

# The landscape of alternative splicing reveals novel events associated with tumorigenesis and the immune microenvironment in gastric cancer

Shenghan Lou<sup>1,\*</sup>, Jian Zhang<sup>2,\*</sup>, Zhao Zhai<sup>1</sup>, Xin Yin<sup>1</sup>, Yimin Wang<sup>1</sup>, Tianyi Fang<sup>1</sup>, Yingwei Xue<sup>1</sup>

<sup>1</sup>Department of Gastroenterological Surgery, Harbin Medical University Cancer Hospital, 150081 Harbin, Heilongjiang Province, China

<sup>2</sup>Department of Thoracic Surgery, Harbin Medical University Cancer Hospital, 150081 Harbin, Heilongjiang Province, China

\*Equal contribution

**Correspondence to:** Yingwei Xue; **email:** [xueyingwei@hrbmu.edu.cn](mailto:xueyingwei@hrbmu.edu.cn)

**Keywords:** gastric cancer, alternative splicing, RNA-seq profile, tumorigenesis, immune microenvironment

**Abbreviations:** GC: gastric cancer; AS: alternative splicing; SF: splicing factor; OS: overall survival; TIDE: tumor immune dysfunction and exclusion

**Received:** August 2, 2020

**Accepted:** October 22, 2020

**Published:** January 10, 2021

**Copyright:** © 2021 Lou et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/3.0/) (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## ABSTRACT

Alternative splicing (AS), contributing to vast protein diversity from a rather limited number of genes in eukaryotic transcripts, has emerged as an important signature for tumor initiation and progression. However, a systematic understanding of its functional impact and relevance to gastric cancer (GC) tumorigenesis is lacking. Differentially expressed AS (DEAS) was verified among GC-associated AS events based on RNA-seq profiles from the TCGA database. Functional enrichment analysis, unsupervised clustering analysis and prognostic models were used to infer the potential roles of DEAS events and their molecular, clinical and immune features. In total, 12,225 AS events were detected from 5,199 genes, among which 314 AS events were identified as DEAS events in GC. The parental genes of the DEAS events were significantly enriched in the regulation of GC-related processes. The splicing correlation network suggested a significant relationship between DEAS events and splicing factors (SFs). Three clusters of DEAS events were identified to be different in prognosis, cancer-specific signatures and immune features between distinct clusters. Univariate and multivariate analyses regarded 3 DEAS events as independent prognostic indicators. Profiling of the AS landscape in GC elucidated the functional roles of the splicing network in GC and might serve as a novel prognostic indicator and therapeutic target.

## INTRODUCTION

With advances in the early screening, diagnosis and treatment of gastric cancer (GC), its annual incidence and mortality have decreased [1]. However, GC is still the fifth most frequently diagnosed cancer and the third leading cause of cancer-related mortality, with almost 1,000,000 new cases and 800,000 deaths each year [2, 3]. Due to indistinctive symptoms, GC often exhibits proliferation, extensive invasion and lymphatic metastasis at the time of diagnosis,

contributing to its poor prognosis [4]. The pathogenesis of GC is complex, as it is controlled by genetic alterations in oncogenes and suppressor genes, and these alterations result in disease heterogeneity [4]. Despite advances in the molecular characteristics of GC, the prospect of tailored individual therapy based on histological and molecular subtypes is still unsatisfactory. The overall 5-year relative survival rate of GC patients is still quite low [5]. Therefore, it is worthwhile to explore the novel biological indicators and molecular mechanisms of GC.

Cancer profiling has allowed more dimensionalities due to advances in the depth and quality of transcriptome sequencing in the era of cancer genomics [6]. Alternative splicing (AS) is one of the most crucial posttranscriptional regulatory mechanisms and modifies more than 90% of human genes [7]. AS significantly enriches protein diversity by generating different RNA isoforms of single genes [8]. AS may result in a series of consequences, such as changing the stability of proteins, adding or deleting structural domains and modifying the interactive relationship between proteins [9]. Abnormal AS events participate in several tumorigenic processes, such as proliferation, apoptosis, angiogenesis and metastasis [10]. The initial unbalanced expression of splice isoforms or the accumulation of incorrect isoforms was believed to be associated with one of the cancer hallmarks, genetic instability and mutation, summarized by Douglas Hanahan and Robert A. Weinberg [11, 12]. Moreover, growing evidence indicates that cancer-specific AS events possess potential applications in cancer therapy, serving as prognostic biomarkers and even therapeutic targets [9, 13].

Here, we repurposed and integrated multi-RNA-seq analysis among entries in the gastric tissue cohort from The Cancer Genome Atlas (TCGA) database to comprehensively analyze AS events. We systematically profiled the genome-wide AS events in GC and identified GC-related AS events. We also explored the potential biological function and underlying regulatory mechanisms of these specific cancer-related AS events in GC. Distinct clusters of GC-related AS events were identified, and the association between the distinct clusters and clinical and immune features was investigated. Finally, we performed survival analyses to identify the prognostic value of AS events.

## RESULTS

### Overview of AS events in the TCGA STAD cohort

The large-scale genome RNA-seq data of 305 stomach adenocarcinoma (STAD) patients from the TCGA database were analyzed, and a total of 12,225 AS events were identified from 5,199 genes. The included population comprised 846 alternate acceptor (AA) site events from 702 genes, 919 alternate donor (AD) site events from 727 genes, 3,137 alternate promoter (AP) events from 1,531 genes, 1,881 alternate terminator (AT) events from 948 genes, 4,417 exon skip (ES) events from 2,465 genes, 86 mutually exclusive exon (ME) events from 82 genes and 939 retained intron (RI) events from 698 genes (Figure 1A). According to the prioritization of proportion among AS events, ES events accounted for 36.13% of all AS events, whereas ME

events accounted for 0.7% of all AS events. Intriguingly, a considerable proportion of genes contained two or more AS events, and 5 different AS events occurred in one single gene (Figure 1B).

### Identification of DEAS events in GC

To identify DEAS events in GC, we compared the percent spliced in (PSI) values of GC and adjacent normal tissues from the TCGA database. After screening, a total of 314 DEAS events from 280 genes were identified. Among these DEAS events, 175 DEAS events from 161 genes were upregulated, and 139 DEAS events from 127 genes were downregulated (Figure 1C, Supplementary Table 1). Intriguingly, GC and normal tissues were clearly separated into two discrete groups using unsupervised hierarchical clustering based on these DEAS events, but unfortunately, two GC tissues were misclassified as normal tissues (Figure 1D). There was an uneven distribution of splicing modes between DEAS events and all AS events (Figure 1B, 1E). Most genes had only one AS event, whereas some genes had up to two different splicing modes (Figure 1E).

Notably, some genes, such as MTMR11, CCL14, HDAC9 and CHN2, showed the opposite trends for the same splicing mode of the parental gene in GC and normal tissues (Figure 2A). Since aberrant AS events might directly affect the expression of parental genes, the relationship between DEAS events and differentially expressed genes (DEGs) was further assessed. By comparing 314 DEAS events from 280 parental genes between GC and normal tissues, a total of 79 DEGs involving 97 DEAS events were observed (Figure 2B, Supplementary Table 2). Spearman's correlation analysis indicated that 69 of 97 (71.13%) DEAS events were significantly correlated with the expression of their parental gene ( $|R| \geq 0.4$  and adjusted  $P < 0.05$ ). Moreover, almost half of these 69 DEAS events were AP events (44.93%) (Supplementary Figure 1, Supplementary Table 3).

### Enrichment and interaction analyses of DEAS events

The potential influence of DEAS events on the parental genes was then assessed with biological function enrichment analysis. The results revealed that certain GO categories, such as protein localization to the cell periphery, cell junction assembly, cell adhesion molecular binding and cadherin binding, were enriched in these 280 parental genes, which are closely associated with GC development (Figure 3A, Supplementary Table 4). Furthermore, certain vital KEGG pathways, such as proteoglycans in cancer, adrenergic signaling in cardiomyocytes, the AMPK

