Prognostic alternative splicing events related splicing factors define the tumor microenvironment and pharmacogenomic landscape in lung adenocarcinoma

Jichang Liu[1,*], Yadong Wang[1,*], Xiaogang Zhao[3], Kai Wang[1], Chao Wang[4], Jiajun Du[1,2]

[1]Institute of Oncology, Shandong Provincial Hospital, Shandong University, Jinan, Shandong, P.R. China
[2]Department of Thoracic Surgery, Shandong Provincial Hospital, Shandong University, Jinan, Shandong, P.R. China
[3]Department of Thoracic Surgery, The Second Hospital of Shandong University, Jinan, Shandong, P.R. China
[4]Department of Respiratory, The First Affiliated Hospital of Shandong First Medical University, Jinan, Shandong, P.R. China
[*]Equal contribution

Correspondence to: Jiajun Du, Chao Wang; **email:** dujiajun@sdu.edu.cn; xiaoxingxingxh@163.com, https://orcid.org/0000-0002-1087-2684

ABSTRACT

Background: Recent studies identified correlations between splicing factors (SFs) and tumor progression and therapy. However, the potential roles of SFs in immune regulation and the tumor microenvironment (TME) remain unknown.

Methods: We used UpSet plots to screen for prognostic-related alternative splicing (AS) events. We evaluated SF patterns in specific immune landscapes. Single sample gene set enrichment analysis (ssGSEA) algorithms were used to quantify relative infiltration levels in immune cell subsets. Principal component analysis (PCA) algorithm-based SFscore were used to evaluate SF patterns in individual tumors with an immune response.

Results: From prognosis-related AS events, 16 prognosis-related SFs were selected to construct three SF patterns. Further TME analyses showed these patterns were highly consistent with immune-inflamed, immune-excluded, and immune-desert landscapes. Based on SFscore constructed using differentially expressed genes (DEGs) between SF patterns, patients were classified into two immune-subtypes associated with differential pharmacogenomic landscapes and cell features. A low SFscore was associated with high immune cell infiltration, high tumor mutation burden (TMB), and elevated expression of immune check points (ICPs), indicating a better immune response.

Conclusions: SFs are significantly associated with TME remodeling. Evaluating different SF patterns enhances our understanding of the TME and improves effective immunotherapy strategies.

INTRODUCTION

Alternative splicing (AS) is a crucial process in most gene expression and generates more than one mRNA from a single gene locus to then generate distinct proteins [1, 2]. More than 90% of human genes undergo AS, with variability between tissues [3]. Studies have shown that AS defects significantly affect cell development and underlie many diseases, including different cancers, as selective AS allows cancer cells to generate isoforms that benefit cell survival [4, 5]. Also, AS affects tumorigenesis by influencing genome instability [6] and gene mutations, other cancer hallmarks [7]. Splicing factor (SF) dysregulation is the

main cause of aberrant AS, with SFs functioning as both oncoproteins and tumor suppressors [8].

Globally, lung cancer has the highest cancer related mortality [9]. Non-small cell lung cancer (NSCLC) accounts for the majority of lung cancers, of which lung adenocarcinoma (LUAD) ratios are continuously increasing [10]. Altered AS occurs in LUAD and mainly results from SF expression, including SRSF and RBM families [3, 11]. It was reported that aberrant AS events in LUAD contribute to different cell functions, such as apoptosis, proliferation, and drug resistance [11]. Although the mechanisms underpinning aberrant AS patterns remain unclear, assessing such patterns to evaluate LUAD risk and predict the effects of LUAD treatment may improve patient survival.

In recent years, LUAD treatments have been continuously developed and improved, but a definitive cure remains elusive [12]. Chemotherapy is the first line of treatment for lung cancer; however, its benefits have plateaued. Treatments targeting epidermal growth factor receptor (EGFR) mutations, anaplastic lymphoma kinase (ALK) fusion oncogenes, and *KRAS* mutations are considered promising but their effects are limited [13].

The tumor microenvironment (TME) plays a vital role in tumor proliferation, invasion, and migration [14]. Consistent with these functions, immunotherapy, including the anti-tumor effects of CTLA-4 and programmed death-ligand 1/programmed death-1 (PD-L1/PD-1) blockade, have become leading and powerful treatments for LUAD treatment [15–17]. TME complexity is mainly reflected by immune infiltration, and is based on the spatial localization of immune cells relative to the tumor and stromal compartments, therefore human tumors are categorized as immune-inflamed, immune-desert, and immune-excluded phenotypes, and present many challenges for tumor immunotherapy [15, 18].

Studies have identified relationships between AS or SFs and the immune microenvironment, including head and neck squamous cell carcinoma, breast cancer, and LUAD [19, 20]. It was reported that *PD-L1* generated a long non-coding RNA by AS to promote LUAD progression [21]. Recent studies also reported interactions between SFs and cancer therapy that some SFs contributed to drug resistance [13, 22]. In terms of cancer immunotherapy, peptides may serve as neoepitopes from tumor specific mRNA splicing events, with a potential to bind to major histocompatibility complex class I molecules [23]. Furthermore, splicing-derived neoantigens may be useful as predictive response biomarkers for immune check point (ICP)

blockade therapies, such as PD-1 or CTLA-4 [23]. There is an urgent need to identify SF related targets for cancer therapy, and recognizing the significance of SFs and TME characteristics, mediated by SFs, will improve our understanding of tumor immunity.

Thanks to the establishment of public databases and bioinformatics, the analysis of tumor expression profiles is both rapid and convenient. In this study, we downloaded gene expression data from The Cancer Genome Atlas (TCGA), Gene Expression Omnibus (GEO), and TCGA SpliceSeq databases. By analyzing SF patterns and SF cluster-related differentially expressed genes (DEGs), we constructed several TME models to predict immunotherapeutic benefits in LUAD. We showed that AS is key for shaping the TME and may have important therapeutic applications for LUAD.

METHODS

Data sources and preprocessing

AS data from LUAD patients were downloaded from the TCGA SpliceSeq database, and included seven different AS events: 1) alternate acceptor site (AA), 2) alternate donor site (AD), 3) alternate promoter (AP), 4) alternate terminator (AT), 5) exon skip (ES), 6) mutually exclusive exons (ME), and 7) retained intron (RI). Specimens were included for further analysis if percentage percent-spliced-in (PSI) value were > 75%. An SF list was identified from a published study [24]. Gene expression and clinical data were respectively retrieved from TCGA (https://portal.gdc.cancer.gov/) and GEO (https://www.ncbi.nlm.nih.gov/geo/) data-bases. In total, six eligible LUAD cohorts (GSE13213 [25], GSE37745 [26], GSE31210 [27], GSE3141 [28], GSE30219 [29], and GSE50081 [30]) were integrated as the training cohort. LUAD patients from TCGA were the testing cohort. Batch effects from non-biological technical biases were corrected using the "ComBat" algorithm of the sva package. Copy number variation (CNV) data from TCGA-LUAD patients were obtained from the UCSC Xena website (https://xena.ucsc.edu/).

Prognosis-related AS events

We eliminated specimens without follow-up or short follow-up information (< 90 days) to exclude the impact of short-term follow-up. AS events were eliminated when the standard deviation of the PSI value among specimens was < 0.01. Prognosis-related AS events were identified using univariate Cox regression analysis, and displayed in a volcano plot and UpSet map. The top 20 AS events were displayed in a bubble chart.

Constructing an AS-SF interaction network

Correlations between prognostic AS events and SFs were evaluated by Spearman's analyses in LUAD patients from the TCGA dataset. Interactions were included when correlation coefficients were > 0.6 and $p < 0.001$. Then, an AS-SF interaction network was constructed and included prognostic AS events and related SFs.

SF-based consensus molecular clustering

In our AS-SF interaction network, we identified 16 SFs, including *CIRBP, CCDC130, CLASRP, LUC7L3, CLK1, CLK4, ALYREF, RBM5, CDK10, SREK1, SNRNP70, RBM15, ARGLU1, SRSF5, SRRM2,* and *SRSF11*. Genes, which were highly correlated with prognosis-related AS events. Four genes not found in the training cohort (integrated GEO cohort) were eliminated. Unsupervised clustering analysis was used to identify distinct patterns using these SFs. A consensus clustering algorithm was used to classify LUAD samples into different SF modification clusters and test corresponding stability. The Consensu-ClusterPlus package [31] was used to perform these steps and 1000 times repetitions were calculated to guarantee corresponding stability.

Gene set variation analysis (GSVA)

To explore biological heterogeneity between different SF patterns, GSVA enrichment was performed in the "GSVA" package [32]. Hallmark gene sets "h.all.v7.4.symbols.gmt" were extracted from the MSigDB database [33] to conduct GSVA. The clusterProfiler R package [34] was used to perform functional annotations for SF-related genes. An adjusted P value < 0.05 was considered statistically significant.

Single sample gene set enrichment analysis (ssGSEA)

We used the ssGSEA algorithm to quantify the relative abundance of 28 infiltrating immune cell types, 13 immune functions, and other related biological processes in LUAD samples. The gene sets for marking each TME infiltration immune cell type stored various immune cell subtypes, including activated B cells, activated CD8 T cells, activated dendritic cells, macrophages, natural killer T cells, and regulatory T cells [35]. We extracted 13 gene sets for other related biological processes from published studies, including (1) immune checkpoints; (2) angiogenesis; (3) antigen processing machinery; (4) CD8 T effectors; (5) epithelial mesenchymal transition (EMT), including EMT1, EMT2, and EMT3; (6) pan-fibroblast TGFb response signatures (Pan-FTBRS); (7) Wnt targets; (8) mismatch repair; (9) DNA damage repair; (10) DNA replication; and (11) nucleotide excision repair [36, 37].

Identifying DEGs between SF distinct clusters

We next determined SF-related DEGs among distinct SF patterns in LUAD. The limma package was used for this, and the filtering criteria for DEGs was an adjusted P value < 0.05.

Constructing SFscore

An SF scoring approach was developed to quantify SF patterns from individual patients based on PCA values. DEGs selected from distinct SF phenotypes underwent prognostic analyses using univariate Cox regression model. Genes with a prognostic significance were extracted for next feature selection using recursive feature elimination in the 'caret' package. PCA analysis was then performed based on finally determined genes. Principal components 1 and 2 served as the signature score. Subsequently, a formula similar to previous studies was constructed to define the SFscore: SFscore $= \sum(PC1_i + PC2_i)$, where i = SF-related signature gene expression [38].

Assessing TMB and predicting ICI therapy responses

Maftools package was used to visualize somatic mutations based on the Mutation Annotation Format file in the TCGA cohort [39]. The Tumor Immune Dysfunction and Exclusion (TIDE) algorithm in TCGA cohort [40], including dysfunction and exclusion of infiltrating cytotoxic T lymphocytes (CTLs), was used to assess immune evasion mechanisms in the TCGA cohort [41]. A high TIDE level indicated a low response to immune checkpoint inhibitor (ICI) therapy. The tumor neoantigen burden, including clonal and sub-clonal neoantigen burdens, was also used to assess immunotherapeutic efficacy in the TCGA cohort [35]. Differential expression analysis of ICP genes was performed in the training cohort.

Prediction of drug sensitivity and small molecular drugs

The R package "pRRophetic" was used to calculate the half maximal inhibitory concentration (IC_{50}) of chemotherapy drugs [42]. DEGs between groups with high and low SFscore were identified by $p < 0.05$ and |logFC|>1 in "limma". SFscore-related small molecular drugs were identified in the Connectivity Map database (CMap; https://clue.io/) after uploading down- and up-regulated genes [43].

Table 1. Multivariate Cox analysis of all prognosis related AS events.

Id	Coef	HR	HR.95L	HR.95H	*p*
TTC39C|44852|AP	1.248442	3.484908	1.391101	8.730198	0.007711
CDKN2A|86004|AP	1.489648	4.435536	2.047892	9.60694	0.000158
C10orf32|12982|RI	−9.15774	0.000105	4.24E-07	0.026199	0.001137
BEST3|23330|AT	1.771023	5.87686	2.010846	17.1756	0.00121
TLE2|46644|AA	−25.1683	1.17E-11	4.38E-16	3.14E-07	1.31E-06
CA5B|98313|ES	−1.1536	0.315499	0.117144	0.849725	0.022483
SDCBP|83930|ES	−13.5109	1.36E-06	8.07E-10	0.002278	0.000363
HNRNPLL|53258|AT	−3.35378	0.034952	0.003746	0.326121	0.003247
MEGF6|315|ES	−1.42332	0.240914	0.095041	0.610679	0.002707

Quantitative real-time PCR

Quantitative real-time PCR Total RNA of carcinoid tissues was extracted with Trizol reagent (A2A0209, Accurate Biotechnology, China). cDNA was amplified with reverse transcription kit (A2A1386) was provided by Accurate Biotechnology Co. (China). The sequences of primers are in Supplementary Table 1. Gene expression levels were assayed by qRT-PCR using the Roche LightCycler® 480 system (Roche, Basel, Switzerland) with the SYBR Green system (A2A1436, Accurate Biotechnology).

Statistical analyses

Statistical tests were performed in R-4.0.2 software, and a two-sided *P* value < 0.05 was considered statistically significant. Statistical significance for normally distributed variables was analyzed using Student's *t*-tests, while non-normally distributed variables were estimated using the Wilcoxon rank-sum test. Kruskal-Wallis tests were used to compare more than two groups, as nonparametric and parametric methods, respectively. Kaplan-Meier survival analyses were conducted using the 'Survminer' package to explore associations between SF patterns and prognoses. Continuous variables were dichotomized for overall survival (OS) before the log-rank test using optimal cutoff values as determined by the "surv_cutpoint" function in the "Survminer" package. After this, that, LUAD samples were categorized into high and low SFscore subgroups.

Availability of data and materials

The dataset supporting the conclusions of this article is available in the TCGA website (https://portal.gdc.cancer.gov/) and GEO database (https://www.ncbi.nlm.nih.gov/geo/).

RESULTS

Total and prognosis-related AS events

A flow chart (Figures 1, 2A) shows the study design and processes. In total, 572 LUAD specimens with 43,948 AS events in 10,366 gene symbols were collected. These events included 16,793 ES in 6,618 genes, 8,546 AT in 3,734 genes, 8,992 AP in 3,605 genes, 3,559 AA in 2,522 genes, 3,057 AD in 2,173 genes, 2,781 RI in 1,866 genes, and 220 ME in 214 genes. An UpSet plot was constructed to depict overlaps in the seven AS event types, which indicated multiple AS events appeared on a single gene (Figure 2B, 2C). ES was the most common ES event type, while ME was the least common.

Univariate Cox regression analysis was used to select prognosis-related AS events, but excluded patients with follow-up < 90 days. In total, 2,692 AS events significantly associated with OS ($p < 0.05$) were selected, including 906 ES, 487 AT, 744 AP, 185 AA, 182 AD, 177 RI, and 11 ME. A volcano map depicted AS events (Supplementary Figure 1A). The top 20 prognosis-related AS events of the seven types are shown (Supplementary Figure 1B–1H). Multivariate Cox analysis was also conducted to identify the independent prognostic factors (Table 1).

