

# RNA processing genes characterize RNA splicing and further stratify lower-grade glioma

Rui-Chao Chai,<sup>1,2,3</sup> Yi-Ming Li,<sup>1,2,4</sup> Ke-Nan Zhang,<sup>1,2</sup> Yu-Zhou Chang,<sup>4</sup> Yu-Qing Liu,<sup>1,2</sup> Zheng Zhao,<sup>1,2</sup> Zhi-Liang Wang,<sup>1,2</sup> Yuan-Hao Chang,<sup>1,2</sup> Guan-Zhang Li,<sup>1,2</sup> Kuan-Yu Wang,<sup>1,2</sup> Fan Wu,<sup>1,2</sup> and Yong-Zhi Wang<sup>1,2,3,4</sup>

<sup>1</sup>Department of Molecular Neuropathology, Beijing Neurosurgical Institute, Capital Medical University, Beijing, China.

<sup>2</sup>Chinese Glioma Genome Atlas, Beijing, China. <sup>3</sup>China National Clinical Research Center for Neurological Diseases and

<sup>4</sup>Department of Neurosurgery, Beijing Tiantan Hospital, Capital Medical University, Beijing, China.

**BACKGROUND.** Aberrant expression of RNA processing genes may drive the alternative RNA profile in lower-grade gliomas (LGGs). Thus, we aimed to further stratify LGGs based on the expression of RNA processing genes.

**METHODS.** This study included 446 LGGs from The Cancer Genome Atlas (training set) and 171 LGGs from the Chinese Glioma Genome Atlas (validation set). The least absolute shrinkage and selection operator (LASSO) Cox regression algorithm was conducted to develop a risk signature. The receiver operating characteristic curves and Kaplan-Meier curves were used to study the prognostic value of the risk signature.

**RESULTS.** Among the tested 784 RNA processing genes, 276 were significantly correlated with the overall survival of LGGs. Further LASSO Cox regression identified a 19-gene risk signature, whose risk score was also an independent prognosis factor ( $P < 0.0001$ , multiplex Cox regression) in the validation data set. The signature had better prognostic value than the traditional factors “age,” “grade,” and “WHO 2016 classification” for 3- and 5-year survival both data sets (AUCs >85%). Importantly, the risk signature could further stratify the survival of LGGs in specific subgroups of WHO 2016 classification. Furthermore, alternative splicing events for genes such as *EGFR* and *FGFR* were found to be associated with the risk score. mRNA expression levels for genes, which participated in cell proliferation and other processes, were significantly correlated to the risk score.

**CONCLUSIONS.** Our results highlight the role of RNA processing genes for further stratifying the survival of patients with LGGs and provide insight into the alternative splicing events underlying this role.

**FUNDING.** The National Natural Science Foundation of China (81773208, 81402052), the Beijing Nova Program (Z16110004916082), and the National Key Research and Development Plan (2016YFC0902500).

**Authorship note:** RCC and YML contributed equally to this work.

**Conflict of interest:** The authors have declared that no conflict of interest exists.

**Copyright:** © 2019, American Society for Clinical Investigation.

**Submitted:** May 28, 2019

**Accepted:** August 6, 2019

**Published:** September 5, 2019.

**Reference information:** JCI Insight. 2019;4(17):e130591.  
<https://doi.org/10.1172/jci.insight.130591>

## Introduction

Diffuse glioma is the most frequent primary malignant brain tumor in adults (1, 2). Although intensive treatment with surgery, radiotherapy, and chemotherapy are conducted, nearly all gliomas relapse. The median overall survival (OS) of patients with glioblastoma (GBM; WHO grade IV) is only approximately 15 months (1, 3, 4). Patients with lower-grade diffuse glioma (LGG, WHO grade II/III) have a relatively favorable prognosis, but these tumors evolve and most relapse as therapy-resistant higher-grade gliomas over time (5–7). LGGs can be classified as WHO grade II or III based on histological features or as oligodendroglioma with isocitrate dehydrogenase–mutant (*IDH*-mutant) and 1p/19q codeletion, astrocytoma with *IDH*-mutant, and astrocytoma with *IDH*-wild-type according to the WHO 2016 integrated analysis (1, 3, 8). However, this classification still does not fully reflect the rate of LGG evolution and, for the OS of patients with LGGs in a stratified subgroup, is rather heterogeneous (7, 8).

Thus, further stratification of LGGs using various methods (9–15), including RNA expression profiles (11, 14, 16–18), has attracted attention.

A dysregulated RNA expression profile is an important hallmark of tumors (19). RNA processing factors, such as RNA-binding proteins and RNA methylation regulators, are the main factors controlling the life cycle of RNA in cells and thus play vital roles in numerous diseases and cancers by determining the final content and isoforms of mature RNA transcripts (6, 20–24). However, studies on RNA processing factors in gliomas have mainly focused on GBM, and the roles of RNA processing factors in stem cell maintenance, cell adhesion, cell migration, cell growth, colony formation, and invasiveness have been demonstrated in GBM (6, 20, 25–29). Recently, some RNA processing factors, such as *HnRNPH1*, *PTB*, *SNRPI*, and *ELAVL*, which are evaluated in the progression and prognosis of GBM, also showed prognostic value for all grades of diffuse gliomas (6), indicating their potential role in LGGs. Considering that abnormal expression of RNA processing genes may drive changes in the dysregulated RNA expression profile in glioma, it is important to systematically examine the roles of RNA processing factors in LGGs.

RNA processing factors also govern the alternative splicing events (ASEs) of individual genes, which can affect several important factors in tumor initiation and progression (30). For instance, our recent study demonstrated the vital role of *MET* exon 14 skipping in the malignant progression and targeted therapy of secondary GBM (31). The Cancer Genome Atlas (TCGA) SpliceSeq data set lists 7 ASE types (32). This makes it possible to link the dysregulation of RNA processing factors with aberrant ASEs in LGGs.

In this study, we systemically analyzed the expression profile of RNA processing genes and their prognostic values in 617 LGGs from TCGA ( $n = 446$ ) and the Chinese Glioma Genome Atlas (CGGA;  $n = 171$ ). We also profiled the ASEs underlying LGGs stratified by the risk signature built with RNA processing genes and identified the corresponding functions.

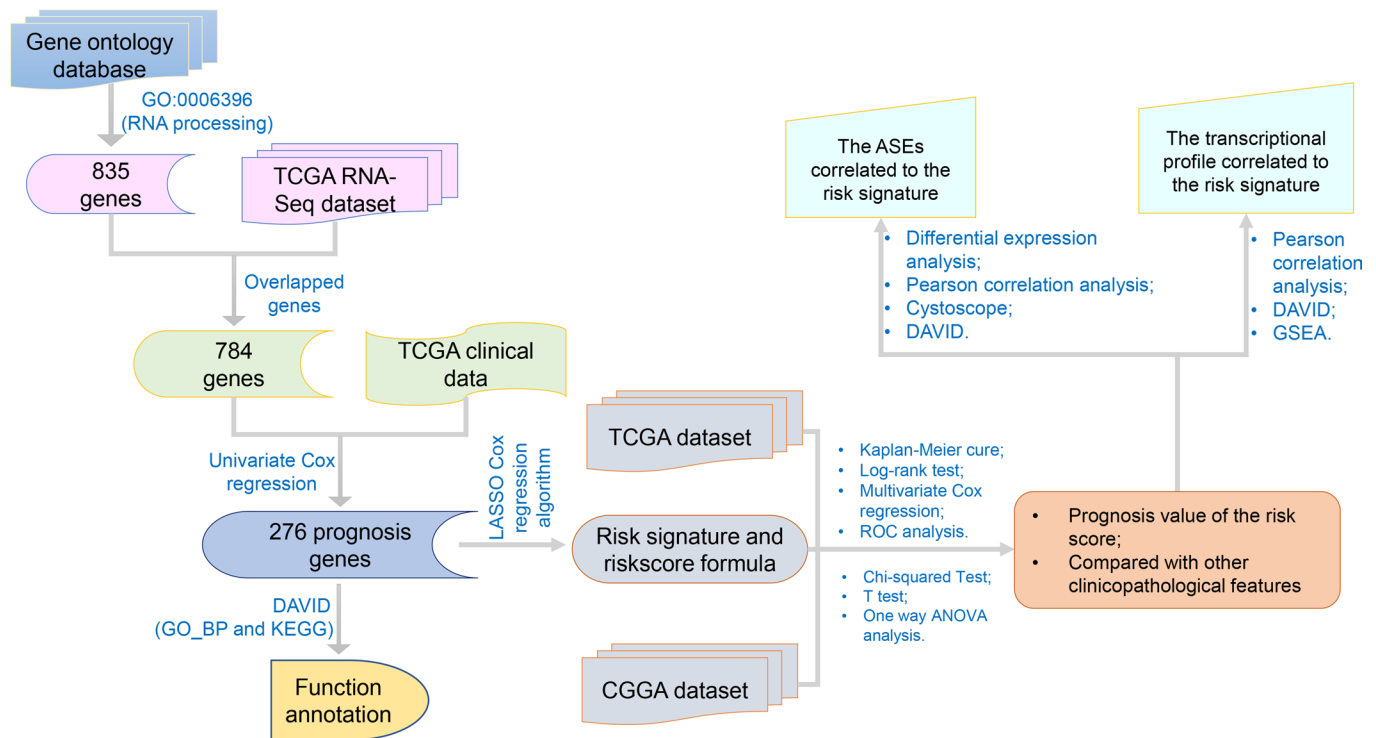
## Results

*Prognostic value of RNA processing genes and their biological function in LGGs.* Our approach and workflow for the selection of RNA processing genes, development and validation of a prognostic risk signature, and the analysis of the ASEs and altered RNA expression profile that correlated to the signature are summarized in Figure 1.

Of the 784 RNA processing genes tested, 276 were significantly correlated with the OS of LGGs (Supplemental Table 1; supplemental material available online with this article; <https://doi.org/10.1172/jci.insight.130591DS1>). A total of 142 of the 276 genes had a HR >1 and were considered risk associated, while the remaining 134 genes had a HR <1 and were considered protection associated. Risk-associated genes showing increasing expression and protection-associated genes showing decreasing expression were associated with increased malignancy in 3 subgroups according to the WHO 2016 classification (Figure 2A). Some of these genes also showed inconsistent expression levels within the same subgroups, suggesting the potential prognostic value of these genes within specific subgroups.

