” “purine ribonucleotide metabolic process,” and “organelle localization by membrane tethering” in biological process (BP), “adherens junction,” “focal adhesion,” and “cell-substrate adherens junction” in cellular component (CC), “cadherin binding,” “cell cadherin molecule binding,” and “retinoic acid receptor binding” in molecular function (MF). Besides, these genes were mainly enriched in “MAPK signaling pathway,” “thermogenesis,” and “human

TABLE 1 | Prognostic signatures for papillary renal cell carcinoma.

| Type | Formula | AUC |
|------|---|-------|
| AA | FKBP8 48446 AA*-8.261326216+ITGB1BP1 52617 AA*-3.645087449+FAM213A 12365 AA*-3.591154025+MYL6 22381 AA*-16.63167012+PILRB 80930 AA*4.461677112+TRPT1 16579 AA*-28.05557115+RPS24 12297 AA*-2.123512398+ENTPD6 58863 AA*-4.538790703+SPATA20 42429 AA3.809741969 | 0.967 |
| AD | TCEB1 84216 AD*-3.227392376+ATP5J 60266 AD*-6.409262477+IRF3 51027 AD*-5.374919151+CFL2 27169 AD*-4.984160824+ATP6V1H 83836 AD*-14.79100606+AKT1S1 51111 AD*-6.962315249+PQB1 89029 AD*-2.068032636 | 0.861 |
| AP | UNG 24277 AP*4.194935004+NHS 88586 AP*-2.822266615+DYNC12 55939 AP*-6.138633492+LIMA1 21688 AP*3.922489143+PMEPA1 59946 AP*1.640017308+RNF220 2555 AP*-4.636337467 | 0.961 |
| AT | CLDN11 67616 AT*1.439647629+SLC25A48 73462 AT*-3.684109126+KIF4A 89373 AT*2.503402505+RBM39 59235 AT*-22.33853542+BCAM 50347 AT*3.282115813+LARP1B 70567 AT*-3.356001505 | 0.959 |
| ES | YIF1A 17008 ES*-20.56958015+RPS24 12295 ES*-1.999160275+C16orf13 265882 ES*-2.503302918+ARHGEF10 82562 ES*-2.479391656+PAX8 55049 ES*-7.098738552+MUM1 46457 ES*2.479075665 | 0.928 |
| ME | AMT 64866 ME*-11.67288726+C14orf2 29528 ME*2.439891378+CBWD5 86504 ME*-9.123153517+MEF2A 32717 ME*-5.089591595+FAIM 67014 ME*-3.726223934+CERS5 21668 ME*-2.23329899+STK36 57557 ME*-15.49762863+FBLN5 28892 ME*10.43894503+GLOD4 123198 ME*-4.401329099 | 0.901 |
| RI | TMUB1 82347 RI*-4.817625749+ELP5 38889 RI*-5.149262596+ZSWIM7 39393 RI*-6.043747899+LRRC29 36982 RI*4.739092712+ZNF276 38138 RI*6.294309106+SNX5 58749 RI*-4.792058193+TTLL3 63208 RI*4.305033369 | 0.963 |
| All | SEC31A 69730 ES*-3.130605352+RPS24 12295 ES*-2.37334211+FKBP8 48446 AA*-10.11459992+UNG 24277 AP*4.542255311+ARHGEF10 82562 ES*-3.009037952 | 0.968 |

Example: "FKBP8|48446|AA*-8.261326216," "FKBP8" is the gene name, "48446" is the AS id, "AA" is the type of AS event, "FKBP8|48446|AA" mean the PSI value of AA event in FKBP8, and "-8.261326216" is the regression coefficient calculated by the multivariate Cox regression analysis.

cytomegalovirus infection" in KEGG. AS events generating from these genes might influence the occurrence and development of pRCC through interfering with the above BPs and pathways.

Prognostic Model for pRCC

Among the survival-associated AS events (Figure 4A), the top 20 significant AS events were shown by bubble chart (Figures 4B–H). These AS events were further treated with LASSO Cox analysis (Figures 5A–G). LASSO Cox analysis of all AS events (ALL) was also shown (Figure 5H). Next, multivariate Cox analysis was performed to construct an independent model with PI. The formula was showed in Table 1. Patients were divided into the low-risk group and the high-risk group based on the median risk score. With the increasing risk score, the patient's survival became worse. The risk plot of ALL AS events was shown in Figure 6. Figure 6C with high resolution were affiliated in the Supplementary Figure 1. The separate risk plots of seven AS event were affiliated in Supplementary Figure 2. The Kaplan-Meier survival analysis suggested that a pRCC patient with a higher risk score might show a worse survival (Figures 7A–H). An AUC value of more than 0.9 was found in all the seven types of AS in pRCC except AD (0.861), which validated the efficiency of these signatures in predicting prognosis. To assess whether this model was an independent predictor of pRCC, univariate analyses was performed between clinical factors and risk score. The results showed that this model could distinguish pRCC patients ($p < 0.001$) (Table 2). Furthermore, by using multivariate analyses, this prognostic model was proved to serve as a moderate and independent prognostic indicator in the AS events of AA, AD, AP, AT, and RI (Figures 8A–H). Finally, to provide a clinically associated quantitative method, we tried to construct a nomogram incorporating riskscore and clinical factors to predict the probabilities of 3- and 5-year OS in pRCC. Since the predict model based on AA event showed the best performance among AS

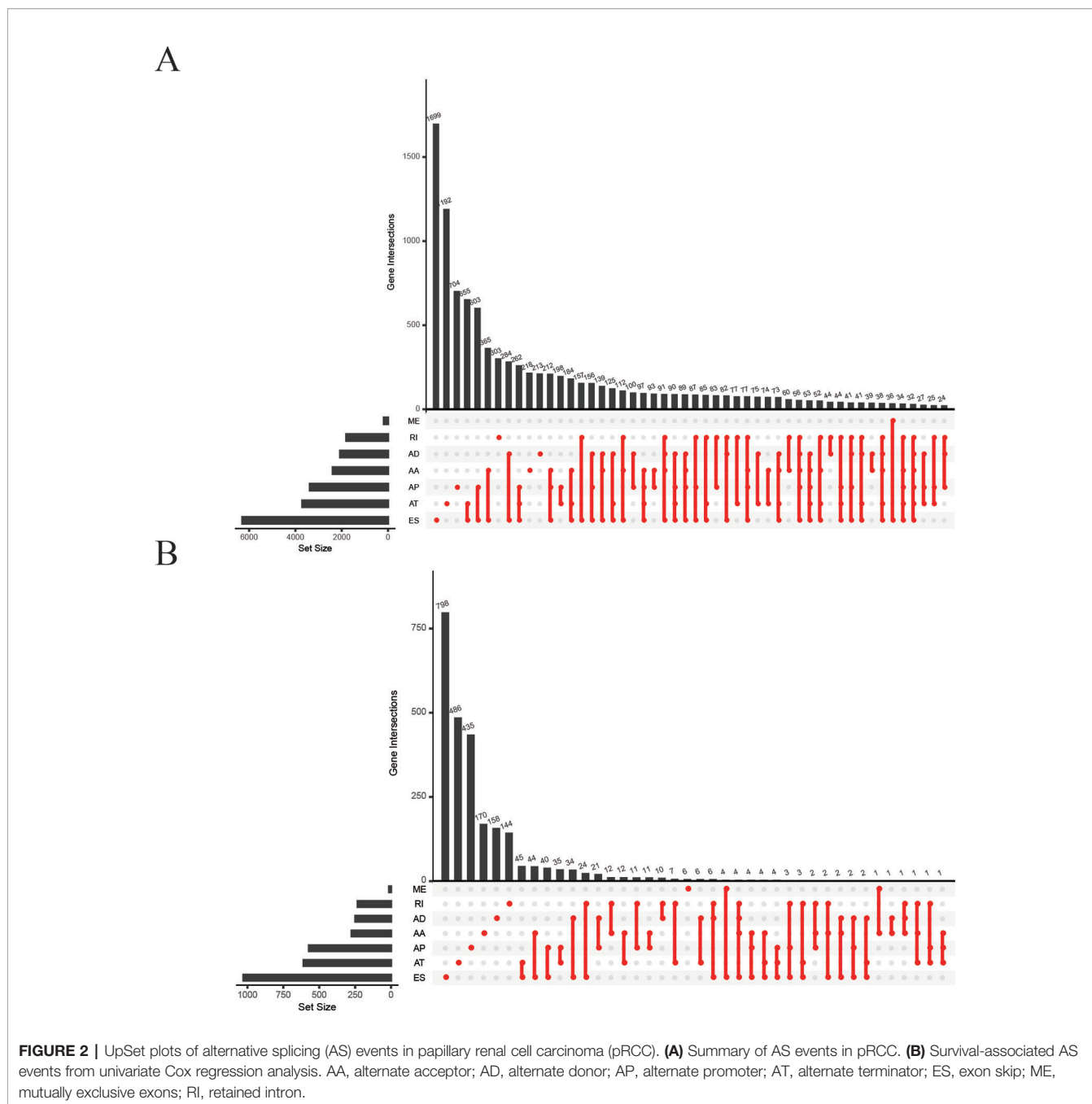
events that proved to serve as a moderate and independent prognostic indicator (AUC = 0.967), the nomogram was constructed based on AA event (Figure 9A). The Harrel's concordance index (C-index) for OS prediction was 0.963, which showed a fairly high prediction accuracy of this nomogram. The calibration curves for the 3- (Figure 9B) OS rates showed good agreement between the prediction and the actual observation, but not so good in 5-year (Figure 9C).

