Survival-Associated Differentially Expressed Alternative Splicing Signatures Reveal Neuron Signaling and Metabolism-Related Processes in Glioblastoma

Yanshan Ge (✉ geyanshan@csu.edu.cn)  
Xiangya Hospital Central South University

Zhaoyu Zhang  
Xiangya Hospital Central South University

Qiang Liu  
Xiangya Hospital Central South University

Minghua Wu  
Xiangya Hospital Central South University

Research

Keywords: Splicing Signatures, Neuron Signaling, Metabolism, Glioblastoma

DOI: https://doi.org/10.21203/rs.3.rs-375966/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background

Alternative splicing (AS) is a molecular event that drives protein diversity by generating multiple mRNA isoforms. Increasing evidence shows that the differential expression of AS is related to tumorigenesis. However, a synthetic analysis for identifying the AS biomarkers attributed to glioblastoma (GBM) is mostly unexplored.

Methods

AS percent-splice-in (PSI) data were obtained from the TCGA SpliceSeq database. We systematically and interactively analyzed differentially expressed alternative splicing (DEAS) events in GBM. Univariate and multivariate Cox regression analysis was successively performed to identify the overall survival (OS)-associated DEAS events, followed by the construction of DEAS predictor through different splicing patterns. Then, a nomogram that combines the final DEAS predictor and clinicopathological characteristics was established. Finally, a splicing regulatory network was created according to the correlation between the DEAS events and the splicing factors (SF). Summary illustrations show Figure1.

Results

A total of 2697 DEAS events detected in 2206 gene symbols consisted of 1054 upregulated DEAS events and 1643 downregulated DEAS events. Notably, these DEAS events are mainly involved in neuron-synapse signaling and metabolism-related processes in GBM. We identified 135 DEAS events that were substantially associated with the overall survival of GBM patients; for example, the alternate promoter (AP) of GSG1L and alternate terminator (AT) of CDKL3 were associated with poor prognosis, whereas AP of TMEM63, exon skip (ES) of CPEB4, AT of FAM13A, AT of OBSL1, AT of HNF1A, and AT of CCDC40 were associated with good prognosis. Following this, we constructed a potential splicing factor (SF)-AS regulatory network and identified nine key SFs, including ELAVL4, CELF3, CELF5, RBFOX1, YWHAH, STXB1, CELF4, ELAVL2, and DNM1, which are mainly involved in RNA metabolism and insulin synthesis.

Conclusions

In summary, we demonstrated the role of DEAS events in the progression of GBM, providing a potential clinical biomarker and therapeutic reference for GBM.

Introduction

Glioblastoma (GBM) remains the most aggressive and malignant form of primary central nervous system (CNS) tumors in adults. According to the 2016 World Health Organization (WHO) molecular diagnosis guidelines for CNS, GBM has high heterogeneity and poor prognosis[1]. The current first-line treatment for newly diagnosed GBM is the maximum safe resection followed by chemoradiation[2, 3]. GBM almost
always recurs, 2-year survival rate of GBM is 25%, and its 5-year survival rate is 5–10%[4]. Consequently, GBM, which is difficult to cure, puts challenging pressure on future work related to the exploration of new therapeutic goals.

Alternative splicing (AS) is the main driving force for the production of multiple proteins, allowing cells to produce a large amount of protein diversity from a limited number of genes[5]. Alternative splicing occurs in approximately 95% of transcripts[6]. Current research indicates that AS events are divided into seven types, and each has a different mechanism of action described as follows: alternate acceptor site (AA), alternate donor site (AD), alternate promoter (AP), alternate terminator (AT), exon skip (ES), mutually exclusive exons (ME), and retained introns (RI). Cancer cells exhibit remarkable transcriptome alterations, partly by adopting cancer-specific splicing isoforms. The aberrations of AS in cancers primarily include four categories: (a) AS alterations in oncogenes or tumor suppressor genes, (b) splicing factor mutations, (c) alterations in the upstream signaling pathways that decontrol splicing factors, and (d) aberrations in cancer-specific spliceosomal components[7]. Alternative splicing can be a target for cancer treatment, and targeting aberrant splicing has provided novel perspectives for clinical therapy strategies[8].