Since these 276 genes are a collection of genes that participated in any process involved in the conversion of one or more primary RNA transcripts into one or more mature RNA molecules. We used Gene Ontology (GO) analysis to study the more specific biological processes that these genes are enriched in, and the results indicated that these survival-associated genes were correlated with terms such as rRNA processing, nuclear-transcribed mRNA catabolic process, nonsense-mediate decay, and mRNA splicing, via spliceosome (Figure 2B). We also further annotated the 142 risk-associated and 134 protection-associated genes through GO and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis, respectively (Supplemental Tables 2 and 3). The results indicated that GO\_BP terms of mRNA splicing via spliceosome and KEGG pathway of spliceosome were the most associated terms for the 142 risk-associated genes and the GO\_BP terms of rRNA processing and KEGG pathway of ribosome were the most associated terms for the 134 protection-associated genes.

*Identification of a panel of 19 RNA processing genes as a risk signature in LGGs.* To easily and reliably stratify the outcomes of LGGs with RNA processing genes, we applied the least absolute shrinkage and selection operator (LASSO) Cox regression algorithm to the 276 genes in the TCGA data set (Supplemental Figure 1). A total of 19 genes were selected to build the risk signature, and the coefficients and normalized expression levels of these genes were used together to calculate risk scores for both the training data set (TCGA data set) and the validation data set (CGGA data set) (Figure 2C and Supplemental Table 4).



**Figure 1. The workflow of the current study.**

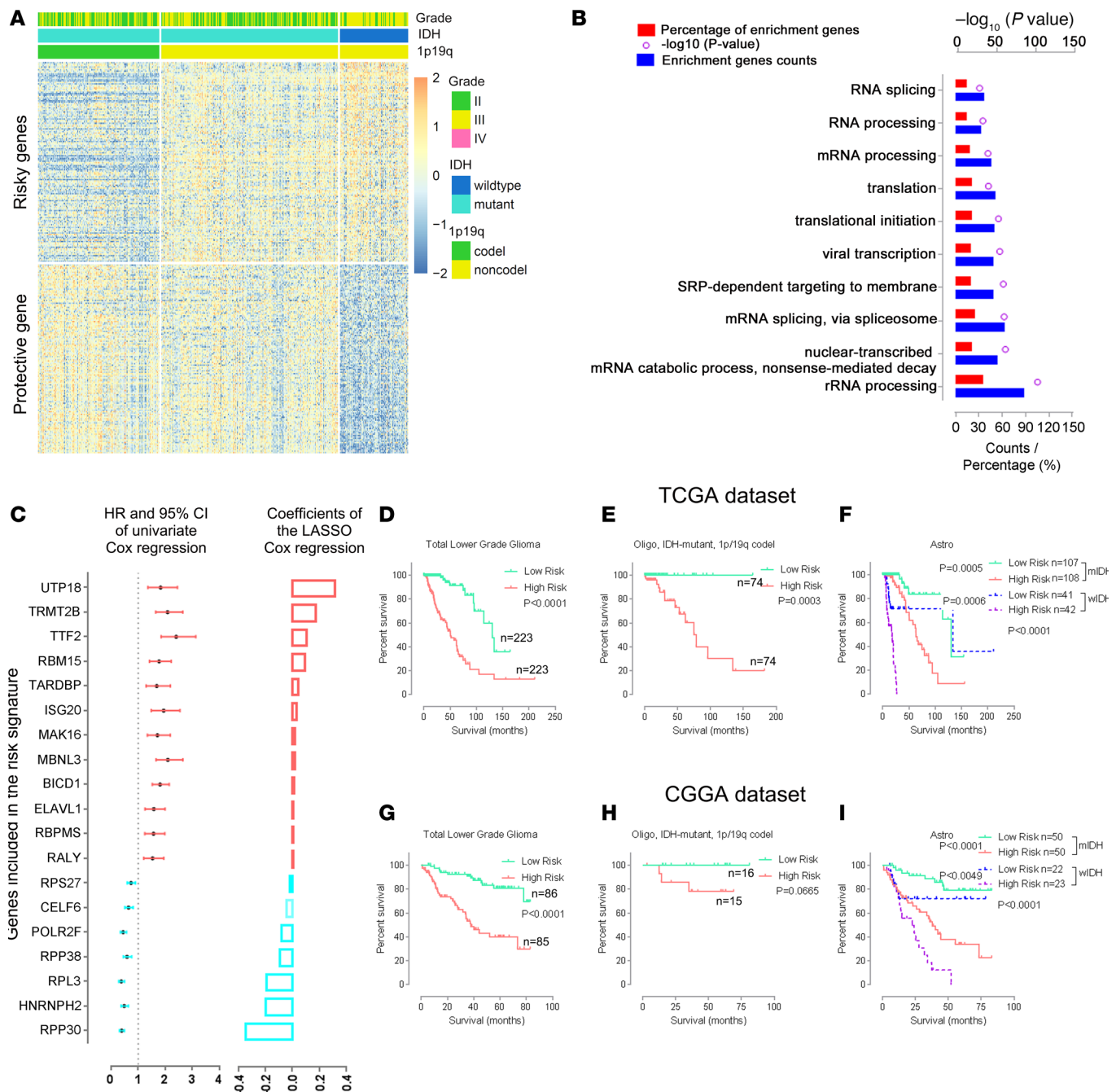
We divided patients into high-risk and low-risk groups in various stratified glioma subtypes using their respective median risk score as a cutoff. We found that patients with low-risk scores had significantly longer OS than patients with high-risk scores in all subtypes evaluated in TCGA data sets: all LGGs (Figure 2D), oligodendroglioma with *IDH*-mutant and 1p/19q codeletion (Figure 2E), and astrocytoma with *IDH*-mutant and *IDH*-wild-type (Figure 2F and Supplemental Figure 2, A and B). Similar results were observed in the CGGA data sets (Figure 2, G–I, and Supplemental Figure 2, C and D).

To investigate the prognostic value of the risk signature and other clinicopathological features, univariate and multivariate Cox regression analysis were performed both in the TCGA (training) and CGGA (validation) data sets. The results revealed that high-risk score was a prognostic factor in both data sets ( $P < 0.0001$ ), independent of grade, age, sex, WHO 2016 subgroups, *IDH* mutation, 1p19q codeletion, and *MGMT* methylation status (Table 1).

We also summarized the race composition in data sets used in this study, as a specific race may represent population-specific exposure and risk factor. In the TCGA data set, 92.60% of cases are White, 3.36% of cases are Black or African American, 1.79% of cases are Asian, 0.22% of cases are American Indian or Alaska Native, and the remaining 2.02% of cases were unknown (Supplemental Figure 3A). All of the cases in the CGGA data set are Asian. We found that patients with low-risk scores had significantly longer OS than patients with high-risk scores in White (Supplemental Figure 3B) and Black or African American (Supplemental Figure 3C) populations of TCGA data set. We cannot compare the OS of the 8 cases of Asian descent in the TCGA data set, as all of them survived at the end of follow-up (Supplemental Figure 3D). However, the signature could stratify the survival of Asian population, since all the cases in the CGGA data set are Asian.

Collectively, these results strongly support the ability of our risk signature to accurately stratify the prognosis of patients with definite subgroups based on WHO 2016 integrated diagnosis.

*The signature risk score is correlated with malignant clinicopathological features of LGGs.* To determine whether the risk signature also reflects the malignant clinicopathological features of LGGs, we determined risk signature expression of 19 genes and clinicopathological features in LGGs with low and high risk in the form of a heatmap (Supplemental Figure 4A). Significant differences were observed between the high- and low-risk groups with respect to the WHO grade ( $P < 0.001$ ), *IDH* status ( $P < 0.001$ ), 1p/19q codeletion status ( $P < 0.001$ ), WHO



**Figure 2. Prognosis-associated RNA processing gene expression profile in LGGs.** (A) Heatmap showing the expression pattern of the 276 RNA processing genes associated with patient survival. (B) GO biological process terms enriched among the 276 survival-associated RNA processing genes.  $P$  value calculated by Fisher Exact test in DAVID annotation system. (C) The 19 genes included in the signature, their HRs, and 95% CIs by univariate Cox regression analysis and the coefficients by multivariate Cox regression analysis using LASSO. (D–I) Kaplan-Meier overall survival (OS) curves for patients from the TCGA (D–F) and CGGA (G–I) data sets with total lower-grade glioma (D and G), oligodendroglioma with *IDH*-mutant and 1p/19q codeletion (E and H), and astrocytoma with *IDH*-mutant or *IDH*-wild-type (F and I).  $P$  values for D–I calculated by log-rank test in Kaplan-Meier curves comparison. All of the sample numbers ( $n$ ) and  $P$  values are labeled in the figure.

2016 subgroups ( $P < 0.001$ ), *MGMT* promoter methylation status ( $P < 0.001$ ), age ( $P < 0.001$ ), and *CIC* ( $P < 0.001$ ) (Supplemental Figure 4A and Supplemental Table 5). We also found that the risk score was significantly different in patients with LGGs stratified by WHO grade, subgroups according to WHO 2016 classification, *IDH* mutation status, 1p/19q codeletion status, *MGMT* promoter methylation status, and age (Supplemental Figure 4, B–G). Moreover, the distributions of risk score among LGGs with and without *TERT* promoter mutation were highly dependent on the *IDH* mutant status (Supplemental Figure 4H). The risk score significantly