A Survival-Associated Network Incorporating SFs and AS Events

SFs play an important part in the occurrence and development of AS events via changing exon selection and splicing site. Therefore, it is necessary to uncover the correlation between SFs and AS events. Survival analyses based on TCGA data was performed to screen out potential SFs. And then, correlation network was constructed using the expression value of SFs and PSI values of prognosis-related AS events (Figure 10A). Figure 10A with high resolution were affiliated in the Supplementary Figure 3. To highlight the key players in this network, we performed LASSO Cox analysis on those SFs (Figure 10B). CDK12, CDK10, SF3A2, SNRNP35, and SNRNA were screened out by multivariate Cox regression analysis (Figure 10C).

Risk score = the expression of CDK12 * 0.260889 + the expression of CDK10 * (-0.06669) + the expression of SF3A2 * 0.041459 + the expression of SNRNP35 * (-0.21053) + the expression of SNRPA * 0.087523

The model formed by these five SFs showed outstanding prognostic efficiency as evidenced by the Kaplan-Meier survival curves (Figure 10D) and ROC curve (Figure 10E). With clinical information, we further analyzed the five key genes with multivariate analyses (Figure 10F). In additionally, only CDK12 and SF3A2 were deemed significant statistically



($p < 0.05$). The HR values of CDK12 and SF3A2 in multivariate analyses were all greater than 1, suggesting that CDK12 and SF3A2 may associate with the poor survival of pRCC patients.

DISCUSSION

Previous studies have presented that AS is a crucial posttranscriptional regulation leading to structural transcript variation and proteome diversity (Zhang and Manley, 2013). Abnormal AS is associated with tumorigenesis (Grosso et al.,

2008). To the best of our knowledge, few systematic AS-related research has been conducted. Prior to our research, Yang XJ et al. identified pRCC into two classes using comparative genomic microarray analysis, one associated with excellent survival and the other with poor prognosis (Yang et al., 2005). Wach S et al. classified pRCC subtypes using microRNA profiles (Wach et al., 2013). Recently, machine learning models have been used to classify stages of PRCC pRCC patients, showing a best performance with area under Precision Recall curve of 0.804, Matthews Correlation Coefficient of 0.711, and accuracy of 88% with Shrunken Centroid classifier on a test dataset based on 80

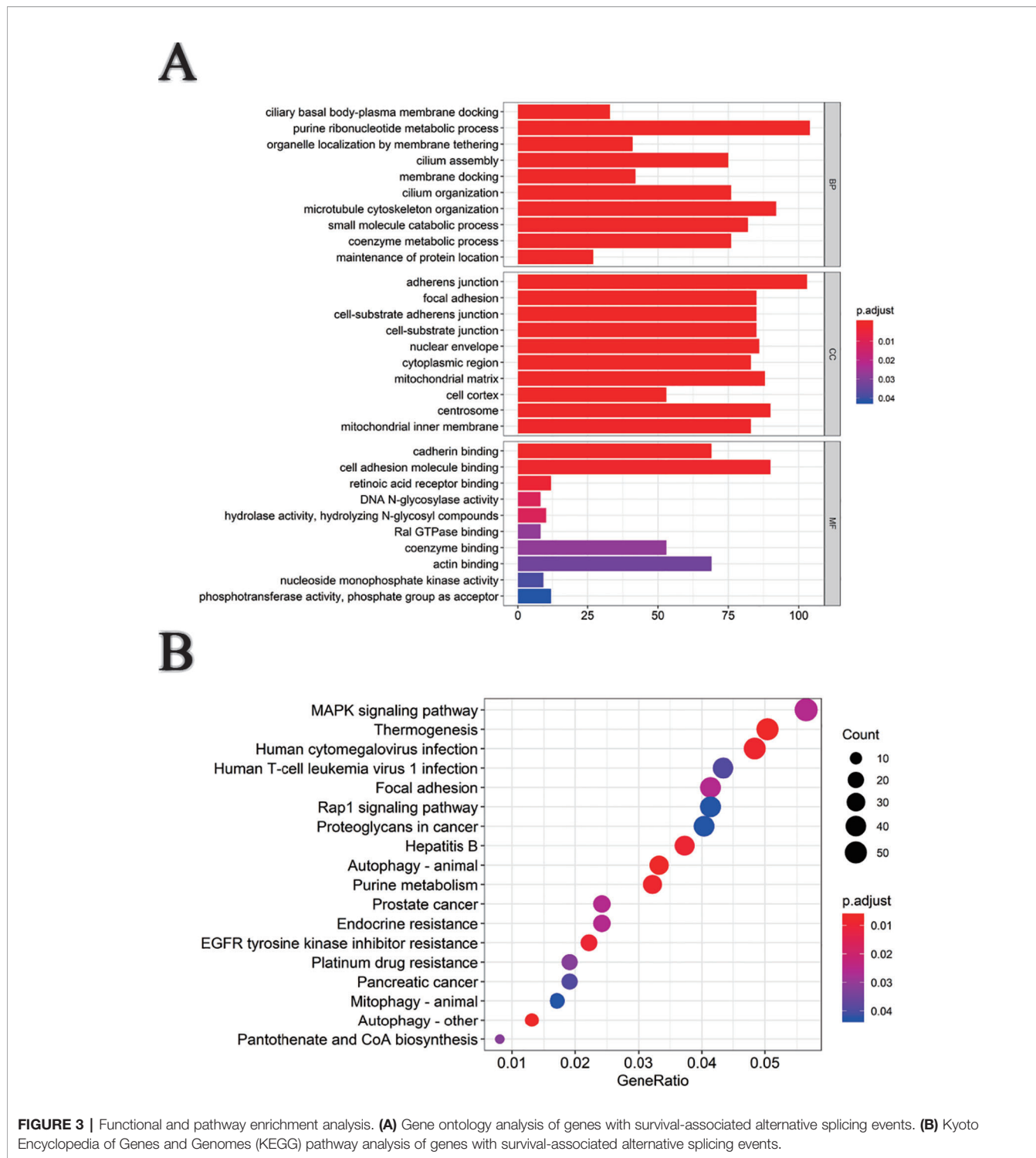
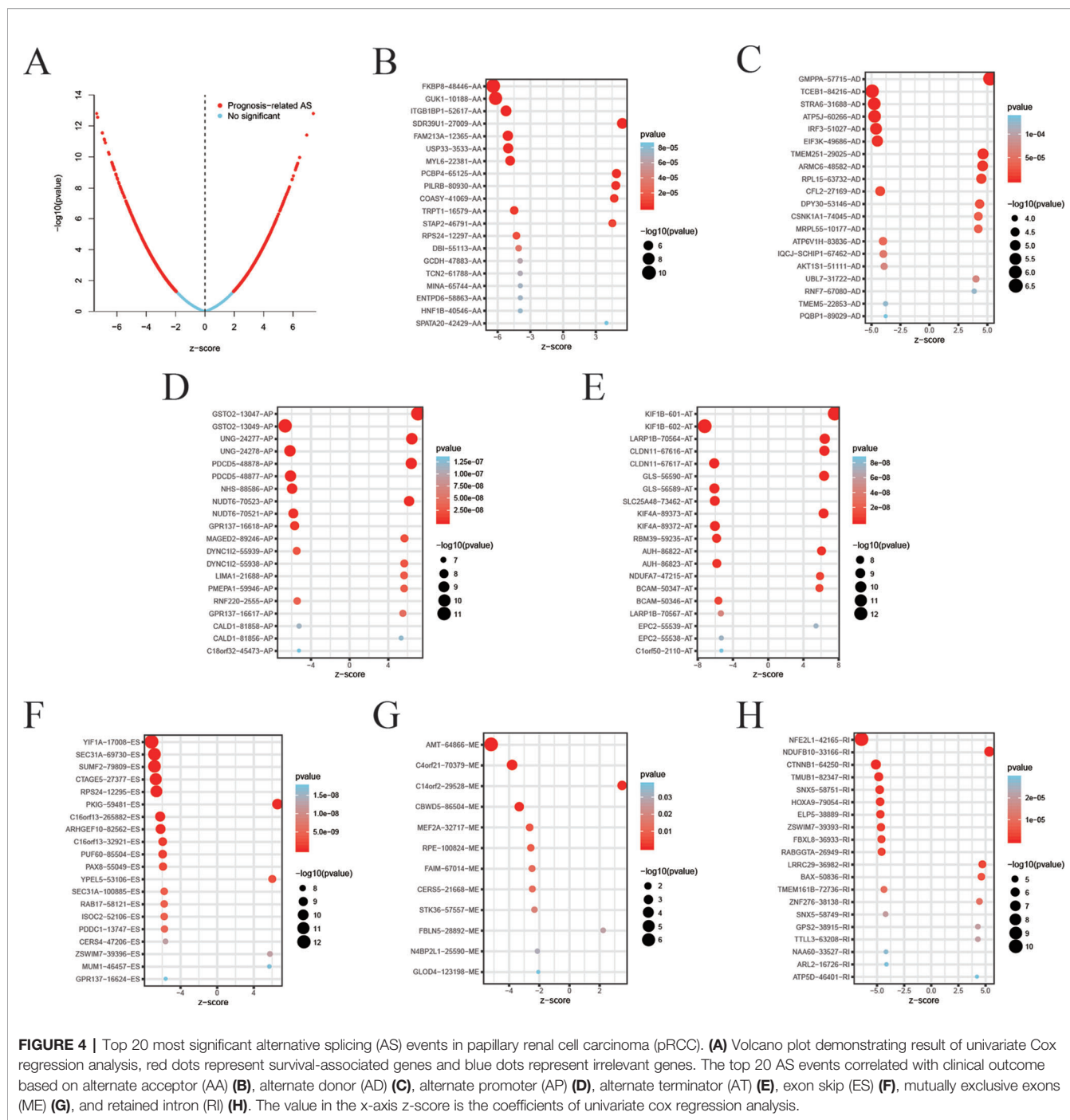