Although Xie et al.[9] and Chen et al.[10] performed a transcriptome-wide analysis of alternative mRNA splicing markers in GBM tissue data from The Cancer Genome Atlas (TCGA), both articles only focused on single AS events. In addition, there is a lack of systematic analysis of differentially expressed AS events (DEAS) between GBM and normal tissues. Although only protein-coding genes have been studied, DEAS can generate differentially expressed transcripts and protein variants.

In this study, we obtained RNA sequences from the Cancer Genome Atlas (TCGA) program and clarified the role of different DEAS patterns in 154 GBM samples and 5 normal control samples and conducted an in-depth analysis of the potential impact of specific DEAS events on the prognosis of GBM patients. Finally, we investigated how splicing factor (SF) gene-regulated DEAS events are involved in the occurrence and development of GBM. The purpose of this study was to explore the role of various DEAS splicing modes in GBM and how SF genes regulate DEAS events, thus providing new ideas for the detection and treatment of GBM.

**Materials And Methods**

**Collecting AS events and GBM-related clinical data**

TCGASpliceSeq (http://bioinformatics.mdanderson.org/TCGASpliceSeq) is a web-based resource used for investigating mRNA AS patterns for 33 different tumor types in TCGA database; it was used to download the splicing profiles of GBM[11]. The profile of the AS events was defined by a percent-spliced-in (PSI) value, which ranged from zero to one. Simultaneously, each splicing event was allocated a unique code consisting of the splicing type, gene symbol, and ID number to ensure accuracy. The splicing profiles of seven types of AS, including alternate acceptor site (AA), alternate donor site (AD), alternate promoter (AP), alternate terminator (AT), exon skip (ES), mutually exclusive exons (ME), and retained introns (RI) were freely available in TCGA database. The expression data for GBM and variable splicing
data were integrated. Finally, 154 GBM samples and five normal control samples were selected for subsequent analyses. The lists of immune-related genes were downloaded from the ImmPort (https://www.immport.org/shared/home) and InnateDB (https://www.innatedb.com/) databases.

**Differentially spliced AS events analysis**

The limma package was used for differential analysis. In this study, a differentially alternative splicing event was carried out between the GBM samples and normal brain samples, and \( P < 0.05 \), and \(| \log FC|>1\) was defined as differential alternative splicing events.

**Gene Ontology (GO) functional and KEGG pathway enrichment analyses of DEAS events**

GO and KEGG pathway enrichment analyses were implemented with the “enrichplot” R package to analyze the function of the DEAS[12]. GO terms included biological process (BP), cellular component (CC), and molecular function (MF). For the analysis results, both P-value and false discovery rate values < 0.05, were defined as statistically significant.

**Prognostic model construction**

We downloaded the clinical parameters of the GBM cohort from TCGA database (https://tcga-data.nci.nih.gov/tcga/). A total of 559 patients with GBM with overall survival (OS) information were included in the study. To analyze the association between DEAS events and OS of patients, we divided patients into two groups based on the median PSI value of each splicing event and performed univariate Cox regression. Multivariate Cox regression was further conducted to determine splicing events that were independent prognostic factors and to build predictive models. Kaplan-Meier curves were used to verify whether the predictive models could differentiate patients with long OS from those with shorter OS. The efficiencies of each predictive model were further assessed by the receiver operator characteristic curves using the survival ROC package in R. To investigate the intersections between different types of AS, we applied UpSetR to visualize intersections and their aggregates, which is more efficient than the traditional Venn diagram when addressing more than five sets[13]. Function annotation, pathway enrichment, and gene interaction network analyses were performed on survival-related AS genes using the Reactome FI plugin in Cytoscape (version 3.6.1)[14].

**Splicing-factor-regulated network construction**

To explore the regulation of splicing factors on prognostic DEAS events, we first curated TCGA level 3 mRNA-seq data of splicing factors in primary mRNA splicing pathways and determined the survival-associated splicing factors. The expression levels of survival-associated splicing factors in GBM and non-tumor tissues were also compared. The Spearman test was then conducted to evaluate the potential relationship between the expression level of splicing factors and PSI value of AS events. The regulation networks were plotted using Cytoscape (version 3.6.1).
Results

Integrating alternative splicing events in GBM

Our research conducts data analysis in sequence according to the flowchart (Figure 1). A total of 45,610 AS events detected in 21,136 gene symbols were collected from 154 GBM samples and five normal control samples. These results comprised 3,827 AAs in 2,684 genes, 3,269 in 2,270 genes, 8,686 in 3,476 genes, 8,456 in 3,695 genes, 18,360 in 6,934 genes, 184 MEs in 180 genes, and 2828 RIs in 1,897 genes (Figure 2A). The exon skip was the most dominant splicing type, followed by the AP splicing type. The number of AS events was larger than the total number of genes, implying that a single gene may have undergone multiple splicing. Some genes can be expressed in up to six AS types. We have described the detailed information of genes of a specific AS type in the Upset diagram (Figure 2A). Compared with a traditional Venn diagram, it can more effectively prove the quantitative results of multiple interaction sets.

