Cell Communication and Signaling

Comprehensive profiling identifies a novel signature with robust predictive value and reveals the potential drug resistance mechanism in glioma

Fan Zeng[1†], Kuanyu Wang[2†], Xiu Liu[2] and Zheng Zhao[1*]

Abstract

Background: Gliomas are the most common and malignant brain tumors. The standard therapy is surgery combined with radiotherapy, chemotherapy, and/or other comprehensive methods. However, the emergence of chemoresistance is the main obstacle in treatment and its mechanism is still unclear.

Methods: We firstly developed a multi-gene signature by integrated analysis of cancer stem cell and drug resistance related genes. The Chinese Glioma Genome Atlas (CGGA, 325 samples) and The Cancer Genome Atlas (TCGA, 699 samples) datasets were then employed to verify the efficacy of the risk signature and investigate its significance in glioma prognosis. GraphPad Prism, SPSS and R language were used for statistical analysis and graphical work.

Results: This signature could distinguish the prognosis of patients, and patients with high risk score exhibited short survival time. The Cox regression and Nomogram model indicated the independent prognostic performance and high prognostic accuracy of the signature for survival. Combined with a well-known chemotherapy impact factor-MGMT promoter methylation status, this risk signature could further subdivide patients with distinct survival. Functional analysis of associated genes revealed signature-related biological process of cell proliferation, immune response and cell stemness. These mechanisms were confirmed in patient samples.

Conclusions: The signature was an independent and powerful prognostic biomarker in glioma, which would improve risk stratification and provide a more accurate assessment of personalized treatment.

Keywords: Glioma, Tumor stemness, Drug resistance, Gene signature, Prognosis

Background

Glioma is the most common primary malignant tumors of the central nervous system, accounting for 30% of all brain tumors and 80% of all malignant brain tumors [1]. It is one of the representative malignant tumors with the characteristics of strong genetic heterogeneity, high mortality and chemotherapy resistance. According to the 2016 WHO classification of central nervous system (CNS) tumors, the diffuse gliomas are mainly divided into five subtypes based on the mutation status of isocitrate dehydrogenase (IDH) and Chromosome 1p/19q status, namely low-grade gliomas (LGG) with IDH-mutant and 1p/19q-intact subtype, IDH-mutant and 1p/19q-codeleted subtype, LGG with IDH-wildtype subtype, glioblastoma (GBM) with IDH-mutant subtype, and GBM with IDH-wildtype subtype [2]. This classification breaks the principle of diagnosis based entirely on histological features and incorporates genetic parameters into the classification of CNS tumors. Therefore, the further exploration of novel and reliable biomarkers for the prediction of glioma may help to elucidate the molecular mechanism of glioma development and progression.

* Correspondence: zhaozheng0503@sina.com
†Fan Zeng and Kuanyu Wang contributed equally to this work.
[1]Department of Molecular Neuropathology, Beijing Neurosurgical Institute, Capital Medical University, No.119 South 4th Ring Road West, Fengtai District, Beijing 100070, China
Full list of author information is available at the end of the article

Temozolomide (TMZ), the most commonly used drug for standard clinical chemotherapy, has improved both the overall survival (OS) and the progression-free survival of patients in glioma. However, the patients usually suffer drug resistance and the underlying mechanism is still elusive. Recently, researchers have found the relationship between cancer stem cells (CSCs) and drug resistance. The CSCs represent a rare population of cancer cells with the capacity to self-renew and initiate tumors. Several biomarkers or relevant molecular markers have been reported differentially expressed in CSCs, suggesting the stemness and unfavorable prognosis [3, 4]. Even therapies that cause complete regression of tumors might spare enough CSCs to allow regrowth of the tumors [5, 6]. Currently, the research on CSC-associated drug resistance is limited, and no method has been established to predict drug resistance through the expression of CSC-related genes, and ultimately predict the prognosis of patients. Therefore, we focused on TMZ therapy, CSCs and diverse related oncogenic drivers within patients by high throughput sequencing method and established gene signature, a statistical model including several biomarkers, to predict the prognosis and chemoresistance of glioma patients [7–9].

In this study, we developed and validated a multigene signature including CSCs and TMZ drug resistant (TMZ-DR)-associated genes to predict patients' prognosis in glioma. Multivariate analysis revealed the correlation between gene signature and malignant progression of cancer and further clarified the potential biological mechanisms. Combined with the methylation state of O^6-methylguanine-DNA methyltransferase (MGMT) promoter, this gene signature could more comprehensively predict the prognosis of patients, exhibiting impressive clinical application value. In addition, we validated the signature in GBM patients that only treated with TMZ in the Chinese Glioma Genome Atlas Network (CGGA) dataset and elucidated the mechanism of TMZ drug resistance to some extent. Our result suggested that the signature integrating comprehensive transcript information would improve risk stratification and provide a more accurate assessment of personalized clinical management in patients with glioma. These results might provide new view for glioma malignancy and individual treatment.

Materials and methods
Patients and databases
A total of 325 glioma samples from the Chinese Glioma Genome Atlas Network (CGGA, http://www.cgga.org.cn) were included in this study [10]. All these samples and clinicopathological information were collected with informed consent. The study was approved by the Tiantan Hospital Institutional Review Board and kept consistent with the principles of the Helsinki Declaration. An independent cohort of 699 patients with clinical and molecular profiling was obtained from TCGA (http://cancergenome.nih.gov) and used as external validation. The GSE23806 dataset from GEO website (https://www.ncbi.nlm.nih.gov/geo/) was downloaded, including 27 glioma stem-like cell lines and 36 conventional glioma cell lines, and used to discover the CSC-associated genes. Besides, a drug screening profile (COSMIC dataset) was obtained from the public website (https://cancer.sanger.ac.uk/cell_lines/) to do the drug sensibility analysis.

Development of signature and analytical approach
The student's t-test was first performed to identify differentially expressed genes in cancer stem cells (GSE23806 dataset) and TMZ-resistant cells (COSMIC dataset) respectively. The common differential genes were then dimensionally reduced by the least absolute shrinkage and selection operator (LASSO) method. These obtained genes finally formed a risk signature that was determined by a linear combination of their expression levels weighted with regression coefficients from univariate Cox regression analyses. The hazard ratio (HR) of each gene figured out in CGGA dataset was used to develop the signature:

$$Signature\ risk\ score = \sum\nolimits_{i=1}^{n} \beta_i x_i$$

where βi indicates the HR for each gene, and x_i indicates the z score transformed relative expression value of each gene.

The Kaplan-Meier survival curves were used to estimate survival distributions. Cox regression was performed to assess the prognostic value of the risk score. The DAVID software (http://david.ncifcrf.gov/) was applied to elucidate the Gene Ontology (GO) biological functions and KEGG pathway. The Gene Set Enrichment Analysis (GSEA, http://www.broadinstitute.org/gsea/index.jsp) was performed to identify gene sets of statistical difference between two groups (high risk score vs. low risk score). Figures were generated by several packages of R software (version 3.2.5), such as 'pheatmap', 'pROC', and 'circlize' [11, 12].