FIGURE 3 | Functional and pathway enrichment analysis. **(A)** Gene ontology analysis of genes with survival-associated alternative splicing events. **(B)** Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of genes with survival-associated alternative splicing events.

selected genes (Singh et al., 2018). Therefore, this study is the first systematic analysis on pRCC-survival-associated AS events. The analysis showed that 4,084 AS events in 2,558 genes were associated with the overall survival (OS) of pRCC patients.

The parent genes of survival-associated AS events were subjected to functional enrichment analyses, and 18 potential

pathways were enriched. Among which, MAPK signaling pathway was the top 1 in the list (**Figure 3B**). Targeted therapy against VEGF was a traditional medical treatment for renal cell carcinoma. Zhang Y et al. reported RKTG to inhibit angiogenesis by suppressing MAPK-mediated autocrine VEGF signaling (Zhang et al., 2010), and MAPK



signaling pathway has also been reported to be involved in the development of renal cell carcinoma by some other molecular regulation (Huang et al., 2008; Huang et al., 2016; Li et al., 2017). AS is closely related to tumor resistance to drugs (Wang et al., 2017; Martinez-Montiel et al., 2018; Siegfried and Karni, 2018). In the pathway that genes enriched, “EGFR tyrosine kinase inhibitor resistance” and “platinum drug resistance” were involved. EGFR (Wei et al., 2013; Robichaux et al., 2018) and platinum (Vaughn et al., 2009; Teo et al., 2017; Pal

et al., 2018) have been reported to be involved in the treatment of cancer, and through these enriched genes we may be able to discover specific mechanisms of drug resistance.

In this paper, we formed prognostic models by these survival-associated AS events. Before our research, Yang XJ et al performed microarray-based microRNA (miRNA) expression profiling of primary ccRCC and pRCC cases, and finally five miRNAs (miR-145, -200c, -210, -502-3p, and let-7c) were screened out to identify the samples with high accuracy (86.5%

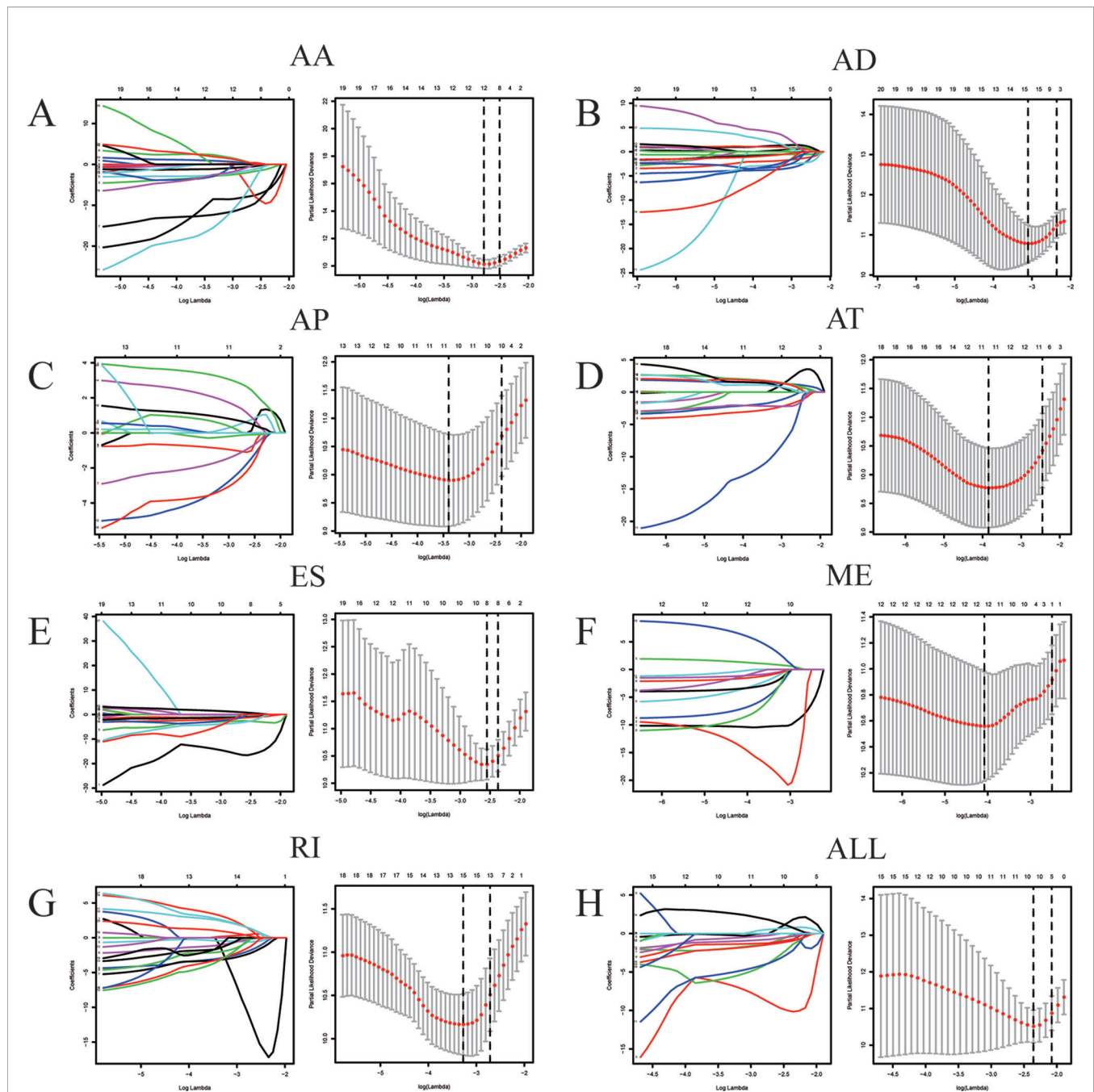


FIGURE 5 | LASSO Cox analysis of alternative splicing (AS) events. LASSO Cox regression model with 10-fold cross-validation was constructed using the top significant survival-associated AS events to screen the key AS features in AA (**A**), AD (**B**), AP (**C**), AT (**D**), ES (**E**), ME (**F**), RI (**G**) and all AS events (**H**).