Identification of differentially expressed AS (DEAS) events

By comparing the differences in AS events between GBM and normal tissues, we identified key DEAS events involved in tumorigenesis. A total of 2,697 DEAS events detected in 2,206 gene symbols were collected, including 1,054 differentially upregulated AS events and 1,643 differentially downregulated AS events. These results comprised 107 AAs in 101 genes, 116 in 110 genes, 695 in 568, 634 in 552, 1,008 in 748, 23 in 23, and 114 in 104, as shown in Figure 2B. These data indicated that in GBMs, ES was the most frequent DEAS event, and ME was the rarest. Furthermore, there were 6,170 upregulated genes and 4,638 downregulated genes in TCGA GBM dataset. We also found that the DEAS events could compensate for the shortcomings of differential expression analysis; for example, we compared the genes involved in the differential AS event with the differentially expressed genes, and found that 234 genes were recognized, as shown in Figure 3A. Through observation, we found that there were no common transcription factors in these 234 genes. However, we used FUNRICH software to analyze and found that 234 genes have some common transcription factors, such as GATA1, POU3F4, MSX1, LHX3, NHLH1, and HENMT1 (p<0.05) (Figure 3B).

DEAS events mainly are enriched in neuron signaling and metabolism-related processes in GBM

Subsequently, 2,697 DEAS events detected in the 2,206 gene were profiled using GO functional enrichment analysis. In biological processes, the top three enriched terms were axonogenesis, modulation of chemical synaptic transmission, and regulation of trans-synaptic signaling. In the cellular component, the top three enriched terms were cell leading edge, asymmetric synapse, and neuron to neuron synapse. Regarding molecular function, the top three enriched terms were actin binding, cadherin binding, and tubulin binding. as shown in Figure 3C. Taken together, the DEAS events were mainly enriched in synaptic genesis and transmission.

A Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis revealed several important pathways. The top five enriched pathways were the HIF-1a signaling pathway, Rap1 signaling pathway, axon guidance,
regulation of actin cytoskeleton, and insulin signaling pathway (Figure 3D). Taken together, the DEAS genes regulated by DEAS events were mainly involved in metabolism-related processes. These results indicate that splicing some core genes involved in neuronal signaling and metabolism-related processes may be important in GBM.

Identification of DEAS events associated with survival of GBM patients

Using the DEAS event profiles in the GBM cohort, we identified 135 DEAS events that were significantly associated with the OS of GBM patients (p<0.05) by univariate Cox regression analysis (Table 1). In particular, we found genes with potentially more than one DEAS event that was significantly associated with patient survival. To better visualize the intersecting sets, an UpSet plot was created, as shown in Figure 3E. Interestingly, our analysis revealed that genes can exhibit up to seven types of DEAS events that were all found to be significantly associated with GBM patient survival. In addition, via FUNRICH software analysis, we found that 135 DEAS events were associated with the OS of GBM patients are not only related to central nervous system development, but are also closely related to other kidney, stomach, breast, skin, hematopoietic, and lymphoid tissues in cosmic analysis. (Figure 3F).

We drew a volcano map based on the obtained data, which clearly showed that DEAS events related to prognosis accounted for the majority (Figures 4A). The bubble plots illustrate the top 20 OS-DEASs in seven types of splicing patterns (Figures 4B–H). In summary, among all the OS-DEAS, SETD6-36663-AA, CSPG5-64534-AD, NKIRAS2-40977-AP, CCDC40-44016-AT, UBXN11-101231-ES, GRIA1-125279-ME, and CNIH2-17000-RI were the most significant events for AA, AD, AP, AT, ES, ME, and RI, respectively. OS-DEAS genes are closely related to the development of multiple organs.

Risk score as a low-risk factor can be an independent prognostic indicator of GBM.