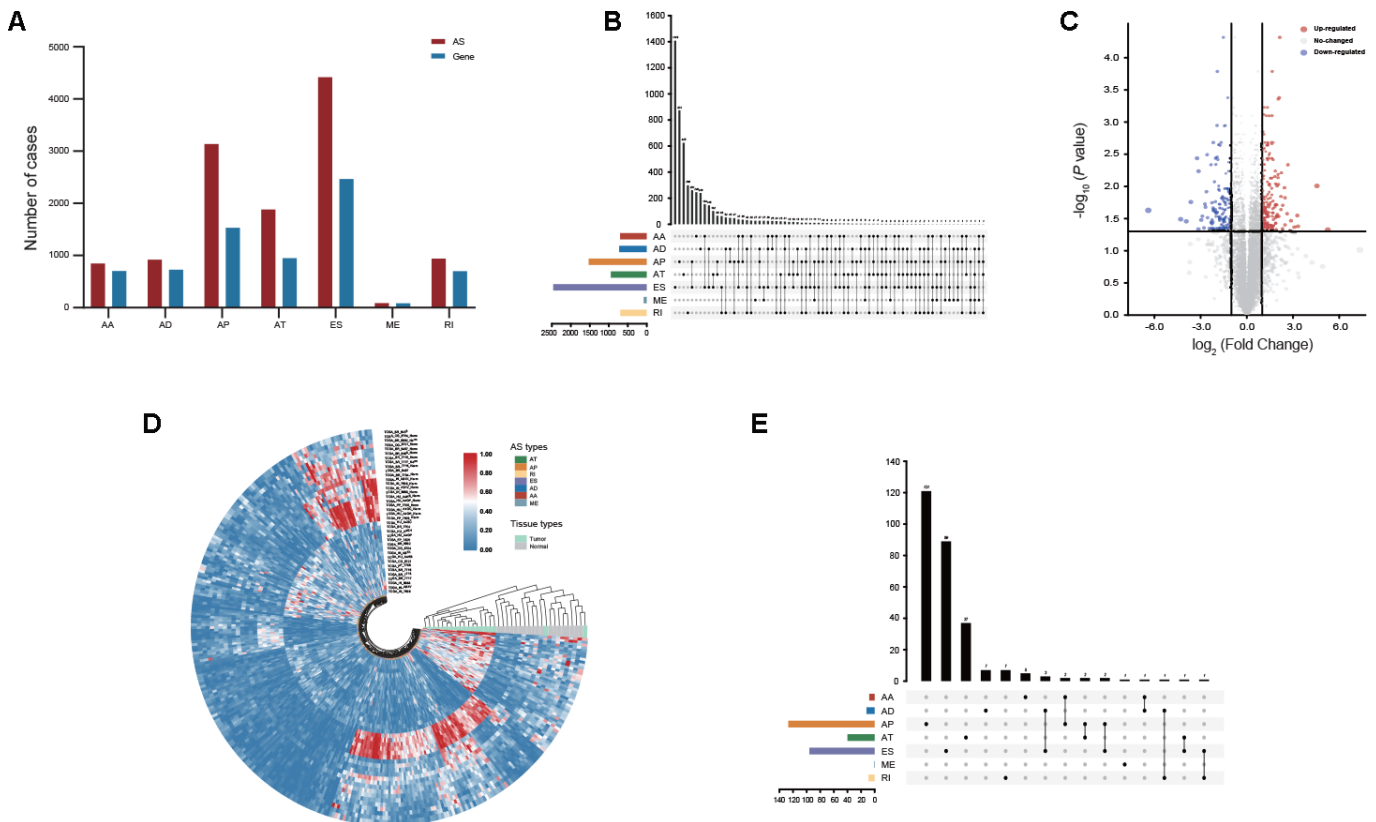
signaling pathway, the HIF-1 signaling pathway and the TNF signaling pathway, were also enriched (Figure 3B, Supplementary Table 5). Consistent with these findings, gene set enrichment analysis (GSEA) revealed that DEAS events in GC were significantly enriched in actin binding, adrenergic signaling in cardiomyocytes and the AMPK signaling pathway (Figure 3C, 3D, Supplementary Tables 6, 7).

Since AS events could inevitably affect protein functions, a protein-protein interaction (PPI) network of the parental genes of DEAS events was constructed to reveal the potential influence of DEAS events at the protein level (Figure 4A). Moreover, hub genes and individual modules were further identified based on the PPI network. Among the PPI network, we found that the top 5 hub genes were RPS6, RPL32, RPL18A, RPS21 and RPS3A (Figure 4B). We also identified two individual modules (Figure 4C, 4D). Consistent with our results from the enrichment analysis, DEAS events in module 1 were enriched in polysomes (GO: 0005844) and ribosomes (KEGG: hsa03010), suggesting that the

function of ribosomes might be frequently affected by DEAS events in GC (Supplementary Table 8). DEAS events in module 2 were enriched in adhesion junction (KEGG: hsa04520), leukocyte transendothelial migration (KEGG: hsa04670) and several other GO terms (Supplementary Table 9).

### Network of DEAS events and SFs

Alternative splicing patterns can be regulated by many types of RNA-binding proteins (RBPs), which are known as splicing factors (SFs) [14]. Among these special proteins, serine-arginine-rich splicing factor (SRSF) and heterogeneous nuclear ribonucleoprotein (hnRNP) are two well-known families of splicing regulatory factors [15]. To explore the potential regulatory relationship between DEAS events and SFs in GC, correlations between the PSI value of 314 DEAS events and the expression of SFs were analyzed in the TCGA STAD cohort. A total of 170 DEAS events were significantly associated with 44 SFs ( $|R| \geq 0.4$  and adjusted  $P < 0.05$ ) (Figure 5A, Supplementary Table 10).



**Figure 1. Profiling of AS events identified in GC.** (A) The number of AS events and their parental genes derived from GC patients was counted based on the AS types. (B) Interactive analysis of seven types of AS events derived from GC patients is shown in an UpSet plot. (C) DEAS events between GC and adjacent normal tissues were visualized in a volcano plot. (D) Distinct DEAS events between GC and adjacent normal tissues were clustered and visualized in sector plots. (E) Interactive analysis of DEAS events between GC and adjacent normal tissues is shown in an UpSet plot.

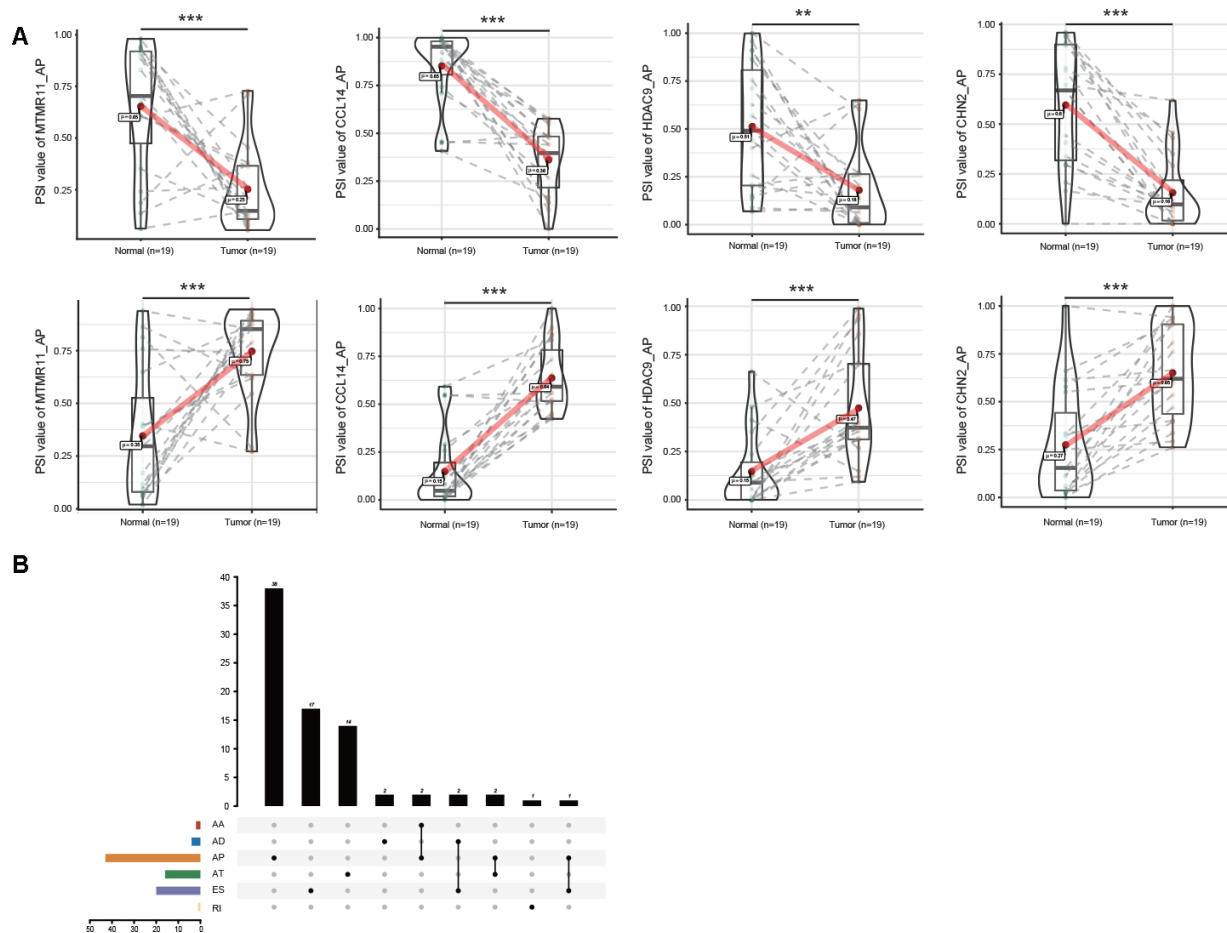
In the SF-associated regulatory network, we observed that most SFs were significantly correlated with more than one DEAS event and that one DEAS event could be regulated by 13 different SFs, all of which characterize the comprehensive regulatory network of cooperative or competitive relationships between DEAS events and SFs. For example, RBM24 showed a positive correlation with LRRFIP1\_ES and SEPT9\_AP (Figure 5B, 5C). On the other hand, RBM24 showed a negative correlation with DIXDC1\_AP and FLNA\_ES (Figure 5D, 5E). Other representative correlations between DEAS events and SFs are presented in scatter plots (Supplementary Figure 2).

### DEAS-based cluster construction and correlations with molecular, clinical and immune features

Because of molecular heterogeneity, GC can be clustered into molecular subtypes based on the distinct expression patterns of genes or proteins [4].

Here, consensus unsupervised analysis was performed on distinct clusters of GC patients according to the DEAS events that varied at the individual level. As shown in the consensus matrix heatmap, 3 clusters of GC patients were identified: cluster 1 (n = 154, 50.5%), cluster 2 (n = 100, 32.78%) and cluster 3 (n = 51, 16.72%) (Figure 6A). Consistent with the consensus clustering results, principal component analysis (PCA) also discerned significant differences among clusters 1, 2 and 3 (Figure 6B). In addition, a gene set variation analysis (GSVA) revealed a set of cancer-specific signatures that differed between the DEAS-based clusters (Figure 6C, Supplementary Table 11). This finding suggested that different clusters were associated with different biological processes, which further validated the reliability of our DEAS-based clusters.

To study the clinical implications of the DEAS-based clusters, univariate survival analysis was conducted to



**Figure 2. Profiling of DEAS EVs identified in GC. (A)** The opposite trend of representative DEGs between GC and adjacent normal tissues is shown in violin plots. Some AS events of these representative DEGs were upregulated in GC tissues (upper). The same AS events of these representative DEGs were downregulated in GC tissues for the different splicing sites (lower). **(B)** Interactive analysis of DEGs between GC and adjacent normal tissues is shown in an UpSet plot.