in tumor/normal classification, 77.6% in ccRCC/pRCC classification, and 86.4% in pRCC type 1/2 classification); Wach S et al. used machine learning models to classify stages of pRCC patients, showing a good performance with area under Precision Recall curve of 0.804, Matthews Correlation Coefficient of 0.711 and accuracy of 88% with Shrunken Centroid classifier on a test dataset based on 80 selected genes. As for the prognostic models in our research, firstly, Kaplan-

Meier survival curves suggested that these models were appropriate methods to stratify pRCC patients into groups of different survivals ($p < 0.01$). Secondly, either single AS event or combined seven AS events performed well in predicting overall survival of pRCC patients ($AUC > 0.9$) with an exception of $AUC = 0.861$ in AD. Meanwhile, PI was proved to be independent in AA, AD, AP, AT, and RI by univariate and multivariate analyses. Compared with previous studies, using the

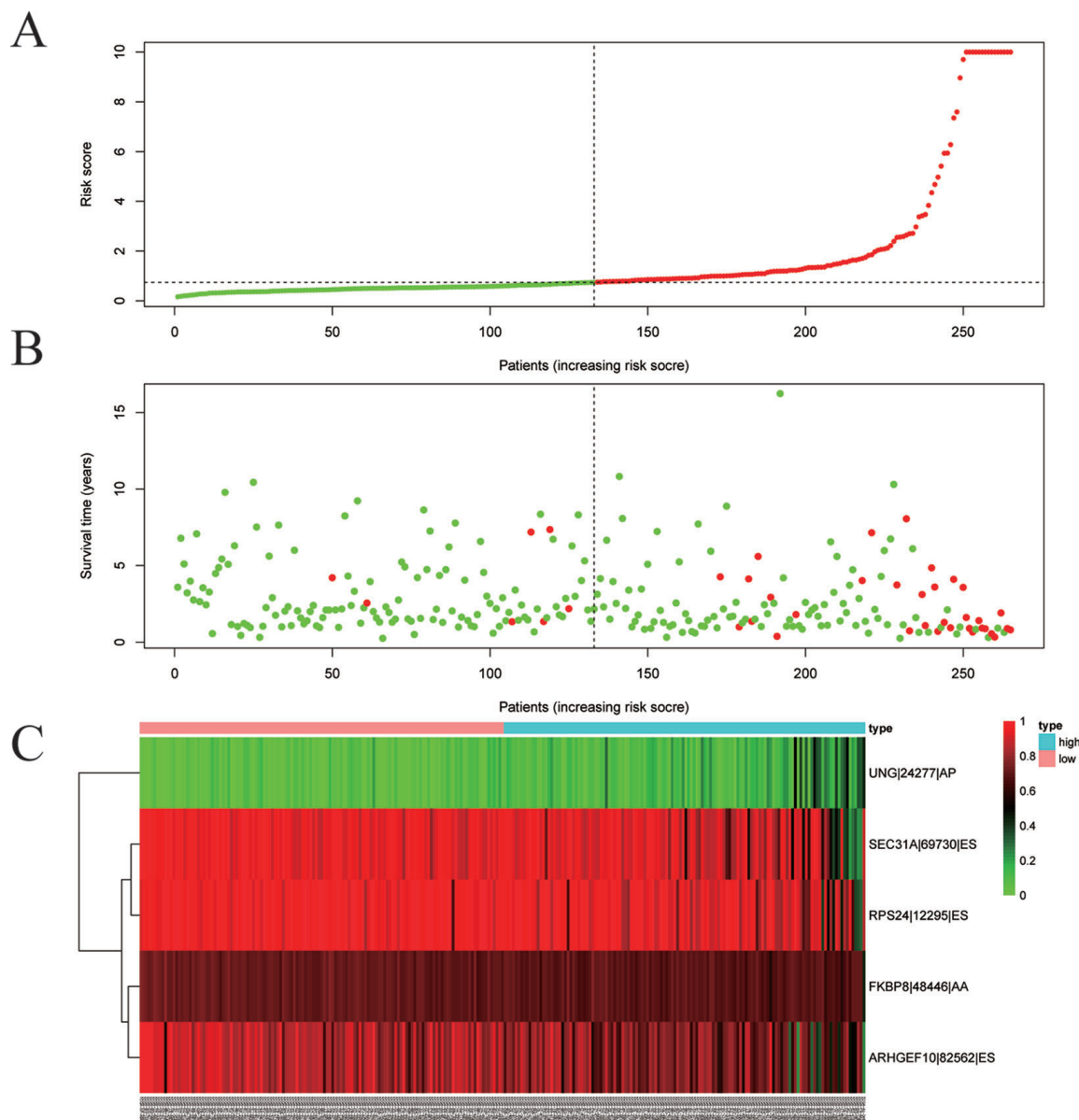


FIGURE 6 | Development of the prognostic index. Risk plot of ALL alternative splicing (AS) event. **(A)** Rank of prognostic index and distribution of groups. Patients with papillary renal cell carcinoma (pRCC) were divided into low- and high-risk subgroups based on the median value of the risk score calculated. **(B)** The survival status and survival time of patients with pRCC ranked by risk score. In **(A)** and **(B)**, green dots represent for patients with a low level of risk score and red dots represent for patients with a high level of risk score. **(C)** Heatmap of included AS event in ALL. Patients were divided into two groups according to risk score. The color from green to red means the Percent Spliced In (PSI) value from 0 to 1.

PSI value of AS events to predict patient's outcome is theoretically more systematic and accurate. At the same time, by searching for articles, we found that there were also related studies based on AS events in other tumors like uteri corpus endometrial carcinoma (Gao et al., 2019) and papillary thyroid cancer (Lin et al., 2019). However, based on the value of AUC in the prognostic model, it seems that the prognostic model based on AS events is more suitable for pRCC. We also built a nomogram for clinical application and validation.

By searching scientific literature, we found that some genes that make up PI have been reported to play an important role in tumors. For example, in the PI model of AT AS events, CLDN11, SLC25A48, KIF4A, RBM39, BCAM, and LARP1B were involved (Table 1, AT). It was reported that inactivation of CLDN11 could promote cell migration in nasopharyngeal carcinoma (Li et al., 2018). KIF4A were identified as prognostic gene or key gene involved in the metastasis of renal cell carcinoma in recent survey (Gu et al., 2017; Wei et al., 2019). Anticancer