Next, we performed a LASSO regression analysis of all types of DEAS events related to prognosis (Figure 5A, B). The number of the other seven AS events is too small, and it is not suitable for LASSO regression analyses. The left mode of each group is the LASSO coefficient curve, and the right mode of each group is the choice of the adjustment parameter (λ) in the LASSO model. The y-axis represents partial likelihood deviations. The lower x-axis indicates the logarithm (λ), and the upper x-axis represents the average number of predictors. The red dot represents the average deviation value for each model with a given lambda, where the location indicates the best data. As a result, the prognostic-relevant AS events were selected in all types of AS events, these AS events included 8 AP events, 5 AT events, 4 ES events, and there was NKIRAS2|40977|AP, TMEM63B|76352|AP, GSG1L|35696|AP, RPL39L|68070|AP, GSG1L|35698|AP, ROBO2|65640|AP, SPATS2L|56734|AP, PAIP1|71958|AP, CCDC40|44016|AT, FAM13A|69905|AT, CDKL3|73367|AT, OBSL1|57730|AT, HNF1A|24784|AT, UBXN11|101231|ES, CPEB4|74594|ES, IP6K2|64757|ES, MTIF2|53607|ES.

To determine the independent prognostic factors associated with GBM patients with DEAS events, we screened the most important DEAS events. A multiple Cox regression model with independent prognostic factors was performed for 17 DEAS events. In our data analysis, we found that the above types of DEAS
events had a strong ability to predict the prognosis of GBM patients (Figure 5C). The red line represents the high-risk subgroup, the blue line represents the low-risk subgroup. The AUC for all types of AS events was 0.826 (Figure 5D).

Figure 5E shows the distributions of survival time, status, and risk scores among the low-risk and high-risk groups. The image is divided into the upper, middle, and lower sections. The top portion indicates the distribution of risk scores, the middle part shows the survival time and status of GBM patients (green indicates alive, red indicates dead), and the bottom part is the heatmap associated with it. The black vertical dashed line in the middle indicates the best dividing line between the low-risk and high-risk groups. DEAS events, including the AP of GSG1L and AT of CDKL3, were associated with poor prognosis. DEAS events including AP of TMEM63, ES of CPEB4, AT of FAM13A, AT of OBSL1, AT of HNF1A, and AT of CCDC40 were associated with good prognosis.

As shown in Figure 5F, we performed univariate and multivariate Cox regression analyses of all groups of prognosis-related AS events in the forest plot. We analyzed the data using Univariate Cox regression analysis showed that age (reference low), IDH mutation (reference high), and risk score (reference low) of the four groups were used as independent prognostic indicators of GBM (P<0.05). Multivariate Cox regression analyses also demonstrated that the risk score (reference low) of the four groups was an independent prognostic indicator of GBM (P<0.05).

**Construction of Potential SF-AS Regulatory Network**

To explore the upstream mechanism of AS regulation, we compared the identified SF genes with the genes obtained from TCGA and extracted the expression data of SF genes from the latter. We screened a total of 49 SF-related genes and constructed a regulatory network of DEAS and SF genes (|cor|> 0.8, P<0.001) using Cytoscape, and integrated the analyses with the risk factors (Figure 6A). Our analysis identified nine key SFs, including ELAVL4, CELF3, CELF5, RBFOX1, YWHAH, STXBP1, CELF4, ELAVL2, and DNM1. Interestingly, we found that 49 SF-related genes were mainly enriched in the regulation of nucleobase, nucleoside, nucleic acid metabolism, and protein metabolism in biological processes. Regarding molecular function, 49 SF-related genes were mainly enriched in the structural constituents of ribosomes and RNA binding. In the biological pathway, 49 SF-related genes were mainly enriched in gene expression, RNA metabolism, and insulin synthesis and processing(Figure 6B). Meanwhile, we use FUNRICH software to find Biological pathway enrichment analysis of SF (Figure 6C) and analysis SF Site of expression(Figure 6D).

**Immune-related hub genes**

By intersecting 135 OS-DEAS with the lists of immune-related genes obtained from ImmPort and InnateDB, two hub immune-related genes, NR1H3|15695|AP and PIK3R2|48396|AT, are related to prognosis(Figure 6D). We found that NR1H3 was highly expressed in GBM(Figure 6E). The cg07298473 methylation site was related to NR1H3 expression(Figure 6F). Among the 25 methylation sites, 18
methylation sites showed low expression and seven methylation sites showed high expression (Figure 6G).