Immunohistochemistry
To verify the significance and potential mechanism of the risk signature, we analyzed immunohistochemical (IHC) protein staining data of CD133, P4HB, IBA1 and CD163 in the glioma samples from CGGA dataset. The IHC expression levels were compared in the low-, medium- and high-risk score groups with a nonparametric test. Briefly, five-micrometer-thick sections were deparaffinized, boiled with EDTA antigen retrieval buffer, and then incubated with the primary antibodies overnight at 4 °C (anti-CD133

antibody, 1:1000 dilution, Proteintech Group; anti-P4HB, 1:1000, Abcam; anti-IBA1, 1:2000, Abcam; anti-CD163, 1: 200, Abcam). Then, the sections were incubated with appropriate secondary antibodies (1:100, ZSGB-Bio, Beijing, China) at room temperature for 1 h. Finally, the stained slides were individually reviewed and evaluated by two investigators. The expression levels of each protein in tumor tissues were defined as the portion of positively stained cells against total counted cells. The difference was assessed by Student-t test.

Construction of an individualized prediction model
A nomogram was established as an individualized prediction model to predict patient's overall survival using the 'rms' package in R language software [13]. The final model was constructed using a backward step-down selection process based on the Akaike information criterion. Concordance index (C-index) and calibration curves were performed to assess predictive accuracy and discriminative ability of the nomogram. The prognostic nomogram model was estimated in the training cohort and then tested in the validation cohort.

Statistical analysis
All statistical analyses were conducted using R language (version 3.4.1), GraphPad Prism (version 7.0) and SPSS (version 16.0). The Student's t-test was used to compare the differences in variables between groups. $P < 0.05$ was regarded as statistically significant.

Results
Gene selection and signature building
CSCs are subpopulations of cancer cells with unlimited self-renewal potential that drive tumorigenesis and exhibit resistance to chemotherapeutics [14, 15]. Elimination of CSCs is the key step to overcome drug resistance. Therefore, our study starts with the analysis of CSCs. The student's t-test was carried out in GSE23806 dataset to identify the different expression genes between glioma stem-like cell lines ($n = 27$) and conventional glioma cell lines ($n = 36$) ($P < 0.05$). A total of 6752 genes were highly expressed in glioma stem-like cell lines and were selected for further analysis (Fig. 1a). The same analysis was applied in COSMIC dataset to reveal the TMZ-DR genes according to the half maximal inhibitory concentration (IC$_{50}$). A total of 447 genes were identified highly expressed in the high IC$_{50}$ group (Fig. 1b). To select bona fide drug-resistant associated genes, 138 genes were identified using intersections of above two gene sets. Lastly, the LASSO Cox regression model was conducted to select the most useful predictive features and identified only seven genes (*ATL1, GRIA3, HPX, IL17D, KLHDC1, NCAM2, TRIM67*) with non-zero regression coefficients (Fig. 1c and d) in CGGA datasets.
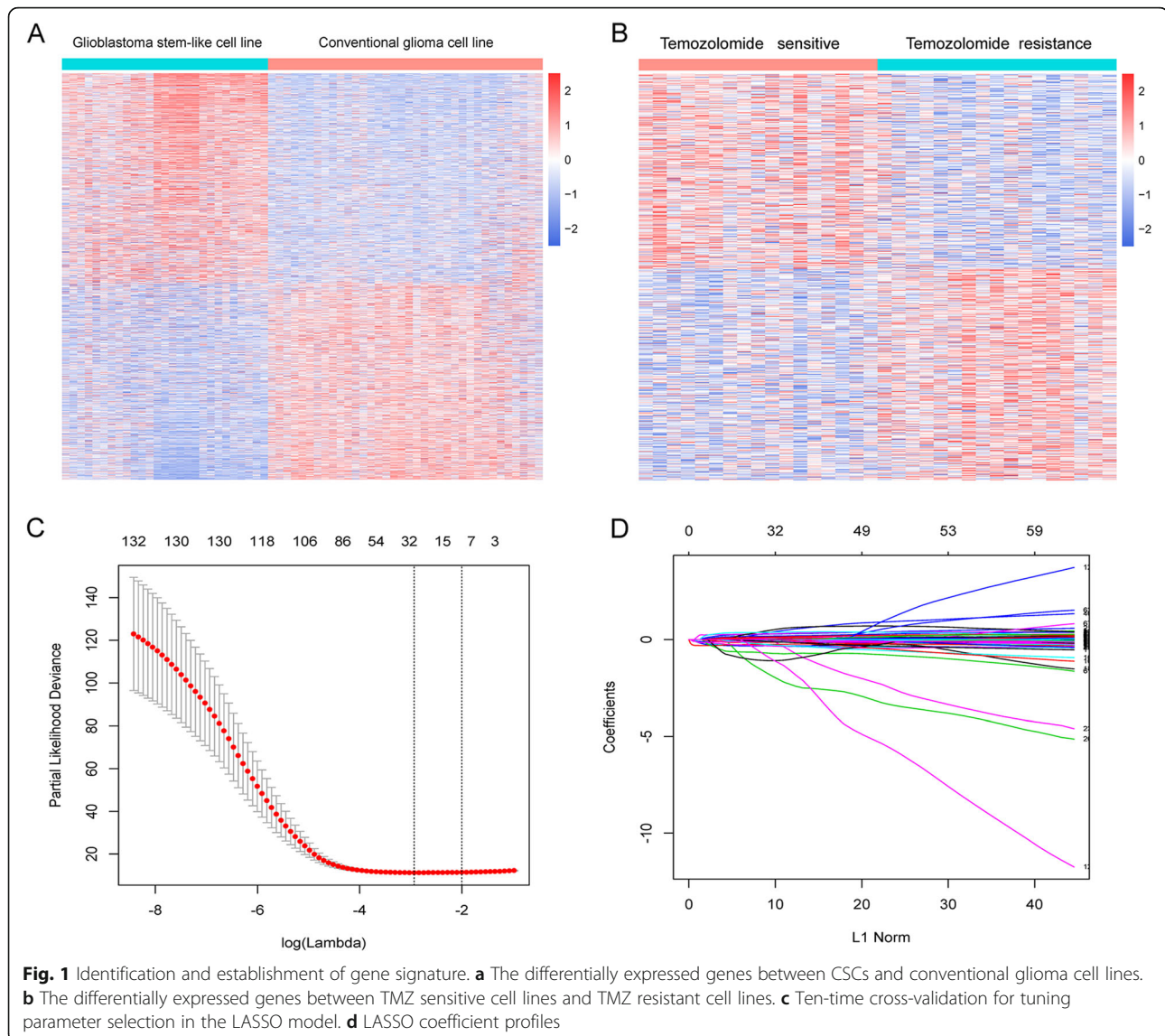
We then established a signature for glioma patients based on the gene expression levels as follows: signature risk score = (− 0.143 × ATL1 expression) + (− 0.094 × GRIA3 expression) + (0.171 × HPX expression) + (− 0.305 × IL17D expression) + (− 0.132 × KLHDC1 expression) + (− 0.096 × NCAM2 expression) + (− 0.124 × TRIM67 expression). The coefficients of each gene indicate the HR. The coefficient of HPX is greater than 1, implying its tumor driving characteristic, while the coefficients of the other four genes are less than 1, revealing their protecting effect.