Construction of an AS-SF interaction network

We performed correlation analysis between prognosis-related AS events and SFs, from which 20 prognosis-related SFs were selected (correlation coefficient > 0.6 and $p < 0.001$). A prognosis-related AS-SF interaction network was constructed, including 12 adverse AS events (red nodes), 28 favorable AS events (blue nodes), and 20 interacting SFs (black nodes) (Figure 3). Most adverse AS were negatively regulated by SFs,

while MCM7-80881-AP was upregulated by ALYREF. In contrast, most favorable AS events were positively regulated by SFs, except MCM7-80880-AP and FLJ27365-62678-AP; they were downregulated by ALYREF and RNU4-1, respectively.

The genetic variation of AS interacting SFs in LUAD

Based on the AS-SF interaction network, we identified 16 SFs (*CIRBP, CCDC130, CLASRP, LUC7L3, CLK1,* *CLK4, ALYREF, RBM5, CDK10, SREK1, SNRNP70, RBM15, ARGLU1, SRSF5, SRRM2,* and *SRSF11*) after excluding four genes (*LUC7L, DDX39B, RNU4-1,* and *RNU5A-1*) not found in training cohort. The differential expression of these SFs between normal and tumor specimens was analyzed in the TCGA cohort (Figure 4A). Most SFs were highly expressed in tumors, while *CIRBP, ARGLU1,* and *SRSF5* were poorly expressed. Also, 13 SFs were identified as significant prognostic risk factors in LUAD (Supplementary Figure 2).
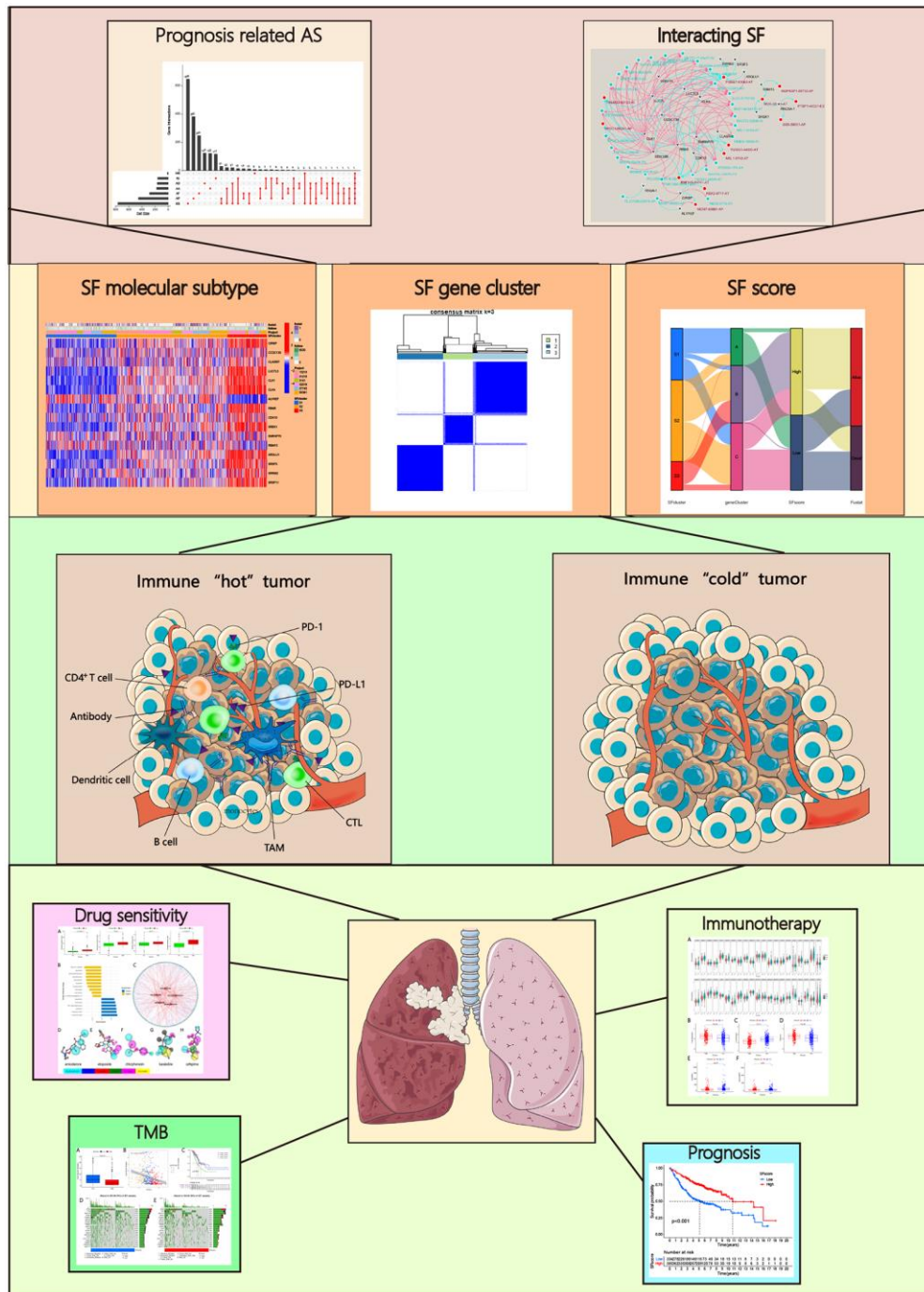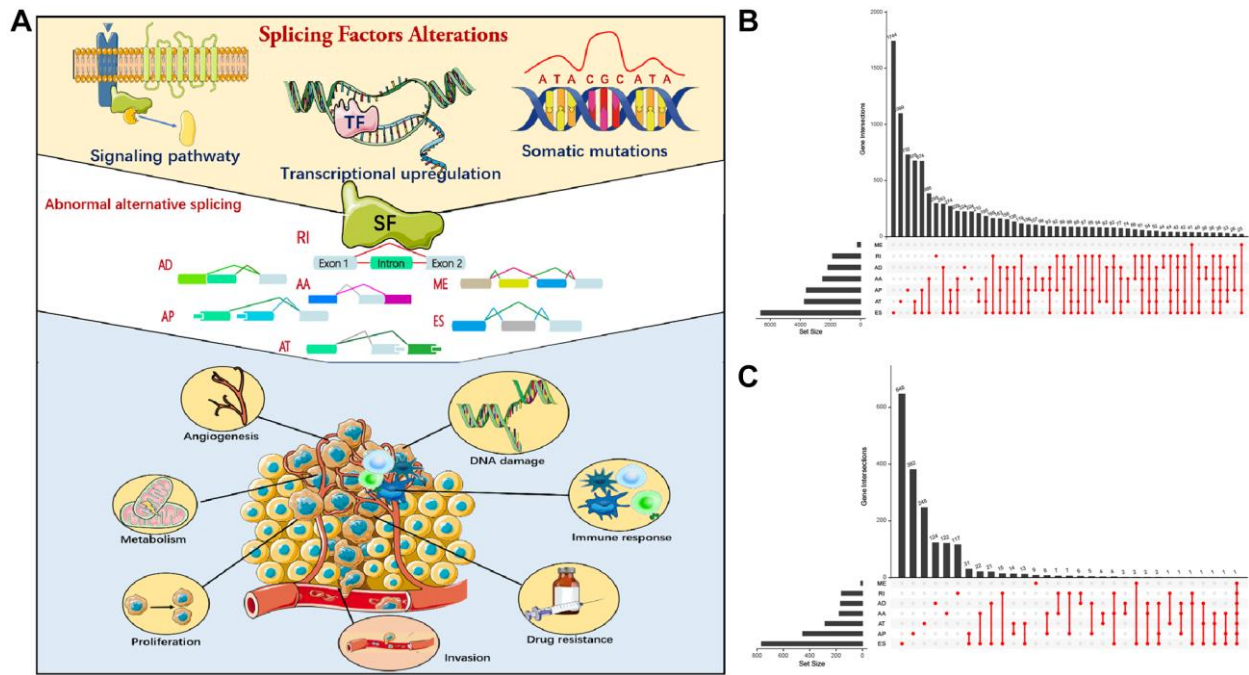

Figure 1. Analysis workflow of this study.

Figure 2. Alternative splicing (AS) events in cancer. (**A**) Mechanism of splicing regulatory factors in regulating RNA alternative splicing and tumor progression. ES (Exon skip) means that an exon is cut from the original transcript. RI (Retained intron): A new exon is formed by the retained Intron and the exons on both sides. AD (Alternate Donor site): The 3'-end splicing sites of different transcripts are the same but the 5'-end splicing sites are different. AA (Alternate acceptor site): The 5'-end splicing sites of different transcripts are the same but the 3'-end splicing sites are different. AP (Alternate promoter): The first exon of the two transcripts is different. AT (Alternate terminator): The last exon of the two transcripts is different. ME (Mutually exclusive exons): Different exons (called inclusive exons) are present in different transcripts. (**B**) Upset plot of all AS events in LUAD. (**C**) Upset plot of prognosis-related AS events in LUAD.
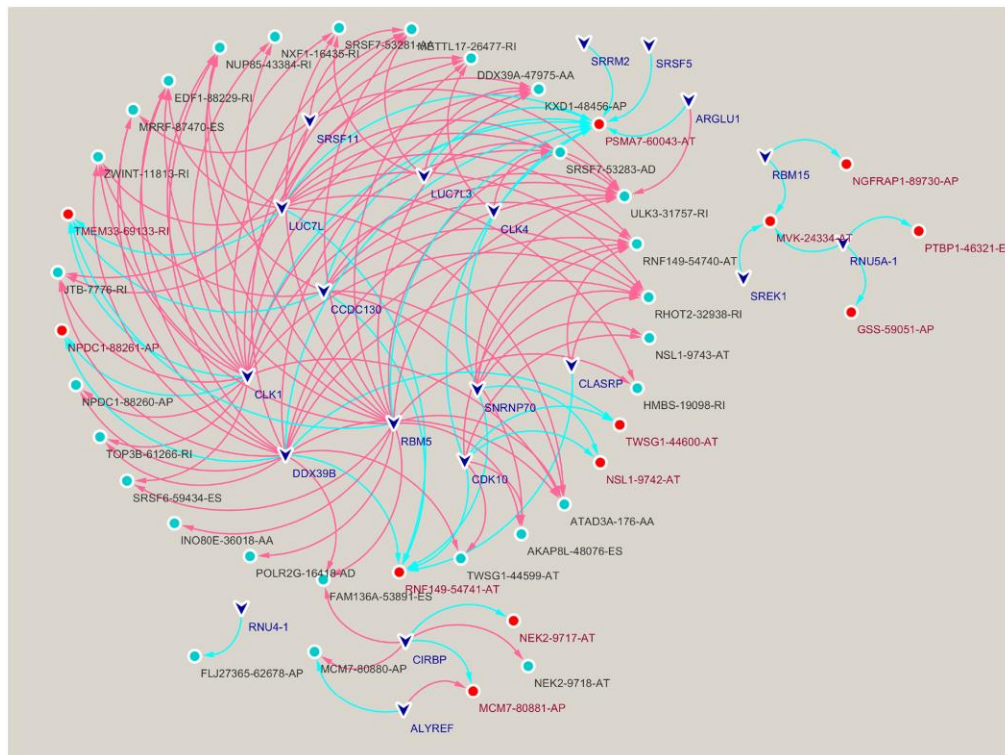
Figure 3. Construction of a splicing factors (SFs)-alternative splicing (AS) events network. Red edges mean positive regulation between AS and SF, while blue edges mean negative.

To understand the genetic variation landscape of SFs in LUAD, we explored the incidence of CNVs and somatic mutations in the 16 SFs. CNV alterations were widespread. *ALYREF, ARGLU1, CLASRP, LUC7L3, CLK4, CLK1,* and *SRRM2* were focused on the amplification in copy number, while the others had a frequency of CNV deletion (Figure 4B). Among the 561 LUAD samples, 55 had mutations in AS interacting SFs (Figure 4C). The chromosomal positions of CNV alteration in the 16 SFs are shown (Figure 4D). Based on the TCGA dataset, aa SF interaction network was constructed to show connections and the prognostic significance of SFs (Figure 4E). This network indicated that most SFs had positive correlation interactions, while *ALYREF* interacted with *CIRBP, SRSF5, CLK1, CLK4* and *RBM5* with negative correlations.

Distinct SF patterns mediated by 16 AS interacting SFs

Six GEO datasets (GSE13213, GSE37745, GSE31210, GSE3141, GSE30219, and GSE50081) were integrated into one meta-cohort. Based on SF expression, LUAD patients in the GEO meta-cohort were stratified into three distinct SF patterns using ConsensusClusterPlus in R (Figure 5A). We termed these patterns, SF clusters S1–S3, which included 229, 362, and 128 patients, respectively. PCA indicated significant distinctions in transcription profiles between patterns (Figure 5B). Survival analysis for these SF clusters showed a significantly prominent survival disadvantage in SF cluster-S1 (Figure 5C). Moreover, most SF genes were upregulated in SF cluster-S3 and downregulated in cluster-S1, while *ALYREF* and *RBM15* showed the opposite trend (Figure 5D, Supplementary Figure 3A).

TME cell infiltration characteristics in distinct SF patterns

To explore biological function differences among the three SF patterns, we performed GSVA enrichment. As shown (Supplementary Figure 3B, 3C), when compared with SF cluster-S3, a series of carcinogenic and stromal pathways were activated in patients in SF-S1 and S2clusters, including MYC signaling, E2F signaling, mTROC1 signaling, and EMT. Additionally,
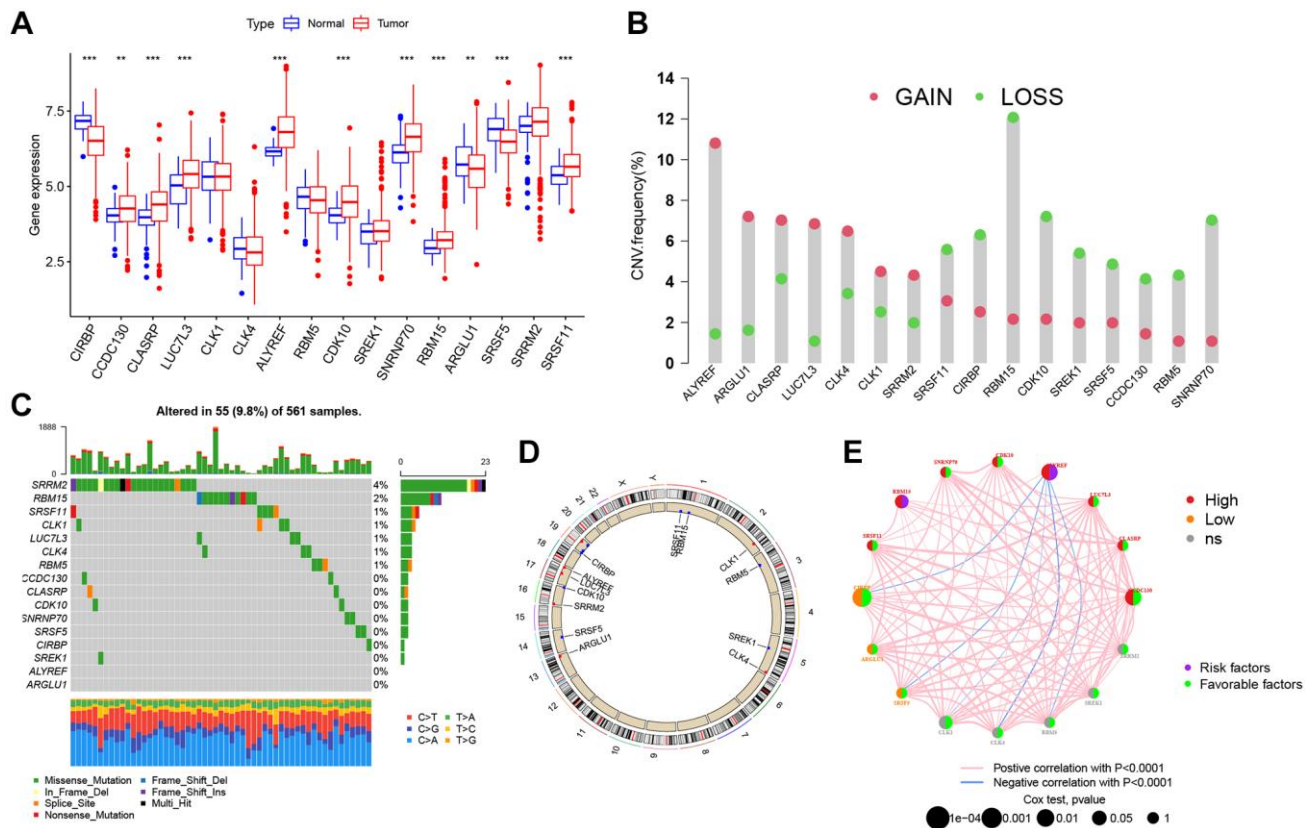

Figure 4. Genetic alteration landscape of 16 splicing factors (SFs) in LUAD. (**A**) Differential mRNA expression of 16 SFs between normal and tumor samples (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$). (**B**) CNV mutation was widespread in 16 SFs. The column represented the alteration frequency. Deletion, green dot; Amplification, pink dot. (**C**) 56 of the 561 LUAD patients showed genetic alterations of 16 SFs. (**D**) The location of CNV alterations of 16 SFs on chromosomes. (**E**) The relationship between SFs in LUAD. The thickness of lines linking SFs showed the correlation strength. Negative correlation, blue; positive correlation, pink. Up-regulated SFs, red; down-regulated SFs, orange; no sense of SFs in differential expression, grey.

immune-related pathways were fully activated in SF cluster-S2 patients, including interferon gamma/alpha responses, IL-6-JAK-STAT3 signaling, allograft rejection, and inflammatory responses. In contrast, SF cluster-S3 patients showed activated TGF-β signaling and suppressed immune functions.