## Discussion

AS is the basic mechanism of gene expression regulation in eukaryotes, which increases the complexity of gene expression, promotes higher transcription efficiency, and protein diversity [15]. It has been found that the dysregulation of DEAS in cancer-related genes is involved in many biological processes of tumors, and these abnormally regulated genes can be used as molecular markers for cancer prognosis and treatment [16, 17]. However, a comprehensive analysis of DEAS signatures in GBM remains unknown.

Through analysis of TCGA database, we identified AS events and regulatory splicing factors in GBM to fully understand their differential RNA splicing patterns. In total, 2697 differentially AS events involving 2206 genes were identified, of which 1054 were upregulated and 1643 were downregulated. We then evaluated the potential functions and pathways of DEAS-associated genes by performing enrichment analysis. The pathways involved in neuronal signaling and metabolism-related processes. It is worth noting that immune-related pathways may be involved in the tumorigenesis of GBM.

To further evaluate whether a specific DEAS can be used as a prognostic indicator of GBM, we established a prognostic model based on the individual AS pattern. A total of 135 DEAS events were found to be significantly associated with OS in GBM patients. Interestingly, among these DEAS events related to OS, our analysis also includes some splice variants that have been identified to play an important role in tumor biology. For instance, NKIRAS2, which was identified as an IκB-interacting small GTPase, has been reported to induce cytokine-induced NF-κB activation [18]. NKIRAS2 is involved in glioma apoptosis [19]. The protein encoded by the gene SETD6 encodes a methyltransferase that adds a methyl group to the histone H2AZ, which is involved in nuclear receptor-dependent transcription [20]. The protein encoded by CSPG5 is a proteoglycan that may function as a neural growth and differentiation factor [21, 22]. GRIA1 is the main excitatory neurotransmitter receptor in the mammalian brain and is activated in a variety of normal neurophysiological processes. These receptors are heteromeric protein complexes with multiple subunits, each subunit has a transmembrane region, and all are arranged to form ligand-gated ion channels [23]. The protein encoded by this gene is an auxiliary subunit of the AMPA ionotropic glutamate receptor. AMPA receptors mediate rapid synaptic neurotransmission in the central nervous system. It has been reported that this protein interacts with the type I AMPA receptor regulatory protein isoform γ-8 to control the assembly of the hippocampal AMPA receptor complex, thereby regulating receptor gating and pharmacological effects. Alternative splicing leads to the development of multiple transcript variants [24, 25]. Splicing events such as CCDC40 and UBXN11 play a key role in tumor biology [26, 27].

In addition, we constructed a predictive model for each splicing type using multivariate Cox regression analysis. The AUC of all types of AS events was 0.826, which could provide patients with new predictive information and facilitate patient assessment. Given that the regulation of key splicing factors may lead
to changes in a wide range of AS events, we further analyzed survival-related AS splicing factors and their related networks, and found that 49 splicing factors are significantly related to the overall survival of patients. Interestingly, it was found that most of the splicing factor expression levels associated with poor survival were upregulated.

Next, we took the intersection of 135 OS-DEAS and the immune-related gene list obtained from the immune port and innate DB, and obtained two hub immune-related genes. Interestingly, we found that NR1H3|15695|AP may affect the expression of NR1H3 through methylation modification.

**Conclusion**

Here, we describe the most comprehensive and up-to-date study to assess prognostic factors and survival outcomes for GBM by analyzing abnormal AS variants, splicing factor regulatory networks, and related molecular cluster subtypes. Although the prognostic significance of these potential therapeutic targets for GBM still requires further functional and clinical trials to validate the findings, this research enhances the potential of new clinical biomarkers and therapeutic approaches to help implement new treatment strategies for GBM.

**Abbreviations**

AS: Alternative splicing; GBM: glioblastoma; DEAS: differentially expressed AA: alternate acceptor site; AD: alternate donor site; AP: alternate promoter; AT: alternate terminator; ME: mutually exclusive exons; ES: exon skip; RI: retained introns; SF: splicing factor; CNS: central nervous system; WHO: World Health Organization; TCGA: The Cancer Genome Atlas; PSI: percent-spliced-in; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; BP: biological process; CC: cellular component; MF: molecular function; OS: overall survival

**Declarations**

**Ethical Approval and Consent to participate**

Not applicable

**Consent for publication**

All authors give consent for the publication of this manuscript in Molecular Cancer

**Availability of supporting data**

All the data used in the manuscript are freely available online.