Molecular characteristics of the risk signature
The novel risk signature were further evaluated and validated in CGGA and TCGA datasets and the clinical information of patients were summarized in Table S1 (Additional file 1). We calculated the risk score of each patient by the formula and further detected its value in patients stratified by grade, subtype, MGMT promoter, IDH and 1p/19q status. As shown in Fig. 2, the risk score was up-regulated along with histological grades, and the increased expression was also observed in Mesenchymal, MGMT promoter unmethylated, LGG IDH-wildtype or GBM IDH-wildtype stratified patients. Similar observations were obtained in TCGA database (Additional file 2: Figure S1). Additionally, the receiving operator characteristic (ROC) curves for risk score and Mesenchymal subtype were performed and the area under the curve (AUC) were up to 89.6% (CGGA, Fig. 2) and 87.7% (TCGA, Additional file 2: Figure S1), respectively. This result suggested that the risk score may serve as a predictor for Mesenchymal subtype in glioma.

Survival analysis and prognostic validity of the risk signature
To comprehensively investigate the risk signature, we evaluated the association of risk score and patient survival. Figure 3a showed the survival overview of each glioma patient in CGGA dataset and patients with high-risk score had poor prognosis. We further analyzed the high and low risk groups separately to clarify the relationship between risk score and molecular characteristics. Results revealed the enrichment of higher grades, wild type-IDH and Mesenchymal subtype in high risk groups, which indicated the consistency between high risk score and malignant molecular characteristics in glioma (Additional file 3: Table S2). Applying to the Kaplan-Meier survival analysis, we found that high risk score conferred reduced overall survival among diffuse glioma patients ($P < 0.001$, Fig. 3b-d). In particular, this risk signature could also be a good predictor of patient survival in GBM, an aggressive glioma subtype. Subsequently, we investigated the predictive accuracy of risk score in predication of 3- and 5-years OS by analyzing the ROC curves. Compared with traditional "age" and

Fig. 1 Identification and establishment of gene signature. **a** The differentially expressed genes between CSCs and conventional glioma cell lines. **b** The differentially expressed genes between TMZ sensitive cell lines and TMZ resistant cell lines. **c** Ten-time cross-validation for tuning parameter selection in the LASSO model. **d** LASSO coefficient profiles

"grade" predictors, the risk signature showed favorable prognostic validity, with higher AUC up to 81.6 and 81.3% for 3- and 5-year survival, respectively, indicating its superior predictive value (Fig. 3e and f). The same results were validated in TCGA dataset (Additional file 3: Table S2 and Additional file 4: Figure S2).
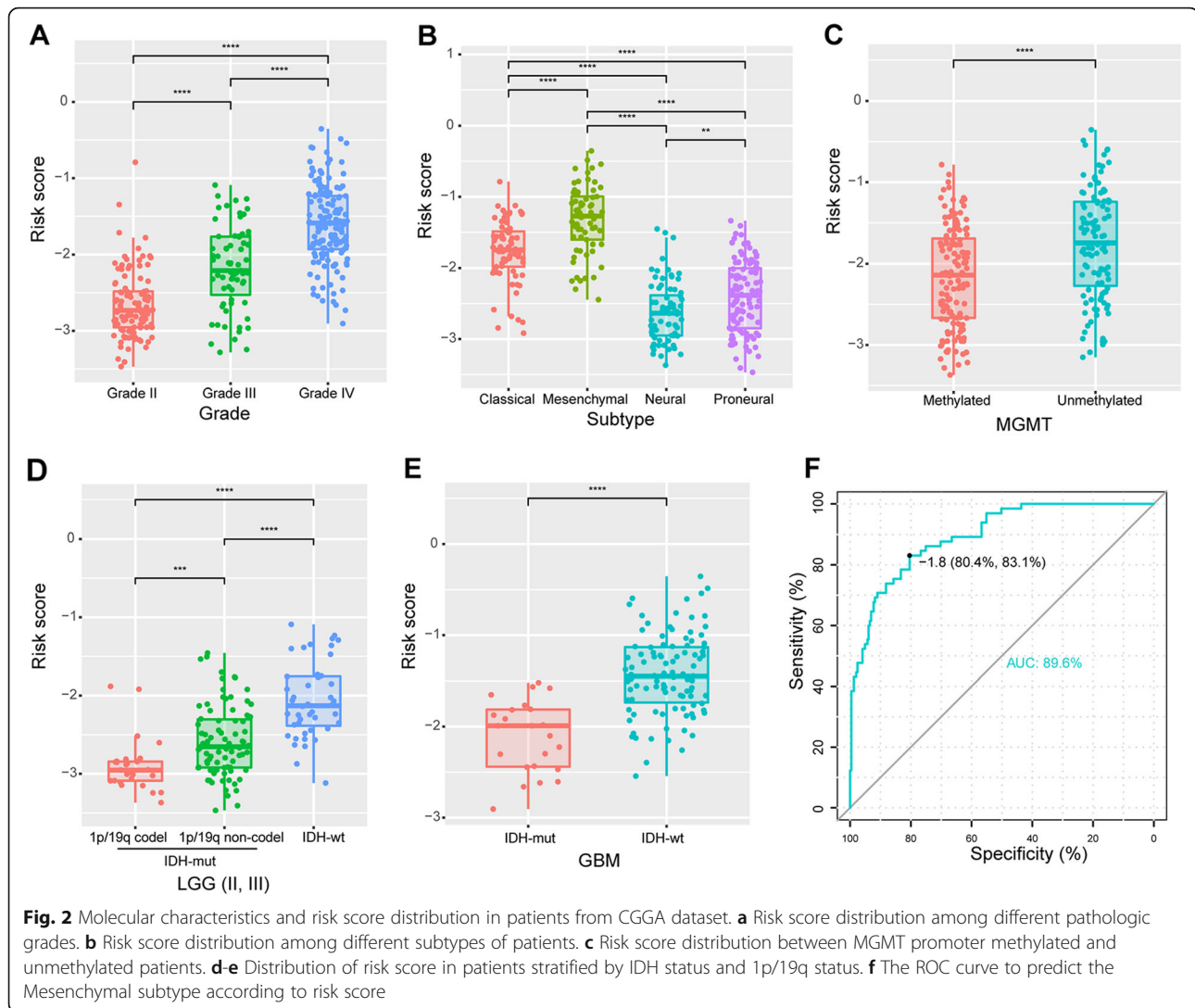
To evaluate the prognostic value of risk signature across diffuse gliomas, we further performed univariate and multivariate Cox regression analysis. As shown in the results, risk score, age, WHO grade, IDH mutation status, MGMT promoter methylation status, 1p/19q deletion status and radiotherapy were significantly associated with patients' survival in CGGA dataset (Fig. 3g). After further multivariable adjustment, the risk score remained an independent factor ($P < 0.001$), implying its powerful ability for predicting OS of glioma patients (Fig. 3h). Similar results were found in TCGA dataset (Additional file 4: Figure S2).