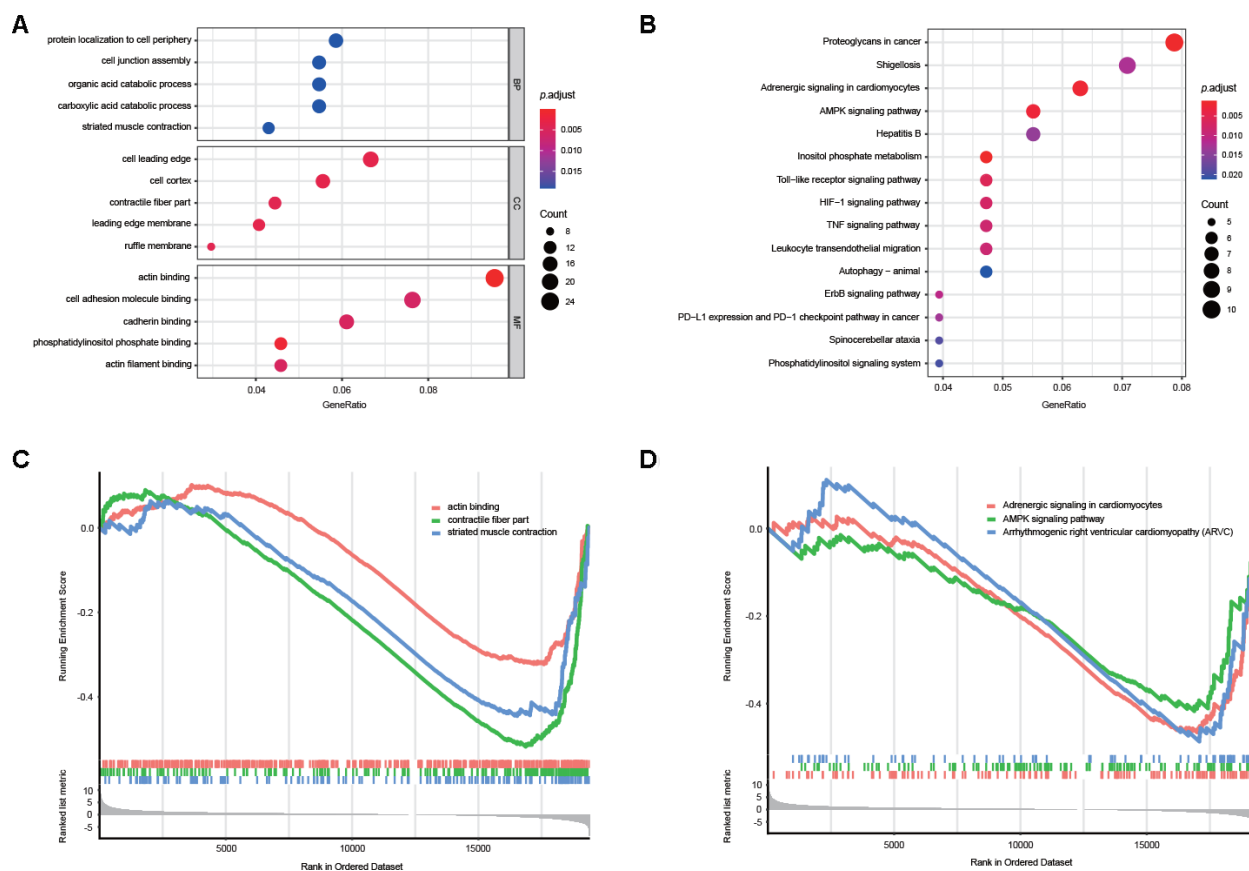
assess the relationship between clusters and survival outcomes. We found that the survival outcome was not randomly distributed across different clusters (Figure 6D). GC patients belonging to cluster 1 had a better prognosis than those belonging to clusters 2 and 3 (Figure 6D). We also assessed the relationship between DEAS-based clusters and clinicopathological factors. The distributions of T stage, histologic grade and microsatellite status among the clusters were significantly different (Figure 7A, Supplementary Table 12).

To comprehensively characterize the immune features based on DEAS events, we investigated differences in the immune microenvironment and found that both the immune and stromal scores were significantly different between the DEAS-based clusters (Figure 7B, 7C). Notably, cluster 1, which was associated with a better prognosis, was associated with lower immune and stromal scores than clusters 2 and 3 (Figure 7B). Moreover, our studies of immune cell infiltration revealed that many types of immune cells were not randomly distributed across the different clusters (Figure 7A, 7C, Supplementary Table 12). Cluster 1 had

a significantly lower proportion of non-immunotherapy-associated cells, such as activated mast cells, and a higher proportion of immunotherapy-associated cells, such as T cell follicular helper cells, compared with clusters 2 and 3 (Figure 7A, 7C Supplementary Table 12). These results provide further insight into immunotherapy, and the tumor immune dysfunction and exclusion (TIDE) algorithm was used to predict the likelihood of a response to immunotherapy. These results suggested that the DEAS-based clusters were correlated with different immune responses (Figure 7A, Supplementary Table 12). Patients in cluster 1 (50.64%, 78/154) presented a higher proportion of immunotherapy-associated cells and a lower proportion of stromal cells; thus, they might be more likely to respond to immunotherapy than patients in cluster 2 (27%, 27/100) and cluster 3 (13.72%, 7/51).

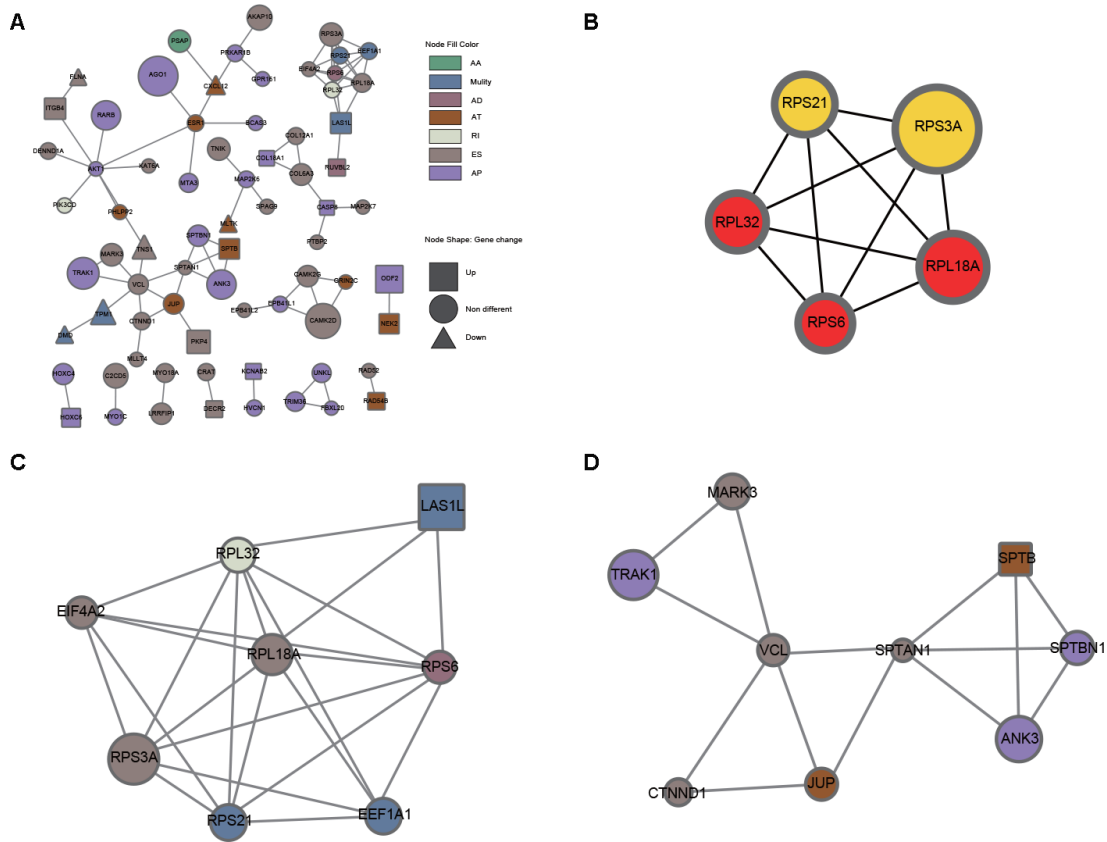
### Prognostic value of DEAS events in GC

To identify the potential clinical value of DEAS events, we investigated the underlying relationship between DEAS events and the prognosis of patients with GC. Among the 314 DEAS events, univariate Cox

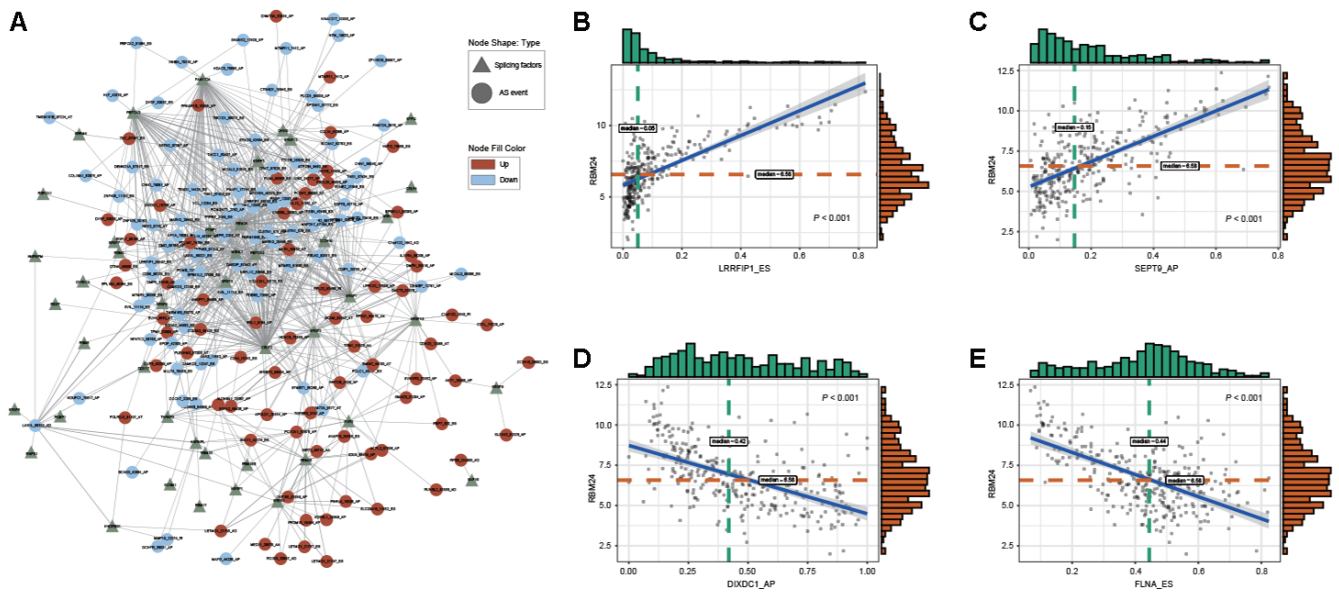


**Figure 3. Potential biological functions of DEAS events.** (A) GO analysis of DEAS events is shown in bubble plots. (B) KEGG analysis of DEAS events is shown in bubble plots. (C, D) GSEA of DEAS events.