In SF patterns, the ssGSEA algorithm was used to determine the abundance of different immune cell types

and immune functions. Surprisingly, the patterns had significantly distinct TME cell-infiltrating and immune function characteristics (Figure 5E, 5F). Cluster-S1 was classified as an immune-inflamed phenotype, characterized by adaptive immune cell infiltration and immune function activation; cluster-S2 was identified as an immune-excluded phenotype, characterized by innate immune cell infiltration; and cluster-S3 was classified as an immune-desert phenotype with prominent immune
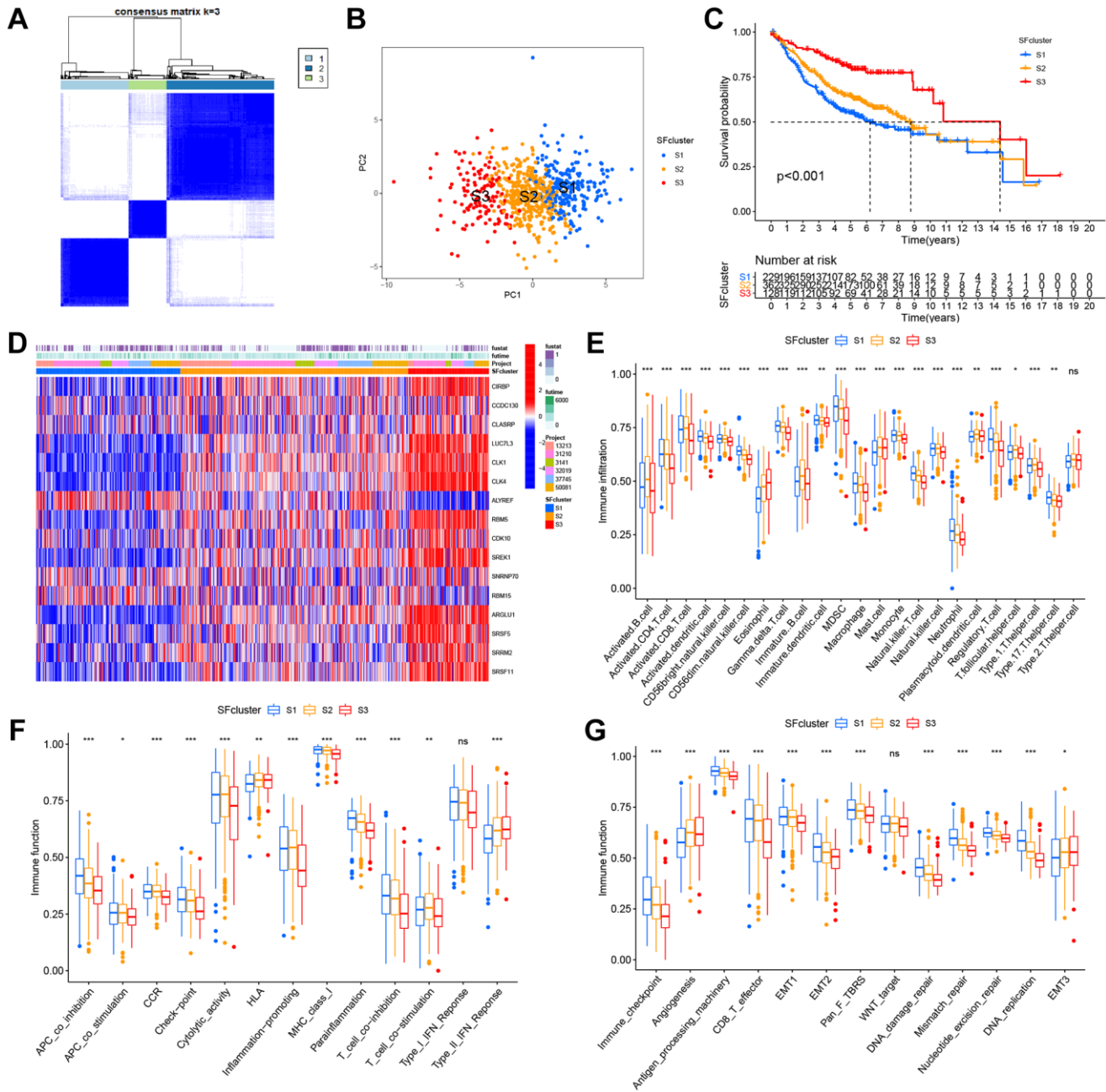
Figure 5. SFs patterns and corresponding TME characteristics. (**A**) Consensus clustering matrix for k = 3. (**B**) Principal component analysis (PCA) was conducted in distinct SFs patterns. (**C**) Survival analyses for three distinct SFs patterns based on six GEO cohorts (GSE13213, GSE37745, GSE31210, GSE3141, GSE30219 and GSE50081). (**D**) Heatmap of 16 SFs expression in LUAD patients. (**E**) TME cell infiltrating abundances in three SFs clusters. (**F**) Difference of immune functions in three SFs clusters. (**G**) Difference of other tumor-related biological processes.

suppression. Significant differences in TME cell-infiltrating composition and immune function were identified between clusters, and indicated these SFs had critical roles in TME remodeling. LUAD samples in cluster-S3 had a significantly lower abundance of all immune cells and immune functions, but a high abundance of eosinophils, mast cells, HLA, and type-II interferon responses. In contrast, clusters-S1 and S2 had a completely opposite landscape in terms of infiltrating cells and immune functions, therefore, cluster-S3 was classified as an immune "cold" phenotype, while Clusters S1 and S2 were immune "hot" phenotypes.

The three SF clusters also exhibited significant differences in distinct tumor-related functions from published research (Figure 5G). Cluster-S1 showed the highest level of most functions, excluding angiogenesis and EMT3, which were enriched in cluster-S3. In contrast, cluster-S3 had the lowest level of most functions, similar to above immune cells and functions, but angiogenesis and EMT3.

Specific correlations between TME infiltrating cell type and each SF were further examined using Spearman's correlation analyses (Supplementary Figure 4). These SFs displayed mainly negative correlations with immune cell-infiltration abundance, except for *ALYREF* and *RBM15,* which were prominently and positively correlated with activated CD4+ T cells, gamma delta T cells, and CD56dim natural killer cells. Most SFs were related to the infiltration of T cells, including activated CD4+ T cells, gamma delta T cells, CD56dim natural killer cells, natural killer T cells, and regulatory T cells.

SF cluster-related DEGs in LUAD

Although LUAD patients were stratified into three SF patterns according to 16 SFs, the underlying genetic alterations were unknown. Therefore, we examined transcriptional changes among different SF patterns and identified 4,819 SF pattern-related DEGs in "limma" (Figure 6A). In total, 2,959 SF-related DEGs were screened out with prognosis significance based on Cox regression analyses. A gene expression heatmap is shown (Supplementary Figure 5A). We next performed unsupervised clustering analyses based on these 2,959 genes to identify different genomic subtypes. We then defined these three phenotypes as gene clusters A, B respectively (Figure 6B, 6C). Subsequent survival analyses showed a prominent prognosis advantage in cluster B (Figure 6D). We also examined the differential expression of the 16 SFs in these gene clusters (Supplementary Figure 5B); significant SF expression differences were observed - cluster B had a lower abundance of all immune cell types and immune functions, with a relevant higher abundance of

eosinophils, mast cells, HLA and type-II interferon responses, while cluster A showed the opposite. Therefore, clusters A and B were classified as immune "hot" and "cold" phenotypes, respectively.

The role of SF clusters in TME immune landscape remodeling

Consistent with immune phenotypes in SF clusters S1–S3, SF gene clusters A–C were characterized as immune-inflamed, immune-desert, and immune-excluded landscapes, respectively (Figure 6E, 6F). SF gene cluster B was identified with a relatively lower immune function and infiltrating abundance of most cells, and could be classified as an immune-inhibition and immune-desert group. In contrast, cluster A was significantly associated with an immune-inflamed status. Cluster C was defined as an immune-excluded phenotype. Similarly, SF gene clusters A–C exhibited a consistent trend in distinct tumor-related functions (Figure 6G).

Construction of the SFscore

Our results showed the non-negligible regulatory role of prognostic AS related SFs in shaping immune landscapes. Nevertheless, considering individual SF heterogeneity and complexity, a scoring system was required to accurately predict SF patterns in individual LUAD patients. Based on SF cluster-related DEGs, we developed SFscore to quantify SF patterns in patients. SF cluster-S3 and gene cluster B had the lowest SFscore (Figure 7A, 7B), thus, low scores were closely linked to immune-inhibition and immune-desert related clusters.

Subsequently, LUAD patients were divided into low or high SFscore groups using a cutoff value determined by the survminer package. A high SFscore was associated with a prominent survival benefit in GEO and TCGA cohorts (Figure 7C, 7D). Similarly, LUAD patients in SF cluster-S3 and gene cluster B had a better prognosis (Figure 7E). Moreover, we identified a prominent correlation between the SFscore and the risk score of a prognostic AS signature (Supplementary Figure 6A, 6B). Thus, a high SFscore was related to a low level AS risk score.

To further understand SFscore characteristics, we performed GSVA (Figure 7F) and showed that patients with high SFscore focused on the inhibition of immune pathways, carcinogenic pathways, and stromal pathways, including interferon gamma/alpha responses, IL-6-JAK-STAT3 signaling, allograft rejection, inflammatory responses, MYC signaling, E2F signaling, mTORC1 signaling, and EMT. Interestingly, bile acid metabolism was activated in the high SFscore group.

We also examined correlations between TME cell infiltration abundance and SFscore (Supplementary Figure 6C). Interestingly, SFscore had significantly negative correlations with most immune infiltration cells, particularly activated CD4+ T cells, CD56dim natural killer cells, gamma delta T cells, MDSCs, natural killer T cells, and regulatory T cells

(Supplementary Figure 6C). Similarly, ssGSEA showed a low abundance of most infiltrating immune cells and immune functions in groups with a high SFscore, except eosinophils and type-II interferon responses (Figure 7G, 7H). Also, a high SFscore was significantly related to lower tumor-related functions, including antigen processes, CD8 effectors, EMT, pan-F-TBRS, Wnt
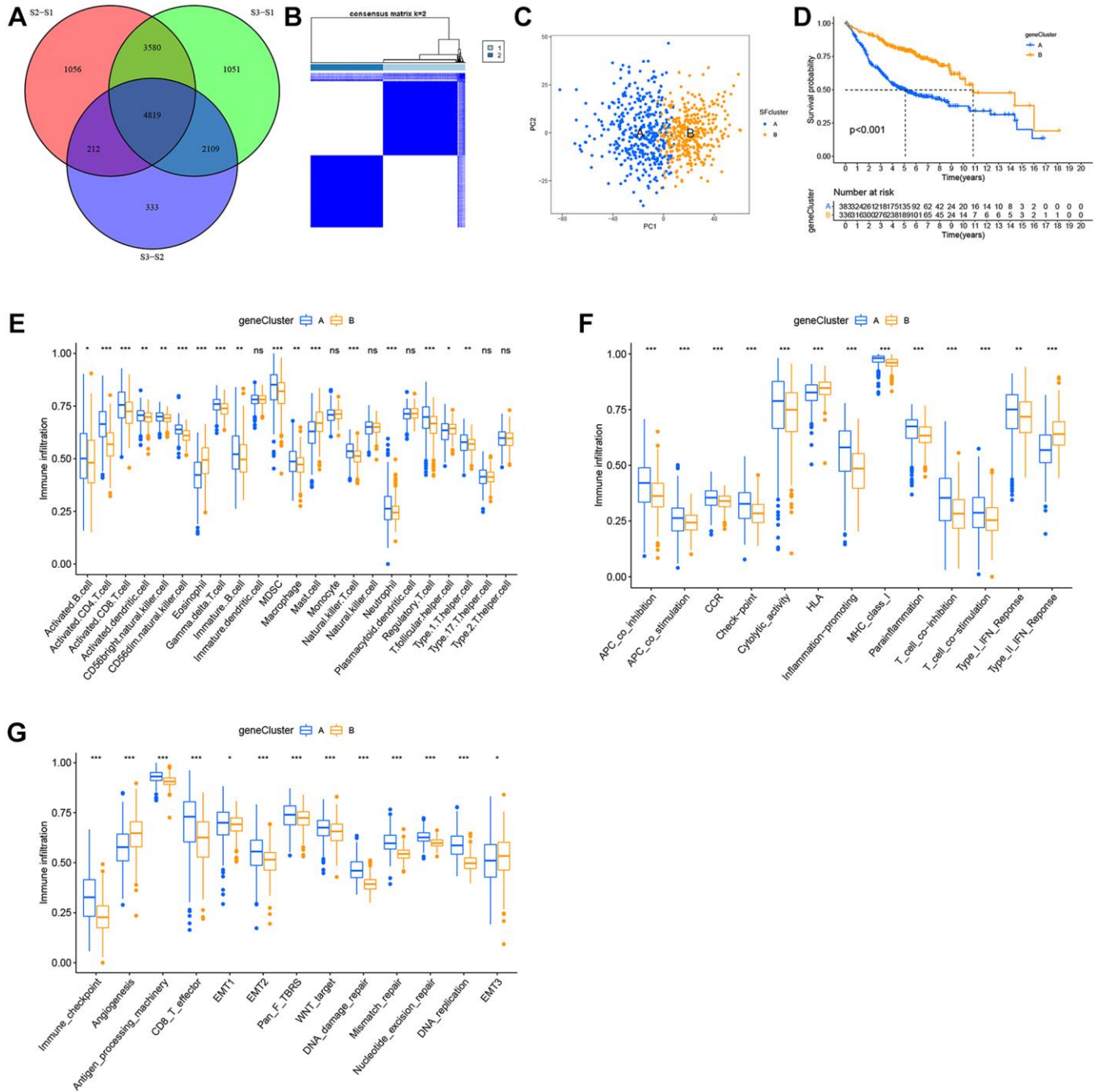
Figure 6. Construction of SFs signatures. (**A**) 4819 SFs-related differentially expressed genes (DEGs) between three distinct SFs patterns were presented in the Venn diagram. (**B**) Consensus clustering matrix for k = 3. (**C**) Principal component analysis (PCA) for the transcriptome profiles of gene cluster A, B. (**D**) Survival analysis for the two gene clusters based on 2959 SFs-related DEGs in GEO cohorts (*P* < 0.001). (**E**) TME cell infiltrating abundances in two gene clusters. (**F**) Difference of immune functions in two gene clusters. (**G**) Difference of other tumor-related biological processes.