Construction of an individualized prediction model

Using a backward stepwise method based on the smallest Akaike information criterion, the prognostic nomograms that integrated independent prognostic parameters were constructed. The results showed that signature risk score contributed the most risk points (range, 0–100), whereas the other clinical variables including grade and radiotherapy exhibited smaller contributions. The C-indices were 0.83 in the CGGA dataset and 0.85 in the TCGA cohort, indicating satisfactory concordance. The calibration curve also demonstrated excellent agreement between prediction and observation in the probabilities of 1-, 3-, and 5-year overall survival in both datasets (Fig. 4 and Additional file 5: Figure S3).

Significant functions and pathway enrichment analysis

To investigate the potential functions related to the gene signature, Pearson correlation analysis was performed to

Fig. 2 Molecular characteristics and risk score distribution in patients from CGGA dataset. **a** Risk score distribution among different pathologic grades. **b** Risk score distribution among different subtypes of patients. **c** Risk score distribution between MGMT promoter methylated and unmethylated patients. **d-e** Distribution of risk score in patients stratified by IDH status and 1p/19q status. **f** The ROC curve to predict the Mesenchymal subtype according to risk score

find out the genes that strongly correlated with risk score (Pearson |R| > 0.5) in CGGA and TCGA datasets. Afterwards the significant related genes were chosen for Gene Ontology analysis with online methods (DAVID, https://david.ncifcrf.gov/). As shown in the figures, the positive related genes were enriched in cell adhesion, proliferation, immune response, *etc* (Fig. 5a, Additional file 6: Figure S4). Both cell adhesion and proliferation can affect the drug resistance and stemness of cancer cells [16, 17]. The negatively related genes were more involved in normal neural development and physiological activity, such as neuron development, neuron differentiation, synaptic transmission, *etc* (Fig. 5a, Additional file 6: Figure S4). Additionally, we found that risk score positive related genes were mainly involved in the biological pathways, such as focal adhesion, regulation of actin cytoskeleton, cell cycle and Jak-STAT signaling pathway (Fig. 5b, Additional file 6: Figure S4).

To further validate the reliability, we investigate the relationships between risk signature and CSCs-related genes [18]. As shown in Fig. 5c, the significant association between risk score and CSCs-related genes were constructed using 'ggcorrplot' in R language. Results showed that the risk score was positively correlated with CSCs-related genes, suggesting that the higher the risk score, the higher the cancer cell stemness or/and drug resistance, which may indicate poor prognosis of glioma patients. The same results were obtained in the TCGA database (Additional file 6: Figure S4). To validate the results of bioinformatics analyses, we conducted IHC protein staining experiments to investigate the association of risk score and cancer cell stemness (*CD133*) and drug resistance (*P4HB*) in CGGA cohort. Representative IHC images were displayed in Fig. 5d. As expected, the expression levels of CD133 and P4HB increased in parallel with the risk score (Fig. 5e). Those results showed
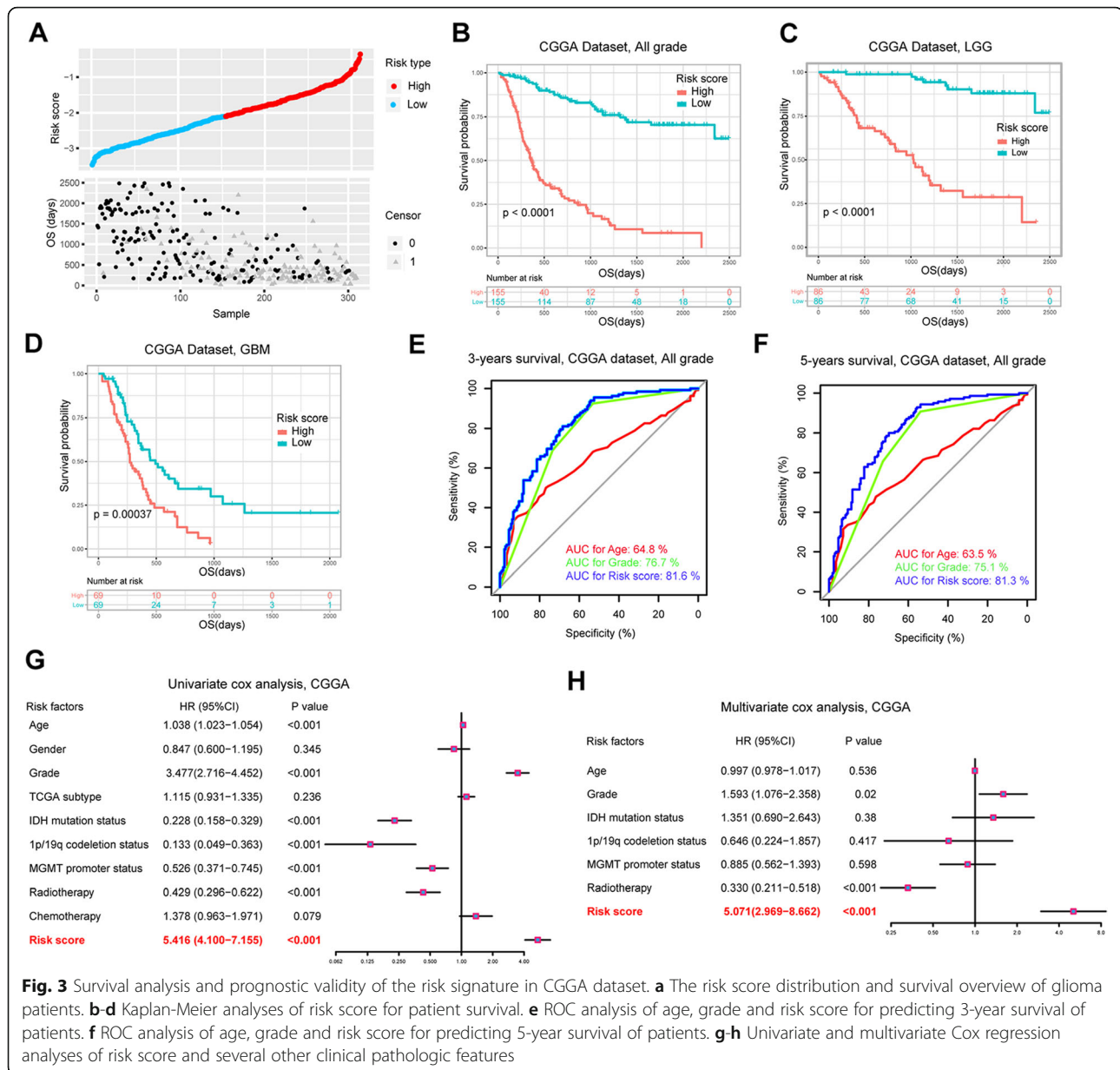
Fig. 3 Survival analysis and prognostic validity of the risk signature in CGGA dataset. a The risk score distribution and survival overview of glioma patients. b-d Kaplan-Meier analyses of risk score for patient survival. e ROC analysis of age, grade and risk score for predicting 3-year survival of patients. f ROC analysis of age, grade and risk score for predicting 5-year survival of patients. g-h Univariate and multivariate Cox regression analyses of risk score and several other clinical pathologic features
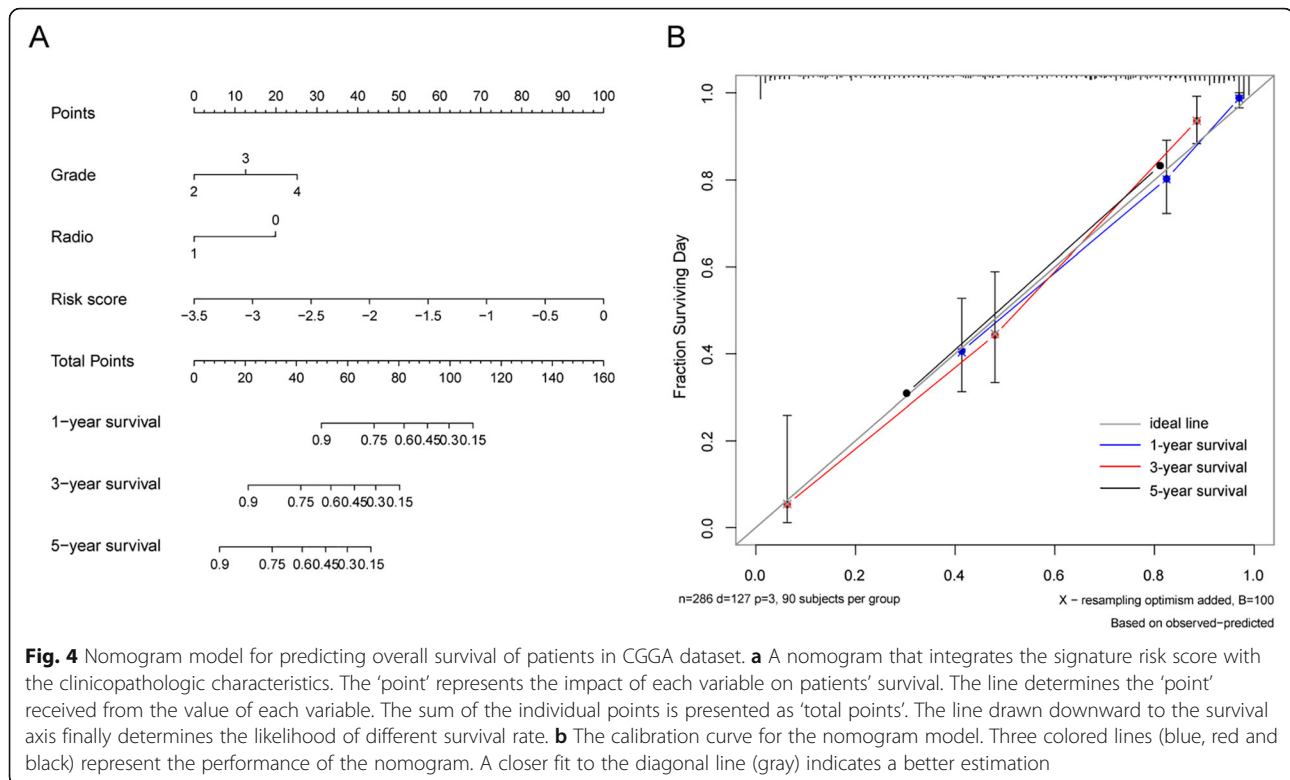
that high risk scores suggested high cell stemness and drug resistance of cancer cells.