**Figure 4. PPI analysis of DEAS events.** (A) The interactome of 280 parental genes for 314 DEAS events is shown in the PPI network. (B) The interactome of the top 5 hub genes. (C) Module 1 was associated with ribosomes. (D) Module 2 was associated with adherens junctions and migration.



**Figure 5. Regulatory splicing correlation network in GC.** (A) The correlation of DEAS events with SFs is shown in network plots. (B, C) Representative positive correlations between DEAS events and SFs are shown in scatter plots. (D, E) Representative negative correlations between DEAS events and SFs are shown in scatter plots.

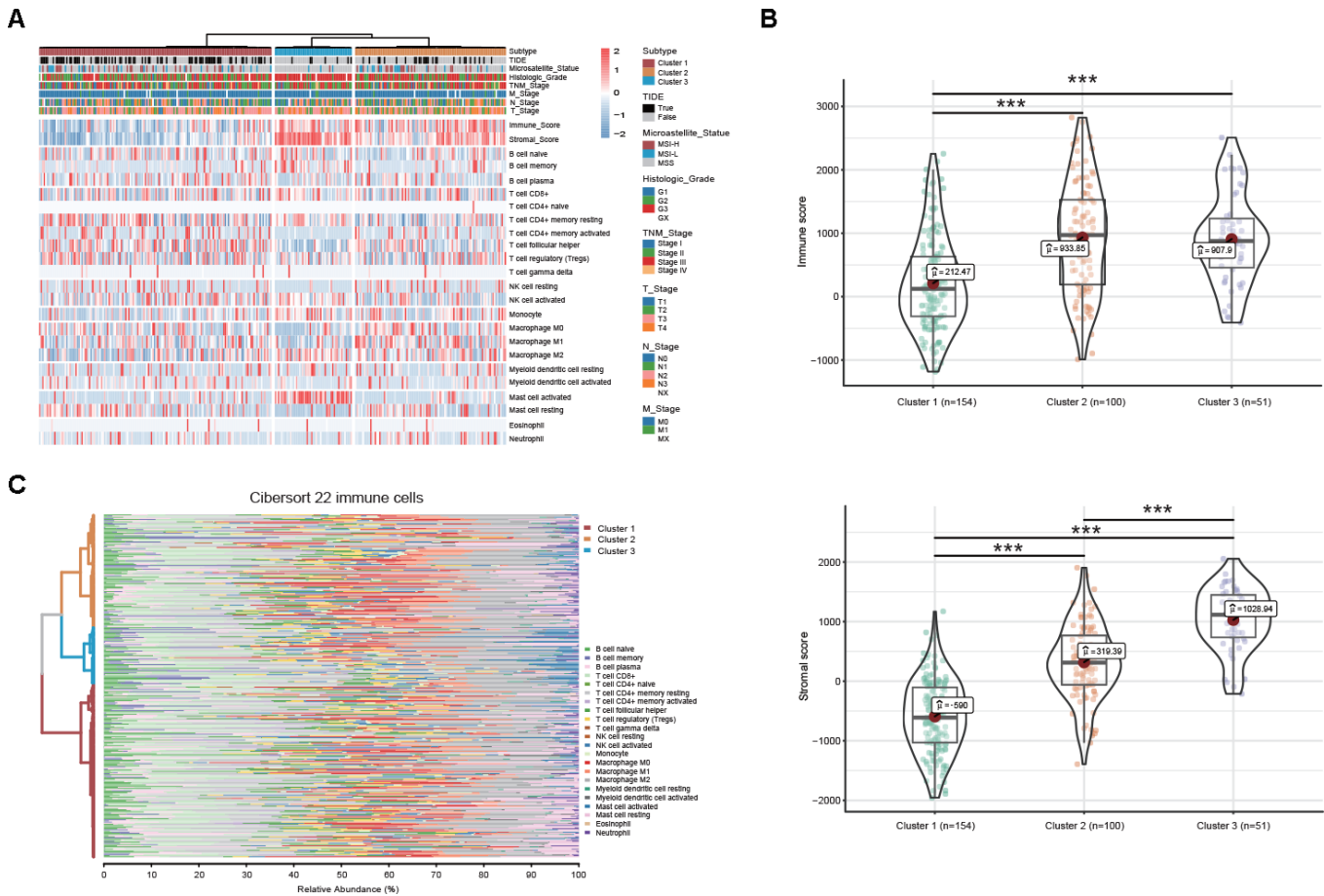


has also proven that the specific dysregulation of AS events plays crucial roles in the initiation and progression of GC [16]. For instance, CD44, MUTYH and WNT2B splice variants participate in the metastasis, carcinogenesis and development of GC [17–19]. Because of the limited sample sizes of one or a few specific AS events in previous studies, the exploration of AS events in GC is far from comprehensive. Hence, we systematically profiled AS events in a large-scale GC cohort to elucidate the landscape of AS events in GC.

Through a rigorous filter, a total of 12,225 AS events of 5,199 genes were identified in 305 GC patients, indicating that AS is a common posttranscriptional modification in GC. Moreover, 314 DEAS events were detected from 280 genes among the GC and normal tissues. The correlation analysis indicated that most of the DEAS events were significantly correlated with the expression of their parental gene, consistent with the

hypothesis that AS acts as an important segment of the posttranscriptional process and alters gene expression [20]. All the above experimentally validated splice variants were also identified by our procedure, suggesting that our results are reliable and that the DEAS events identified in our study are ubiquitous in GC. In addition, we found that GC shares some common DEAS events with those in colorectal and head-neck squamous cell cancers, indicating that certain AS events are common in cancer tumorigenesis and development [21, 22].

To explore the potential mechanism of DEAS events in GC, we performed functional enrichment analysis and generated a PPI network of their parental genes. In the functional enrichment analysis, the parental genes were mainly enriched in pathways related to the cytoplasm and extracellular matrix, such as “cell junction organization and assembly”, “cell adhesion molecular binding and cadherin binding” and “AMPK signaling