targets, DNA replication, DDR, mismatch repair, and nucleotide excision repair (Figure 7I). Nevertheless, angiogenesis was activated in groups with high SFscore. Finally, patients with high SFscore had low immune scores, stromal scores, ESTIMATE scores, and high tumor purity (Supplementary Figure 6D). In general, high SFscore were identified with immune-desert and stromal-inhibited landscapes, i.e., the immune "cold" group.
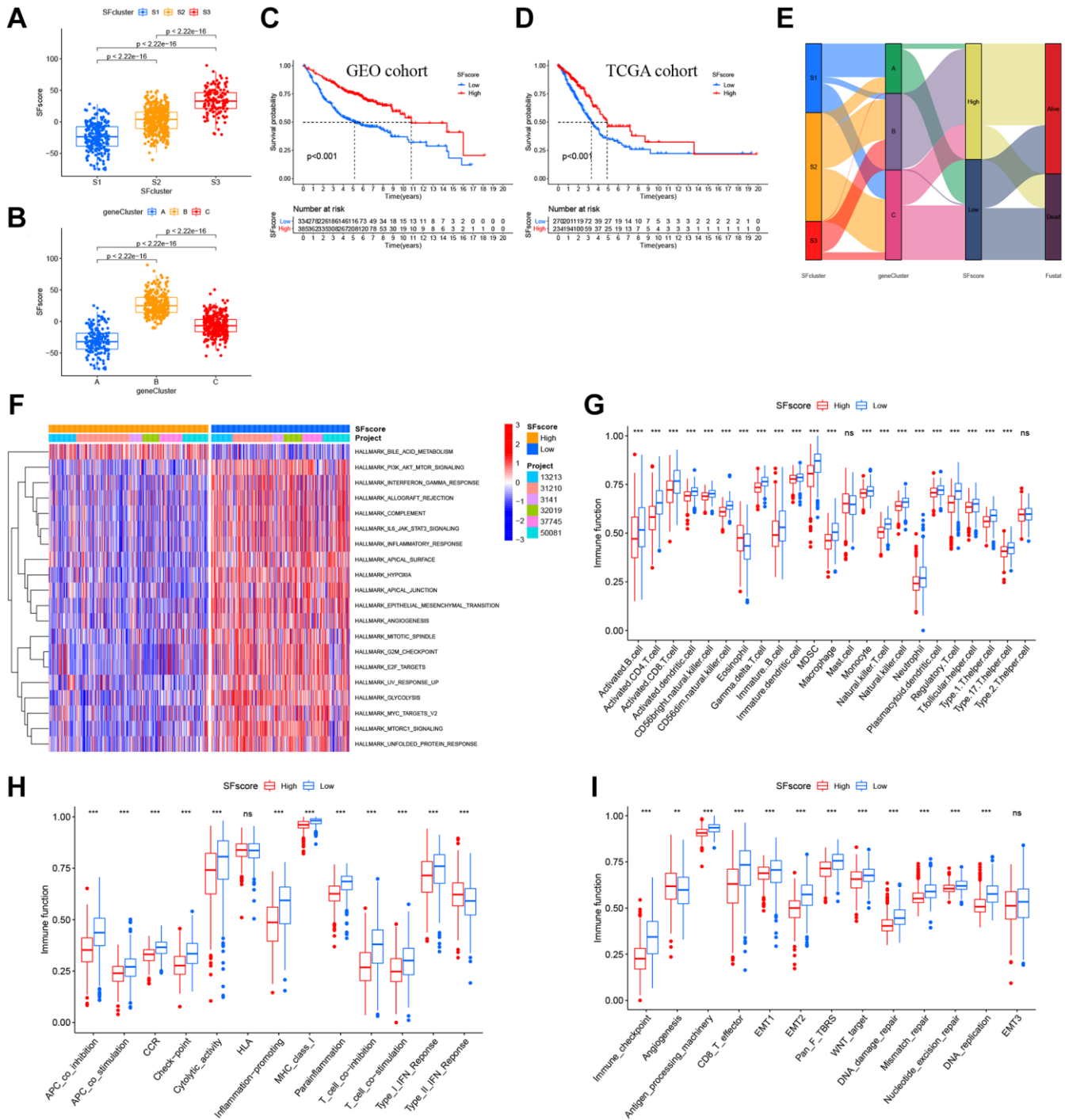
Figure 7. Characteristics of SFscore in prognosis and TME landscapes. (**A**) Differences in the SFscore between three SF clusters in LUAD ($P < 2.22e-16$). (**B**) Differences in the SFscore between two gene clusters in LUAD ($P < 2.22e-16$). (**C**) Kaplan-Meier curves for low and high SFscore patient groups in GEO cohorts ($P < 0.001$, Log-rank test). (**D**) Kaplan-Meier curves for low and high SFscore patient groups were validated in TCGA cohorts ($P < 0.001$, Log-rank test). (**E**) Alluvial diagram showing the changes of survival status, SFs clusters, gene clusters and SFscore. (**F**) GSVA enrichment analyses between groups with low/high SFscore. (**G**) TME cell infiltrating abundances in low/high SFscore groups. (**H**) Difference of immune functions in low/high SFscore groups. (**I**) Difference of other tumor-related biological processes in low/high SFscore groups.

We also performed correlation analyses between SFscore and SF-related prognostic AS events in the network (Supplementary Figure 6A). Four AS events in two genes exhibited significant correlations with SFscore, including NEK2|9717|AT (cor = −0.65), NEK2|9718|AT (cor = 0.65), MCM7|80880|AP (cor = 0.63), and MCM7|80881|AP (cor = −0.63). Interestingly, the four AS events showed strong correlations with each other, and a distinct AS in a single gene showed an opposite correlation with the SFscore. Consistently, (Supplementary Figure 6B) these four AS events were significantly associated with multi-type immune cells, in particular activated CD4 T cells (cor > 0.5). Also, gamma delta T cells were widely associated with these selected AS events.

Characteristic SFscore in tumor somatic mutations

Several studies have highlighted associations between tumor somatic mutations and immunotherapeutic response [1, 44]. We explored the distribution patterns of tumor mutation burden in distinct SFscore groups and found that the low SFscore group had a higher TMB (Figure 8A). Additionally, the SFscore had a predominantly negative correlation with TMB across different gene clusters (Figure 8B). Survival analysis showed prominent differences among different TMB and SFscore combinations (Figure 8C). We next performed significantly mutated gene (SMG) analysis on LUAD samples in low SFscore versus high SFscore subgroups. LUAD samples with a low SFscore had significantly higher SMG rates (Figure 8D, 8E) and provided novel insights on SFscore in tumor somatic mutations, TME remodeling, and ICI therapy.

**The role of SFscore in predicting immunotherapeutic benefits**

Immunotherapies, represented by ICP blockade, have emerged as major breakthroughs in cancer therapy.
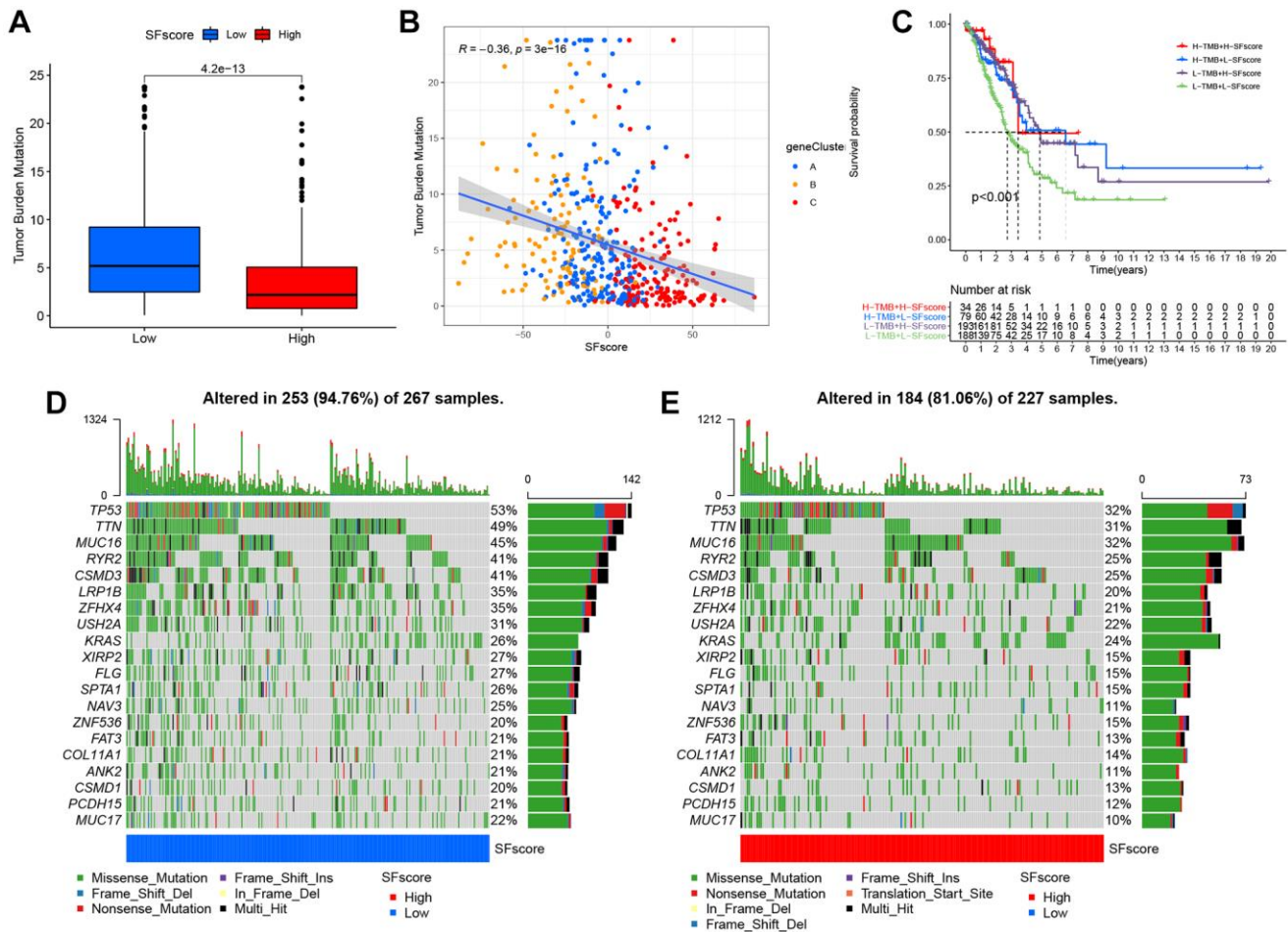
Figure 8. Characteristics of SFscore in tumor somatic mutation. (**A**) Differences in tumor mutation burden (TMB) between low and high SFscore groups of TCGA cohort (*P* < 0.001). (**B**) Correlations between SFscore and TMB using Spearman analysis (R = −0.36, *P* < 0.001). (**C**) Survival analyses for subgroup patients stratified by both SFscore and TMB using Kaplan-Meier curves. H-TMB, high TMB; L-TMB, low TMB (*P* < 0.001, Log-rank test). (**D**) The waterfall plot of tumor somatic mutation established by those with low SFscore. (**E**) The waterfall plot of tumor somatic mutation based on LUAD samples with high SFscore.

Currently, immunotherapy is considered a first-line treatment for patients with LUAD. We extracted 42 ICPs, including PD-1, PD-L1, and CTLA-4; the mRNA expression levels of most ICPs from low SFscore subtypes were significantly higher than those in high

SFscore subtypes (Figure 9A). In addition to well-known ICPs, newly identified predictors, such as TIDE, are widely used and strongly recommended to evaluate immune response. We showed (Figure 9B–9D) that low SFscore patients were distinguished by a high level of
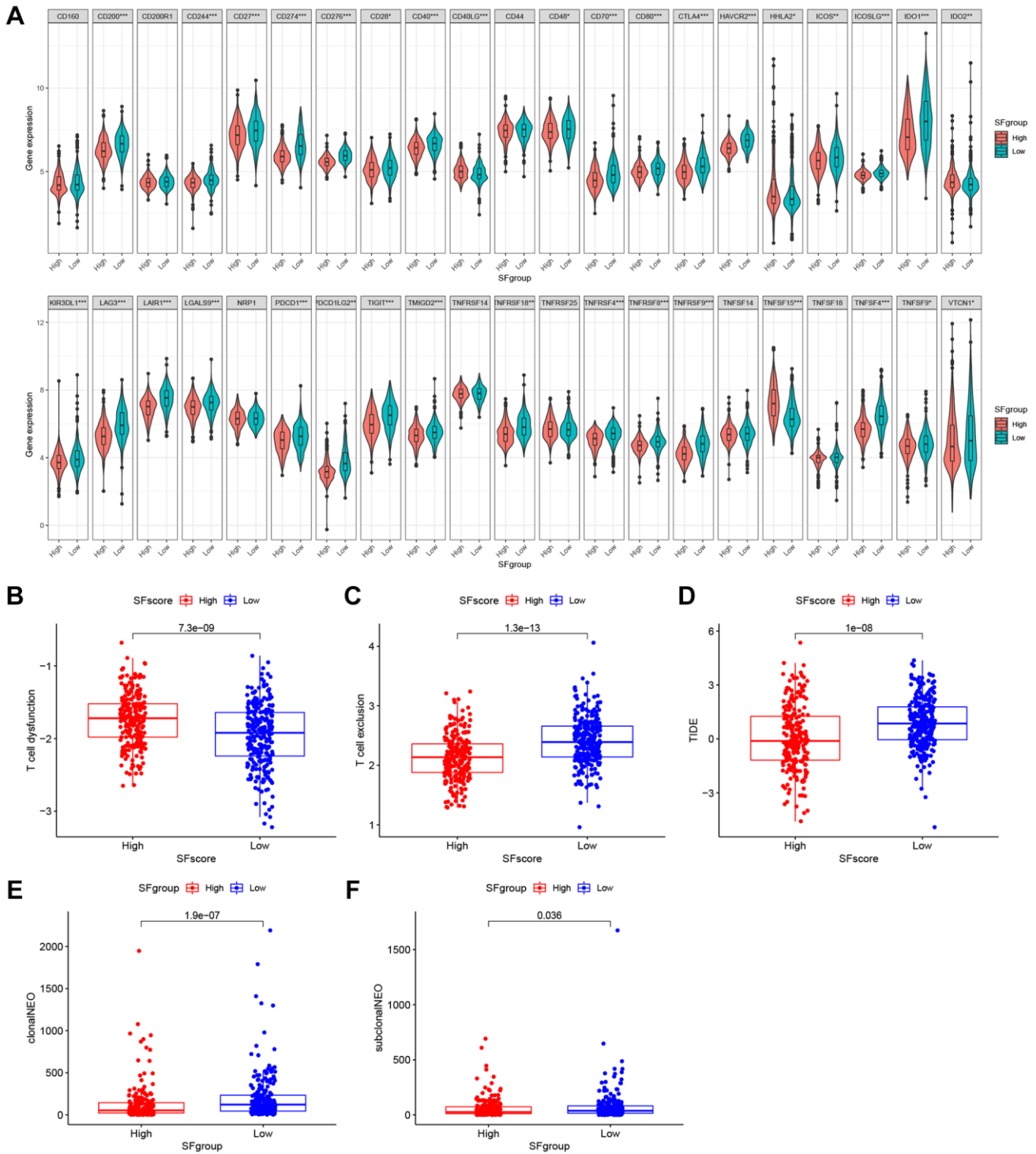

Figure 9. The SFscore predicts immunotherapeutic benefits. (**A**) The relative mRNA expression level of 42 immune checkpoints was compared between SFscore high versus low groups. (**B–D**) The relative distribution of T cell exclusion score, T cell dysfunction score and TIDE score were compared between SFscore high versus low groups in TCGA-LUAD cohort ($P < 0.001$). (**E**, **F**) The relative distribution of clonal neoantigens and sub-clonal neoantigens were compared between SFscore high versus low groups in TCGA-LUAD cohort ($P < 0.05$).