**Table 1. Univariate and multivariate Cox regression in the TCGA (training) and CGGA (validation) data sets**

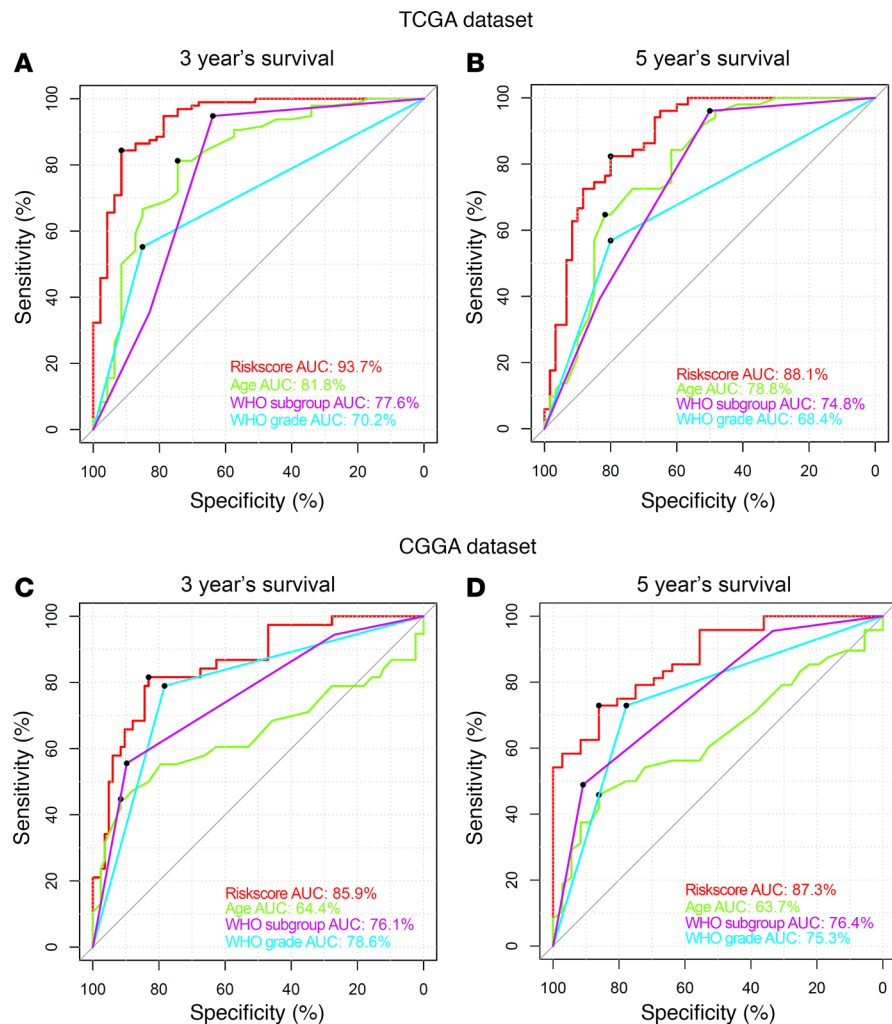
		Univariate cox analysis				Multivariate cox analysis			
		<i>P</i> value	HR	HR (95% CI)		<i>P</i> value	HR	HR (95% CI)	
The TCGA data set	Grade (II vs. III)	$3.86 \times 10^{-6}$	0.306	0.185	0.505	$2.19 \times 10^{-1}$	0.698	0.394	1.239
	Age (age <40 vs. age ≥40)	$2.31 \times 10^{-12}$	1.068	1.049	1.088	$9.53 \times 10^{-6}$	1.048	1.026	1.07
	Sex (female vs. male)	$8.78 \times 10^{-1}$	1.036	0.662	1.621				
	IDH (mutant vs. wild-type)	$4.49 \times 10^{-14}$	0.161	0.1	0.258	$8.15 \times 10^{-1}$	0.902	0.382	2.129
	1p/19q (codel vs. noncodel)	$2.95 \times 10^{-3}$	0.425	0.241	0.747	$2.42 \times 10^{-1}$	0.68	0.356	1.297
	WHO 2016 subgroups								
	Oligo, 1p/19q codel, IDH-mutant	Reference							
	Astro, IDH-mutant	$2.71 \times 10^{-1}$	1.412	0.764	2.61				
	Astro, IDH-wild-type	$2.04 \times 10^{-10}$	7.754	4.124	14.579				
	MGMT (methylation vs. unmethylation)	$1.52 \times 10^{-3}$	0.444	0.269	0.733	$3.38 \times 10^{-1}$	1.401	0.703	2.789
The CGGA data set	Risk score (high vs. low)	$4.85 \times 10^{-26}$	4.7	3.526	6.265	$6.68 \times 10^{-10}$	3.589	2.392	5.384
	Grade (II vs. III)	$9.97 \times 10^{-9}$	0.166	0.09	0.307	$2.40 \times 10^{-3}$	0.578	0.406	0.823
	Increasing age	$9.21 \times 10^{-4}$	1.048	1.019	1.078	$1.19 \times 10^{-1}$	1.024	0.994	1.055
	Sex (female vs. male)	$7.83 \times 10^{-1}$	0.923	0.521	1.634				
	IDH (mutant vs. wild-type)	$2.36 \times 10^{-6}$	0.253	0.143	0.447	$9.05 \times 10^{-1}$	25.541	$2.04 \times 10^{-22}$	$3.19 \times 10^{24}$
	1p/19q (codel vs. noncodel)	$8.08 \times 10^{-3}$	0.147	0.036	0.607	$9.01 \times 10^{-1}$	0.001	$7.55 \times 10^{-50}$	$1.82 \times 10^{43}$
	WHO 2016 subgroups								
	Oligo, 1p/19q codel, IDH-mutant	Reference							
	Astro, IDH-mutant	$4.41 \times 10^{-2}$	4.426	1.04	18.833				
	Astro, IDH-wild-type	$2.19 \times 10^{-4}$	15.542	3.629	66.568				
	MGMT (methylation vs. unmethylation)	$5.30 \times 10^{-3}$	0.249	0.094	0.662	$4.64 \times 10^{-1}$	0.787	0.416	1.492
	Risk score (high vs. low)	$1.13 \times 10^{-13}$	11.578	6.065	22.1	2.23E-05	7.554	2.967	19.233

IDH, isocitrate dehydrogenase; codel, codeletion; Oligo, oligodendrocytoma; Astro, astrocytoma. *P* value obtained by univariate and multivariate Cox regression analysis. Oligo, 1p/19q codel, IDH-mutant is the reference for Astro, IDH-mutant and Astro, IDH-wild-type in the WHO 2016 subgroups.

increased in *TERT*-mutant LGGs with *IDH*-wild-type ( $P < 0.01$ ), while the risk score decreased in *TERT*-mutant LGGs with *IDH*-mutant ( $P < 0.01$ ). A similar distribution of risk score was observed in the CGGA data set (all  $P < 0.05$ , except for *TERT* status, and  $P = 0.07$  for *TERT* status, Supplemental Figure 5). These results indicated that the signature was associated with the malignant clinicopathological features of gliomas.

The receiver operating characteristic (ROC) curve analysis showed that the signature risk score had the best efficiency (compared with age, WHO 2016 subgroups, and WHO grade) for predicting the 3-year and 5-year survival of patients with LGGs both in the TCGA and CGGA data sets (Figure 3). Areas under the curve of risk score, age, WHO 2016 subgroups, and WHO grade were 93.7%, 81.8%, 77.6%, and 70.2%, respectively, for 3-year survival in the TCGA data set; 88.1%, 78.8%, 74.8%, and 68.4%, respectively, for 5-year survival in the TCGA data set; 85.9%, 64.4%, 76.1%, and 78.6%, respectively, for 3-year survival in the CGGA data set; and 87.3%, 63.7%, 76.4%, and 75.3%, respectively, for 5-year survival in the CGGA data set.

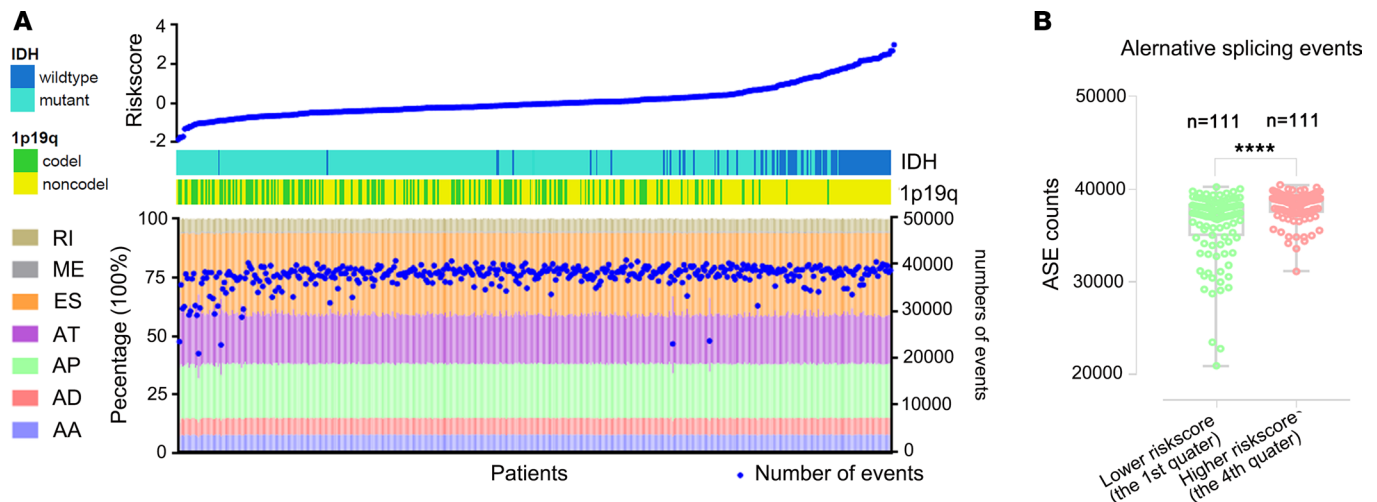
*The signature risk score is closely correlated with the ASEs.* RNA splicing activities are governed by RNA processing genes, and our findings showed that worse survival-associated RNA processing genes were most enriched in RNA splicing-related activities. Thus, we also systematically characterized ASEs in LGGs with different risk scores. In total, tens of thousands of 7 ASE types, 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), were detected in each LGG (Supplemental Figure 6A). Furthermore, the proportion of differential ASE types in LGGs varied widely (from 0.5% to approximately 35%). Although all LGGs shared similar patterns of ASE types (ES and AP were the most frequently observed ASE types, while ME and RI were the least frequently observed ASE types), the total number of detected differential ASEs gradually increased along with the increasing risk score (Figure 4A), and the total number of ASEs was significantly smaller in LGGs with a lower-risk (first quarter,  $n = 111$ ) score compared with those with a higher-risk (fourth quarter  $n = 111$ ) score (Figure 4B).



**Figure 3. Comparisons of survival predictive efficiencies between the risk score and clinicopathological characteristics.** ROC curves showed the predictive efficiencies of risk scores, age, subgroups according to WHO 2016 classification, and WHO grade on 3-year and 5-year survival in the TCGA (**A**,  $n = 143$ ; **B**,  $n = 111$ ) and CGGA (**C**,  $n = 121$ ; **D**,  $n = 84$ ) data sets.