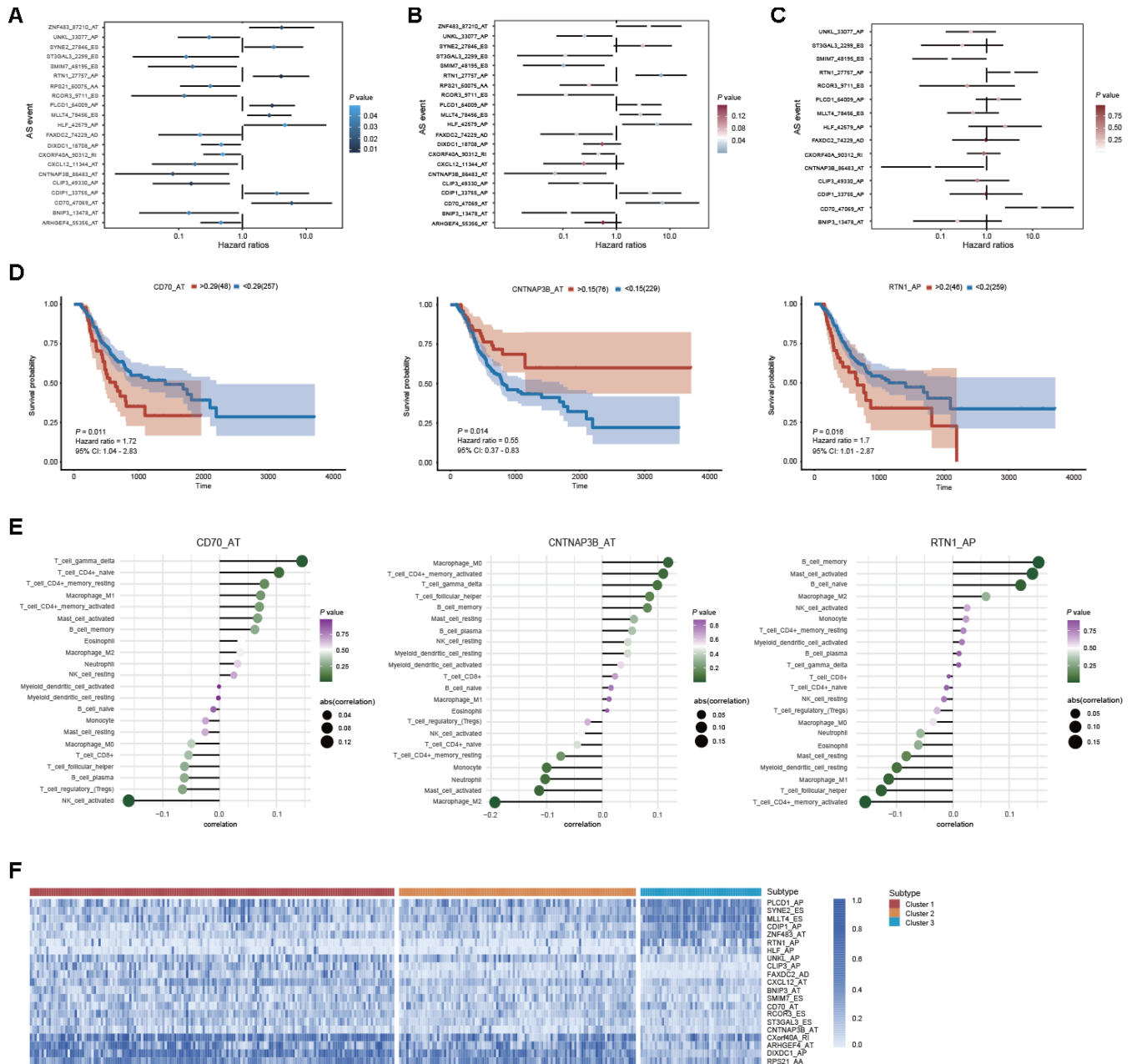


**Figure 7. DEAS-based clusters associated with clinicopathological characteristics and immune microenvironment features.** (A) A total of 305 DEAS events ordered by distinct clusters with annotations associated with clinicopathological characteristics and immune microenvironment features were visualized in a matrix heatmap. (B) Immune and stromal scores of each DEAS-based cluster. (C) Percentage matrix heatmap of immune cell infiltration in the tumor microenvironment between distinct clusters.



pathway". The cytoplasm is the major location of AMPK, and the AMPK signaling pathway is considered a vital factor in the development and treatment of GC [23–26]. Abnormal cell adhesion and cell junctions are also associated with the progression and metastasis of GC, which further supports the accuracy and reliability of our enrichment analysis [27–29]. Correspondingly,

the DEAS events influenced the above potential pathophysiological processes in GC. By constructing the interaction network of the parental genes, we found that most were correlated with each other. Two individual modules were highlighted in the PPI network, suggesting potential molecular complexes. Of note, all top 5 hub genes belonging to the ribosomal



**Figure 8. Prognostic value of DEAS events in GC.** (A) Univariate analysis of DEAS events associated with OS is shown in forest plots of hazard ratios. (B) Univariate analysis of DEAS events associated with OS after adjusting for relative clinical covariates is shown in forest plots of hazard ratios. (C) Multivariate analysis of DEAS events associated with OS is shown in forest plots of hazard ratios. (D) Kaplan-Meier survival analysis of the independent prognostic DEAS events on OS. (E) Correlation analysis of the independent prognostic DEAS events and immune cells. (F) Differential expression of representative prognostic DEAS events between distinct clusters was visualized in a matrix heatmap.

protein family comprised one individual module, which implies that the functions of ribosomes might be affected by DEAS events in GC.

SFs have been reported to participate in the precise regulation of RNA splicing by binding to specific RNA sequences [30]. Thus, we performed an integrated analysis of SFs and DEAS events to address the underlying mechanism of the splicing pathway in GC. The splicing correlation network showed distinguished interactions between DEAS events and SFs. Of note, we found that the same SF might play dual roles in the regulation of AS events and that the same AS event could be synergistically or antagonistically regulated by different SFs. This suggests that complex regulatory mechanisms may influence SFs and AS events. Indeed, the real regulatory mechanisms involved in RNA splicing could be more sophisticated than we revealed in this study. The relationship between DEAS events and SFs should be considered a dynamic interaction network instead of a simple “one to one” pattern.

Increasing evidence has demonstrated that AS events play an indispensable role in immune micro-environment formation [13]. Actual identification of the distinct patterns of AS events in the tumor microenvironment will contribute to GC therapy. Here, we revealed three distinct DEAS-based patterns among 314 DEAS events, and these three patterns presented distinct molecular, clinical and immune features. We found that most of the cancer hallmark pathways were significantly different between the 3 DEAS-based clusters and that pathogenetic pathways were broadly different regarding tumor biology. This further indicates that the DEAS-based clusters identified herein are reliable. According to our results, GC patients belonging to cluster 1 were significantly associated with a low histologic grade, a low T stage and a favorable survival outcome.

Moreover, cluster 1 was characterized by high tumor purity and was significantly associated with low immune and stromal scores. Consistent with our results, previous studies have also determined that low tumor purity is associated with a poor prognosis and an enhanced immune phenotype [31–33]. Referring to the amount of immune cell infiltration, the relative proportion of M2 macrophages and neutrophils is substantially inversely correlated with tumor purity [31–33]. Consistent with this finding, we found that the relative proportion of multiple types of immune cells was significantly low in cluster 1. According to the TIDE algorithm, the low proportion of immune cell types indicated that tumor purity was an essential factor of immunotherapy in GC. In addition, cluster 1

presented a higher proportion of immunotherapy-associated cells and a lower proportion of stromal cells, which was beneficial for immunotherapy. GC patients belonging to cluster 1 might obtain a better response than patients in other clusters. These findings suggest that the DEAS events in GC confer essential biological and clinical implications.

Due to the potential significance of AS events in tumor biology, their clinical relevance in GC was further assessed by survival analysis. A total of 21 DEAS events were found to be associated with OS. Among these survival-associated DEAS events, some genes were determined to play crucial roles in tumor biology. For instance, *PLCD1* inhibits tumor formation in breast cancer by inducing apoptosis [34]. *HLF* transactivates *c-Jun* to promote tumor-initiating cell generation in hepatocellular carcinoma [35]. The inactivation of *BNIP3* likely plays a vital role in the progression of colorectal cancer and GC [36]. Moreover, 15 of 21 survival-associated DEAS events also showed a significant association with OS, even after adjusting for multiple clinical factors, including sex, age and TNM stage. Finally, 3 of 21 survival-associated DEAS events were revealed as independent prognostic factors, and these 3 DEAS events performed well in stratifying GC patients into groups based on survival through Kaplan-Meier survival curves. In addition, each of these 3 DEAS events showed a significant association with several types of immune cells. The AS landscape profile in GC was mapped, thus providing an overview of this research, and the findings elucidated the functional roles of the splicing network in GC. Unfortunately, a limited number of validation experiments were performed; however, we intend to perform additional validation experiments in the future.

Our study depicted a comprehensive landscape of AS events in GC patients, and the implementation of strict criteria ensured the identification of abnormal AS events related to GC. Our functional analysis indicated that 314 DEAS events identified in our study might play essential roles in the development of GC. The SF-AS regulatory network further clarified the underlying mechanism of the splicing-associated pathway. Moreover, based on the DEAS events, the comprehensive clustering analysis of GC revealed the intrinsic relevance of molecular alterations and immune features, which indicated the value of AS events in predicting the clinical outcomes of GC patients. In addition, survival-related DEAS events might not only serve as prognostic indicators and therapeutic targets for GC patients but also help to decipher the mechanism of AS events in GC oncogenesis.

## MATERIALS AND METHODS

### Data acquisition and curation

Patients who met the following criteria were included in the TCGA GC cohort: (1) patients who were histologically confirmed as having primary GC; (2) patients with RNA-seq data; (3) patients with alternative RNA splicing data; (4) patients with detailed clinicopathological and follow-up information, including sex, age, race, TNM stage, histologic grade, microsatellite status and overall survival (OS) status; and (5) patients with an OS time of over 90 days. The corresponding RNA-seq data and clinical information of the GC cohort were downloaded from the TCGA data portal using the “GDCRNATools” package [37]. The corresponding alternative RNA splicing data of the GC cohort were downloaded from the TCGA SpliceSeq dataset [38]. Splicing events in the dataset were divided into seven categories: ES, RI, AP, AT, AD, AA and ME. Each splicing event was quantified by the PSI value [39], which ranges from 0 to 1 and represents the ratio of normalized read counts to indicate the inclusion of a transcript element over the total normalized reads for that event. To generate a reliable set of AS events, we implemented a series of stringent filters, which included “percentage of samples with a PSI value  $\geq 75\%$ ”, “average PSI value  $\geq 0.05$ ” and “standard deviation of the PSI value  $\geq 0.1$ ”. Only AS events meeting the above criteria were included in the present analysis. Moreover, we filled in the missing PSI values using the k nearest neighbors algorithm [40].

### DEAS event identification and functional analysis

To identify the DEAS events in GC, we applied the paired sample *t*-test to compare the PSI values between tumor tissues and matched adjacent normal tissues, and the *P*-value was adjusted by the Benjamini-Hochberg (BH) method. AS events with an absolute value of  $\log_2$ -fold change  $\geq 1$  and an adjusted *P*-value  $< 0.05$  were considered statistically significant. Then, we subjected the parental genes of the DEAS events to a biological function enrichment analysis, which was performed with the “clusterProfiler” package [41]. Gene Ontology (GO) terms, including “biological process (BP)”, “cellular component (CC)” and “molecular function (MF)”, and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways with adjusted *P* values  $< 0.05$  were selected for further analysis. In addition, the parental genes of these DEAS events were mapped to coding proteins, and the PPIs between these coding proteins were downloaded from the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database. A minimum required interaction score of 0.9

was used to identify reliable interaction results. The PPI network was further visualized with Cytoscape [42]. Hub genes and specific modules of the PPI network were identified by CytoHubba and the Molecular Complex Detection (MCODE) plug-in of Cytoscape [43, 44].

### SFs and splicing correlation network

A total of 78 genes that participated in the process of alternative RNA splicing (GO: 0000380) were obtained from the Molecular Signatures Database (MSigDB) [45]. The raw count value of these splicing genes was then extracted from the RNA-seq data and normalized by trimmed means of M (TNM) values [46]. We implemented the “voom” method to further transform the raw counts [47]. Spearman’s correlation analysis was conducted to explore the correlations between the expression of the splicing genes and the PSI values of the DEAS events. The *P*-value was adjusted by the BH method, and a correlation coefficient  $\geq 0.4$  and an adjusted *P*-value  $< 0.05$  were considered statistically significant. The correlation network of the DEAS events and splicing genes was visualized with Cytoscape [42].