T cell exclusion scores, TIDE and low T cell dysfunction scores. Tumor neoantigen burden, which is closely related to immunotherapeutic efficacy, was also assessed; LUAD patients with a low SFscore had a higher clonal and sub-clonal neoantigen burden (Figure 9E, 9F). The distinct distribution of these markers indicated that low SFscore patients could benefit from immunotherapy, especially ICIs. Thus, a crucial role for

SFscore in mediating immune responses in patients was identified.

Drug sensitivity analysis and small molecule drug screening based on SFscore

We identified four chemotherapy drugs and estimated IC$_{50}$ levels. As shown (Figure 10A), all four exhibited
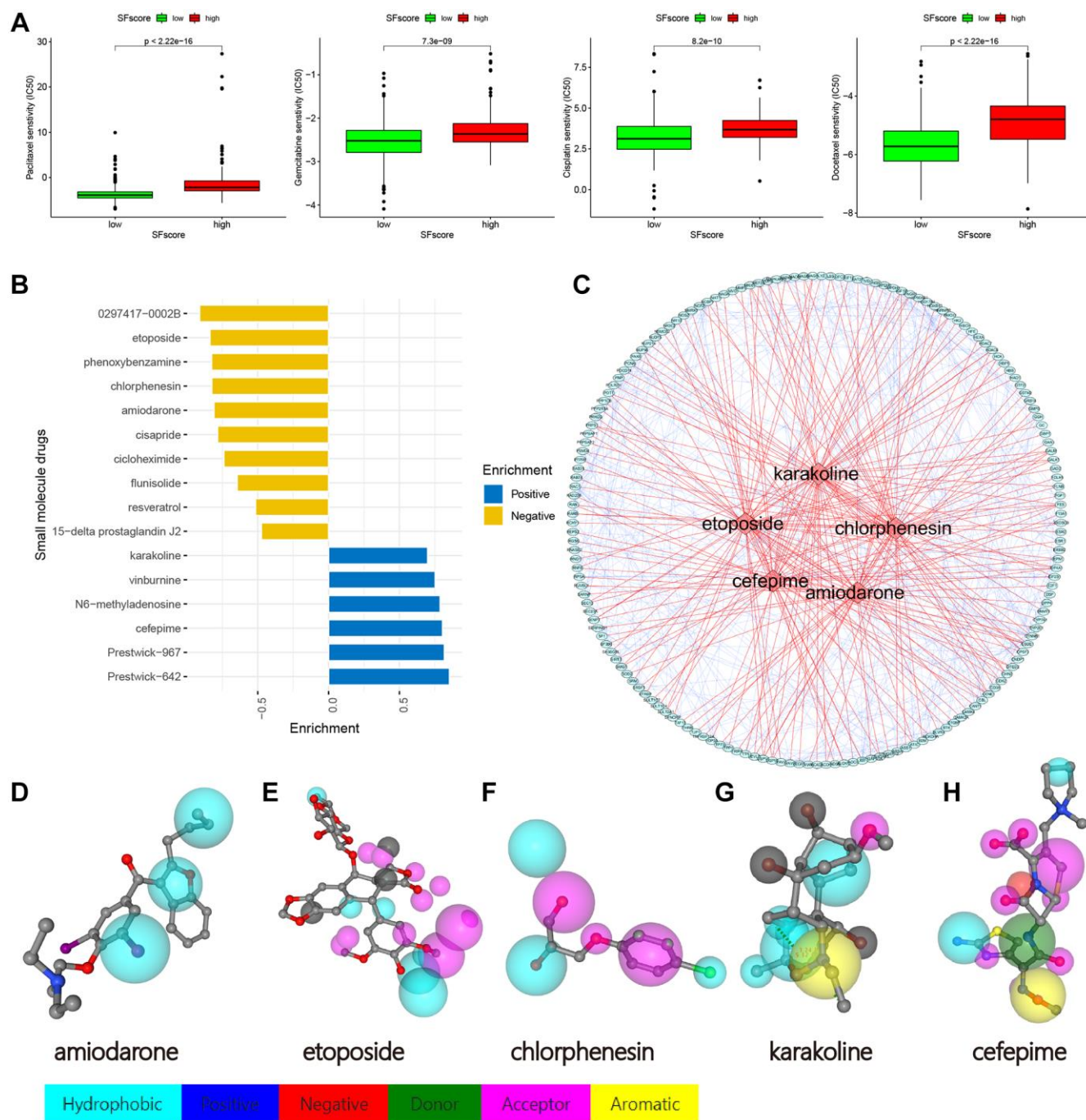
Figure 10. Drug sensitivity analysis and small molecule drugs screening. (**A**) Low SFscore is more sensitive to sorafenib, paclitaxel, sunitinib, cisplatin, docetaxel, Etoposide, vinorelbine, gemcitabine, gefitinib and vorinostat ($p < 0.001$). (**B**) Screening of small molecule drugs based on SFscore. (**C**) Network construction of target protein and five small molecule drugs. (**D–H**) Druggable pharmacophore models of five small molecule drugs.

significantly low IC$_{50}$ levels in the low SFscore group, including paclitaxel ($p < 2.22e-16$), cisplatin ($p < 8.2e-10$), docetaxel ($p < 2.22e-16$), and gemcitabine ($p < 7.3e-09$). Thus, low SFscore patients were more sensitive to the ten selected drugs. To further screen for specific drugs for patients with distinct SFscore, we analyzed DEGs between groups with distinct SFscore, including 111 up- and 106 down-regulated genes. DEGs were uploaded to the CMap database, and 16 small molecule drugs were identified ($p < 0.01$, |mean|>0.4) (Figure 10B). Among these, ten drugs targeting lower SFscore patients were selected with negative enrichment and six small molecule drugs were selected for higher SFscore patients (with positive enrichment). Amiodarone ($p = 0.00056$, enrichment = −0.809), etoposide ($p = 0.00123$, enrichment = −0.837), chlorphenesin ($p = 0.00185$, enrichment = −0.824), karakoline ($p = 0.00201$, enrichment = −0.699), and cefepime ($p = 0.00277$, enrichment = 0.804) were identified, and a network constructed to identify drug interactions with target proteins (Figure 10C). Also, druggable pharmacophore models of these five aforementioned drugs were analyzed and visualized in PharmMapper (Figure 10D–10H).

DISCUSSION

While the LUAD incidence in NSCLC is increasing, the treatment did not receive promising effect [9, 10]. TME has a vital role in LUAD tumorigenesis, progression, metastasis, and therapeutic responses [14]. In the lung TME, tumor-infiltrating stromal cells are reprogrammed by malignant cancer cells, which in turn contribute to carcinogenesis [14]. However, the precise mechanistic details remain elusive.

Practically all protein-coding genes undergo one or more AS processes, including ES, AT, AP, AA, AD, RI, and ME. These AS forms require the spliceosome, which includes small nuclear ribonucleoprotein molecules (snRNPs, U1, U2, U4/U6, and U5) and other proteins, to generate different protein isoforms [45, 46]. SFs bind to pre-mRNAs which activate AS processes, to ultimately regulate cell function [47, 48].In our study, we constructed an AS-SF interaction network to identify SFs associated with AS events, and summarize interactions between OS-related AS events and SFs. This network included 12 adverse AS events (red nodes), 28 favorable AS events (blue nodes), and 20 interacting SFs (black nodes) (Figure 3), which suggested these SFs may be promising therapeutic targets.

It was reported that RBPs include two families: serine/arginine-rich proteins and heterogeneous RNPs (hnRNPs) which promote exon inclusion and exon skipping, respectively [46, 49]. Also, recurrent somatic mutations in SF genes may also directly affect cancer progression. Patients carrying the *SRRM2* missense genetic variant exhibited mis-splicing in specific cassette exons, while the variant segregated with familial papillary thyroid carcinoma [8, 50]. *RBM5, RBM6,* and *RBM10* are commonly deleted, mutated, and/or under-expressed or overexpressed genes in many cancers [8, 44, 51]; however, they have different roles in *in vitro* colony formation assays, partially due to their antagonistic regulation of NUMB alternative splicing [51].

Our LUAD data were divided into three patterns based on the expression of 16 SFs (Figure 5A, 5B). These patterns showed significant differences in OS (Figure 5C) and oncogenic pathways (Figure 5A, 5B), and indicated the important regulatory role of SFs in cancer progression.

Importantly, we matched these patterns with three major immunophenotypes, immune-inflamed, immune-desert, and immune-excluded phenotypes, based on the abundance of different immune cell types and immune functions (Figure 5E, 5F). Clusters S1−3 were classified as immune-inflamed, immune-excluded, and immune-desert phenotypes, respectively. Among clusters, there were apparently differences whether TME cell infiltration compositions or immune functions, which indicated the vital role of SFs in TME remodeling.

Previous studies demonstrated that TME contexture, including tumor-infiltrating CD4+/CD8+ T cells, macrophage M1, natural killer cells, and inflammatory cytokines, played vital roles in tumor progression and immunotherapeutic efficacy [18, 52]. We also showed that S3 patterns were significantly related to lower tumor-infiltrating CD8+ T cells, macrophage while elevated tumor-infiltrating eosinophil, HLA and type-II interferon response, which supported the potential value on immunotherapy and tumor progression.

To explore the underlying mechanisms of TME remodeling as mediated by SFs, we filtered all DEGs between the three SF patterns and screened out 2,959 DEGs which were significantly associated with OS (Figure 6D, Supplementary Figure 5A). Based on these DEGs, a different genomic subtype was identified (Figure 6B, 6C). Two transcriptomic subtypes, based on SF pattern genes, were significantly associated with different survival outcomes and TME landscape. Interestingly, similar to SF patterns, SF cluster-S3 and gene cluster B, classified as immune-desert and immune "cold" phenotypes, respectively, demonstrated better prognoses. As shown (Figure 5G and Figure 6G), these clusters were associated with many cancer-related

phenotypes, including EMT and Wnt targets. Also, we observed significant differences in immune characteristics between SF clusters; however, this did not mean no immune cell infiltration or immune-related cell factors were present (Figure 5E, 5F and Figure 6E, 6F). In fact, the three immune infiltration phenotypes, which represented the degree of immune cell infiltration in this study, was a relative description. Also, the cancer-immune cycle assumes that the best anti-cancer immune response depends on cooperation between immune cells, host factors, and tumor antigens, and that this cycle must be efficiently initiated and precisely maintained, otherwise, the cancer-immunity cycle is broken and replaced by inflammation-promoted tumor progression [53–55]. In our study, the inflammation potentially occurring in relative immune "hot" phenotypes also exerted part of the immunosuppressive function, such as the up-regulation of immuno-suppressive cells MDSC, Treg, etc. This may cause dysregulated anti-tumor immunity and mediate malignant progression. To a certain extent, this explains why SF cluster-S3 and gene cluster B had a relatively better prognosis.

We also constructed a quantification system "SFscore" to define different SF patterns and evaluate therapeutic effects and outcomes for LUAD patients. Our analyses highlighted that the SFscore was a prognostic biomarker for LUAD. Consequently, the immune "cold" phenotype showed the highest SFscore indicating a relatively low risk, while patterns characterized by the immune "hot" phenotype showed lower SFscore but with higher risk. As shown (Supplementary Figure 6D), the group with the highest SFscore was accompanied by a lower ImmuneScore, StromalScore, ESTIMATE score, and higher TumorPurity, indicating a lower level of immune cell infiltration. Similar to the SF cluster-S3 in hallmark LUAD, most signaling was downregulated in the high SF score group, including IL6/JAK/STAT3 signaling and inflammatory responses, and this downregulation was consistent with immune-desert characteristics.

It was reported that eosinophils have antitumorigenic roles in many cancers, including breast, colorectal, esophageal, and gastric cancers [56]. Activated eosinophils inhibited prostate cancer cell growth *in vitro* by secreting interleukin-10 (IL-10) and IL-12, and increasing E-cadherin expression, which putatively suppressed metastatic seeding [57]. Also, eosinophil-mediated cytotoxicity was reported in several co-culture studies of mouse or human eosinophils grown with cancer cells, including hepatocellular carcinoma cells, fibrosarcoma, melanoma, and CRC cells [58–61]. Interestingly, several mediators were shown to augment eosinophil-mediated killing, including interferon-γ [56].

In our study, type-II interferon was elevated in SF cluster-S3 and gene cluster B, with high SFscore and low risk, and bound the interferon-γ receptor to activate cell signaling pathways (involving JAK/STAT signaling). Consistently, human and mouse *in vitro* eosinophil activation with interferon-γ (but not TLR ligands) potentiated the eosinophil-mediated killing of CRC cells. Other studies suggested that interferon-γ and interferon-γ-induced pathways functioned as key regulators of eosinophil antitumorigenic activities [62, 63]. A collective view of these data indicated that eosinophil and type-II interferon had vital roles in antitumorigenic processes and, to a certain extent, explained the better prognoses in the group with elevated eosinophil and type-II interferon levels [56].

It was reported that gamma delta T cells had important roles in the development and progression of lung cancer and local microbiota by activating lung-resident gamma delta T cells and promoting inflammation associated with LUAD [64, 65]. One interesting finding from our study (Supplementary Figure 6B) showed that gamma delta T cells were negatively associated with all selected AS events. In fact, some AS events, such as NEK2|9717|AT and MCM7|80881|AP [66], are the most common form of gene expression that can perform right functions, and in this way, we may conclude that gamma delta T cell is closely related to gene AS and its function maybe affected by the special form of gene splicing. Interferon-γ induces CD8+ T cells to antigen (Ag)-specific CTLs and CD4+ T cell differentiation [67, 68], however, continuous exposure to interferon-γ may induce T cell exhaustion and tumor progression [69] by inducing immune escape mediators, including PD-L1, STAT3, and IDO1 [70, 71].