Percent spliced in (PSI) is the ratio of reads indicating the presence of a transcript element versus the total reads covering the event (32). We also analyzed the PSI levels for ASEs in all LGGs and LGGs with lower- and higher-risk scores. The results revealed that events with lower PSI levels ( $\text{PSI} \leq 0.2$ ) and higher PSI levels ( $\text{PSI} > 0.8$ ) constituted most types of ASEs in all LGGs and LGGs with lower and higher-risk scores (Supplemental Figure 6, B–H). These observations indicated that transcripts with high splice-out alternative exons or high splice-in alternative exons were the predominantly transcribed form of most genes.

We further identified differentially expressed RNA splicing genes ( $P < 0.05$  and fold-change  $\geq 2$  or  $\leq 0.5$ ) and ASEs with significantly different PSI ( $P < 0.05$  and fold change  $\geq 2$  or  $\leq 0.5$ ) in LGGs with lower- (first quarter,  $n = 111$ ) and higher-risk (fourth quarter,  $n = 111$ ) scores (Figure 5A and Supplemental Tables 6 and 7). There were 257 ASEs for 247 genes with decreased PSI levels in LGGs with higher-risk score and 604 ASEs for 527 genes with increased PSI levels in LGGs with higher-risk scores (Supplemental Figure 7A). The proportion of the AP type of ASEs dramatically increased (from 23.14% to 47.7%) in these ASEs with markedly different PSI. In LGGs with higher-risk score, AP (45.1%) and ES (26.5%) showed the largest proportion in ASEs with decreased PSI, while AP (49.3%) and AT (25.0%) constituted a larger proportion of ASEs with increased PSI (Supplemental Figure 7B). This suggests that ES transcripts with high splice-out exons are more likely to be present in LGGs with higher-risk scores and the AP type of ASEs appeared to be more relevant to the prognosis of LGGs.

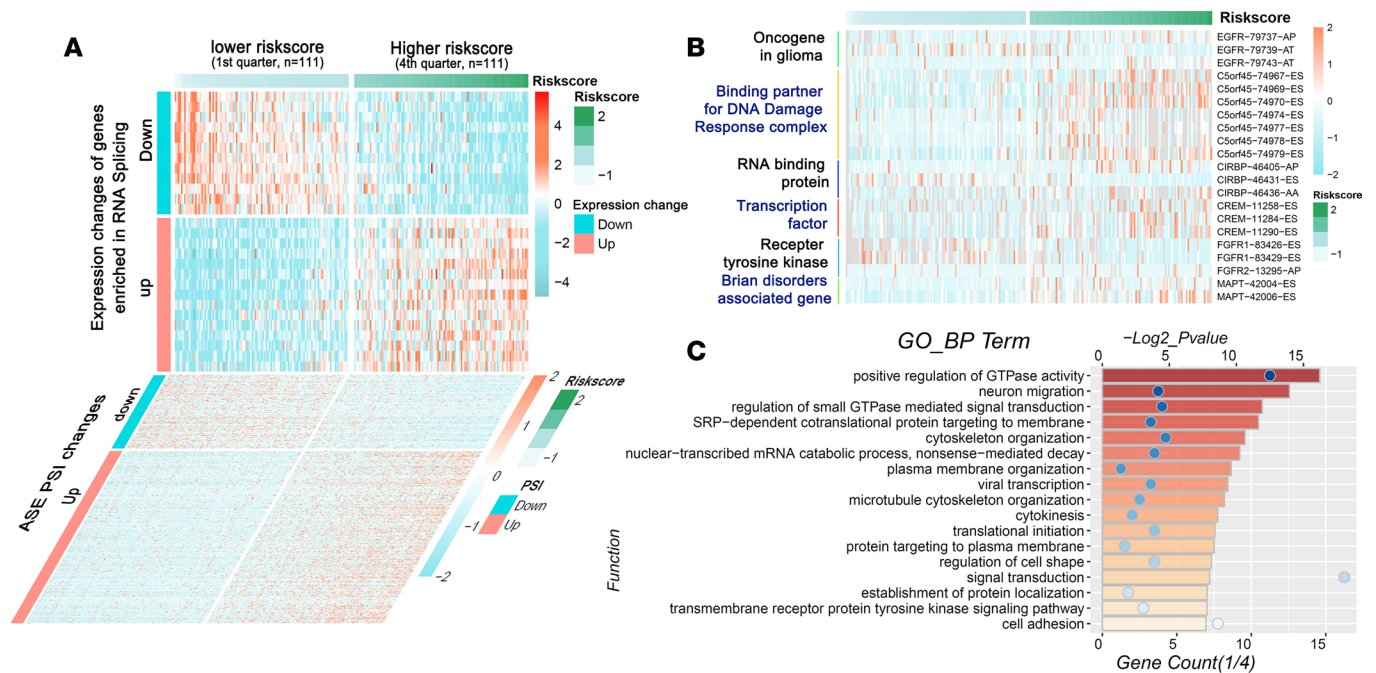


**Figure 4. Alternative splicing profile analysis in LGGs with lower- or higher-risk scores. (A)** Differentially spliced events in 446 LGGs patients with increased risk score. Bars indicate the proportion of each ASE type. Dots indicate the number of differentially spliced events in each patient. The IDH status and 1p/19q codeletion status of these LGGs are also presented. **(B)** The absolute numbers of all alternative spliced events were compared in LGGs with lower- ( $n = 111$ ) or higher-risk ( $n = 111$ ) scores. \*\*\*\* $P < 0.0001$ .  $P$  value obtained by unpaired  $t$  test.

We found that genes involved in the receptor tyrosine kinase signaling pathway (*EGFR*, *FGFR1*, and *FGFR2*), DNA damage response (*C5orf45*), RNA binding (*CIRBP*), transcription regulation (*CREM*), and brain disease development (*MAPT*) were differentially spliced among LGGs with lower and higher-risk scores (Figure 5B). To further explore the function of alternative splicing in the malignancy of LGGs, GO analysis was performed for all genes that were differentially spliced in LGGs with lower- and higher-risk scores. In general, genes with differential PSI were mainly enriched in biological processes, such as “positive regulation of GTPase activity,” “SRP-dependent co-translational protein targeting to membrane,” “cytoskeleton,” “cytokinesis,” “transmembrane receptor protein tyrosine kinase signaling pathway,” “cell adhesion,” and others (Figure 5C and Supplemental Figure 8). Our analysis indicated that differential ASEs participate in many cancer-related biological processes, particularly signal transduction pathways, suggesting that ASEs are a critical mechanism underlying the prognostic value of RNA processing genes in LGGs.

We further investigated the potential regulatory networks among the significantly changed 28 RNA splicing genes and 862 ASEs (see Methods). Finally, a network of the 28 differential RNA splicing genes and 221 of their significantly correlated differential ASEs was constructed (Figure 6A and Supplemental Table 8). The 28 RNA splicing genes regulated distinct amounts of ASEs, ranging from 4 to 104 (Figure 6B). For RNA splicing genes with increased expression in LGGs with higher-risk scores, *TTF2* and *MBNL3* were found to regulate a larger number of ASEs, while *GPATCH1* and *DHX15* regulated fewer ASEs (Figure 6B). In RNA splicing genes with decreased expression in LGGs with higher-risk scores, *POLR2F* and *PQBPI* regulated a larger number of ASEs, while *GPKOW* and *SAP18* regulated fewer ASEs (Figure 6B). The proportion of the AP type of ASE was still the largest one in these 221 ASEs.

*The function analysis of genes with mRNA expression that correlated with the signature risk score in LGGs.* Since RNA processing genes are the main factors controlling the life cycle of RNA in cells, we also evaluate the RNA expression profile influenced by the differentially expressed RNA processing genes. We identified genes positively (Pearson coefficient  $>0.5$  and Bonferroni corrected  $P < 0.01$ ) or negatively (Pearson coefficient  $<-0.5$ , and Bonferroni corrected  $P < 0.01$ ) correlated with the risk score and then annotated their functions using GO analysis for biological processes and KEGG pathways analysis (Figure 7). Overall, the results indicated that the genes that positively correlated with risk scores were mainly enriched in regulation of chromosome division and DNA stability, cell division, DNA replication, cell cycle, DNA repair, angiogenesis, and other malignancy-related biological processes (Figure 7A). The corresponding pathways could also be observed in the KEGG pathway analysis (Figure 7B), suggesting that the altered expression of RNA processing genes may be associated with the increased expression



**Figure 5. Differential ASEs in LGGs with lower- or higher-risk scores.** (A) Heatmaps showing the expression level of RNA splicing genes (top) and PSI of ASEs with significant differences between LGG groups with lower- and higher-risk scores. (B) Representative ASEs with differential PSI between LGG groups with lower- and higher-risk scores. (C) GO\_BP terms of spliced genes with differential PSI between LGGs with lower- and higher-risk scores. P value calculated by Fisher Exact test in DAVID annotation system.