### Cluster analysis and correlation with molecular, clinical and immune features

Based on the identified DEAS events ( $n = 314$ ), unsupervised clustering of the TCGA STAD cohort was performed via hierarchical consensus clustering with the “ConsensusClusterPlus” package [48]. Associations between clusters and clinicopathological variables (pathological T stage, pathological N stage, pathological M stage, pathological TNM stage, histologic grade and microsatellite status), survival status (OS), immune features (immune score, stromal score and immune cell infiltration) and immunotherapy responses between clusters were analyzed. Immune and stromal scores were calculated based on the ESTIMATE algorithm [49]. Immune cell infiltration was analyzed by CIBERSORT [50]. The TIDE algorithm was used to predict the clinical response to immunotherapy [51, 52]. GSVA was performed by the “clusterProfiler” package to investigate differences in the biological processes between the distinct DEAS-based clusters, and those with an adjusted *P*-value  $< 0.05$  were considered statistically significant [41, 53].

### Survival analysis

To determine the survival-associated DEAS events in GC, we performed a univariate Cox proportional hazards regression analysis to estimate the PSI value of DEAS events with OS. DEAS events with a *P*-value  $< 0.05$  in the univariate Cox regression analysis were

selected as candidate prognostic AS events. To investigate whether these candidate prognostic AS events were independent clinical indicators, a multivariate Cox regression analysis was then applied and adjusted for the following clinically relevant covariates: sex, age, pathological T stage, pathological N stage, pathological M stage, pathological TNM stage, histologic grade and microsatellite status. Furthermore, based on the PSI value of each independent prognostic AS event, patients were divided into low-risk and high-risk subgroups according to the optimal cutoff value, which was determined by the “survminer” package [54]. Kaplan-Meier survival analyses with log-rank tests were performed to demonstrate the survival probability variation over time between the low-risk and high-risk subgroups, and *P*-values < 0.05 were considered statistically significant.

## AUTHOR CONTRIBUTIONS

Conception and design: Yingwei Xue and Shenghan Lou. Data acquisition, assembly and experiments: Shenghan Lou and Jian Zhang. Data analyse and interpretation: Shenghan Lou, Jian Zhang, Zhao Zhai, Xin Yin, Yimin Wang and Tianyi Fang. Financial support: Yingwei Xue. All authors contributed to the manuscript. All authors read and approved the final manuscript.

## CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

## FUNDING

This work was supported by funding from the Project Nn10 of Harbin Medical University Cancer Hospital (Grant Number Nn102017–03).

## REFERENCES

1. Yako-Suketomo H, Katanoda K. Comparison of time trends in stomach cancer mortality (1990-2006) in the world, from the WHO mortality database. *Jpn J Clin Oncol.* 2009; 39:622–23. <https://doi.org/10.1093/jjco/hyp107> PMID:19713318
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018; 68:394–424. <https://doi.org/10.3322/caac.21492> PMID:30207593
3. 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. <https://doi.org/10.3322/caac.21262> PMID:25651787
4. Röcken C. Molecular classification of gastric cancer. *Expert Rev Mol Diagn.* 2017; 17:293–301. <https://doi.org/10.1080/14737159.2017.1286985> PMID:28118758
5. Gomceli I, Demiriz B, Tez M. Gastric carcinogenesis. *World J Gastroenterol.* 2012; 18:5164–70. <https://doi.org/10.3748/wjg.v18.i37.5164> PMID:23066309
6. Shibata T. Current and future molecular profiling of cancer by next-generation sequencing. *Jpn J Clin Oncol.* 2015; 45:895–99. <https://doi.org/10.1093/jjco/hyv122> PMID:26292697
7. Feng H, Qin Z, Zhang X. Opportunities and methods for studying alternative splicing in cancer with RNA-seq. *Cancer Lett.* 2013; 340:179–91. <https://doi.org/10.1016/j.canlet.2012.11.010> PMID:23196057
8. Matera AG, Wang Z. A day in the life of the spliceosome. *Nat Rev Mol Cell Biol.* 2014; 15:108–21. <https://doi.org/10.1038/nrm3742> PMID:24452469
9. Lee SC, Abdel-Wahab O. Therapeutic targeting of splicing in cancer. *Nat Med.* 2016; 22:976–86. <https://doi.org/10.1038/nm.4165> PMID:27603132
10. Climente-González H, Porta-Pardo E, Godzik A, Eyraş E. The functional impact of alternative splicing in cancer. *Cell Rep.* 2017; 20:2215–26. <https://doi.org/10.1016/j.celrep.2017.08.012> PMID:28854369
11. Oltean S, Bates DO. Hallmarks of alternative splicing in cancer. *Oncogene.* 2014; 33:5311–18. <https://doi.org/10.1038/onc.2013.533> PMID:24336324
12. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011; 144:646–74. <https://doi.org/10.1016/j.cell.2011.02.013> PMID:21376230
13. Silva AL, Faria M, Matos P. Inflammatory microenvironment modulation of alternative splicing in cancer: a way to adapt. *Adv Exp Med Biol.* 2020; 1219:243–58. [https://doi.org/10.1007/978-3-030-34025-4\\_13](https://doi.org/10.1007/978-3-030-34025-4_13) PMID:32130703
14. Ule J, Blencowe BJ. Alternative splicing regulatory networks: functions, mechanisms, and evolution. *Mol Cell.* 2019; 76:329–45. <https://doi.org/10.1016/j.molcel.2019.09.017> PMID:31626751
15. Latorre E, Harries LW. Splicing regulatory factors, ageing and age-related disease. *Ageing Res Rev.* 2017;



- 36:165–70.  
<https://doi.org/10.1016/j.arr.2017.04.004>  
PMID:[28456680](https://pubmed.ncbi.nlm.nih.gov/28456680/)
16. Li Y, Yuan Y. Alternative RNA splicing and gastric cancer. *Mutat Res*. 2017; 773:263–73.  
<https://doi.org/10.1016/j.mrrev.2016.07.011>  
PMID:[28927534](https://pubmed.ncbi.nlm.nih.gov/28927534/)
17. Chen XY, Wang ZC, Li H, Cheng XX, Sun Y, Wang XW, Wu ML, Liu J. Nuclear translocations of beta-catenin and TCF4 in gastric cancers correlate with lymph node metastasis but probably not with CD44 expression. *Hum Pathol*. 2005; 36:1294–301.  
<https://doi.org/10.1016/j.humpath.2005.09.003>  
PMID:[16311123](https://pubmed.ncbi.nlm.nih.gov/16311123/)
18. Kobayashi K, Shida A, Yamada H, Ishibashi Y, Nakayama R, Toriumi Y, Mitsumori N, Kashiwagi H, Yanaga K. Frequent splicing aberration of the base excision repair gene hMYH in human gastric cancer. *Anticancer Res*. 2008; 28:215–21.  
PMID:[18383848](https://pubmed.ncbi.nlm.nih.gov/18383848/)
19. Katoh M, Kirikoshi H, Terasaki H, Shiokawa K. Wnt2B2 mRNA, up-regulated in primary gastric cancer, is a positive regulator of the WNT- beta-catenin-TCF signaling pathway. *Biochem Biophys Res Commun*. 2001; 289:1093–98.  
<https://doi.org/10.1006/bbrc.2001.6076>  
PMID:[11741304](https://pubmed.ncbi.nlm.nih.gov/11741304/)
20. Zhou R, Moshgabadi N, Adams KL. Extensive changes to alternative splicing patterns following allopolyploidy in natural and resynthesized polyploids. *Proc Natl Acad Sci U S A*. 2011; 108:16122–27.  
<https://doi.org/10.1073/pnas.1109551108>  
PMID:[21900601](https://pubmed.ncbi.nlm.nih.gov/21900601/)
21. Xiong Y, Deng Y, Wang K, Zhou H, Zheng X, Si L, Fu Z. Profiles of alternative splicing in colorectal cancer and their clinical significance: a study based on large-scale sequencing data. *EBioMedicine*. 2018; 36:183–95.  
<https://doi.org/10.1016/j.ebiom.2018.09.021>  
PMID:[30243491](https://pubmed.ncbi.nlm.nih.gov/30243491/)
22. Li ZX, Zheng ZQ, Wei ZH, Zhang LL, Li F, Lin L, Liu RQ, Huang XD, Lv JW, Chen FP, He XJ, Guan JL, Kou J, et al. Comprehensive characterization of the alternative splicing landscape in head and neck squamous cell carcinoma reveals novel events associated with tumorigenesis and the immune microenvironment. *Theranostics*. 2019; 9:7648–65.  
<https://doi.org/10.7150/thno.36585> PMID:[31695792](https://pubmed.ncbi.nlm.nih.gov/31695792/)
23. Kodiha M, Rassi JG, Brown CM, Stochaj U. Localization of AMP kinase is regulated by stress, cell density, and signaling through the MEK→ERK1/2 pathway. *Am J Physiol Cell Physiol*. 2007; 293:C1427–36.  
<https://doi.org/10.1152/ajpcell.00176.2007>  
PMID:[17728396](https://pubmed.ncbi.nlm.nih.gov/17728396/)
24. Zhao Y, Liu Y, Lin L, Huang Q, He W, Zhang S, Dong S, Wen Z, Rao J, Liao W, Shi M. The lncRNA MACC1-AS1 promotes gastric cancer cell metabolic plasticity via AMPK/Lin28 mediated mRNA stability of MACC1. *Mol Cancer*. 2018; 17:69.  
<https://doi.org/10.1186/s12943-018-0820-2>  
PMID:[29510730](https://pubmed.ncbi.nlm.nih.gov/29510730/)
25. Chang HR, Nam S, Kook MC, Kim KT, Liu X, Yao H, Jung HR, Lemos R Jr, Seo HH, Park HS, Gim Y, Hong D, Huh I, et al. HNF4α is a therapeutic target that links AMPK to Wnt signalling in early-stage gastric cancer. *Gut*. 2016; 65:19–32.  
<https://doi.org/10.1136/gutjnl-2014-307918>  
PMID:[25410163](https://pubmed.ncbi.nlm.nih.gov/25410163/)
26. Han G, Gong H, Wang Y, Guo S, Liu K. AMPK/mTOR-mediated inhibition of survivin partly contributes to metformin-induced apoptosis in human gastric cancer cell. *Cancer Biol Ther*. 2015; 16:77–87.  
<https://doi.org/10.4161/15384047.2014.987021>  
PMID:[25456211](https://pubmed.ncbi.nlm.nih.gov/25456211/)
27. Chen CN, Chang CC, Lai HS, Jeng YM, Chen CI, Chang KJ, Lee PH, Lee H. Connective tissue growth factor inhibits gastric cancer peritoneal metastasis by blocking integrin α3β1-dependent adhesion. *Gastric Cancer*. 2015; 18:504–15.  
<https://doi.org/10.1007/s10120-014-0400-0>  
PMID:[24985492](https://pubmed.ncbi.nlm.nih.gov/24985492/)
28. Xie W, Chen C, Han Z, Huang J, Liu X, Chen H, Zhang T, Chen S, Chen C, Lu M, Shen X, Xue X. CD2AP inhibits metastasis in gastric cancer by promoting cellular adhesion and cytoskeleton assembly. *Mol Carcinog*. 2020; 59:339–52.  
<https://doi.org/10.1002/mc.23158> PMID:[31989722](https://pubmed.ncbi.nlm.nih.gov/31989722/)
29. Guo B, Zhang J, Li Q, Zhao Z, Wang W, Zhou K, Wang X, Tong D, Zhao L, Yang J, Huang C. Hypermethylation of miR-338-3p and impact of its suppression on cell metastasis through N-cadherin accumulation at the cell-cell junction and degradation of MMP in gastric cancer. *Cell Physiol Biochem*. 2018; 50:411–25.  
<https://doi.org/10.1159/000494153> PMID:[30308487](https://pubmed.ncbi.nlm.nih.gov/30308487/)
30. Zhu LY, Zhu YR, Dai DJ, Wang X, Jin HC. Epigenetic regulation of alternative splicing. *Am J Cancer Res*. 2018; 8:2346–58.  
PMID:[30662796](https://pubmed.ncbi.nlm.nih.gov/30662796/)
31. Zhang C, Cheng W, Ren X, Wang Z, Liu X, Li G, Han S, Jiang T, Wu A. Tumor purity as an underlying key factor in glioma. *Clin Cancer Res*. 2017; 23:6279–91.  
<https://doi.org/10.1158/1078-0432.CCR-16-2598>  
PMID:[28754819](https://pubmed.ncbi.nlm.nih.gov/28754819/)
32. Mao Y, Feng Q, Zheng P, Yang L, Liu T, Xu Y, Zhu D, Chang W, Ji M, Ren L, Wei Y, He G, Xu J. Low tumor