Recent studies also identified associations between high TMB and clinical benefit in NSCLC, melanoma, and bladder cancer patients when treated with PD-1/PD-L1 inhibitors or CTLA4 blockade [1, 44]. We analyzed correlations between TMB and SFscore; a higher SFscore was associated with lower TMB (Figure 8A, 8B) and the TMB of almost all common mutated genes in lung cancer decreased in the high SFscore cohort (Figure 8D, 8E). The L-TMB+L-SFscore cohort had the worst prognosis (Figure 8C) and no significant differences were identified between H-TMB+H-SFscore and L-TMB+H-SFscore cohorts. Therefore, TMB may be effective at predicting survival benefit in groups with relatively high degrees of immune infiltration, but not in the immune-desert phenotype.

When compared with the low SFscore cohort, almost all ICPs were decreased in the high SFscore cohort, and accompanied by low T cell exclusion, higher T cell dysfunction, and lower TIDE scores (Figure 9). These

data suggested that SFscore may provide critical guidelines for clinical immunotherapy.

Apart from immunotherapy, we also investigated correlations between SFscore and the effects of other common chemotherapy drugs; we observed significantly different effects between high and low SFscore cohorts ($p < 0.001$).

Importantly, we identified different tumor immune phenotypes based on different AS events, provided new insights for improving patient clinical responses to immunotherapy, and promoted personalized cancer immunotherapy in the future.

CONCLUSIONS

We demonstrated some of the extensive regulatory mechanisms underpinning AS events in the TME. Difference in AS events and SF patterns between patients are factors that cannot ignored and may cause TME heterogeneity and complexity in individuals. A comprehensive evaluation of individual tumor AS events and SF patterns will enhance our understanding of TMEs and guide more effective therapeutic strategies for lung cancer.

Abbreviations

AS: Alternative splicing; NSCLC: Non-small Cell Lung Cancer; LUAD: Lung Adenocarcinoma; AA: Alternate Acceptor site; AD: Alternate Donor site; AP: Alternate Promoter; AT: Alternate Terminator; ES: Exon Skip; ME: Mutually Exclusive exons; RI: Retained Intron; TME: Tumor Microenvironment; ICI: Immune Check-point Inhibitors; DEGs: Differentially Expressed Genes; ICP: Immune Checkpoint; PCA: Principal Component Analysis; GSVA: Gene Set Variation Analysis; CNV: Copy Number Variations; TMB: Tumor Mutation Burden; IPS: Immunophenoscore; TIDE: Tumor Immune Dysfunction and Exclusion; ROC: Receiver Operating Characteristic; RFE: Recursive Feature Elimination; MAF: Mutation Annotation Format; CTLs: Cytotoxic T Lymphocytes; EMT: Epithelial-Mesenchymal Transition; SMG: Significantly Mutated Gene.

AUTHOR CONTRIBUTIONS

Jiajun Du and Chao Wang contributed to the designing and supervising the study, and correspondence. Co-authors Jichang Liu and Yadong Wang analyzed the data and completed the manuscript. Xiaogang Zhao, Kai Wang, helped to search for references and offered guidelines of statistical methods. All authors have read and approved the final version to be published.

REFERENCES

1. Samstein RM, Lee CH, Shoushtari AN, Hellmann MD, Shen R, Janjigian YY, Barron DA, Zehir A, Jordan EJ, Omuro A, Kaley TJ, Kendall SM, Motzer RJ, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. Nat Genet. 2019; 51:202–6.
https://doi.org/10.1038/s41588-018-0312-8
PMID:30643254

2. Keren H, Lev-Maor G, Ast G. Alternative splicing and evolution: diversification, exon definition and function. Nat Rev Genet. 2010; 11:345–55.
https://doi.org/10.1038/nrg2776
PMID:20376054

3. Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. Alternative isoform regulation in human tissue transcriptomes. Nature. 2008; 456:470–6.
https://doi.org/10.1038/nature07509
PMID:18978772

4. Bonnal SC, López-Oreja I, Valcárcel J. Roles and mechanisms of alternative splicing in cancer - implications for care. Nat Rev Clin Oncol. 2020; 17:457–74.
https://doi.org/10.1038/s41571-020-0350-x
PMID:32303702

5. Sciarrillo R, Wojtuszkiewicz A, Assaraf YG, Jansen G, Kaspers GJL, Giovannetti E, Cloos J. The role of alternative splicing in cancer: From oncogenesis to drug resistance. Drug Resist Updat. 2020; 53:100728.

https://doi.org/10.1016/j.drup.2020.100728
PMID:33070093

6. Jiménez M, Urtasun R, Elizalde M, Azkona M, Latasa MU, Uriarte I, Arechederra M, Alignani D, Bárcena-Varela M, Álvarez-Sola G, Colyn L, Santamaría E, Sangro B, et al. Splicing events in the control of genome integrity: role of SLU7 and truncated SRSF3 proteins. Nucleic Acids Res. 2019; 47:3450–66.
https://doi.org/10.1093/nar/gkz014
PMID:30657957

7. 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

8. Dvinge H, Kim E, Abdel-Wahab O, Bradley RK. RNA splicing factors as oncoproteins and tumour suppressors. Nat Rev Cancer. 2016; 16:413–30.
https://doi.org/10.1038/nrc.2016.51
PMID:27282250

9. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. CA Cancer J Clin. 2021; 71:7–33.
https://doi.org/10.3322/caac.21654
PMID:33433946

10. Huang C, Qu X, Du J. Proportion of lung adenocarcinoma in female never-smokers has increased dramatically over the past 28 years. J Thorac Dis. 2019; 11:2685–8.
https://doi.org/10.21037/jtd.2019.07.08
PMID:31463093

11. Coomer AO, Black F, Greystoke A, Munkley J, Elliott DJ. Alternative splicing in lung cancer. Biochim Biophys Acta Gene Regul Mech. 2019; 1862:194388.
https://doi.org/10.1016/j.bbagrm.2019.05.006
PMID:31152916

12. Hirsch FR, Scagliotti GV, Mulshine JL, Kwon R, Curran WJ Jr, Wu YL, Paz-Ares L. Lung cancer: current therapies and new targeted treatments. Lancet. 2017; 389:299–311.
https://doi.org/10.1016/S0140-6736(16)30958-8
PMID:27574741

13. Aleksakhina SN, Kashyap A, Imyanitov EN. Mechanisms of acquired tumor drug resistance. Biochim Biophys Acta Rev Cancer. 2019; 1872:188310.
https://doi.org/10.1016/j.bbcan.2019.188310
PMID:31442474

14. Altorki NK, Markowitz GJ, Gao D, Port JL, Saxena A, Stiles B, McGraw T, Mittal V. The lung microenvironment: an important regulator of tumour growth and metastasis. Nat Rev Cancer. 2019; 19:9–31.
https://doi.org/10.1038/s41568-018-0081-9
PMID:30532012

15. Hegde PS, Chen DS. Top 10 Challenges in Cancer Immunotherapy. Immunity. 2020; 52:17–35.
https://doi.org/10.1016/j.immuni.2019.12.011
PMID:31940268

16. Steven A, Fisher SA, Robinson BW. Immunotherapy for lung cancer. Respirology. 2016; 21:821–33.
https://doi.org/10.1111/resp.12789
PMID:27101251

17. Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. Nature. 2017; 541:321–30.
https://doi.org/10.1038/nature21349
PMID:28102259

18. Galon J, Bruni D. Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. Nat Rev Drug Discov. 2019; 18:197–218.
https://doi.org/10.1038/s41573-018-0007-y
PMID:30610226

19. 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

20. Yu S, Hu C, Liu L, Cai L, Du X, Yu Q, Lin F, Zhao J, Zhao Y, Zhang C, Liu X, Li W. Comprehensive analysis and establishment of a prediction model of alternative splicing events reveal the prognostic predictor and immune microenvironment signatures in triple negative breast cancer. J Transl Med. 2020; 18:286.
https://doi.org/10.1186/s12967-020-02454-1
PMID:32723333

21. Qu S, Jiao Z, Lu G, Yao B, Wang T, Rong W, Xu J, Fan T, Sun X, Yang R, Wang J, Yao Y, Xu G, et al. PD-L1 lncRNA splice isoform promotes lung adenocarcinoma progression via enhancing c-Myc activity. Genome Biol. 2021; 22:104.
https://doi.org/10.1186/s13059-021-02331-0
PMID:33849634

22. Tripathi V, Shin JH, Stuelten CH, Zhang YE. TGF-β-induced alternative splicing of TAK1 promotes EMT and drug resistance. Oncogene. 2019; 38:3185–200.
https://doi.org/10.1038/s41388-018-0655-8
PMID:30626936

23. Frankiw L, Baltimore D, Li G. Alternative mRNA splicing in cancer immunotherapy. Nat Rev Immunol. 2019; 19:675–87.
https://doi.org/10.1038/s41577-019-0195-7
PMID:31363190

24. Piva F, Giulietti M, Burini AB, Principato G. SpliceAid 2: a database of human splicing factors expression data and RNA target motifs. Hum Mutat. 2012; 33:81–5.
https://doi.org/10.1002/humu.21609
PMID:21922594

25. Tomida S, Takeuchi T, Shimada Y, Arima C, Matsuo K, Mitsudomi T, Yatabe Y, Takahashi T. Relapse-related molecular signature in lung adenocarcinomas identifies patients with dismal prognosis. J Clin Oncol. 2009; 27:2793–9.
https://doi.org/10.1200/JCO.2008.19.7053
PMID:19414676

26. Jabs V, Edlund K, König H, Grinberg M, Madjar K, Rahnenführer J, Ekman S, Bergkvist M, Holmberg L, Ickstadt K, Botling J, Hengstler JG, Micke P. Integrative analysis of genome-wide gene copy number changes and gene expression in non-small cell lung cancer. PLoS One. 2017; 12:e0187246.
https://doi.org/10.1371/journal.pone.0187246
PMID:29112949

27. Yamauchi M, Yamaguchi R, Nakata A, Kohno T, Nagasaki M, Shimamura T, Imoto S, Saito A, Ueno K, Hatanaka Y, Yoshida R, Higuchi T, Nomura M, et al. Epidermal growth factor receptor tyrosine kinase defines critical prognostic genes of stage I lung adenocarcinoma. PLoS One. 2012; 7:e43923.
https://doi.org/10.1371/journal.pone.0043923
PMID:23028479

28. Bild AH, Yao G, Chang JT, Wang Q, Potti A, Chasse D, Joshi MB, Harpole D, Lancaster JM, Berchuck A, Olson JA Jr, Marks JR, Dressman HK, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature. 2006; 439:353–7.
https://doi.org/10.1038/nature04296
PMID:16273092

29. Sokolove J, Bromberg R, Deane KD, Lahey LJ, Derber LA, Chandra PE, Edison JD, Gilliland WR, Tibshirani RJ, Norris JM, Holers VM, Robinson WH. Autoantibody epitope spreading in the pre-clinical phase predicts progression to rheumatoid arthritis. PLoS One. 2012; 7:e35296.
https://doi.org/10.1371/journal.pone.0035296
PMID:22662108

30. Der SD, Sykes J, Pintilie M, Zhu CQ, Strumpf D, Liu N, Jurisica I, Shepherd FA, Tsao MS. Validation of a histology-independent prognostic gene signature for early-stage, non-small-cell lung cancer including stage IA patients. J Thorac Oncol. 2014; 9:59–64.
https://doi.org/10.1097/JTO.0000000000000042
PMID:24305008

31. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics. 2010; 26:1572–3.
https://doi.org/10.1093/bioinformatics/btq170
PMID:20427518

32. 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

33. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics. 2011; 27:1739–40.
https://doi.org/10.1093/bioinformatics/btr260
PMID:21546393

34. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, Fu X, Liu S, Bo X, Yu G. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (Camb). 2021; 2:100141.
https://doi.org/10.1016/j.xinn.2021.100141
PMID:34557778

35. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, Hackl H, Trajanoski Z. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. Cell Rep. 2017; 18:248–62.
https://doi.org/10.1016/j.celrep.2016.12.019
PMID:28052254

36. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, Kadel EE III, Koeppen H, Astarita JL, Cubas R, Jhunjhunwala S, Banchereau R, Yang Y, et al. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. Nature. 2018; 554:544–8.
https://doi.org/10.1038/nature25501
PMID:29443960

37. Şenbabaoğlu Y, Gejman RS, Winer AG, Liu M, Van Allen EM, de Velasco G, Miao D, Ostrovnaya I, Drill E, Luna A, Weinhold N, Lee W, Manley BJ, et al. Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. Genome Biol. 2016; 17:231.
https://doi.org/10.1186/s13059-016-1092-z
PMID:27855702

38. Zhang B, Wu Q, Li B, Wang D, Wang L, Zhou YL. m6A regulator-mediated methylation modification patterns and tumor microenvironment infiltration characterization in gastric cancer. Mol Cancer. 2020; 19:53.
https://doi.org/10.1186/s12943-020-01170-0
PMID:32164750

39. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res. 2018; 28:1747–56.
https://doi.org/10.1101/gr.239244.118
PMID:30341162

40. 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

41. 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; 24:1550–8.
https://doi.org/10.1038/s41591-018-0136-1
PMID:30127393

42. Geeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. PLoS One. 2014; 9:e107468.
https://doi.org/10.1371/journal.pone.0107468
PMID:25229481

43. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet JP, Subramanian A, Ross KN, Reich M, Hieronymus H, Wei G, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science. 2006; 313:1929–35.
https://doi.org/10.1126/science.1132939
PMID:17008526

44. Imielinski M, Berger AH, Hammerman PS, Hernandez B, Pugh TJ, Hodis E, Cho J, Suh J, Capelletti M, Sivachenko A, Sougnez C, Auclair D, Lawrence MS, et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell. 2012; 150:1107–20.
https://doi.org/10.1016/j.cell.2012.08.029
PMID:22980975

45. 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

46. Chen M, Manley JL. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nat Rev Mol Cell Biol. 2009; 10:741–54.
https://doi.org/10.1038/nrm2777
PMID:19773805

47. Wan L, Yu W, Shen E, Sun W, Liu Y, Kong J, Wu Y, Han F, Zhang L, Yu T, Zhou Y, Xie S, Xu E, et al. SRSF6-regulated alternative splicing that promotes tumour progression offers a therapy target for colorectal cancer. Gut. 2019; 68:118–29.
https://doi.org/10.1136/gutjnl-2017-314983
PMID:29114070

48. Liu F, Dai M, Xu Q, Zhu X, Zhou Y, Jiang S, Wang Y, Ai Z, Ma L, Zhang Y, Hu L, Yang Q, Li J, et al. SRSF10-mediated IL1RAP alternative splicing regulates cervical cancer oncogenesis via mIL1RAP-NF-κB-CD47 axis. Oncogene. 2018; 37:2394–409.
https://doi.org/10.1038/s41388-017-0119-6
PMID:29429992

49. Zhang J, Manley JL. Misregulation of pre-mRNA alternative splicing in cancer. Cancer Discov. 2013; 3:1228–37.
https://doi.org/10.1158/2159-8290.CD-13-0253
PMID:24145039

50. Tomsic J, He H, Akagi K, Liyanarachchi S, Pan Q, Bertani B, Nagy R, Symer DE, Blencowe BJ, de la Chapelle A. A germline mutation in SRRM2, a splicing factor gene, is implicated in papillary thyroid carcinoma predisposition. Sci Rep. 2015; 5:10566.
https://doi.org/10.1038/srep10566
PMID:26135620