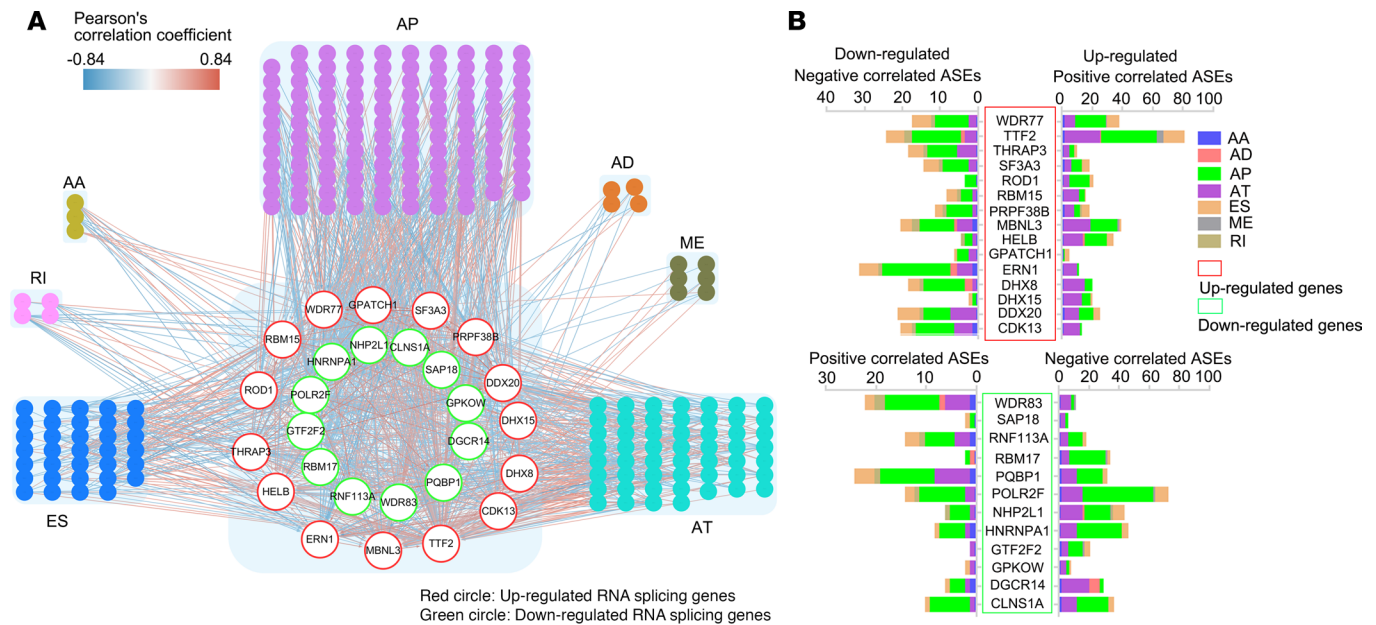
of genes enriched in these processes. Interesting, we observed that the genes that negatively correlated with signature risk scores were mainly involved in the GO processes of translation and rRNA processing and the KEGG pathway of ribosome (Figure 7). This finding is consistent with the finding that the GO\_BP terms of rRNA processing and KEGG pathway of ribosome are the most associated terms for the 134 protection-associated RNA processing genes (Supplemental Table 2).

Furthermore, gene set enrichment analysis (GSEA) revealed the hallmarks of malignant tumors, including epithelial-mesenchymal transition, angiogenesis, G2M checkpoint, mitotic spindle, inflammatory response, complement, *PI3K/AKT/MTOR* signaling, and *KRAS* signaling (Figure 8), as significantly enriched in LGGs with risk scores higher than median risk score. These findings indicate that the risk score of the RNA processing signature reflects the expression alterations of genes involved in malignant biological processes, signaling pathways, and malignant hallmarks in LGGs.

*RNA expression profile influenced by knockdown of mRNA expression of TTF2.* To study whether the change of RNA processing genes is a driver or merely correlative to the dysregulated RNA profile, we selected 3 specific siRNAs to knockdown the mRNA expression of TTF2. This gene is risk associated in the signature (with highest HR), and it is also included in the GO term of RNA splicing. The results indicated that all 3 siRNAs could significantly downregulate the mRNA expression of TTF2 in glioma cell line LN229 cells, and TTF2 siRNA2 appeared to be the most effective siRNA (Figure 9A).

We collected whole transcriptome data of LN229 cells at 48 hours after transfection of TTF2 siRNA2 ( $n = 3$ ) and scramble siRNA ( $n = 3$ ), respectively. Compared with the scrambled siRNA group, there were 2676 genes with increased expression and 2710 genes with decreased expression in the group of TTF2 siRNA, and the mRNA expression of TTF2 gene was also significantly decreased (Figure 9B and Supplemental Table 9). We found that the 2710 genes with decreased expression after knockdown of TTF2 mRNA were significantly enriched in the biological processes of chromosome segregation, cell division, cell cycle, mRNA metabolic, and RNA splicing (Figure 9C and Supplemental Table 10). KEGG pathway analysis showed that these genes were also enriched in the pathways of "cell cycle," "spliceosome," "DNA replication," "RNA transport," "proteasome," "mismatch repair," and others (Figure 9D and Supplemental Table 11). These findings were largely consistent with the result of function analysis of genes with mRNA expression positively correlated to the risk signature,





**Figure 6. The ASE networks and RNA splicing genes.** (A) Labeled circles in the center represent RNA splicing genes. Red circles indicate upregulated RNA splicing genes in LGGs with higher-risk scores, while green circles indicate downregulated splicing genes. Colored circles connected to splicing genes by red or blue lines are distinct types of differential AEs. The red connecting lines represent positive correlation, while blue connecting lines represent negative correlation. (B) The numbers of ASEs significantly correlated to upregulated (top) or downregulated (bottom) RNA splicing genes.

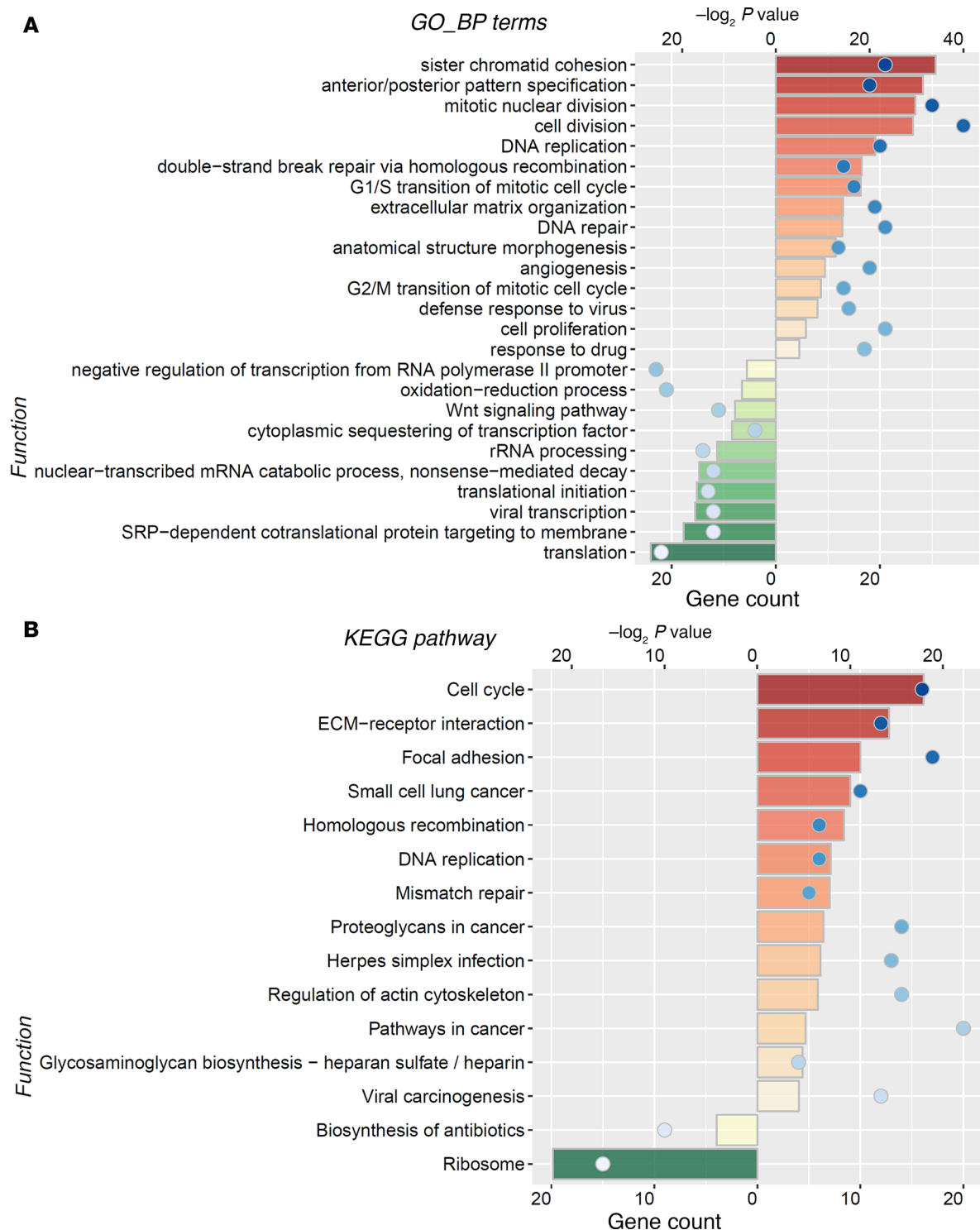
indicating that changes of RNA processing gene TTF2 maybe a driver of these dysregulated RNA profiles in glioma.

As mentioned above, functions of genes with mRNA expression negatively correlated to the risk signature mainly enrich in the biological processes of “translation” and “SRP-dependent co-translational protein targeting to membrane,” the KEGG pathways of “Ribosome.” We also noticed that functions of the 2676 upregulated genes were mainly involved in protein-processing activities, such as the biological processes of “extracellular matrix organization,” “response to endoplasmic reticulum stress,” and others and the KEGG pathways of “lysosome,” “protein processing in endoplasmic reticulum,” and others (Figure 9, C and D, and Supplemental Tables 12 and 13).”

## Discussion

In this study, we found that the general expression pattern of RNA processing genes is correlated with the malignancy features of LGGs and identified RNA processing genes significantly associated with the prognosis of LGGs. We further built a risk signature that not only further predicted the prognosis of stratified LGGs but also could perfectly reflect the malignancy-correlated clinicopathological features, biological processes, key signaling pathways, and hallmarks. Moreover, we systemically analyzed ASEs, underlying LGGs with lower- and higher-risk scores, and identified the corresponding functions. Our results highlight the role of RNA processing genes in further stratifying the survival of patients with LGGs, and the developed signature shows a potential to become an effective supplement for the integrated diagnosis criteria of WHO 2016 in further stratifying the prognosis of LGGs. Moreover, we also reveal the associated ASEs and biological processing that were associated with the risk scores of the developed signature.