Application of the signature across diffuse gliomas
In the clinical treatment of glioma, the MGMT promoter methylation status is an important prognostic indicator, because patients with MGMT promoter methylation are prone to benefit from TMZ treatment. Therefore, we combined the methylation status of MGMT promoter to analyze the survival status of patients. As shown in Fig. 6a-c and Additional file 7: Figure S5, the survival analysis based on risk signature and MGMT promoter methylation status demonstrated remarkable stratification of the clinical courses into four

subgroups. Patients with MGMT promoter unmethylation and high risk score had the worst prognosis, while patients with MGMT promoter methylation and low risk score had the best prognosis.

Especially, we subsequently focused on the GBM patients that treated with TMZ from CGGA dataset. According to the risk score and MGMT promoter methylation status, patients were divided into four subgroups. The risk score and overall survival distribution are shown in Fig. 6d. As shown, the risk score could well distinguish the prognosis of patients and a significant reduction in survival was observed in the high-risk group (Fig. 6e). We next conducted univariate and multivariate Cox analysis and revealed the independent correlation

Zeng *et al. Cell Communication and Signaling*   (2020) 18:2

Page 7 of 13

Fig. 4 Nomogram model for predicting overall survival of patients in CGGA dataset. **a** A nomogram that integrates the signature risk score with the clinicopathologic characteristics. The 'point' represents the impact of each variable on patients' survival. The line determines the 'point' received from the value of each variable. The sum of the individual points is presented as 'total points'. The line drawn downward to the survival axis finally determines the likelihood of different survival rate. **b** The calibration curve for the nomogram model. Three colored lines (blue, red and black) represent the performance of the nomogram. A closer fit to the diagonal line (gray) indicates a better estimation