**Funding**
This work was supported by the National Science Foundation of China under Grant number 82073096 and the Hunan Province Key Research and Development Project under 2020SK2053.

Conflict of interests

The authors declare that they have no competing interests.

Author contributions

WMH and GYS collected the related paper and finished the manuscript and figures. WMH, LQ gave constructive guidance and made critical revisions. GYS, ZZY participated in the design of this review. All authors read and approved the final manuscript.

Authors' information

Not applicable

References

1. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, Ohgaki H, Wiestler OD, Kleihues P, Ellison DW: The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta Neuropathol 2016, 131:803-820.

2. Fabian D, Guillermo PEM, Alnahhas I, Sebastian N, Giglio P, Puduvalli V, Gonzalez J, Palmer JD: Treatment of Glioblastoma (GBM) with the Addition of Tumor-Treating Fields (TTF): A Review. Cancers (Basel) 2019, 11.

3. Bureta C, Saitoh Y, Tokumoto H, Sasaki H, Maeda S, Nagano S, Komiya S, Taniguchi N, Setoguchi T: Synergistic effect of arsenic trioxide, vismodegib and temozolomide on glioblastoma. Oncol Rep 2019, 41:3404-3412.

4. Stupp R, Taillibert S, Kanner A, Read W, Steinberg D, Lhermitte B, Toms S, Idbaih A, Ahluwalia MS, Fink K, et al: Effect of Tumor-Treating Fields Plus Maintenance Temozolomide vs Maintenance Temozolomide Alone on Survival in Patients With Glioblastoma: A Randomized Clinical Trial. JAMA 2017, 318:2306-2316.

5. Leivonen SK, Taskinen M, Cervera A, Karjalainen-Lindsberg ML, Delabie J, Holte H, Lehtonen R, Hautaniemi S, Leppa S: Alternative splicing discriminates molecular subtypes and has prognostic impact in diffuse large B-cell lymphoma. Blood Cancer J 2017, 7:e596.

6. Zhu LY, Zhu YR, Dai DJ, Wang X, Jin HC: Epigenetic regulation of alternative splicing. Am J Cancer Res 2018, 8:2346-2358.

7. El ME, Younis I: The Cancer Spliceome: Reprograming of Alternative Splicing in Cancer. Front Mol Biosci 2018, 5:80.

8. Lin JC: Therapeutic Applications of Targeted Alternative Splicing to Cancer Treatment. Int J Mol Sci 2017, 19.
9. Xie ZC, Wu HY, Dang YW, Chen G: Role of alternative splicing signatures in the prognosis of glioblastoma. *Cancer Med* 2019, 8:7623-7636.

10. Chen X, Zhao C, Guo B, Zhao Z, Wang H, Fang Z: Systematic Profiling of Alternative mRNA Splicing Signature for Predicting Glioblastoma Prognosis. *Front Oncol* 2019, 9:928.

11. Ryan M, Wong WC, Brown R, Akhani R, Su X, Broom B, Melott J, Weinstein J: TCGASpliceSeq a compendium of alternative mRNA splicing in cancer. *Nucleic Acids Res* 2016, 44:D1018-D1022.

12. Yu G, Wang LG, Han Y, He QY: clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* 2012, 16:284-287.

13. Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H: UpSet: Visualization of Intersecting Sets. *IEEE Trans Vis Comput Graph* 2014, 20:1983-1992.

14. Wu G, Feng X, Stein L: A human functional protein interaction network and its application to cancer data analysis. *Genome Biol* 2010, 11:R53.

15. Lee SC, Abdel-Wahab O: Therapeutic targeting of splicing in cancer. *Nat Med* 2016, 22:976-986.

16. Oltean S, Bates DO: Hallmarks of alternative splicing in cancer. *Oncogene* 2014, 33:5311-5318.

17. Lin C, Yu B, Zhang M, Chen Y, Li L, Zhao D: Systematic Analyses of the Differentially Expressed Alternative Splicing Events in Gastric Cancer and Its Clinical Significance. *Front Genet* 2020, 11:522831.

18. Tago K, Funakoshi-Tago M, Sakinawa M, Mizuno N, Itoh H: KappaB-Ras is a nuclear-cytoplasmic small GTPase that inhibits NF-kappaB activation through the suppression of transcriptional activation of p65/RelA. *J Biol Chem* 2010, 285:30622-30633.