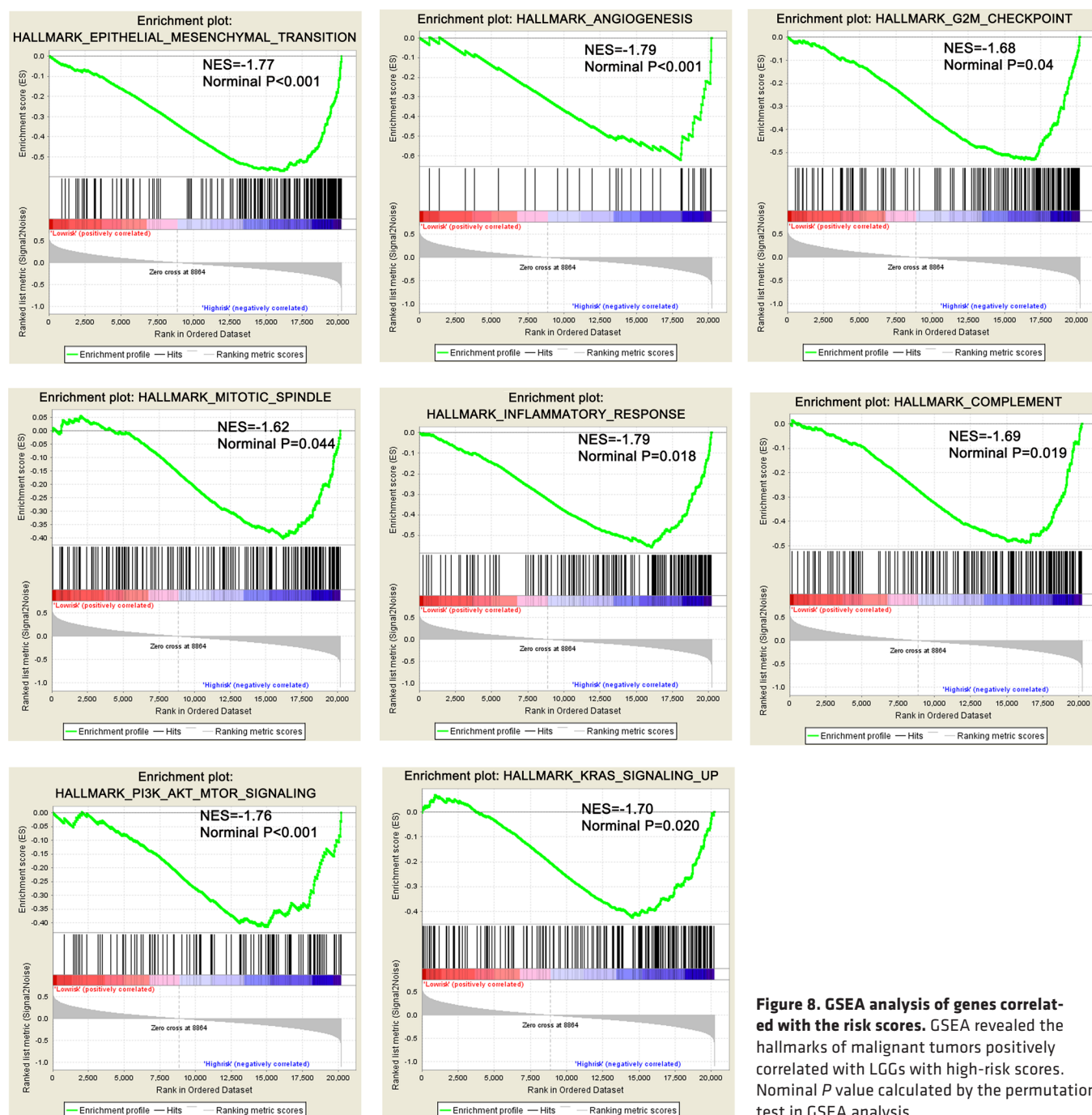
Genetic alterations, such as *IDH* mutation and 1p/19q codeletion, have been included in the integrated diagnosis criteria of WHO 2016 (3). Very large differences in prognosis exist in LGGs within the definite WHO 2016 subgroup (16). Several important biomarkers have been identified to further stratify LGGs, including genetic alterations in *CDKN2A/B* in astrocytoma with *IDH*-mutant (11, 33); *EGFR* amplification, chromosome 7 gain and chromosome 10 loss, and *TERT* promoter mutation in astrocytoma with *IDH*-wild-type (11). Importantly, “the *EGFR* amplification, or chromosome 7 gain and chromosome 10 loss, or *TERT* promoter mutation” had been recommended as the diagnostic criteria for “diffuse astrocytic glioma, *IDH*-wild-type, with molecular features of GBM, WHO grade IV” by the Consortium to Inform Molecular



**Figure 7. Functions of genes correlated with risk scores.** (A and B) Functional analysis of genes positively (red bar chart) or negatively (green bar chart) correlated with the risk score using GO terms of biological processes (A) and the KEGG pathway (B). *P* value calculated by Fisher Exact test in DAVID annotation system.

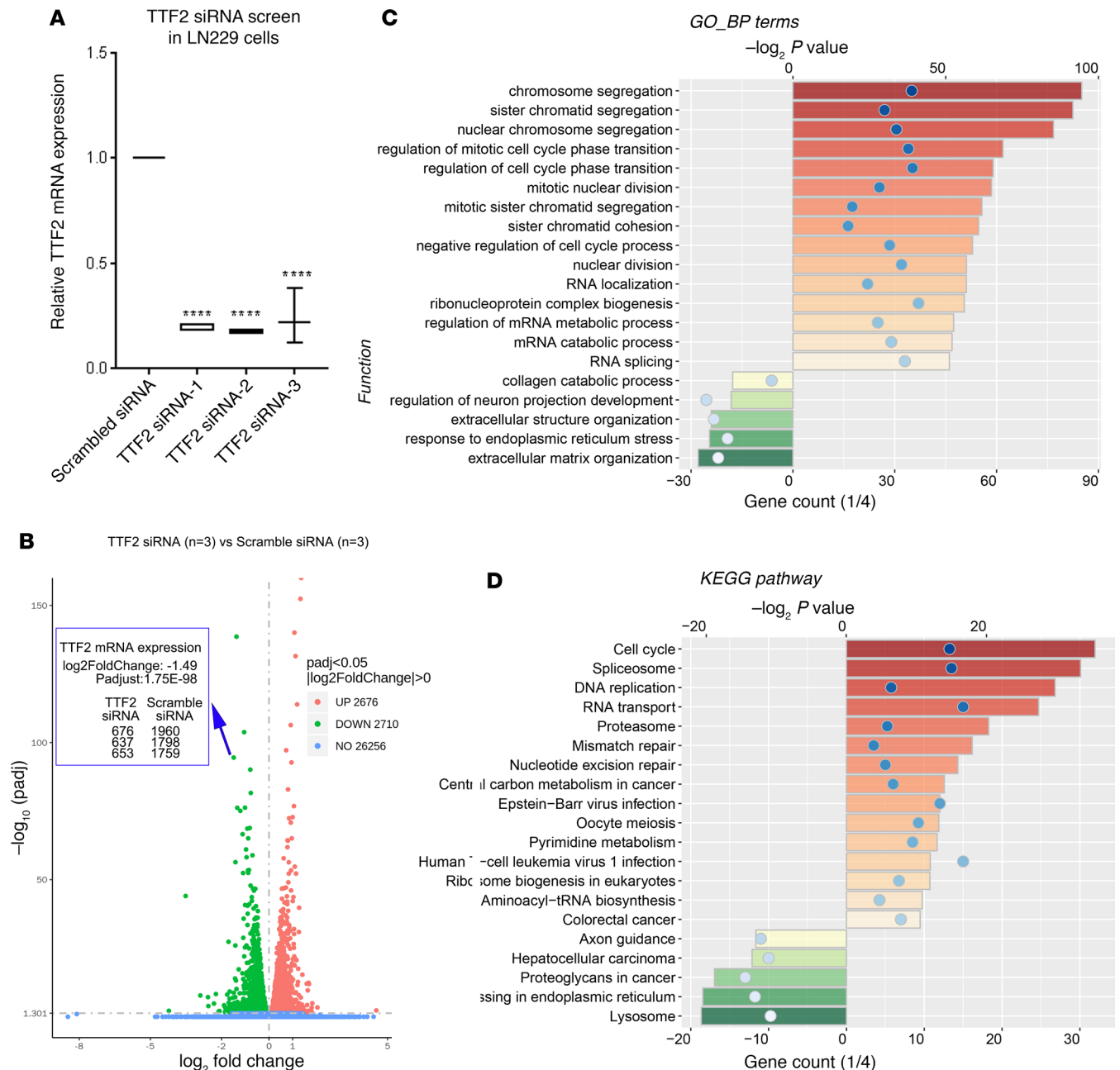
and Practical Approaches to Central Nervous System Tumor Taxonomy (34). Thus, further stratification of LGs with definite *IDH* and 1p/19q status is urgently needed.

Apart from these genetic alteration markers, RNA expression is also valuable for predicting the outcomes of tumors with low rates of mutation, and the RNA expression profile has become an indispensable element for classifying medulloblastoma (35). Compared with GBM, LGs have relative lower mutation rates and the



**Figure 8. GSEA analysis of genes correlated with the risk scores.** GSEA revealed the hallmarks of malignant tumors positively correlated with LGGs with high-risk scores. Nominal *P* value calculated by the permutation test in GSEA analysis.

RNA expression of a set of genes was suggested to be useful for predicting the prognosis of patients with 1p19q co-deletion diffuse glioma (16). In this study, we confirmed the prognostic value of a signature built with 19 RNA processing genes in each stratified subgroup of LGGs (Figure 2, D–I). Compared with traditional stratifying factors (age, WHO grade, subgroups of WHO 2016 classification), the risk score of this signature also had the best predictive efficiency on both the 3-year and 5-year survival (Figure 3). Though the risk score of the signature is highly associated with traditional clinicopathological features, we confirmed that the risk score is an independent prognosis factor in both the training and validation data set (Table 1). A previous study identified 104 key genes that were prognostic for LGGs. However, the average area under the ROC curve (AUC) values ranged from 0.7 to 0.8 (16), which was lower than the performance of our signature (AUC >0.85). These data indicate that the RNA processing gene signature is a powerful tool for further predicting prognosis of LGGs, which were stratified by IDH and 1p/19q status based on the 2016 WHO classification guidelines.



**Figure 9. Transcriptome changes of LN229 cells after knockdown mRNA expression of TTF2.** (A) TTF2 mRNA expression changes at 48 hours after transfection of TTF2-specific siRNA. \*\*\*\* $P < 0.0001$ ,  $n = 3$ .  $P$  value calculated by the one-way ANOVA. (B) The volcano plot of differentially expression genes between TTF2 siRNA ( $n = 3$ ) and scramble siRNA ( $n = 3$ ).  $P$  value calculated by two-tailed unpaired  $t$  test, adjusted by Bonferroni correction. (C and D) Functional analysis of genes downregulated (red) or upregulated (green) after TTF2 mRNA knockdown using GO terms of biological processes (C) and the KEGG pathway (D).  $P$  value calculated by Fisher Exact test in DAVID annotation system.