with OS and risk score, confirming its power for independently predicting prognosis (Table 1). The GSEA result showed the enrichment of epithelial mesenchymal transition, IL6-JAK-STAT3 pathway and TNFA signaling via NF-kB in the high-risk group (Fig. 6f), suggesting a potential mechanism of TMZ resistance. These findings were in accordance with the previous studies that the inhibition of NF-kB pathway or activated Stat3 could re-sensitize the resistant cells to therapeutic drugs [19, 20]. Collectively, we proved that this risk signature had a strong predictive effect on the prognosis of patients and could guide the individual treatment.
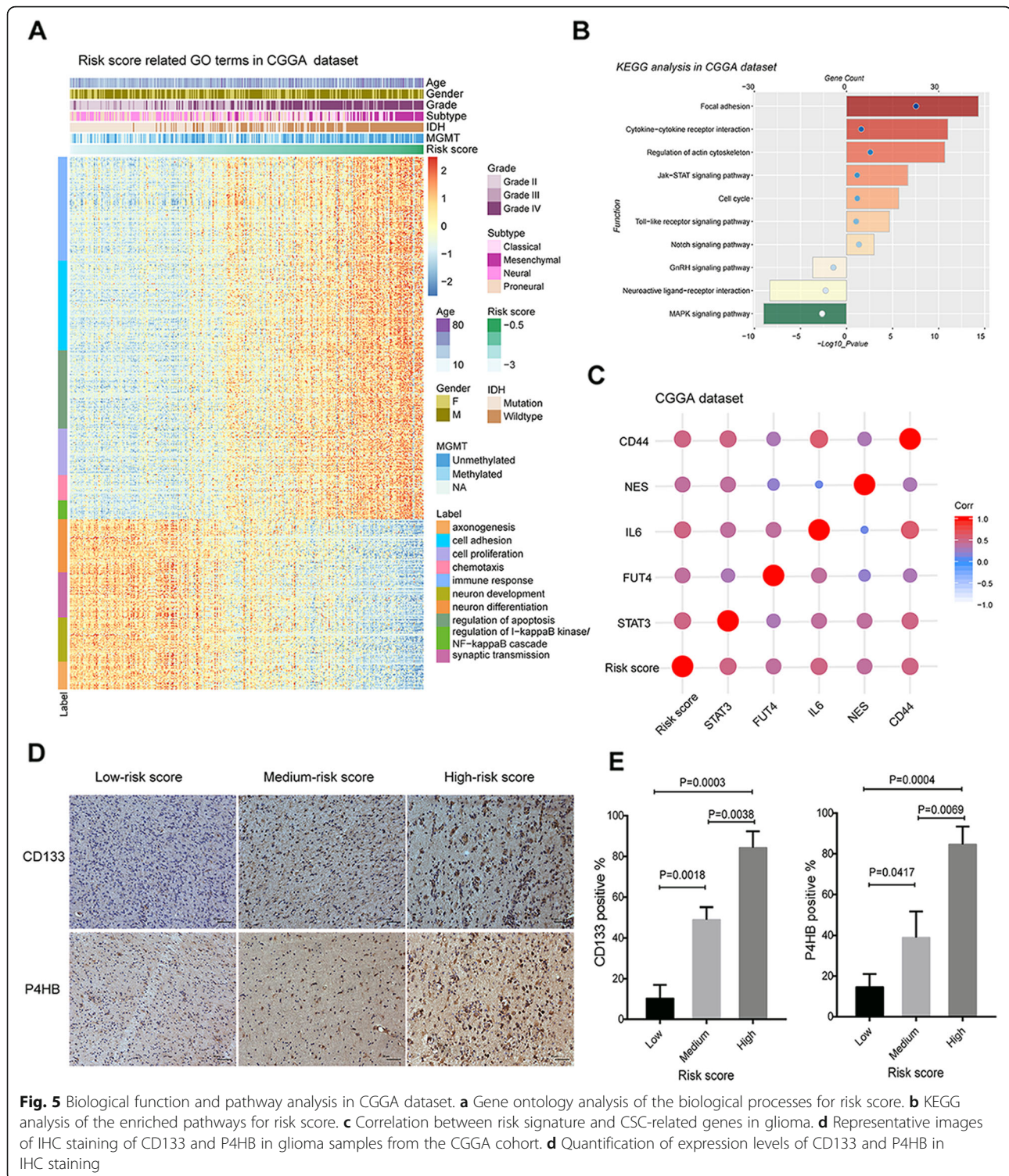
Potential mechanisms of drug resistance in glioma

In the previous studies on the biological mechanism of this risk signature (Fig. 5a, Additional file 6: Figure S4) and patients after temozolomide treatment (Fig. 6f), it was suggested that drug resistance was related to immune response in glioma patients. It has also been reported that immune response and immune genes like B7H3 and Tim3 can contribute to the stemness and drug resistance [21, 22]. Then, the relationship between risk signature and immune checkpoints, including *B7H3, B7H4, Tim3, PD1*, and *PDL1* were analyzed [23, 24]. Our results demonstrated a tight connection between the risk score and *Tim3/B7H3* (Fig. 7a-b), which may partly explain the poor prognosis of patients and provide some clues for potential immunotherapy. We further estimated the abundance of

various types of immune cell with CIBERSORT for CGGA and TCGA database (Fig. 7c-d). Samples with higher risk score exhibited apparent concordance with encirclement of macrophages in M0 phase, which have been reported to secrete immunosuppressive factors like tumor-supportive M2 macrophages in glioblastoma [25]. The results are consistent with the previously mentioned analysis that the enriched pathway containing immune factors after treatment with temozolomide (Fig. 6f). To further confirm these initial findings, we next studied the intratumoral immune cell infiltrates from glioma patients with IHC analysis and compared their immune cell profiles relative to the risk signature. As shown in the results, the expression levels of CD163 and IBA1 increased in parallel with the risk score (Fig. 7e-f). These results may explain the mechanism of drug resistance to some extent and provide new ideas for clinical treatment.

Discussion

The origination and development of glioma is a complex biological process of multiple factors and steps. Due to the heterogeneity of glioma cells, the therapeutic efficacy of conventional treatment including surgery, chemotherapy and radiotherapy is limited [26]. After chemotherapy with temozolomide, the residual cancer stem cells and changes in gene expression of cancer cells may lead to drug resistance, thus leading to the failure of chemotherapy. In the present study, we developed a seven-gene-based signature

Fig. 5 Biological function and pathway analysis in CGGA dataset. **a** Gene ontology analysis of the biological processes for risk score. **b** KEGG analysis of the enriched pathways for risk score. **c** Correlation between risk signature and CSC-related genes in glioma. **d** Representative images of IHC staining of CD133 and P4HB in glioma samples from the CGGA cohort. **d** Quantification of expression levels of CD133 and P4HB in IHC staining

to predict patients' prognosis by integration analysis of TMZ-DR genes and CSC-related genes, and further reveal the potential mechanism of drug resistance. The risk signature is proved to be associated with poor prognosis of patients and its role as an independent prognosis predictor and individual survival estimation have been revealed.