51. Bechara EG, Sebestyén E, Bernardis I, Eyras E, Valcárcel J. RBM5, 6, and 10 differentially regulate NUMB alternative splicing to control cancer cell proliferation. Mol Cell. 2013; 52:720–33.
https://doi.org/10.1016/j.molcel.2013.11.010
PMID:24332178

52. Topalian SL, Taube JM, Anders RA, Pardoll DM. Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. Nat Rev Cancer. 2016; 16:275–87.
https://doi.org/10.1038/nrc.2016.36
PMID:27079802

53. Hou J, Karin M, Sun B. Targeting cancer-promoting inflammation - have anti-inflammatory therapies come of age? Nat Rev Clin Oncol. 2021; 18:261–79.
https://doi.org/10.1038/s41571-020-00459-9
PMID:33469195

54. Chen DS, Mellman I. Oncology meets immunology: the cancer-immunity cycle. Immunity. 2013; 39:1–10.
https://doi.org/10.1016/j.immuni.2013.07.012
PMID:23890059

55. Grivennikov SI, Greten FR, Karin M. Immunity, inflammation, and cancer. Cell. 2010; 140:883–99.
https://doi.org/10.1016/j.cell.2010.01.025
PMID:20303878

56. Grisaru-Tal S, Itan M, Klion AD, Munitz A. A new dawn for eosinophils in the tumour microenvironment. Nat

Rev Cancer. 2020; 20:594–607.
https://doi.org/10.1038/s41568-020-0283-9
PMID:32678342

57. Furbert-Harris PM, Parish-Gause D, Hunter KA, Vaughn TR, Howland C, Okomo-Awich J, Forrest K, Laniyan I, Abdelnaby A, Oredipe OA. Activated eosinophils upregulate the metastasis suppressor molecule E-cadherin on prostate tumor cells. Cell Mol Biol (Noisy-le-grand). 2003; 49:1009–16.
PMID:14682382

58. Kataoka S, Konishi Y, Nishio Y, Fujikawa-Adachi K, Tominaga A. Antitumor activity of eosinophils activated by IL-5 and eotaxin against hepatocellular carcinoma. DNA Cell Biol. 2004; 23:549–60.
https://doi.org/10.1089/dna.2004.23.549
PMID:15383175

59. Lucarini V, Ziccheddu G, Macchia I, La Sorsa V, Peschiaroli F, Buccione C, Sistigu A, Sanchez M, Andreone S, D'Urso MT, Spada M, Macchia D, Afferni C, et al. IL-33 restricts tumor growth and inhibits pulmonary metastasis in melanoma-bearing mice through eosinophils. Oncoimmunology. 2017; 6:e1317420.
https://doi.org/10.1080/2162402X.2017.1317420
PMID:28680750

60. Gatault S, Delbeke M, Driss V, Sarazin A, Dendooven A, Kahn JE, Lefèvre G, Capron M. IL-18 Is Involved in Eosinophil-Mediated Tumoricidal Activity against a Colon Carcinoma Cell Line by Upregulating LFA-1 and ICAM-1. J Immunol. 2015; 195:2483–92.
https://doi.org/10.4049/jimmunol.1402914
PMID:26216891

61. Simson L, Ellyard JI, Dent LA, Matthaei KI, Rothenberg ME, Foster PS, Smyth MJ, Parish CR. Regulation of carcinogenesis by IL-5 and CCL11: a potential role for eosinophils in tumor immune surveillance. J Immunol. 2007; 178:4222–9.
https://doi.org/10.4049/jimmunol.178.7.4222
PMID:17371978

62. Reichman H, Itan M, Rozenberg P, Yarmolovski T, Brazowski E, Varol C, Gluck N, Shapira S, Arber N, Qimron U, Karo-Atar D, Lee JJ, Munitz A. Activated Eosinophils Exert Antitumorigenic Activities in Colorectal Cancer. Cancer Immunol Res. 2019; 7:388–400.
https://doi.org/10.1158/2326-6066.CIR-18-0494
PMID:30665890

63. Carretero R, Sektioglu IM, Garbi N, Salgado OC, Beckhove P, Hämmerling GJ. Eosinophils orchestrate cancer rejection by normalizing tumor vessels and enhancing infiltration of CD8(+) T cells. Nat Immunol. 2015; 16:609–17.
https://doi.org/10.1038/ni.3159
PMID:25915731

64. Jin C, Lagoudas GK, Zhao C, Bullman S, Bhutkar A, Hu B, Ameh S, Sandel D, Liang XS, Mazzilli S, Whary MT, Meyerson M, Germain R, et al. Commensal Microbiota Promote Lung Cancer Development via γδ T Cells. Cell. 2019; 176:998–1013.e16.
https://doi.org/10.1016/j.cell.2018.12.040
PMID:30712876

65. Fleming C, Morrissey S, Cai Y, Yan J. γδ T Cells: Unexpected Regulators of Cancer Development and Progression. Trends Cancer. 2017; 3:561–70.
https://doi.org/10.1016/j.trecan.2017.06.003
PMID:28780933

66. Wu W, Baxter JE, Wattam SL, Hayward DG, Fardilha M, Knebel A, Ford EM, da Cruz e Silva EF, Fry AM. Alternative splicing controls nuclear translocation of the cell cycle-regulated Nek2 kinase. J Biol Chem. 2007; 282:26431–40.
https://doi.org/10.1074/jbc.M704969200
PMID:17626005

67. Smyth MJ, Hayakawa Y, Takeda K, Yagita H. New aspects of natural-killer-cell surveillance and therapy of cancer. Nat Rev Cancer. 2002; 2:850–61.
https://doi.org/10.1038/nrc928
PMID:12415255

68. Farhood B, Najafi M, Mortezaee K. CD8$^+$ cytotoxic T lymphocytes in cancer immunotherapy: A review. J Cell Physiol. 2019; 234:8509–21.
https://doi.org/10.1002/jcp.27782
PMID:30520029

69. Zhang Y, Wu L, Li Z, Zhang W, Luo F, Chu Y, Chen G. Glycocalyx-Mimicking Nanoparticles Improve Anti-PD-L1 Cancer Immunotherapy through Reversion of Tumor-Associated Macrophages. Biomacromolecules. 2018; 19:2098–108.
https://doi.org/10.1021/acs.biomac.8b00305
PMID:29634251

70. Huang Q, Xia J, Wang L, Wang X, Ma X, Deng Q, Lu Y, Kumar M, Zhou Z, Li L, Zeng Z, Young KH, Yi Q, et al. miR-153 suppresses IDO1 expression and enhances CAR T cell immunotherapy. J Hematol Oncol. 2018; 11:58.
https://doi.org/10.1186/s13045-018-0600-x
PMID:29685162