The ASEs of individual genes can affect several important players in tumor initiation and progression (6, 30, 36–38). For instance, the  $\delta$  isoform of Max could promote glioma cell proliferation by enhancing functions of *MYC* in GBM harboring *EGFRvIII* mutation (39); 2 *CD97* isoforms, *EGF*-1,2,5 and *EGF*-1,2,3,5, could promote growth, migration, metastasis, and angiogenesis in GBM (40); and an aberrant splicing of cyclin-dependent kinase-associated protein phosphatase, KAP, increases proliferation and migration in GBM (30). Currently, genome-wide analyses of exon expression arrays have begun to reveal the roles of ASEs correlated with the progression and prognosis of GBM (41). However, the role of alterations of ASEs in LGGs is not well understood. Our study not only revealed the landscape of ASEs in LGGs, but also identified 861 ASEs correlated with the prognosis of LGGs



(Supplemental Table 5), although more specific studies are needed to thoroughly determine the functions of these identified ASEs. Moreover, we identified RNA splicing genes correlated with differential ASEs in LGGs with distinct prognosis.

RNA plays crucial roles in biological functions in cells not only by passing genetic information from DNA to protein but also by regulating various biological processes (42). Dysregulation of RNA profiles, including hypoxia-associated gene sets (43), vascular gene sets (44), WNT/ $\beta$ -catenin pathway-related genes (45), genes in the NF- $\kappa$ B signaling pathway (46), microRNA (18), and long noncoding RNA (47), have been shown to play crucial roles in the malignant progression and prognosis of LGGs. Here, we highlight the stratification ability of RNA processing genes in LGGs. Additionally, we found that genes with expression that was significantly correlated with the risk score of the signature built with 19 RNA processing genes were enriched in the biological processes/pathways of cell cycle, cell division, DNA replication and repair, angiogenesis, cell proliferation, and pathways in cancer (Figure 6A). Among the 19 genes included in the signature, *RPL3* and *BICD1* have been suggested to be associated with temozolomide resistance in GBM (48, 49) and *ELAVL1* was reported participated in the cell proliferation, migration, and glioma stem-like cell maintenance (50, 51). The RNA expression profile and RNA fate are highly dependent on RNA processing factors responsible for precise temporal and spatial coordinating gene expression (6, 20). In addition, our findings also confirmed that knockdown of the expression of *TTF2* genes could alter the expression of genes involved in the activities of the cell cycle, RNA splicing, DNA replication, mismatch repair, extracellular matrix organization, and others. All of these indicate that abnormal expression of RNA processing genes is a potential driver of the previously reported dysregulated RNA profiles in LGGs.

In summary, our study highlighted the prognostic value of RNA processing genes in LGGs and revealed a signature with 19 RNA processing genes for further stratifying the outcomes of LGGs with definite WHO subgroups. Clinicopathological features, ASEs, biological processes, signaling pathways, and hallmarks of tumors correlated with the risk signature were also identified. These results provide fundamental information for understanding the roles of RNA processing and indicate the potential clinical implications of RNA processing genes in LGGs.

## Methods

**Patients.** A total of 457 LGG samples were collected from a publication using TCGA data (52). Of these TCGA samples, we excluded 11 samples that do not contain RNA-seq data or molecular pathological information or useful OS information; the other 446 LGG samples with RNA-seq transcriptome data and corresponding clinical and molecular pathological information were obtained from TCGA (<http://cancergenome.nih.gov/>) and used for systematic analysis. Another 171 LGG samples with RNA-seq transcriptome and corresponding clinicopathological information from CGGA (<http://www.cgga.org.cn/>) were used to validate the performance of the risk signature. Clinicopathological information regarding the TCGA and CGGA data sets was summarized in Supplemental Table 14.

**Selection of RNA processing genes.** We first collated a list of 835 genes that participated in any process involved in the conversion of 1 or more primary RNA transcripts into 1 or more mature RNA molecules (<http://amigo.geneontology.org/amigo/term/GO:0006396>). Then we used the 784 genes with available RNA expression data in TCGA data sets for further analysis.

**Identification of the risk signature.** We performed univariate Cox regression analyses of the expression of RNA processing genes to identify genes significantly correlated with the prognosis of patients with LGGs in TCGA data set. Next, we used the LASSO Cox regression algorithm to build an optimal risk signature with the minimum number of genes (53). Finally, a set of genes and their coefficients were determined by minimum criteria, which involves selecting the best penalty parameter  $\lambda$  associated with the smallest 10-fold cross validation within the training set. The risk score for the gene signature was calculated using the following formula:

$$\text{Risk score} = \sum_{i=1}^n \text{Coef}_i * x_i$$

where *Coef* is the coefficient and  $x_i$  is the Z-score-transformed relative expression value of each selected gene.

Patients were divided into high-risk and low-risk groups using the median risk score as the cutoff value. Patients could also be compared between groups with lower-risk score (first quarter) and higher-risk score (fourth quarter).

**Bioinformatic analysis.** GO and KEGG pathway enrichment analyses with the Database for Annotation, Visualization, and Integrated Discovery (<http://david.abcc.ncifcrf.gov/home.jsp>) to functionally annotate genes with prognosis value in LGGs, significantly correlated to the risk score, and differentially spliced between patients with lower (first quarter) and higher (fourth quarter) risk score. GSEA was performed to investigate the functions of genes that significantly correlated to the risk score.

**Construction of regulation networks between RNA splicing genes and ASEs in LGGs.** The RNA splicing data were collected from the online database (<http://bioinformatics.mdanderson.org/TCGASpliceSeq>). The ASEs could be quantified by the PSI value, which represents the ratio of included transcript reads in total transcript reads; detailed information on the PSI calculation has been previously reported (32).

RNA splicing genes (GO:0008380) showing significant changes in expression levels were predicted to be correlated with the differential PSI level of ASEs between LGGs with lower- (first quarter) and higher-risk (fourth quarter) scores. We calculated the Pearson correlation for each RNA splicing gene–ASE pair and RNA splicing gene–ASE pair with correlation coefficients greater than 0.4 or less than –0.4; and the corresponding *P* value of less than 0.05 was considered significantly correlated (Bonferroni correction was performed to adjust the *P* value for multiple comparisons). The predicted regulatory network was visualized with Cytoscape (<http://www.cytoscape.org>).

**Cell lines.** The human glioma cell line LN229 was purchased from ATCC. We performed the reauthentication by short-tandem repeat analysis in July 2017 in our laboratory. Cells were routinely cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (HyClone) and 100 units/ml penicillin and 100 mg/ml streptomycin (Invitrogen) at 37°C in a humidified atmosphere of 5% CO<sub>2</sub> as in our previous study (7).

**siRNA and quantitative PCR.** The scrambled siRNA and KIF2C-specific siRNA were synthesized from the Genepharma Corporation. The sequences for TTF2 siRNA are as follows: TTF2 siRNA1, 5'-GGCCCC-CAGGUA AUGCUAATT-3'; TTF2 siRNA2, 5'-CCAAGAUCACGUUCAUGCATT-3'; TTF2 siRNA3, 5'-GGAAAGAGCUUCUACGUGUTT-3'.

Quantitative PCR was performed with SYBR Select Master Mix (Thermo Fisher) in an ABI 7500 (Thermo Fisher). The sequences for the TTF2 primers are as follows: 5'-ATTTACCGAGTAGGGCAGCA-3' (forward primer) and 5'-GGACTCTGAGGTCAGCCAAG-3' (reverse primer). The sequences for the GAPDH primers are as follows: 5'-GGTGGTCTCTCTGACTTCAACA-3' (forward primer) and 5'-GTTGCTGTAGCCAAATTCGTTGT-3' (reverse primer).

**Transcriptome sequencing.** The total RNA of LN229 cells was extracted by the Trizol reagent (Thermo Fisher). Then, the RNA-Seq library and sequencing were finished by Novogene using the Illumina Novaseq 6000 System (Illumina). After quality control and data mapping, quantitative analysis of gene expression was finished by the package of subread. Differential analysis of gene expression was performed by the DESeq2 (54). The quantitative RNA expression data for genes in LN229 with or without *TTF2* knockdown had been uploaded to figshare (doi: 10.6084/m9.figshare.9164399).

**Statistics.** Patients were divided into high-risk and low-risk groups using the median risk score as the cutoff value in both the training and validation data sets. A nonparametric test was used to compare the distribution of age between the 2 risk groups, and  $\chi^2$  tests were used to compare the distribution of other clinicopathological features. One-way ANOVA test was performed to compare the risk scores in patients grouped by WHO 2016 subgroup or combination of *TERT* promoter and *IDH* status. Two-tailed Student's *t* test was performed to compare the risk scores in patients grouped by other clinical or molecular pathological characteristics, and a *P* value of less than 0.05 was considered significant.

The prediction efficiency of the signature risk scores, age, WHO grade, and subgroups according to WHO 2016 classification for 3-year survival and 5-year survival were examined using ROC curves. Univariate and multivariate Cox regression analysis were performed to determine the prognostic value of the risk score and various clinical and molecular pathological characteristics.

The Kaplan-Meier method with a 2-sided log-rank test was used to compare the OS of patients in the different RNA processing clusters or in the high- and low-risk groups. All statistical analyses were conducted using R v3.4.1 (<https://www.r-project.org/>), SPSS 16.0 (SPSS Inc.), and Prism 7 (GraphPad Software Inc.).

**Study approval.** This retrospective study was approved and reviewed by the institutional review board of Beijing Tiantan Hospital, and informed consent was waived.

**Abbreviations.** The abbreviations used in this manuscript are listed in Supplemental Table 15.

## Author contributions

RCC and YML were responsible for study design, data analysis, and manuscript drafting. RCC and YZC performed the cell biological experiments and analyzed transcriptional data. KNZ, YZC, YQL, and ZZ illustrated data. ZLW, YHC, GZL, and KYW acquired data. FW and YZW designed and supervised the study.

## Acknowledgments

This study was supported by the National Key Research and Development Program of China (2018YFC0115604), the National Natural Science Foundation of China (81773208), the National Natural Science Foundation of China (NSFC)/Research Grants Council (RGC) Joint Research Scheme (81761168038), and Beijing Municipal Administration of Hospitals' Mission Plan (SML20180501).