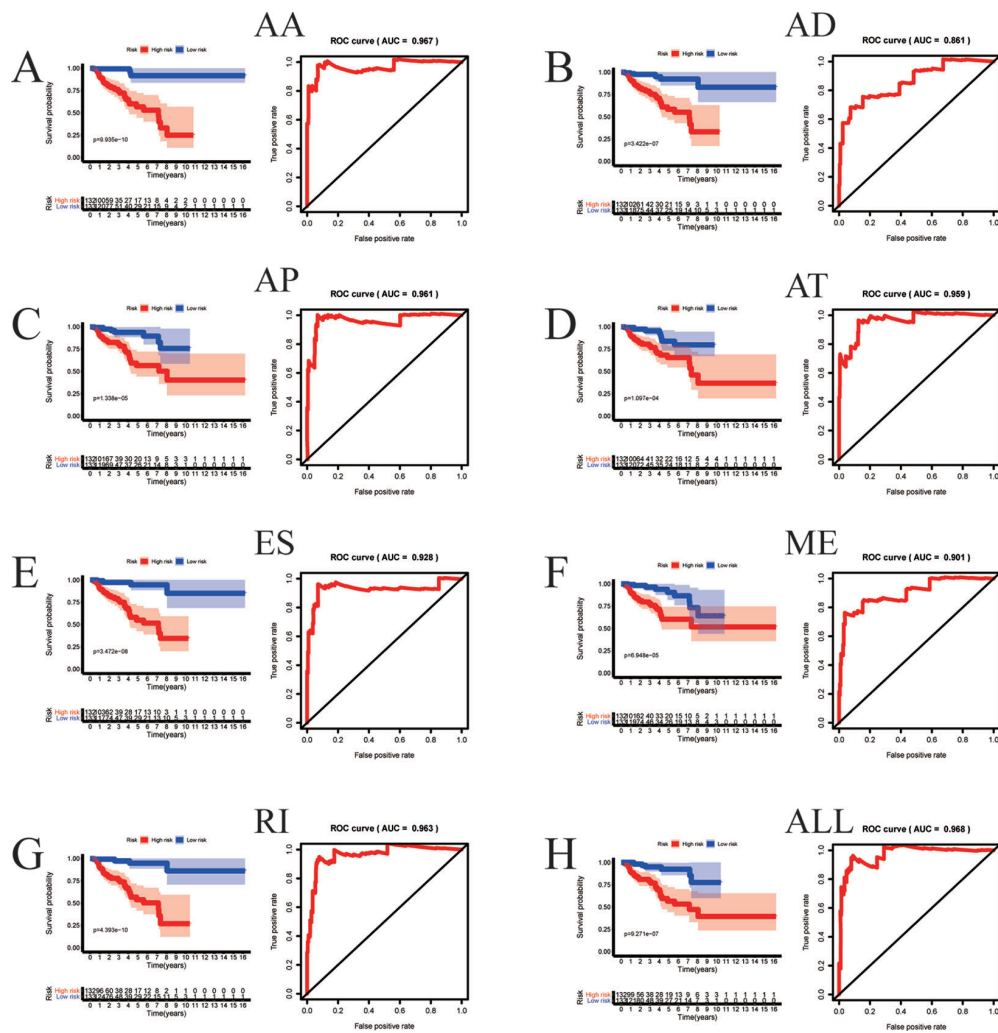


FIGURE 7 | The prognostic value of PI presented by overall survival (OS) and ROC curves. Kaplan-Meier plot depicting the survival probability over time for prognostic predictor of seven types of AS events (A–G) and all AS event (H) with high (red) and low (blue) risk group, respectively. The ROC curve to which the respective model belongs is located to the right of the KM curve.

TABLE 2 | Univariate analysis between clinical parameters and alternative splicing (AS) events in The Cancer Genome Atlas (TCGA) cohort of papillary renal cell carcinoma (pRCC) patients.

| id | HR | HR.95L | HR.95H | pvalue |
|-------------------------|-----------|-----------|-----------|------------------|
| Age | 1.8095854 | 0.8102357 | 4.041539 | 0.1479857 |
| Gender | 0.7549775 | 0.3182836 | 1.7908277 | 0.5236151 |
| Stage | 12.273306 | 5.4549378 | 27.614253 | 1.36E-09 |
| T(Tumor) | 5.3163151 | 2.4726062 | 11.430533 | 1.89E-05 |
| M(Metastasis) | 41.108075 | 12.143928 | 139.15381 | 2.33E-09 |
| N (Lymph Node) | 12.51509 | 5.7257522 | 27.354916 | 2.39E-10 |
| Risk score based on AA | 1.0053074 | 1.0032742 | 1.0073448 | 2.98E-07 |
| Risk score based on AD | 1.0162305 | 1.0078678 | 1.0246627 | 0.0001341 |
| Risk score based on AP | 1.0088837 | 1.0050934 | 1.0126883 | 4.12E-06 |
| Risk score based on AT | 1.0223319 | 1.0147675 | 1.0299526 | 5.58E-09 |
| Risk score based on ES | 1.0106936 | 1.0065517 | 1.0148524 | 3.84E-07 |
| Risk score based on ME | 1.0523824 | 1.0325915 | 1.0725527 | 1.36E-07 |
| Risk score based on RI | 1.0257308 | 1.0171857 | 1.0343477 | 2.65E-09 |
| Risk score based on ALL | 1.0076827 | 1.0039129 | 1.0114666 | 6.28E-05 |

Font bold indicates $P < 0.05$.

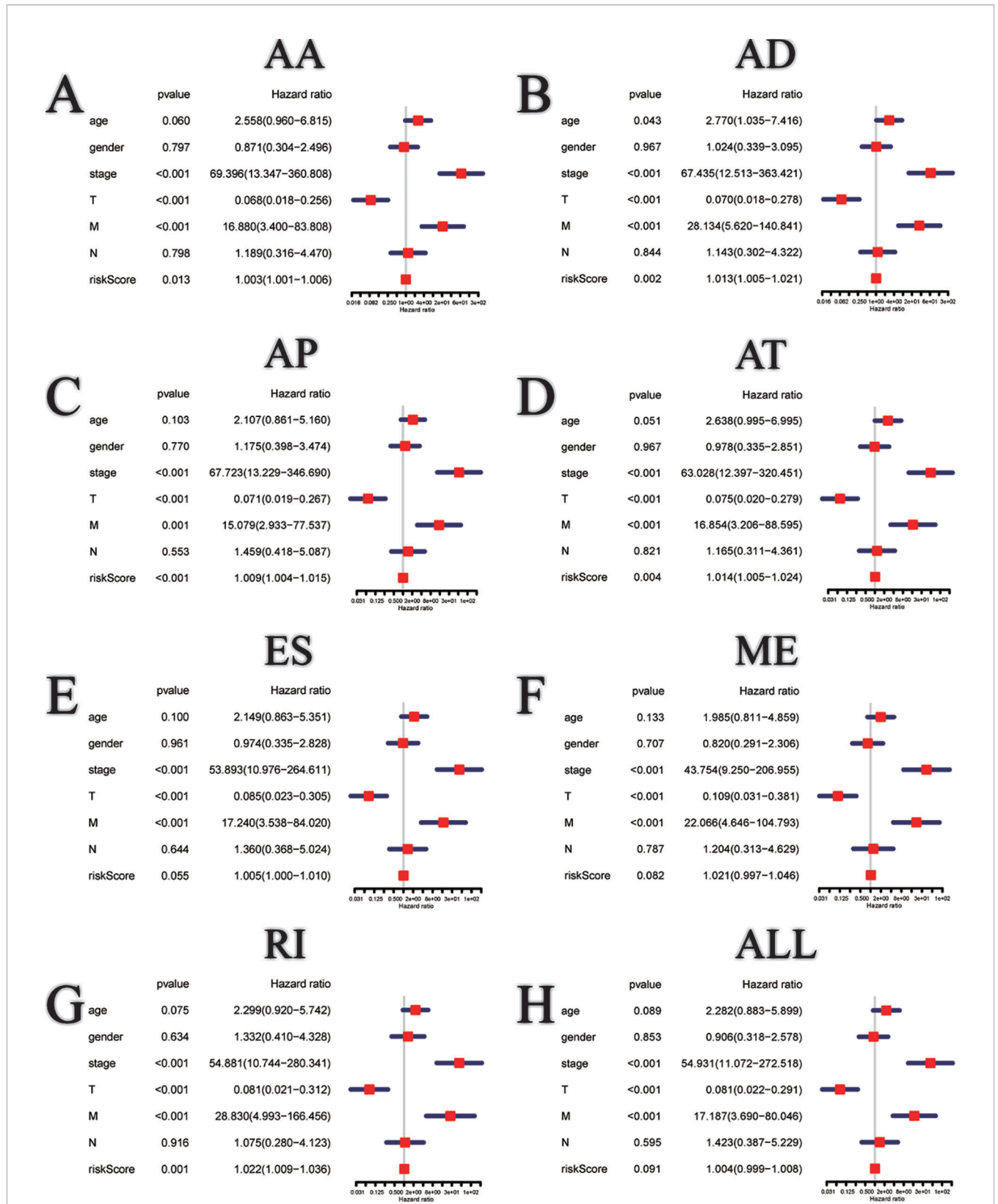
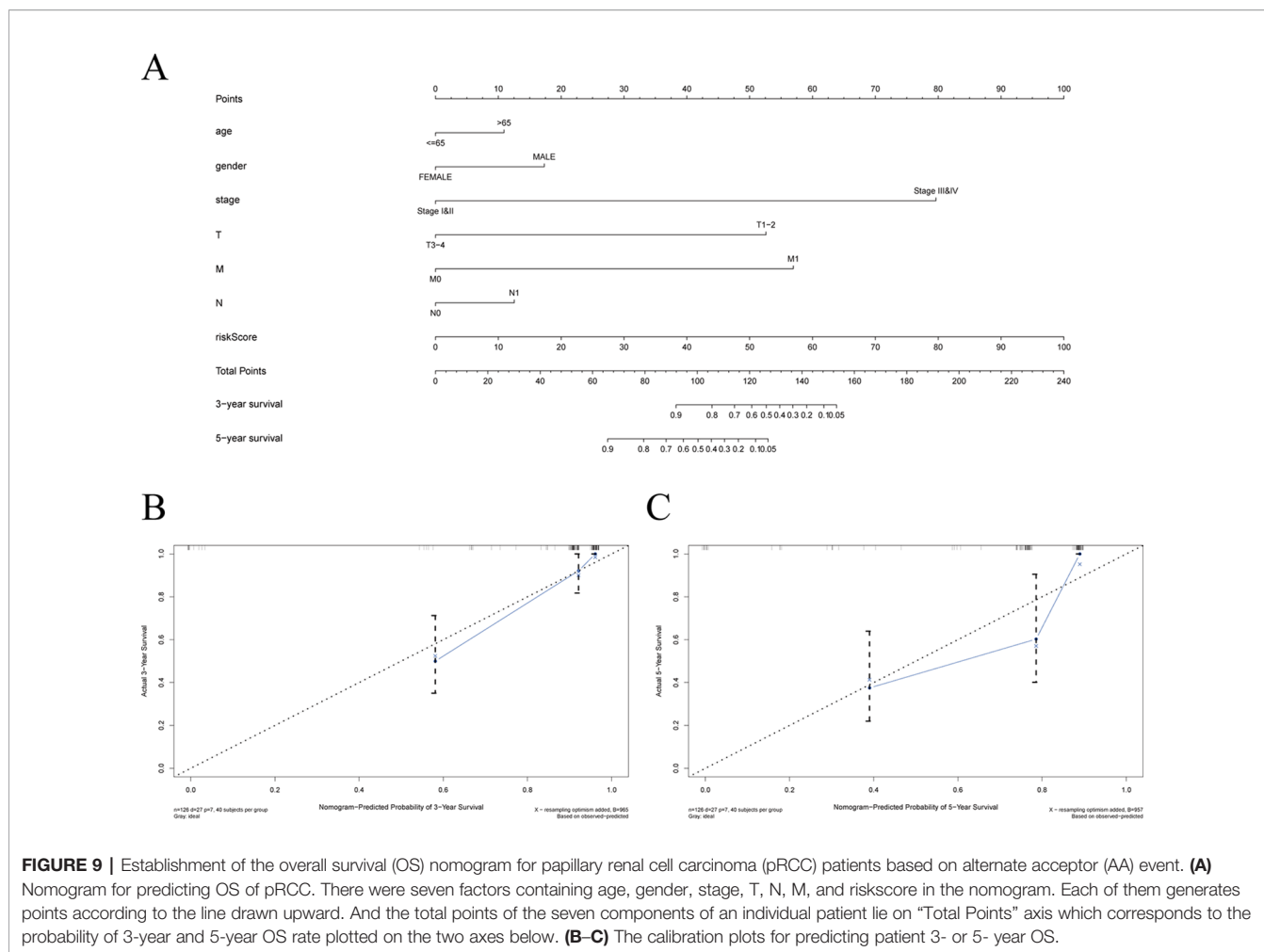