71. Attili I, Karachaliou N, Bonanno L, Berenguer J, Bracht J, Codony-Servat J, Codony-Servat C, Ito M, Rosell R. STAT3 as a potential immunotherapy biomarker in oncogene-addicted non-small cell lung cancer. Ther Adv Med Oncol. 2018; 10:1758835918763744.
https://doi.org/10.1177/1758835918763744
PMID:29636826
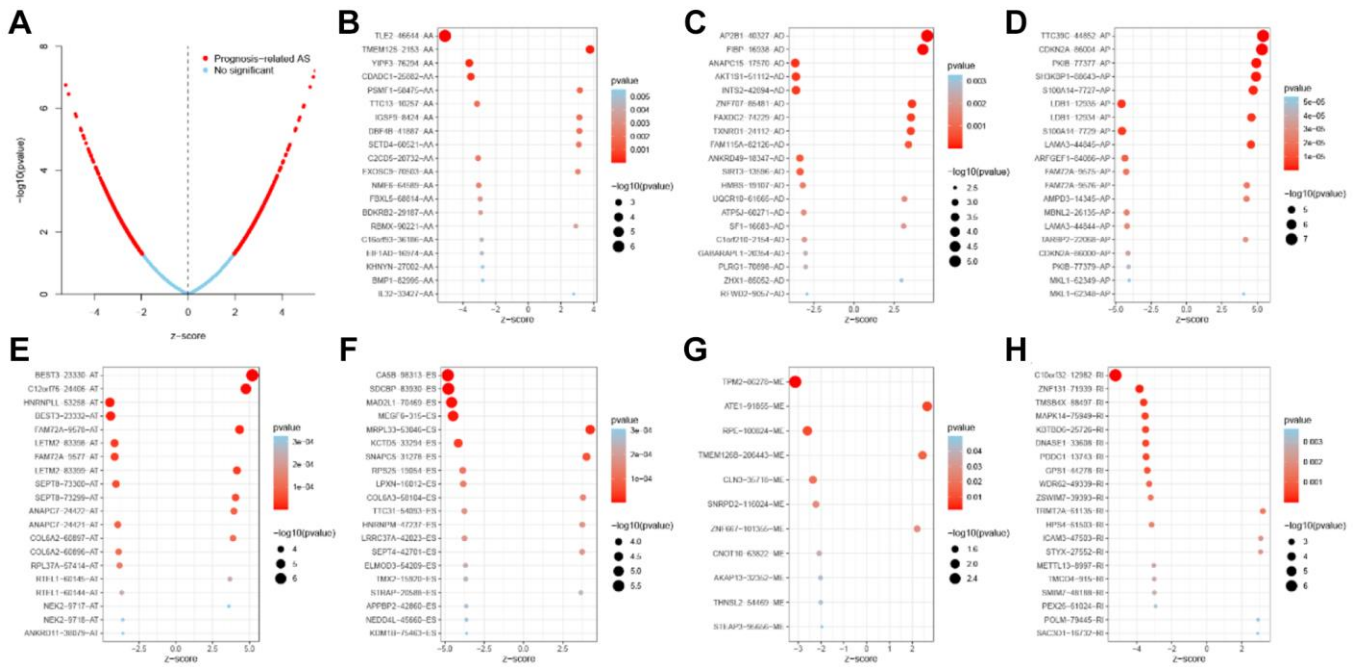
Supplementary Figure 1. Prognosis-related AS events in LUAD patients. (**A**) Prognosis-related AS events distributions in volcano plot. (**B**–**H**) Top 20 prognosis-related AS events of seven types in forest plot.
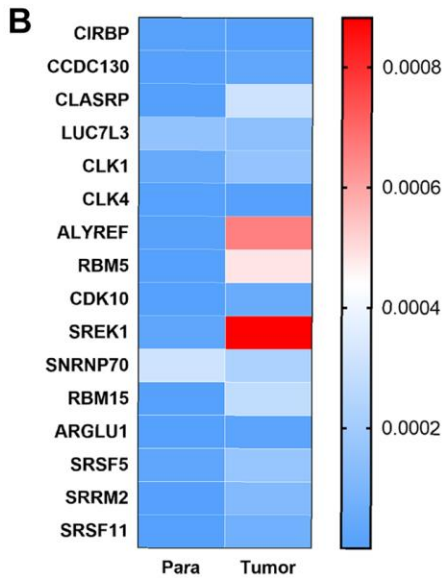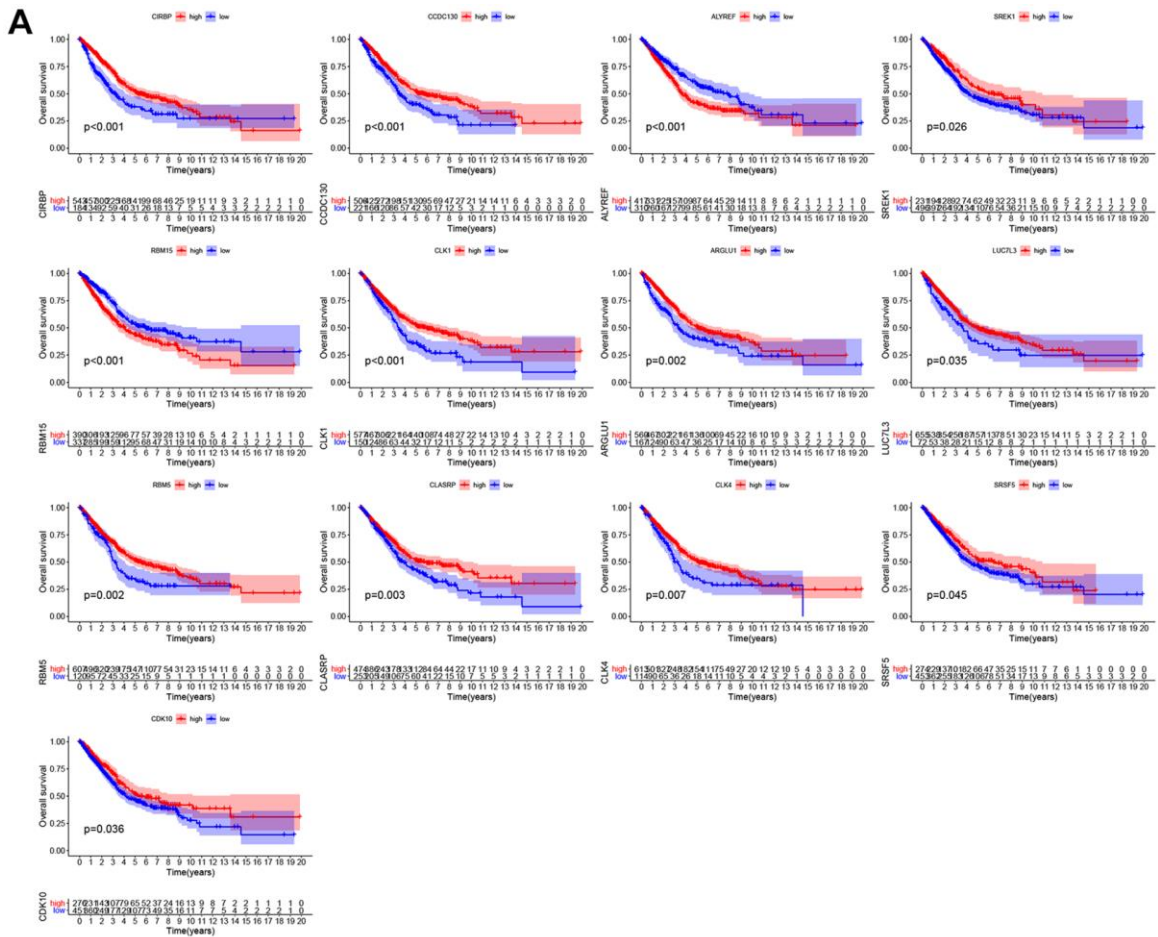
Supplementary Figure 2. (**A**) Kaplan-Meier survival analyses of SFs with prognostic significance; (**B**) expression level of SFs in surgical resected samples.
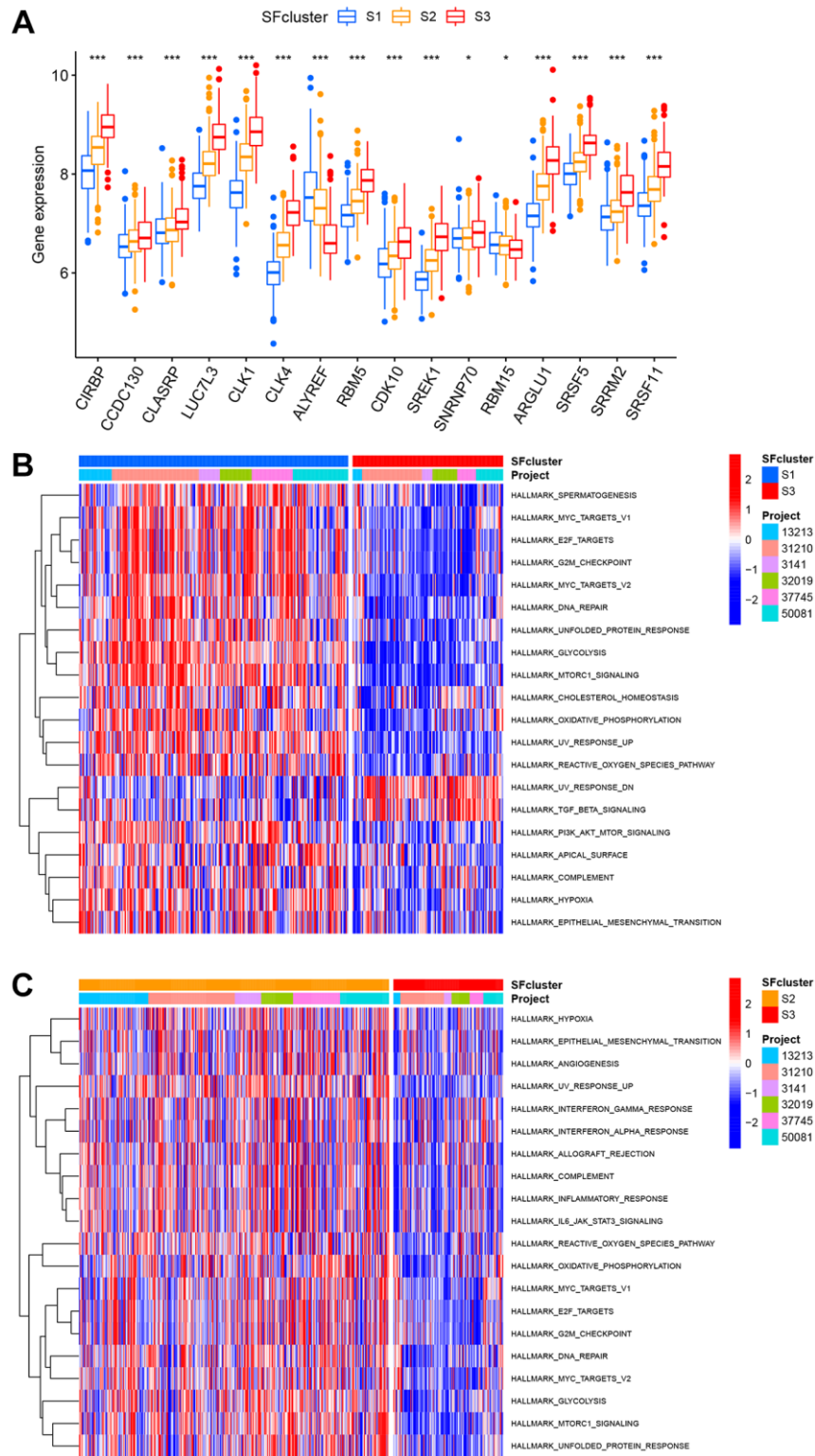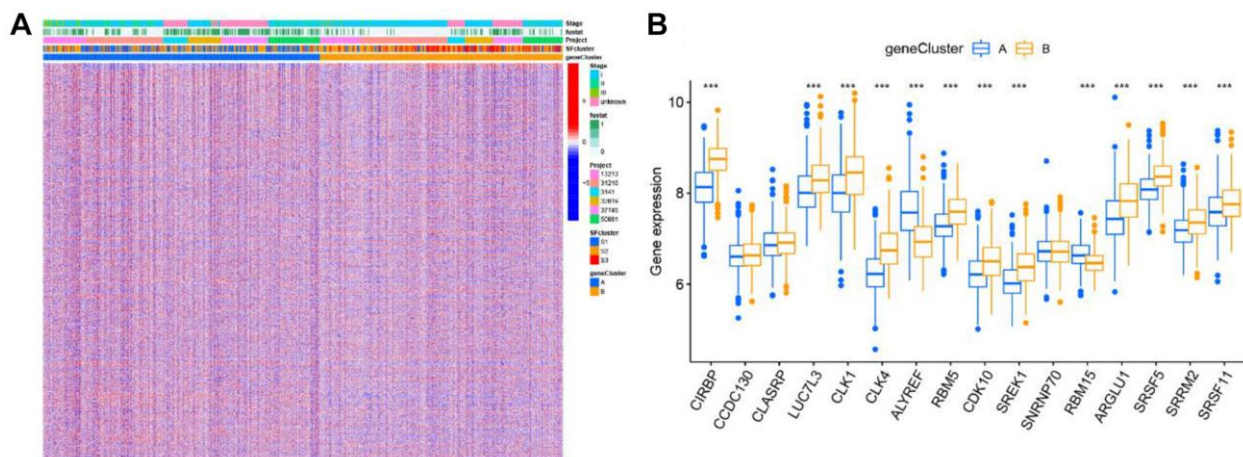
Supplementary Figure 3. (**A**) expression level of SFs in three SFs patterns; GSVA enrichment analyses based on the Hallmark gene set showed the states of biological processes of SFs cluster S1 vs. SFs cluster S3 (**B**) and SFs cluster S2 vs. SFs cluster S3 (**C**).
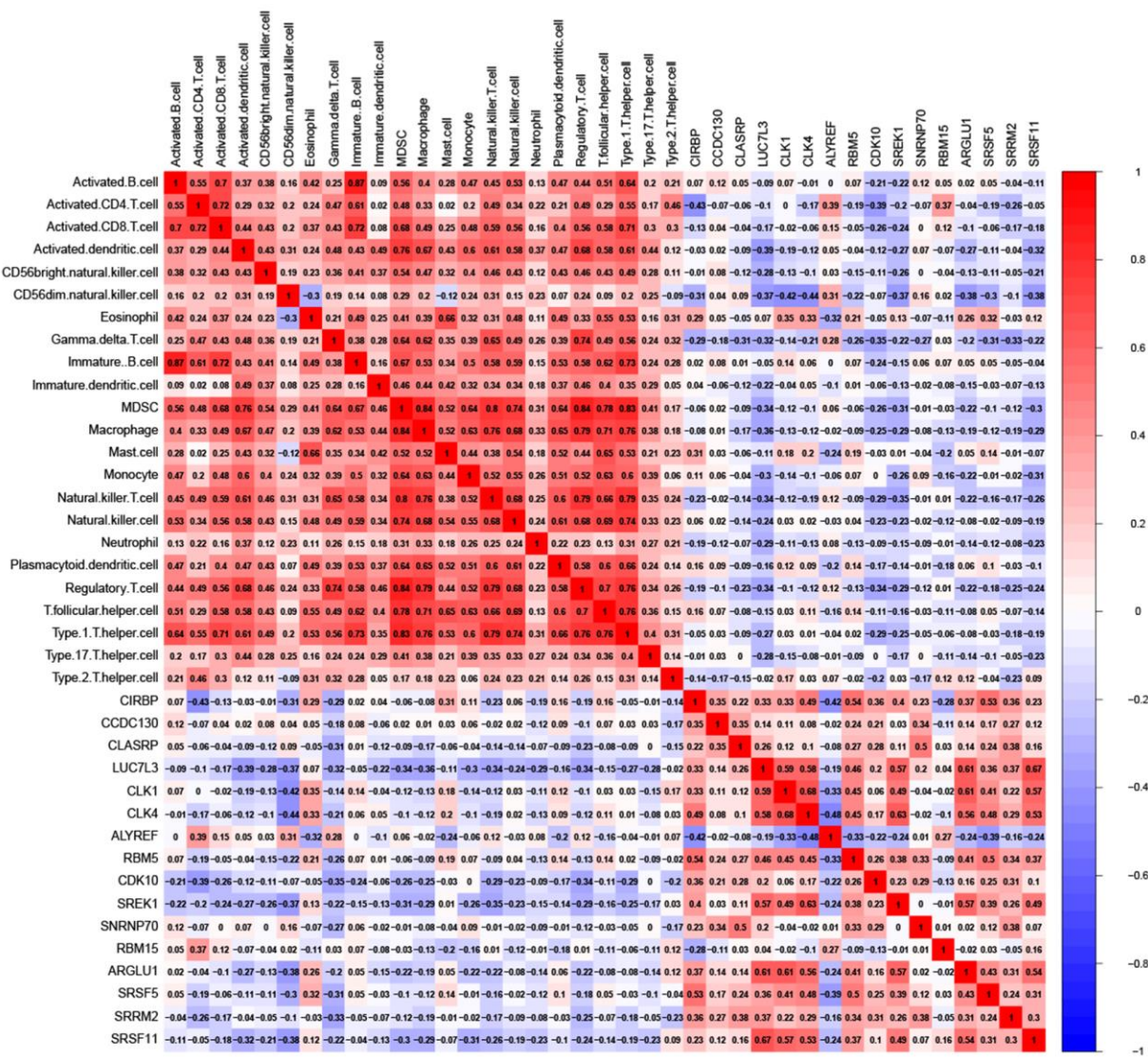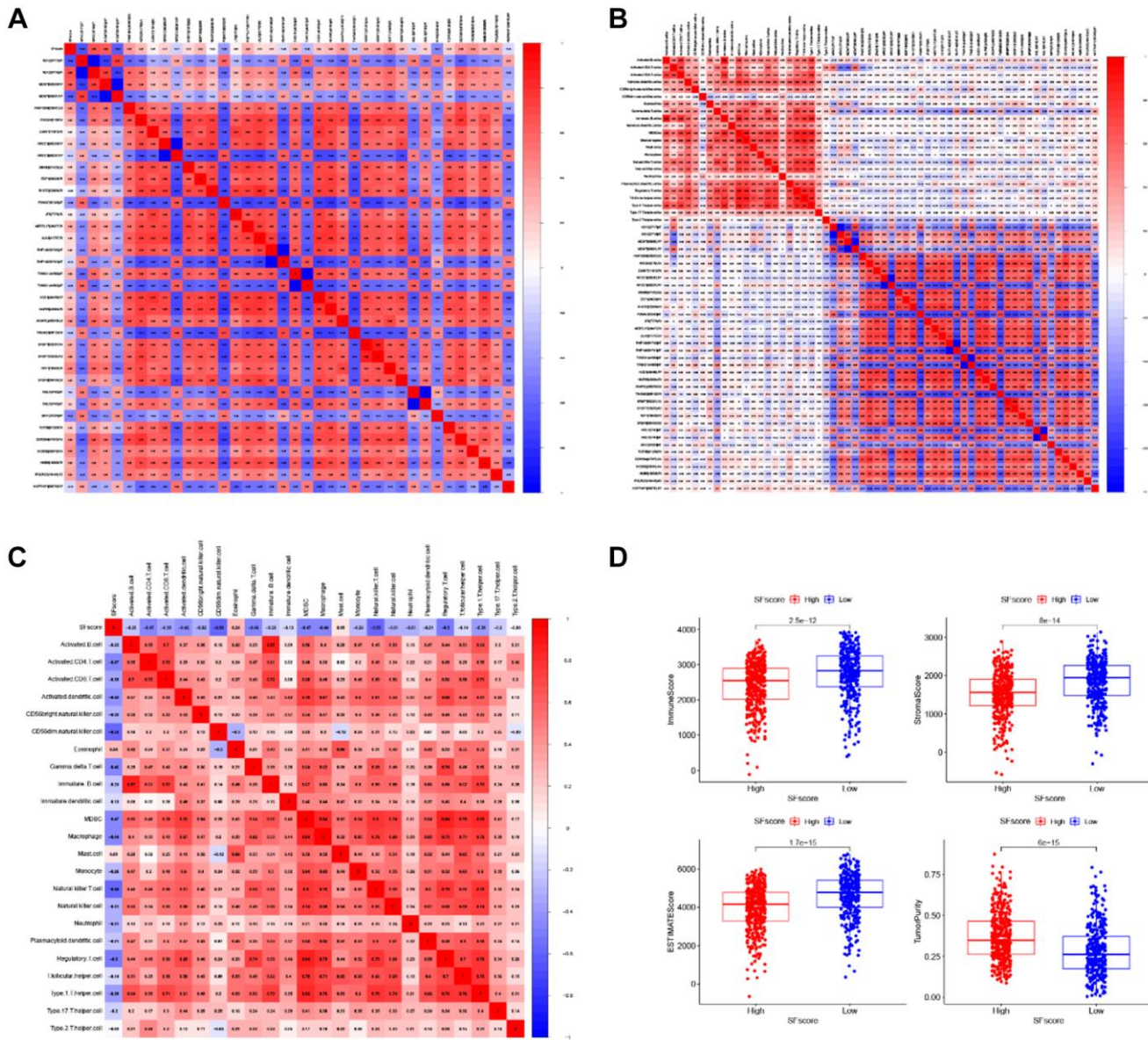
Supplementary Figure 4. Correlations between 16 SFs and 28 TME infiltrating cells in LUAD using spearman analysis.

Supplementary Figure 5. (**A**) Unsupervised clustering of 2959 SFs -related DEGs with prognostic significance in GEO cohorts. (**B**) The expression level of SFs in two gene clusters.

Supplementary Figure 6. (**A**) Correlations between SFscore and SFs interacting AS events. (**B**) Correlations between SFs interacting AS events and 28 TME infiltrating cells. (**C**) Correlation analysis between SFscore and TME infiltrating cells. (**D**) Difference of immune score, stromal score, ESTIMATE score and tumor purity in low/high SFscore groups.

Supplementary Table

Supplementary Table 1. Sequences of primers.

Gene	Forward	Reverse
CIRBP	AGACTACTATAGCAGCCGGAGT	AAGGTGAACCGAGCTCCC
CCDC130	GCATCCTCATCATCCGATTCG	GTAGTAATTGCCCACCTTCTTCTT
CLASRP	TTTCATTTCCGCTTCCGGTG	ATGTTGTTGGTGTCCCCCTG
LUC7L3	TACTGTTTGACAGTGCAGCG	CTAATGGGGGCACGAGTCTG
CLK1	ATTTTGTTGTTGGTGCGCGA	TGAGTGTCTCATCGTCCTGG
CLK4	TGGCATGCATGTAGCAGTGA	GCATCTGGACACATCGGAAGA
ALYREF	AGAGCGTAAACAGAGGTGGC	ACTGGTGTCCATTCTCGCAT
RBM5	CCGGAGAGGACAGTGGATTG	TCAGACATGCTTGACCCACC
CDK10	ACGACCCTAAGAAAAGGGCG	GATCGGAAGACCTGGTGTGG
SREK1	GTCGTAGACGTTGGGGAGC	TCACCTGAATCACCGACGTG
SNRNP70	CGAGACATGCACTCCGCTTA	CTCTCATCGTAGCGGGAGGT
RBM15	ATCATTGTCCGTGGGTTTGGT	ACTATAACAGGGTCAGCGCC
ARGLU1	CGGCAGCGAAAAATTCGACA	CGCTCTAGCTCCTCACGTTT
SRSF5	GACCCCGTCCGGTAGGAAG	GACCGAGCCCTAGCATGTTC
SRRM2	CTCCGATACTTCCCGCAGTC	GTCGAGTTGCAGATTTCTCCT
SRSF11	CCTCTTCCCCCTCCTTCTCA	AAAGGCGAATCATCCGGCG