Address correspondence to: Yong-Zhi Wang or Fan Wu, Beijing Neurosurgical Institute, Beijing Tiantan Hospital, Capital Medical University, Number 119 South Fourth Ring West Road, Beijing 100070, China. Phone: 861067096794; Email: yongzhiwang\_bni@163.com (YZW); wufan0510284@163.com (FW).

- Jiang T, et al. CGCG clinical practice guidelines for the management of adult diffuse gliomas. *Cancer Lett.* 2016;375(2):263–273.
- Ostrom QT, et al. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2009–2013. *Neuro-oncology.* 2016;18(suppl\_5):v1–v75.
- Louis DN, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol.* 2016;131(6):803–820.
- Chai RC, et al. A novel analytical model of MGMT methylation pyrosequencing offers improved predictive performance in patients with gliomas. *Mod Pathol.* 2019;32(1):4–15.
- Baysan M, et al. Detailed longitudinal sampling of glioma stem cells in situ reveals Chr7 gain and Chr10 loss as repeated events in primary tumor formation and recurrence. *Int J Cancer.* 2017;141(10):2002–2013.
- Marcelino Meliso F, Hubert CG, Favoretto Galante PA, Penalva LO. RNA processing as an alternative route to attack glioblastoma. *Hum Genet.* 2017;136(9):1129–1141.
- Chai RC, et al. Systematically characterize the clinical and biological significances of 1p19q genes in 1p/19q non-codeletion glioma. *Carcinogenesis.* 2019;bgz102.
- GLASS Consortium. Glioma through the looking GLASS: molecular evolution of diffuse gliomas and the Glioma Longitudinal Analysis Consortium. *Neuro-oncology.* 2018;20(7):873–884.
- Pekmezci M, et al. Adult infiltrating gliomas with WHO 2016 integrated diagnosis: additional prognostic roles of ATRX and TERT. *Acta Neuropathol.* 2017;133(6):1001–1016.
- Aibaidula A, et al. Adult IDH wild-type lower-grade gliomas should be further stratified. *Neuro-oncology.* 2017;19(10):1327–1337.
- Aoki K, et al. Prognostic relevance of genetic alterations in diffuse lower-grade gliomas. *Neuro-oncology.* 2018;20(1):66–77.
- Eckel-Passow JE, et al. Glioma groups based on 1p/19q, IDH, and TERT promoter mutations in tumors. *N Engl J Med.* 2015;372(26):2499–2508.
- Gleize V, et al. CIC inactivating mutations identify aggressive subset of 1p19q codeleted gliomas. *Ann Neurol.* 2015;78(3):355–374.
- Kamoun A, et al. Integrated multi-omics analysis of oligodendroglial tumours identifies three subgroups of 1p/19q co-deleted gliomas. *Nat Commun.* 2016;7:11263.
- Thon N, et al. Dynamic 18F-FET PET in suspected WHO grade II gliomas defines distinct biological subgroups with different clinical courses. *Int J Cancer.* 2015;136(9):2132–2145.
- Hu X, et al. Multigene signature for predicting prognosis of patients with 1p19q co-deletion diffuse glioma. *Neuro-oncology.* 2017;19(6):786–795.
- Ames HM, Yuan M, Vizcaino MA, Yu W, Rodriguez FJ. MicroRNA profiling of low-grade glial and glioneuronal tumors shows an independent role for cluster 14q32.31 member miR-487b. *Mod Pathol.* 2017;30(2):204–216.
- Rao SA, Santosh V, Somasundaram K. Genome-wide expression profiling identifies deregulated miRNAs in malignant astrocytoma. *Mod Pathol.* 2010;23(10):1404–1417.
- Jaffrey SR, Kharas MG. Emerging links between m<sup>6</sup>A and misregulated mRNA methylation in cancer. *Genome Med.* 2017;9(1):2.
- Coppin L, Leclerc J, Vincent A, Porchet N, Pigny P. Messenger RNA life-cycle in cancer cells: emerging role of conventional and non-conventional RNA-binding proteins? *Int J Mol Sci.* 2018;19(3):E650.
- Geles KG, et al. Upregulation of RNA processing factors in poorly differentiated lung cancer cells. *Transl Oncol.* 2016;9(2):89–98.
- Correa BR, et al. Functional genomics analyses of RNA-binding proteins reveal the splicing regulator SNRNP as an oncogenic candidate in glioblastoma. *Genome Biol.* 2016;17(1):125.
- Kechavarzi B, Janga SC. Dissecting the expression landscape of RNA-binding proteins in human cancers. *Genome Biol.* 2014;15(1):R14.
- Chai RC, et al. m<sup>6</sup>A RNA methylation regulators contribute to malignant progression and have clinical prognostic impact in gliomas. *Aging (Albany NY).* 2019;11(4):1204–1225.
- Yuan M, Eberhart CG, Kai M. RNA binding protein RBM14 promotes radio-resistance in glioblastoma by regulating DNA repair and cell differentiation. *Oncotarget.* 2014;5(9):2820–2826.
- Lefave CV, et al. Splicing factor hnRNPH drives an oncogenic splicing switch in gliomas. *EMBO J.* 2011;30(19):4084–4097.
- Fontana L, et al. Suggestive evidence on the involvement of polypyrimidine-tract binding protein in regulating alternative splicing of MAP/microtubule affinity-regulating kinase 4 in glioma. *Cancer Lett.* 2015;359(1):87–96.

28. Su R, et al. R-2HG exhibits anti-tumor activity by targeting FTO/m<sup>6</sup>A/MYC/CEBPA signaling. *Cell*. 2018;172(1-2):90–105.e23.
29. Cui Q, et al. m<sup>6</sup>A RNA methylation regulates the self-renewal and tumorigenesis of glioblastoma stem cells. *Cell Rep*. 2017;18(11):2622–2634.
30. Yu Y, Jiang X, Schoch BS, Carroll RS, Black PM, Johnson MD. Aberrant splicing of cyclin-dependent kinase-associated protein phosphatase KAP increases proliferation and migration in glioblastoma. *Cancer Res*. 2007;67(1):130–138.
31. Hu H, et al. Mutational landscape of secondary glioblastoma guides MET-targeted trial in brain tumor. *Cell*. 2018;175(6):1665–1678.e18.
32. Ryan M, et al. TCGASpliceSeq a compendium of alternative mRNA splicing in cancer. *Nucleic Acids Res*. 2016;44(D1):D1018–D1022.
33. Shirahata M, et al. Novel, improved grading system(s) for IDH-mutant astrocytic gliomas. *Acta Neuropathol*. 2018;136(1):153–166.
34. Brat DJ, et al. cImpact-NOW update 3: recommended diagnostic criteria for “Diffuse astrocytic glioma, IDH-wildtype, with molecular features of glioblastoma, WHO grade IV.” *Acta Neuropathol*. 2018;136(5):805–810.
35. Pomeroy SL, et al. Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature*. 2002;415(6870):436–442.
36. Wang ET, et al. Alternative isoform regulation in human tissue transcriptomes. *Nature*. 2008;456(7221):470–476.
37. Wan L, et al. SRSF6-regulated alternative splicing that promotes tumour progression offers a therapy target for colorectal cancer. *Gut*. 2019;68(1):118–129.
38. Li S, Hu Z, Zhao Y, Huang S, He X. Transcriptome-wide analysis reveals the landscape of aberrant alternative splicing events in liver cancer. *Hepatology*. 2019;69(1):359–375.
39. Babic I, et al. EGFR mutation-induced alternative splicing of Max contributes to growth of glycolytic tumors in brain cancer. *Cell Metab*. 2013;17(6):1000–1008.
40. Safaee M, et al. Proportional upregulation of CD97 isoforms in glioblastoma and glioblastoma-derived brain tumor initiating cells. *PLoS ONE*. 2015;10(2):e0111532.
41. Yu F, Fu WM. Identification of differential splicing genes in gliomas using exon expression profiling. *Mol Med Rep*. 2015;11(2):843–850.
42. Fu Y, Dominissini D, Rechavi G, He C. Gene expression regulation mediated through reversible m<sup>6</sup>A RNA methylation. *Nat Rev Genet*. 2014;15(5):293–306.
43. Dao Trong P, et al. Identification of a prognostic hypoxia-associated gene set in IDH-mutant glioma. *Int J Mol Sci*. 2018;19(10):E2903.
44. Zhang L, et al. IDH mutation status is associated with distinct vascular gene expression signatures in lower-grade gliomas. *Neuro-oncology*. 2018;20(11):1505–1516.
45. Kim SA, et al. Promoter methylation of WNT inhibitory factor-1 and expression pattern of WNT/ $\beta$ -catenin pathway in human astrocytoma: pathologic and prognostic correlations. *Mod Pathol*. 2013;26(5):626–639.
46. Ius T, et al. An NF- $\kappa$ B signature predicts low-grade glioma prognosis: a precision medicine approach based on patient-derived stem cells. *Neuro-oncology*. 2018;20(6):776–787.
47. Li J, Liang R, Song C, Xiang Y, Liu Y. Prognostic and clinicopathological significance of long non-coding RNA in glioma. *Neurosurg Re* [published online ahead of print March 21, 2018]. doi: 10.1007/s10143-018-0965-x.
48. Huang SP, et al. BICD1 expression, as a potential biomarker for prognosis and predicting response to therapy in patients with glioblastomas. *Oncotarget*. 2017;8(69):113766–113791.
49. Yi GZ, et al. Identification of key candidate proteins and pathways associated with temozolomide resistance in glioblastoma based on subcellular proteomics and bioinformatical analysis. *Biomed Res Int*. 2018;2018:5238760.
50. Visvanathan A, et al. Essential role of METTL3-mediated m<sup>6</sup>A modification in glioma stem-like cells maintenance and radioresistance. *Oncogene*. 2018;37(4):522–533.
51. Lei W, Wang ZL, Feng HJ, Lin XD, Li CZ, Fan D. Long non-coding RNA SNHG12 promotes the proliferation and migration of glioma cells by binding to HuR. *Int J Oncol*. 2018;53(3):1374–1384.
52. Ceccarelli M, et al. Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. *Cell*. 2016;164(3):550–563.
53. Bøvelstad HM, et al. Predicting survival from microarray data—a comparative study. *Bioinformatics*. 2007;23(16):2080–2087.
54. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.