FIGURE 8 | Multivariate Cox regression analysis of clinical parameters and different PI models constructed by AA (A), AD (B), AP (C), AT (D), ES (E), ME (F), RI (G) and all AS events (H) in pRCC patients.



sulfonamides functions by inducing RBM39 degradation (Anticancer Sulfonamides Induce Splicing Factor RBM39 Degradation, 2017; Han et al., 2017). BCAM was reported to mediate recognition between tumor cells and the endothelium in KRAS-Mutant colorectal cancer. In the present study, the role of SLC25A48 and LARP1B was still unclear, and our analysis may guide the direction of future research on pRCC.

SFs are implicated in the process of alternative mRNA splicing (Venables et al., 2009). Splicing abnormalities arise owing to aberrant expression and/or mutations of SFs (Venables, 2006). Hence, SFs have a tight link with AS events. In this paper, survival-associated SFs were screened out by survival analyses. Correlation network was then formed to describe the interactions between SFs and AS events. Both positive and negative correlations were observed between one SF and multiple survival-associated AS events, or between one survival-associated AS event and multiple SFs. By performing LASSO Cox analysis and multivariate Cox regression analysis, CDK12, CDK10, SF3A2, and SNRNA were screened out. However, only CDK12 and SF3A2 were deemed significant statistically ($p < 0.05$) according to multivariate analyses. CDK12 contains an arginine–serine-rich (RS) domain, and can regulate the splicing of a minigene construct (Chen et al., 2006). CDK12 may be inactivated in patients with metastatic

castration-resistant prostate cancer, and may make tumors more responsive to PD-1 inhibitors (CDK12 Changes Telling in Prostate Cancer, 2018). According to the regulatory network in our research, we found a negative correlation between CDK12 and SP100-57896-AT. It has been reported that SP100 could reduce malignancy of human glioma cells (Held-Feindt et al., 2011). The HR value of SF3A2 was 1.051, and SF3A2 was also reported to be associated with the metastasis and recurrence of osteosarcoma (Zhang et al., 2019). These results indicated that these altered SFs, as independent molecules, can construct a regulatory network in the carcinogenesis and progression in pRCC. However, this network may be optimized with more molecules. Besides, only 129 pRCC patients were involved in our analysis due to the restricted standard of OS time more than 90 days and requirement for complete clinical data. PI was proved to be independent in AA, AD, AP, AT, and RI by univariate and multivariate analyses in this paper, but when we brought all the AS events together, $p > 0.05$ in multivariate analyses (**Figure 8H**) which indicated that there were still some key factors that were not considered in our analysis and certain errors were inevitable due to the heterogeneity of patients.

In conclusion, our study created an efficient prognostic model based on survival-associated AS events for pRCC, which may