The Gene Ontology analysis in our study have revealed that the positive genes associated with risk score were involved in cell proliferation, adhesion, chemotaxis and immune response, which have been demonstrated to induce the relapse of cancers [27, 28]. Moreover, the KEGG analysis has showed that Jak-stat signaling
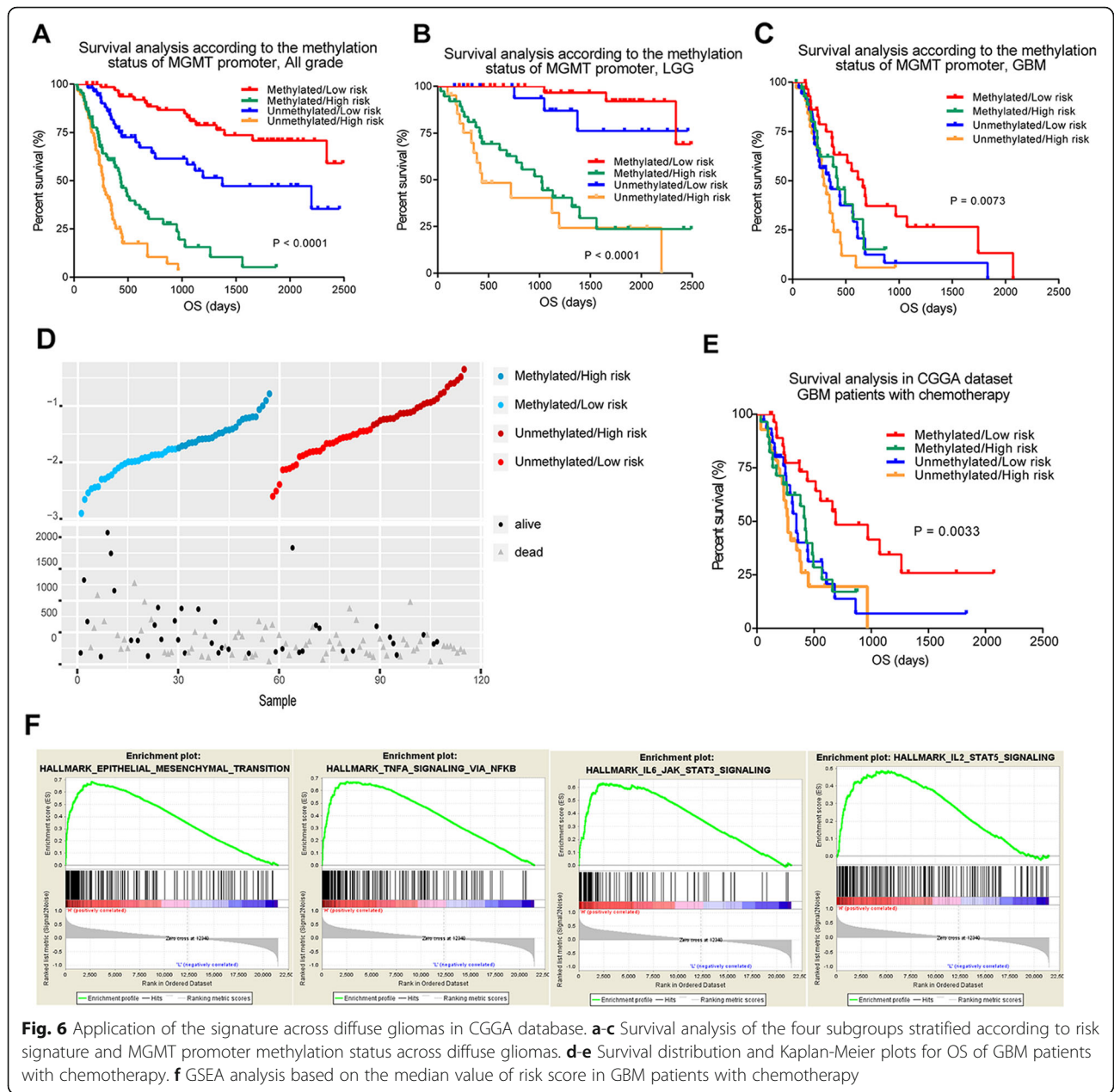
Fig. 6 Application of the signature across diffuse gliomas in CGGA database. **a-c** Survival analysis of the four subgroups stratified according to risk signature and MGMT promoter methylation status across diffuse gliomas. **d-e** Survival distribution and Kaplan-Meier plots for OS of GBM patients with chemotherapy. **f** GSEA analysis based on the median value of risk score in GBM patients with chemotherapy

Table 1 Univariate and multivariate analysis of OS in CGGA RNA sequencing dataset, GBM patients wtith chemotherapy

Variables	Univariate analysis		Multivariate analysis	
	HR (95% CI)	*p* value	HR (95% CI)	p value
Risk score	2.533 (1.348–4.760)	0.004	2.578 (1.217–5.462)	0.013
Age at diagnosis	1.001 (0.974–1.028)	0.954	–	–
Gender	0.675 (0.353–1.293)	0.236	–	–
TCGA subtype	1.014 (0.789–1.304)	0.912	–	–
MGMT promoter status	0.539 (0.295–0.985)	0.045	0.759 (0.393–1.463)	0.410
IDH mutation status	0.657 (0.325–1.328)	0.242	–	–
Radiotherapy	0.303 (0.157–0.584)	< 0.001	0.759 (0.393–1.463)	< 0.001

Fig. 7 Relationship between drug resistance and immunity in patients with glioma. **a-b** Correlation between risk score and immune checkpoints in glioma from CGGA and TCGA datasets. **c-d** Component types of the immune cells infiltrated into glioma are analyzed with CIBERSORT. **e** Representative images of IHC staining of CD163 and IBA1 in glioma samples from the CGGA cohort. **f** Quantification of expression levels of CD163 and IBA1 in IHC staining

pathway and high risk score are closely related. The high level of stat activation is associated with Pan-cancers, especially for Stat3 and Stat5, which is mostly related to more dangerous tumors [29, 30]. The interaction network of risk signature and CSC-related genes indicated

the relationship between cancer cell stemness and drug resistance. The hyperactivation of IL6/STAT3 pathway was reported to be essential for CSC self-renewal and tumorigenic [31, 32]. While other biomarkers, such as CD44 and Nestin, were correlated with tumor staging,

tumor size and lymph node metastasis [33, 34]. The correlations between risk signature and cancer cell stemness and drug resistance have been confirmed by IHC assay. These findings may explain the poor prognosis of patients with high risk score to some extent. In addition, the further GSEA analysis in GBM patients with chemotherapy have found the enrichment of epithelial mesenchymal transition, IL6-JAK-STAT3 pathway and IL2-STAT5 pathway in high-risk group, indicating the important role of immune factors in tumor drug resistance. Further studies have shown the relationship between drug resistance and immune response, including immune checkpoints and immune cell infiltration in glioma [35–37]. Our findings have been validated in patient tissues and provide alternative therapeutic options for drug resistant patients in clinical application. These results are consistent with previous studies that drug resistance in glioma is associated with mesenchymal characteristics and immune responses [38–41].

The seven selected genes in signature have been studied and reported previously [42–44]. For instance, *TRIM67*, the most evolutionarily conserved member of tripartite motif (TRIM) family, was reported to function as a Ras inhibitor. The down regulation of *TRIM67* may contribute to the over-activation of Ras signal pathway and overgrowth of non-small cell lung cancer cells, leading to the malignant progression of tumor [45]. Moreover, *TRIM67* has been proved to play a necessary role in appropriate brain development and behavior [46]. In addition, Heyde et al. have investigated the resistance mechanism of monoclonal antibody drug trastuzumab in HER2-positive breast cancer. *NCAM2* and five other genes were significantly lower expressed in HCC1954 (drug resistance) than in BT474 (drug sensitivity) cell lines [47]. Other risk factors, such as *HPX* and *GRIA3*, have been reported in many cancers, including glioma, breast cancer and pancreatic cancer [48, 49]. These findings reported previously are consistent with our results in biological function analysis, implying the potential role of gene signature in brain functions and tumor progression.