19. Haemmig S, Baumgartner U, Gluck A, Zbinden S, Tschan MP, Kappeler A, Mariani L, Vajtai I, Vassella E: miR-125b controls apoptosis and temozolomide resistance by targeting TNFAIP3 and NKIRAS2 in glioblastomas. *Cell Death Dis* 2014, 5:e1279.

20. Weil LE, Shmidov Y, Kublanovsky M, Morgenstern D, Feldman M, Bitton R, Levy D: Oligomerization and Auto-methylation of the Human Lysine Methyltransferase SETD6. *J Mol Biol* 2018, 430:4359-4368.

21. Aono S, Tokita Y, Yasuda Y, Hirano K, Yamauchi S, Shuo T, Matsui F, Keino H, Kashiwai A, Kawamura N, et al: Expression and identification of a new splice variant of neuroglycan C, a transmembrane chondroitin sulfate proteoglycan, in the human brain. *J Neurosci Res* 2006, 83:110-118.

22. Kinugasa Y, Ishiguro H, Tokita Y, Oohira A, Ohimoto H, Higashiyama S: Neuroglycan C, a novel member of the neuregulin family. *Biochem Biophys Res Commun* 2004, 321:1045-1049.

23. Zhao LX, Ge YH, Li JB, Xiong CH, Law PY, Xu JR, Qiu Y, Chen HZ: M1 muscarinic receptors regulate the phosphorylation of AMPA receptor subunit GluA1 via a signaling pathway linking cAMP-PKA and PI3K-Akt. *Faseb J* 2019, 33:6622-6631.

24. Coombs ID, Soto D, Zonouzi M, Renzi M, Shelley C, Farrant M, Cull-Candy SG: Cornichons modify channel properties of recombinant and glial AMPA receptors. *J Neurosci* 2012, 32:9796-9804.
25. Boudkkazi S, Brechet A, Schwenk J, Fakler B: Comichon2 dictates the time course of excitatory transmission at individual hippocampal synapses. *Neuron* 2014, **82**:848-858.

26. Wang J, Lin J, Chang Y, Li P, Yang Y: MCM3AP, a novel HBV integration site in hepatocellular carcinoma and its implication in hepatocarcinogenesis. *J Huazhong Univ Sci Technolog Med Sci* 2010, **30**:425-429.

27. Maccalli C, Di Cristanziano V, Fodale V, Corsi D, D’Agostino G, Petrangeli V, Laurenti L, Guida S, Mazzocchi A, Arienti F, et al: Induction of both CD8+ and CD4+ T-cell-mediated responses in colorectal cancer patients by colon antigen-1. *Clin Cancer Res* 2008, **14**:7292-7303.

### Tables

#### Table 1: The detailed information of the top 20 OS-DEAS events.

| Symbol  | AS type | z     | HR   | HR.95L  | HR.95H  | pvalue |
|---------|---------|-------|------|---------|---------|--------|
| TMEM63B | AP      | 3.787 | 39.655 | 5.903   | 266.372 | 0.000  |
| HDAC9   | AP      | 3.552 | 5.669 | 2.176   | 14.767  | 0.000  |
| GSG1L   | AP      | -3.546 | 0.270 | 0.131   | 0.557   | 0.000  |
| RPL39L  | AP      | -3.374 | 0.182 | 0.068   | 0.490   | 0.001  |
| GSG1L   | AP      | 3.335  | 5.367 | 1.650   | 6.872   | 0.001  |
| UBXN11  | ES      | 3.182  | 7.709 | 2.191   | 27.121  | 0.001  |
| CPEB4   | ES      | 3.097  | 14.981 | 2.702   | 83.056  | 0.002  |
| FAM13A  | AT      | -3.077 | 0.013 | 0.001   | 0.205   | 0.002  |
| CDKL3   | AT      | 3.099  | 26.547 | 3.136   | 224.748 | 0.003  |
| MAT2B   | AP      | -3.007 | 0.026 | 0.002   | 0.282   | 0.003  |
| IP6K2   | ES      | 2.927  | 23.170 | 2.824   | 190.096 | 0.003  |
| MTIF2   | ES      | 2.924  | 20.596 | 2.711   | 156.492 | 0.003  |
| HNF1A   | AT      | -2.919 | 0.014 | 0.001   | 0.244   | 0.004  |
| ROBO2   | AP      | 2.882  | 5.269 | 1.702   | 16.312  | 0.004  |
| SPAT52L | AP      | -2.868 | 0.015 | 0.001   | 0.262   | 0.004  |
| ERC1    | ES      | 2.845  | 10.922 | 2.104   | 56.701  | 0.004  |
| ARPP21  | AT      | 2.841  | 6.502 | 1.795   | 24.277  | 0.005  |
| CLEC16A | ES      | 2.794  | 35.946 | 2.913   | 443.528 | 0.005  |
| ROBO2   | AP      | -2.743 | 0.202 | 0.064   | 0.633   | 0.006  |
| ZNF195  | ES      | 2.733  | 19.203 | 2.306   | 159.900 | 0.006  |