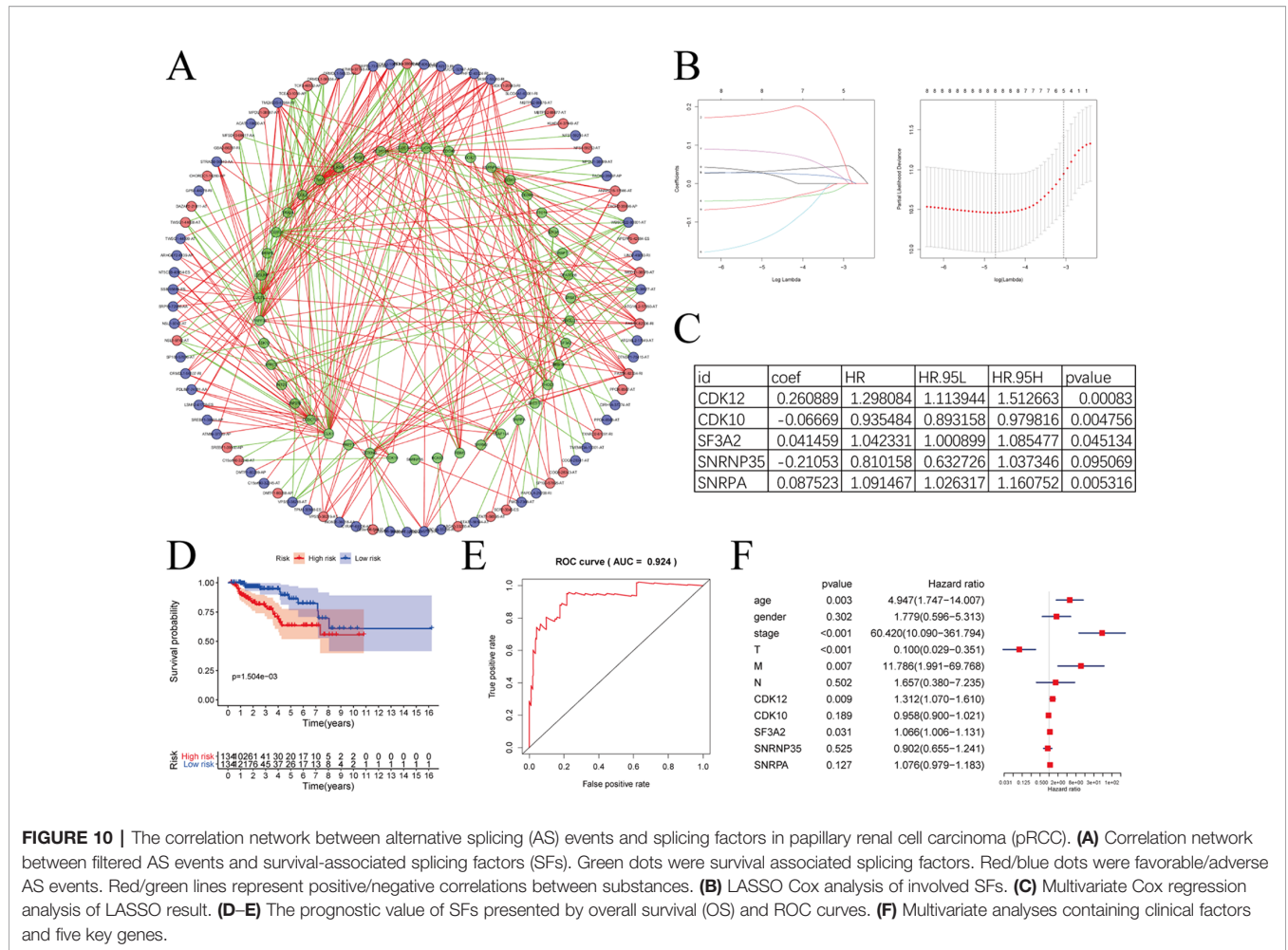


FIGURE 10 | The correlation network between alternative splicing (AS) events and splicing factors in papillary renal cell carcinoma (pRCC). **(A)** Correlation network between filtered AS events and survival-associated splicing factors (SFs). Green dots were survival associated splicing factors. Red/blue dots were favorable/adverse AS events. Red/green lines represent positive/negative correlations between substances. **(B)** LASSO Cox analysis of involved SFs. **(C)** Multivariate Cox regression analysis of LASSO result. **(D-E)** The prognostic value of SFs presented by overall survival (OS) and ROC curves. **(F)** Multivariate analyses containing clinical factors and five key genes.

help clinicians in selecting reliable prognostic indicators and understanding the mechanism of pRCC.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: <https://portal.gdc.cancer.gov/repository>. AS files could be downloaded from TCGASpliceSeq: <https://bioinformatics.mdanderson.org/TCGASpliceSeq/>.

AUTHOR CONTRIBUTIONS

Conception and design: XX and CC. Collection and assembly of data: ZW and JL. Data analysis and interpretation: RS and DC.

REFERENCES

Agrawal, A. A., Yu, L., Smith, P. G., and Buonomi, S. (2018). Targeting splicing abnormalities in cancer. *Curr. Opin. Genet. Dev.* 48, 67-74. doi: 10.1016/j.gde.2017.10.010

Anticancer Sulfonamides Induce Splicing Factor RBM39 Degradation. (2017). *Cancer Discovery*. doi: 10.1158/2159-8290.CD-RW2017-08

Manuscript writing: ZW and KW. Final approval of manuscript: All authors.

ACKNOWLEDGMENTS

The authors would like to thank the TCGA databases and TCGASpliceseq for the availability of the data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2019.01333/full#supplementary-material>

Armstrong, A. J., Halabi, S., Eisen, T., Broderick, S., Stadler, W. M., Jones, R. J., et al. (2016). Everolimus versus sunitinib for patients with metastatic non-clear cell renal cell carcinoma (ASPEN): a multicentre, open-label, randomised phase 2 trial. *Lancet Oncol.* 17 (3), 378-388. doi: 10.1016/S1470-2045(15)00515-X

Bowler, E., Porazinski, S., Uzor, S., Thibault, P., Durand, M., Lapointe, E., et al. (2018). Hypoxia leads to significant changes in alternative splicing and elevated