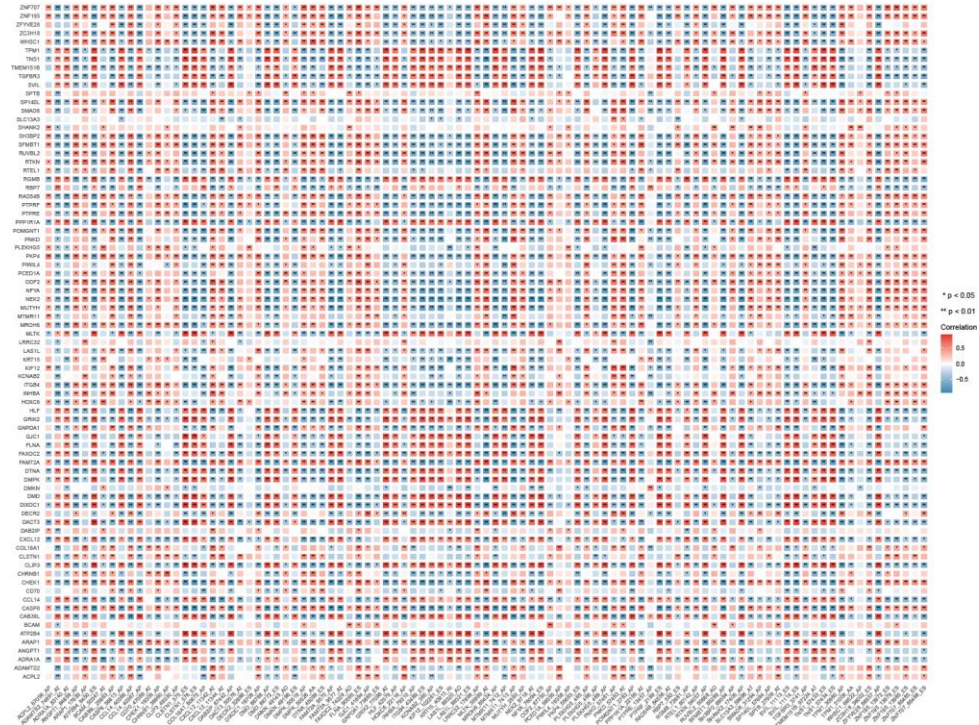
- purity is associated with poor prognosis, heavy mutation burden, and intense immune phenotype in colon cancer. *Cancer Manag Res*. 2018; 10:3569–77.  
<https://doi.org/10.2147/CMAR.S171855>  
 PMID:30271205
33. Rhee JK, Jung YC, Kim KR, Yoo J, Kim J, Lee YJ, Ko YH, Lee HH, Cho BC, Kim TM. Impact of tumor purity on immune gene expression and clustering analyses across multiple cancer types. *Cancer Immunol Res*. 2018; 6:87–97.  
<https://doi.org/10.1158/2326-6066.CIR-17-0201>  
 PMID:29141981
  34. Mu H, Wang N, Zhao L, Li S, Li Q, Chen L, Luo X, Qiu Z, Li L, Ren G, Xu Y, Zhou X, Xiang T. Methylation of PLCD1 and adenovirus-mediated PLCD1 overexpression elicits a gene therapy effect on human breast cancer. *Exp Cell Res*. 2015; 332:179–89.  
<https://doi.org/10.1016/j.yexcr.2015.01.017>  
 PMID:25655282
  35. Xiang DM, Sun W, Zhou T, Zhang C, Cheng Z, Li SC, Jiang W, Wang R, Fu G, Cui X, Hou G, Jin GZ, Li H, et al. Oncofetal HLF transactivates c-Jun to promote hepatocellular carcinoma development and sorafenib resistance. *Gut*. 2019; 68:1858–71.  
<https://doi.org/10.1136/gutjnl-2018-317440>  
 PMID:31118247
  36. Murai M, Toyota M, Suzuki H, Satoh A, Sasaki Y, Akino K, Ueno M, Takahashi F, Kusano M, Mita H, Yanagihara K, Endo T, Hinoda Y, et al. Aberrant methylation and silencing of the BNIP3 gene in colorectal and gastric cancer. *Clin Cancer Res*. 2005; 11:1021–27.  
 PMID:15709167
  37. Li R, Qu H, Wang S, Wei J, Zhang L, Ma R, Lu J, Zhu J, Zhong WD, Jia Z. GDCRNATools: an r/bioconductor package for integrative analysis of lncRNA, miRNA and mRNA data in GDC. *Bioinformatics*. 2018; 34:2515–17.  
<https://doi.org/10.1093/bioinformatics/bty124>  
 PMID:29509844
  38. Ryan M, Wong WC, Brown R, Akbani R, Su X, Broom B, Melott J, Weinstein J. TCGASpliceSeq a compendium of alternative mRNA splicing in cancer. *Nucleic Acids Res*. 2016; 44:D1018–22.  
<https://doi.org/10.1093/nar/gkv1288> PMID:26602693
  39. Wang ET, Sandberg R, Luo S, Khrebtkova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. Alternative isoform regulation in human tissue transcriptomes. *Nature*. 2008; 456:470–76.  
<https://doi.org/10.1038/nature07509> PMID:18978772
  40. Beretta L, Santaniello A. Nearest neighbor imputation algorithms: a critical evaluation. *BMC Med Inform Decis Mak*. 2016 (Suppl 3); 16:74.  
<https://doi.org/10.1186/s12911-016-0318-z>  
 PMID:27454392
  41. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012; 16:284–87.  
<https://doi.org/10.1089/omi.2011.0118>  
 PMID:22455463
  42. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003; 13:2498–504.  
<https://doi.org/10.1101/gr.1239303> PMID:14597658
  43. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol*. 2014 (Suppl 4); 8:S11.  
<https://doi.org/10.1186/1752-0509-8-S4-S11>  
 PMID:25521941
  44. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics*. 2003; 4:2.  
<https://doi.org/10.1186/1471-2105-4-2>  
 PMID:12525261
  45. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015; 1:417–25.  
<https://doi.org/10.1016/j.cels.2015.12.004>  
 PMID:26771021
  46. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010; 26:139–40.  
<https://doi.org/10.1093/bioinformatics/btp616>  
 PMID:19910308
  47. Law CW, Chen Y, Shi W, Smyth GK. Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*. 2014; 15:R29.  
<https://doi.org/10.1186/gb-2014-15-2-r29>  
 PMID:24485249
  48. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics*. 2010; 26:1572–73.  
<https://doi.org/10.1093/bioinformatics/btq170>  
 PMID:20427518
  49. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, Carter SL, Getz G, Stemke-Hale K, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. 2013; 4:2612.  
<https://doi.org/10.1038/ncomms3612>

PMID:[24113773](https://pubmed.ncbi.nlm.nih.gov/24113773/)