### Figures
Figure 1

Flow chart for analyzing GBM alternative splicing in large-scale RNA-Seq data. The RNA sequence data and the corresponding clinical information of the GBM cohort can be downloaded from TCGA data portal website. After excluding patients with incomplete clinical data and repeated data, we combined these data sets into a large-scale GBM cohort, and then further analyzed them via SpliceSeq. Based on the PSI value of each AS event, we determined the DEAS between GBM and normal tissues, and further studied
the parental genes of these AS events through enrichment analysis. Next, we visually analyzed the interaction network of these parental genes and the regulatory network of DEAS and splicing factors. Finally, to evaluate the prognostic value of each DEAS event in GBM, Cox regression and Kaplan-Meier analysis were used to determine their impact on OS.

Figure 2
A. The number of AS events and gene symbols for GBM shown in an UpSet and venn plot. B. The number of DEAS events and gene symbols for GBM shown in an UpSet and venn plot. (A) Illustration of the splicing patterns of seven types of AS events, including alternate acceptor site (AA), alternate donor site (AD), alternate promoter (AP), alternate terminator (AT), exon skip (ES), mutually exclusive exons (ME), and retained intron (RI).
Functional analyses of parent genes from 234 DEAS events in GBM, including GO and KEGG analyses. A. Intersection of DEAS and DEGS. B. Genes (234) have some common transcription factors. C. Gene Ontology (GO) analyses. The enrichment analysis of gene symbols in 234 alternative splicing events. The top 10 enriched terms are displayed. D. KEGG pathway enrichment analysis of gene symbols in 234 alternative splicing events. The top 20 enriched terms are displayed. E. UpSet plots in GBM, showing the interactions among the seven types of OS-associated DEAS events. One gene may have up to seven types of DEAS events. F. COSMIC analysis of 234 genes.
Figure 4

Top 20 most significant AS events in GBM. (A) The volcano plots of prognosis-connected AS events. The top 20 AS events associated with survival outcome for GBM in different splice patterns, including (B) AA, alternate acceptor site; (C) AD, alternate donor site; (D) AP, alternate promoter; (E) AT, alternate terminator; (F) ES, exon skip; (G) ME, mutually exclusive exons; and (H) RI, retained intron.

Figure 5
A and B. LASSO regression analyses were then performed in prognosis-related AS events in (A) all types of DEAS events. C. Kaplan-Meier curve of prognostic predictors constructed with all seven AS types in the GBM cohort. Red line indicates high-risk subgroup, whereas blue line indicates low-risk subgroup. D. ROC curves with calculated AUC values of prognostic predictors constructed with all seven AS types in the GBM cohort. ALL: all seven AS types combined. E. Patients were divided into high- and low-risk subgroups based on the median of risk scores based on the GBM prognostic predictor. The upper part is the heatmap of DEAS events involved in the prognostic predictor; color transition from green to red indicates the increasing PSI score of corresponding DEAS event from low to high. The middle part is the survival status and survival time of each individual. Color of each plot represents the survival status of each patient. The bottom part is risk score of each individual. F. Univariate and multivariate Cox regression analyses of prognosis-related AS events in all types of AS events groups were performed in forest plots of all types of AS events.
Figure 6

A. SF-AS correlation network. Blue triangles were OS-associated splicing factors. Red/green lines represent positive/negative correlations between substances. Red/green dots represent adverse/favorable AS events. B. Biological pathway enrichment analysis of SF. C. analysis SF Site of expression. D. Intersecting 135 OS-DEAS with the lists of immune-related genes obtained from ImmPort
and InnateDB. E. the expression of NR1H3 mRNA. F. NR1H3 expression is related to cg07298473 methylation site. G. Trend relationship between NR1H3 expression and its methylation sites.