- expression of CLK splice factor kinases in PC3 prostate cancer cells. *BMC Cancer* 18 (1), 355. doi: 10.1186/s12885-018-4227-7
- CDK12 Changes Telling in Prostate Cancer. (2018). *Cancer Discovery* 8, 1055. doi: 10.1158/2159-8290.CD-NB2018-093
- Chen, H. H., Wang, Y. C., and Fann, M. J. (2006). Identification and characterization of the CDK12/cyclin L1 complex involved in alternative splicing regulation. *Mol. Cell Biol.* 26 (7), 2736–2745. doi: 10.1128/MCB.26.7.2736-2745.2006
- Conway, J. R., Lex, A., and Gehlenborg, N. (2017). UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* 33 (18), 2938–2940. doi: 10.1093/bioinformatics/btx364
- Delahunt, B., and Eble, J. N. (1997). Papillary renal cell carcinoma: a clinicopathologic and immunohistochemical study of 105 tumors. *Mod. Pathol.* 10 (6), 537–544.
- Frankish, A., Diekhans, M., Ferreira, A. M., Johnson, R., Jungreis, I., Loveland, J., et al. (2019). GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* 47 (D1), D766–D773. doi: 10.1093/nar/gky955
- Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* 33 (1), 1–22.
- Gao, L., Xie, Z. C., Pang, J. S., Li, T. T., and Chen, G. (2019). A novel alternative splicing-based prediction model for uteri corpus endometrial carcinoma. *Aging (Albany NY)* 11 (1), 263–283. doi: 10.18632/aging.101753
- Grosso, A. R., Martins, S., and Carmo-Fonseca, M. (2008). The emerging role of splicing factors in cancer. *EMBO Rep.* 9 (11), 1087–1093. doi: 10.1038/embor.2008.189
- Gu, Y., Lu, L., Wu, L., Chen, H., Zhu, W., and He, Y. (2017). Identification of prognostic genes in kidney renal clear cell carcinoma by RNAseq data analysis. *Mol. Med. Rep.* 15 (4), 1661–1667. doi: 10.3892/mmr.2017.6194
- Han, T., Goralski, M., Gaskill, N., Capota, E., Kim, J., Ting, T. C., et al. (2017). Anticancer sulfonamides target splicing by inducing RBM39 degradation via recruitment to DCAF15. *Science* 356. doi: 10.1126/science.aal3755
- Held-Feindt, J., Hattermann, K., Knerlich-Lukoschus, F., Mehdorn, H. M., and Mentlein, R. (2011). SP100 reduces malignancy of human glioma cells. *Int. J. Oncol.* 38 (4), 1023–1030. doi: 10.3892/ijo.2011.927
- Huang, D., Ding, Y., Luo, W. M., Bender, S., Qian, C. N., Kort, E., et al. (2008). Inhibition of MAPK kinase signaling pathways suppressed renal cell carcinoma growth and angiogenesis *in vivo*. *Cancer Res.* 68 (1), 81–88. doi: 10.1158/0008-5472.CAN-07-5311
- Huang, B., Fu, S. J., Fan, W. Z., Wang, Z. H., Chen, Z. B., Guo, S. J., et al. (2016). PKCepsilon inhibits isolation and stemness of side population cells via the suppression of ABCB1 transporter and PI3K/Akt, MAPK/ERK signaling in renal cell carcinoma cell line 769P. *Cancer Lett.* 376 (1), 148–154. doi: 10.1016/j.canlet.2016.03.041
- Jonasch, E., Gao, J., and Rathmell, W. K. (2014). Renal cell carcinoma. *BMJ* 349, g4797. doi: 10.1136/bmj.g4797
- Kahles, A., Lehmann, K. V., Toussaint, N. C., Huser, M., Stark, S. G., Sachsenberg, T., et al. (2018). Comprehensive analysis of alternative splicing across tumors from 8,705 patients. *Cancer Cell* 34 (2), 211–224. e216. doi: 10.1016/j.ccell.2018.07.001
- Li, J. K., Chen, C., Liu, J. Y., Shi, J. Z., Liu, S. P., Liu, B., et al. (2017). Long noncoding RNA MRCCAT1 promotes metastasis of clear cell renal cell carcinoma via inhibiting NPR3 and activating p38-MAPK signaling. *Mol. Cancer* 16 (1), 111. doi: 10.1186/s12943-017-0681-0
- Li, H. P., Peng, C. C., Wu, C. C., Chen, C. H., Shih, M. J., Huang, M. Y., et al. (2018). Inactivation of the tight junction gene CLDN11 by aberrant hypermethylation modulates tubulins polymerization and promotes cell migration in nasopharyngeal carcinoma. *J. Exp. Clin. Cancer Res.* 37 (1), 102. doi: 10.1186/s13046-018-0754-y
- Lin, P., He, R. Q., Huang, Z. G., Zhang, R., Wu, H. Y., Shi, L., et al. (2019). Role of global aberrant alternative splicing events in papillary thyroid cancer prognosis. *Aging (Albany NY)* 11 (7), 2082–2097. doi: 10.18632/aging.101902
- Malouf, G. G., Su, X., Zhang, J., Creighton, C. J., Ho, T. H., Lu, Y., et al. (2016). DNA methylation signature reveals cell ontogeny of renal cell carcinomas. *Clin. Cancer Res.* 22 (24), 6236–6246. doi: 10.1158/1078-0432.CCR-15-1217
- Martinez-Montiel, N., Rosas-Murrieta, N. H., Anaya Ruiz, M., Monjaraz-Guzman, E., and Martinez-Contreras, R. (2018). Alternative Splicing as a Target for Cancer Treatment. *Int. J. Mol. Sci.* 19 (2). doi: 10.3390/ijms19020545
- O'leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., Mcveigh, R., et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* 44 (D1), D733–D745. doi: 10.1093/nar/gkv1189
- Pal, S. K., Hoffman-Censits, J., Zheng, H., Kaiser, C., Tayama, D., and Bellmunt, J. (2018). Atezolizumab in platinum-treated locally advanced or metastatic urothelial carcinoma: clinical experience from an expanded access study in the United States. *Eur. Urol.* 73 (5), 800–806. doi: 10.1016/j.eururo.2018.02.010
- Perteau, M., and Salzberg, S. L. (2010). Between a chicken and a grape: estimating the number of human genes. *Genome Biol.* 11. doi: 10.1186/gb-2010-11-5-206
- Perteau, M., Shumate, A., Perteau, G., Varabyou, A., and Salzberg, S. (2018). Thousands of large-scale RNA sequencing experiments yield a comprehensive new human gene list and reveal extensive transcriptional noise. *Piva, F., Giuliotti, M., Burini, A. B., and Principato, G. (2012). SpliceAid 2: a database of human splicing factors expression data and RNA target motifs. Hum. Mutat.* 33 (1), 81–85. doi: 10.1002/humu.21609
- Robichaux, J. P., Elamin, Y. Y., Tan, Z., Carter, B. W., Zhang, S., Liu, S., et al. (2018). Mechanisms and clinical activity of an EGFR and HER2 exon 20-selective kinase inhibitor in non-small cell lung cancer. *Nat. Med.* 24 (5), 638–646. doi: 10.1038/s41591-018-0007-9
- Ryan, M., Wong, W. C., Brown, R., Akbani, R., Su, X., Broom, B., et al. (2016). TCGASpliceSeq a compendium of alternative mRNA splicing in cancer. *Nucleic Acids Res.* 44 (D1), D1018–D1022. doi: 10.1093/nar/gkv1288
- Siefried, Z., and Karni, R. (2018). The role of alternative splicing in cancer drug resistance. *Curr. Opin. Genet. Dev.* 48, 16–21. doi: 10.1016/j.gde.2017.10.001
- Singh, N. P., Bapi, R. S., and Vinod, P. K. (2018). Machine learning models to predict the progression from early to late stages of papillary renal cell carcinoma. *Comput. Biol. Med.* 100, 92–99. doi: 10.1016/j.combiomed.2018.06.030
- Tang, J. Y., Lee, J. C., Hou, M. F., Wang, C. L., Chen, C. C., Huang, H. W., et al. (2013). Alternative splicing for diseases, cancers, drugs, and databases. *ScientificWorldJournal* 2013, 703568. doi: 10.1155/2013/703568
- Teo, M. Y., Bambury, R. M., Zabor, E. C., Jordan, E., Al-Ahmadie, H., Boyd, M. E., et al. (2017). DNA damage response and repair gene alterations are associated with improved survival in patients with platinum-treated advanced urothelial carcinoma. *Clin. Cancer Res.* 23 (14), 3610–3618. doi: 10.1158/1078-0432.CCR-16-2520
- Vaughn, D. J., Srinivas, S., Stadler, W. M., Pili, R., Petrylak, D., Sternberg, C. N., et al. (2009). Vinflumine in platinum-pretreated patients with locally advanced or metastatic urothelial carcinoma: results of a large phase 2 study. *Cancer* 115 (18), 4110–4117. doi: 10.1002/ncr.24460
- Venables, J. P., Klinck, R., Koh, C., Gervais-Bird, J., Bramard, A., Inkel, L., et al. (2009). Cancer-associated regulation of alternative splicing. *Nat. Struct. Mol. Biol.* 16 (6), 670–676. doi: 10.1038/nsmb.1608
- Venables, J. P. (2006). Unbalanced alternative splicing and its significance in cancer. *Bioessays* 28, 378–386. doi: 10.1002/bies.20390
- Wach, S., Nolte, E., Theil, A., Stohr, C., T. T. R., Hartmann, A., et al. (2013). MicroRNA profiles classify papillary renal cell carcinoma subtypes. *Br. J. Cancer* 109 (3), 714–722. doi: 10.1038/bjc.2013.313
- Wan, L., Yu, W., Shen, E., Sun, W., Liu, Y., Kong, J., et al. (2019). SRSF6-regulated alternative splicing that promotes tumor progression offers a therapy target for colorectal cancer. *Gut* 68 (1), 118–129. doi: 10.1136/gutjnl-2017-314983
- Wang, E. T., Sandberg, R., Luo, S., Khrebtkova, I., Zhang, L., Mayr, C., et al. (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* 456 (7221), 470–476. doi: 10.1038/nature07509
- Wang, B. D., Ceniccola, K., Hwang, S., Andrawis, R., Horvath, A., Freedman, J. A., et al. (2017). Alternative splicing promotes tumour aggressiveness and drug resistance in African American prostate cancer. *Nat. Commun.* 8, 15921. doi: 10.1038/ncomms15921
- Wei, Y., Zou, Z., Becker, N., Anderson, M., Sumpter, R., Xiao, G., et al. (2013). EGFR-mediated Beclin 1 phosphorylation in autophagy suppression, tumor progression, and tumor chemoresistance. *Cell* 154 (6), 1269–1284. doi: 10.3892/ol.2019.10130
- Wei, W., Lv, Y., Gan, Z., Zhang, Y., Han, X., and Xu, Z. (2019). Identification of key genes involved in the metastasis of clear cell renal cell carcinoma. *Oncol. Lett.* 17 (5), 4321–4328. doi: 10.1016/j.cell.2013.08.015
- Yang, X. J., Tan, M. H., Kim, H. L., Ditlev, J. A., Betten, M. W., Png, C. E., et al. (2005). A molecular classification of papillary renal cell carcinoma. *Cancer Res.* 65 (13), 5628–5637. doi: 10.1158/0008-5472.CAN-05-0533

- Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16 (5), 284–287. doi: 10.1089/omi.2011.0118
- Zhang, J., and Manley, J. L. (2013). Misregulation of pre-mRNA alternative splicing in cancer. *Cancer Discovery* 3 (11), 1228–1237. doi: 10.1158/2159-8290.CD-13-0253
- Zhang, Y., Jiang, X., Qin, X., Ye, D., Yi, Z., Liu, M., et al. (2010). RKTG inhibits angiogenesis by suppressing MAPK-mediated autocrine VEGF signaling and is downregulated in clear-cell renal cell carcinoma. *Oncogene* 29 (39), 5404–5415. doi: 10.1038/onc.2010.270
- Zhang, Y., Zhao, H., Xu, W., Jiang, D., Huang, L., and Li, L. (2019). High expression of PQBP1 and low expression of PCK2 are associated with metastasis and recurrence of osteosarcoma and unfavorable survival outcomes of the patients. *J. Cancer* 10(9), 2091–2101. doi: 10.7150/jca.28480
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Copyright © 2020 Wu, Liu, Sun, Chen, Wang, Cao and Xu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*