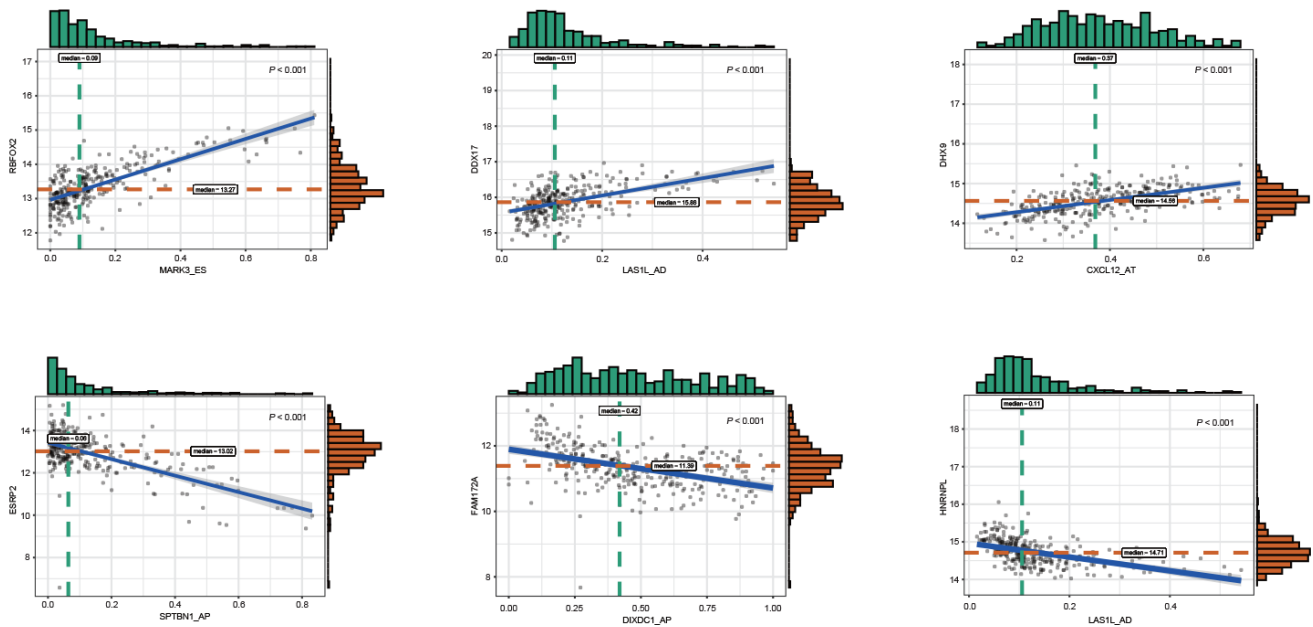
50. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015; 12:453–57.  
<https://doi.org/10.1038/nmeth.3337> PMID:[25822800](https://pubmed.ncbi.nlm.nih.gov/25822800/)
51. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, Li Z, Traugh N, Bu X, Li B, Liu J, Freeman GJ, Brown MA, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med*. 2018; 24:1550–58.  
<https://doi.org/10.1038/s41591-018-0136-1>  
PMID:[30127393](https://pubmed.ncbi.nlm.nih.gov/30127393/)
52. Fu J, Li K, Zhang W, Wan C, Zhang J, Jiang P, Liu XS. Large-scale public data reuse to model immunotherapy response and resistance. *Genome Med*. 2020; 12:21.  
<https://doi.org/10.1186/s13073-020-0721-z>  
PMID:[32102694](https://pubmed.ncbi.nlm.nih.gov/32102694/)
53. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. 2013; 14:7.  
<https://doi.org/10.1186/1471-2105-14-7>  
PMID:[23323831](https://pubmed.ncbi.nlm.nih.gov/23323831/)
54. Li S, Chen S, Wang B, Zhang L, Su Y, Zhang X. A robust 6-lncRNA prognostic signature for predicting the prognosis of patients with colorectal cancer metastasis. *Front Med (Lausanne)*. 2020; 7:56.  
<https://doi.org/10.3389/fmed.2020.00056>  
PMID:[32211413](https://pubmed.ncbi.nlm.nih.gov/32211413/)

SUPPLEMENTARY MATERIALS

Supplementary Figures



Supplementary Figure 1. Correlation heatmap of DEAS events and their parental genes.



Supplementary Figure 2. Representative correlation between DEAS events and SFs.



## Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 1–4, 6–12.

**Supplementary Table 1. DEAS events.**

**Supplementary Table 2. DEGs of the 280 parental genes.**

**Supplementary Table 3. Spearman’s correlation analysis of DEGs and DEAS events.**

**Supplementary Table 4. GO enrichment analysis of DEAS events.**

**Supplementary Table 5. KEGG enrichment analysis of DEAS events.**

KEGG ID	Description	GeneRatio	BgRatio	Adjusted P-value	gene ID
hsa00562	Inositol phosphate metabolism	6/127	74/8032	0.0010682	MTMR1/MTMR3/MTMR4/ALDH6A1/PIK3CD/PLCD1
hsa05205	Proteoglycans in cancer	10/127	204/8032	0.0014152	CD44/ESR1/WNT2B/RPS6/AKT1/CAMK2D/CAMK2G/ANK3/PIK3CD/FLNA
hsa04261	Adrenergic signaling in cardiomyocytes	8/127	149/8032	0.0024234	ATP2B4/ADRA1A/PPP1R1A/TPM1/AKT1/CAMK2D/CAMK2G/CACNA2D4
hsa04152	AMPK signaling pathway	7/127	120/8032	0.0028412	TBC1D1/CAB39L/LEPR/ADRA1A/AKT1/SCD5/PIK3CD
hsa04620	Toll-like receptor signaling pathway	6/127	104/8032	0.0059612	CASP8/MAP2K7/MAP2K6/AKT1/TIRAP/PIK3CD
hsa04066	HIF-1 signaling pathway	6/127	109/8032	0.0074647	RPS6/AKT1/CAMK2D/CAMK2G/ANGPT1/PIK3CD
hsa04668	TNF signaling pathway	6/127	112/8032	0.0084888	CASP8/MAP2K7/MAP2K6/DAB2IP/AKT1/PIK3CD
hsa04670	Leukocyte transendothelial migration	6/127	113/8032	0.0088515	CD99/VCL/CXCL12/AFDN/PIK3CD/CTNND1
hsa04012	ErbB signaling pathway	5/127	85/8032	0.0109392	MAP2K7/AKT1/CAMK2D/CAMK2G/PIK3CD
hsa05131	Shigellosis	9/127	236/8032	0.0123094	CD44/VCL/AKT1/WIP1/PIK3CD/BNIP3/SEPTIN9/PLCD1/FNBP1
hsa05235	PD-L1 expression and PD-1 checkpoint pathway in cancer	5/127	89/8032	0.0131715	NFATC2/MAP2K6/AKT1/TIRAP/PIK3CD
hsa05161	Hepatitis B	7/127	162/8032	0.0141447	CASP8/MAP2K7/NFATC2/MAP2K6/AKT1/TIRAP/PIK3CD
hsa05017	Spinocerebellar ataxia	5/127	98/8032	0.0192755	NFYA/AKT1/WIP1/GRIN2C/PIK3CD
hsa04070	Phosphatidylinositol signaling system	5/127	99/8032	0.0200512	MTMR1/MTMR3/MTMR4/PIK3CD/PLCD1
hsa04140	Autophagy - animal	6/127	137/8032	0.0212426	MTMR3/MTMR4/AKT1/WIP1/PIK3CD/BNIP3
hsa04720	Long-term potentiation	4/127	67/8032	0.0212478	PPP1R1A/CAMK2D/CAMK2G/GRIN2C
hsa04664	Fc epsilon RI signaling pathway	4/127	68/8032	0.0223072	MAP2K7/MAP2K6/AKT1/PIK3CD
hsa05152	Tuberculosis	7/127	180/8032	0.0237873	NFYA/CASP8/LSP1/AKT1/CAMK2D/CAMK2G/TIRAP
hsa04625	C-type lectin receptor signaling pathway	5/127	104/8032	0.0242368	CASP8/NFATC2/LSP1/AKT1/PIK3CD
hsa04520	Adherens junction	4/127	71/8032	0.0256768	VCL/AFDN/PTPRF/CTNND1
hsa05214	Glioma	4/127	75/8032	0.0306229	AKT1/CAMK2D/CAMK2G/PIK3CD
hsa05412	Arrhythmogenic right ventricular cardiomyopathy (ARVC)	4/127	77/8032	0.0332926	ITGB4/CACNA2D4/JUP/DMD
hsa04724	Glutamatergic synapse	5/127	114/8032	0.0342061	SLC38A1/GLUL/GRIN2C/SHANK2/GRIK2
hsa04510	Focal adhesion	7/127	199/8032	0.0380996	VCL/TNC/ITGB4/AKT1/COL6A3/PIK3CD/FLNA
hsa03010	Ribosome	6/127	158/8032	0.0389509	MRPL22/RPL18A/RPS6/RPL32/RPS3A/RPS21
hsa04722	Neurotrophin signaling pathway	5/127	119/8032	0.0400213	MAP2K7/AKT1/CAMK2D/CAMK2G/PIK3CD
hsa05135	Yersinia infection	5/127	120/8032	0.0412523	MAP2K7/NFATC2/MAP2K6/AKT1/PIK3CD

**Supplementary Table 6. GO GSEA between tumor and normal samples.**

**Supplementary Table 7. KEGG GSEA between tumor and normal samples.**

**Supplementary Table 8. Functional enrichment analysis for module 1.**

**Supplementary Table 9. Functional enrichment analysis for module 2.**

**Supplementary Table 10. Spearman’s correlation analysis of DEAS events and SFs.**

**Supplementary Table 11. Kruskal-Wallis test of GSVA scores among the three clusters.**

**Supplementary Table 12. Clinical and immune features among the 3 DEAS-based clusters.**