The methylation status of MGMT promoter is an important factor in the clinical treatment of glioma patients. Combined with MGMT promoter methylation status, the signature can further divide patients into four subtypes. This classification is more conducive to clinical treatment. Additionally, this risk signature exhibited a good predictive value for the prognosis of GBM patients with chemotherapy. These results may provide some ideas for the follow-up treatment of patients. The current classification of gliomas is mainly based on grade, subtype, IDH mutation status and 1p/19q status. In this study, the signature could provide a perpendicular look on molecular characteristics to some extent, which means that high risk score implies malignant phenotypes, including higher grade,

wild type-IDH and Mesenchymal subtype. Thus it might be reasonable to assume that we could reclassify gliomas based on the signature solely, where previous molecular characteristics have been taken into account in fact. However, our study is limited because it is retrospective, and should be further validated by prospective studies.

Conclusions

In conclusion, we successfully built a novel gene signature using seven genes relevant to CSCs and TMZ resistance and reliably validated its molecular characteristics and clinical significance in two large sample-sized glioma datasets. The signature was proved to be an independent predictor in prognosis as the high risk scores always accompanied with shorter survival and more malignant phenotypes. Further functional analysis provided a clue for the underlying mechanisms in tumor progression and drug resistance. Our study revealed that the signature might be a potential prognostic biomarker and/or be used as therapeutic target in personalized clinical treatment in glioma patients.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12964-019-0492-6.

Additional file 1 : Table S1. Clinical and molecular characteristics of patients in CGGA and TCGA datasets

Additional file 2 : Figure S1. Molecular characteristics and risk score distribution in patients from TCGA dataset. (A) Risk score distribution among different pathologic grades. (B) Risk score distribution among different subtypes of patients. (C) Risk score distribution between MGMT promoter methylated and unmethylated patients. (D-E) Distribution of risk score in patients stratified by IDH status and 1p/19q status. (F) The ROC curve to predict the Mesenchymal subtype according to risk score.

Additional file 3 : Table S2. Molecular characteristics of patients stratified by risk score in CGGA and TCGA datasets.

Additional file 4 : Figure S2. Survival analysis and prognostic validity of the risk signature in TCGA dataset. (A) The risk score distribution and survival overview of glioma patients. (B-D) Kaplan-Meier analyses of risk score for patient survival. (E) ROC analysis of age, grade and risk score for predicting 3-year survival of patients. (F) ROC analysis of age, grade and risk score for predicting 5-year survival of patients. (G-H) Univariate and multivariate Cox regression analyses of risk score and several other clinical pathologic features.

Additional file 5 : Figure S3. Nomogram model for predicting overall survival of patients in TCGA dataset. (A) A nomogram that integrates the signature risk score with the clinicopathologic characteristics. The 'point' represents the impact of each variable on patients' survival. The line determines the 'point' received from the value of each variable. The sum of the individual points is presented as 'total points'. The line drawn downward to the survival axis finally determines the likelihood of different survival rate. (B) The calibration curve for the nomogram model. Three colored lines (blue, red and black) represent the performance of the nomogram. A closer fit to the diagonal line (gray) indicates a better estimation.

Additional file 6 : Figure S4. Biological function and pathway analysis in TCGA dataset. (A) Gene ontology analysis of the biological processes for risk score. (B) KEGG analysis of the enriched pathways for risk score. (C) Correlation between risk signature and CSC-related genes in glioma.

Zeng *et al. Cell Communication and Signaling*   (2020) 18:2

Page 12 of 13

Additional file 7 : Figure S5. Survival analysis of the four subgroups stratified according to risk signature and MGMT promoter methylation status in TCGA database.

Abbreviations

AUC: The area under the curve; CGGA: Chinese Glioma Genome Atlas; CNS: Central nervous system; CSCs: Cancer stem cells; DAVID: Database for annotation,visualization and integrated discovery; GBM: Glioblastoma; GO: Gene ontology; GSEA: Gene set enrichment analysis; IC_{50}: Half maximal inhibitory concentration; IDH: Isocitrate dehydrogenase; LASSO: Least absolute shrinkage and selection operator; LGG: Low-grade gliomas; OS: Overall survival; ROC: Receiver operating characteristic; TCGA: The Cancer Genome Atlas; TMZ: Temozolomide; TRIM: Tripartite motif

Author details

[1]Department of Molecular Neuropathology, Beijing Neurosurgical Institute, Capital Medical University, No.119 South 4th Ring Road West, Fengtai District, Beijing 100070, China. [2]Department of Neurosurgery, Beijing Neurosurgical Institute, Capital Medical University, Beijing 100070, China.

References

1. Louis DN. Molecular pathology of malignant gliomas. Annu Rev Pathol. 2006;1:97–117.
2. 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;131:803–20.
3. Lathia JD, Gallagher J, Heddleston JM, Wang J, Eyler CE, Macswords J, Wu Q, Vasanji A, McLendon RE, Hjelmeland AB, Rich JN. Integrin alpha 6 regulates glioblastoma stem cells. Cell Stem Cell. 2010;6:421–32.
4. Gangemi RM, Griffero F, Marubbi D, Perera M, Capra MC, Malatesta P, Ravetti GL, Zona GL, Daga A, Corte G. SOX2 silencing in glioblastoma tumor-initiating cells causes stop of proliferation and loss of tumorigenicity. Stem Cells. 2009;27:40–8.
5. Roos A, Ding Z, Loftus JC, Tran NL. Molecular and microenvironmental determinants of Glioma stem-like cell survival and invasion. Front Oncol. 2017;7:120.
6. Beier D, Schulz JB, Beier CP. Chemoresistance of glioblastoma cancer stem cells--much more complex than expected. Mol Cancer. 2011;10:128.
7. Zhao S, Cai J, Li J, Bao G, Li D, Li Y, Zhai X, Jiang C, Fan L. Bioinformatic profiling identifies a glucose-related risk signature for the malignancy of Glioma and the survival of patients. Mol Neurobiol. 2017;54:8203–10.
8. Zhang ZL, Zhao LJ, Chai L, Zhou SH, Wang F, Wei Y, Xu YP, Zhao P. Seven LncRNA-mRNA based risk score predicts the survival of head and neck squamous cell carcinoma. Sci Rep. 2017;7:309.
9. Cheng W, Ren X, Zhang C, Cai J, Liu Y, Han S, Wu A. Bioinformatic profiling identifies an immune-related risk signature for glioblastoma. Neurology. 2016;86:2226–34.
10. Zhao Z, Meng F, Wang W, Wang Z, Zhang C, Jiang T. Comprehensive RNA-seq transcriptomic profiling in the malignant progression of gliomas. Sci Data. 2017;4:170024.
11. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Muller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011;12:77.
12. Gu Z, Gu L, Eils R, Schlesner M, Brors B. Circlize implements and enhances circular visualization in R. Bioinformatics. 2014;30:2811–2.
13. Feng LH, Dong H, Lau WY, Yu H, Zhu YY, Zhao Y, Lin YX, Chen J, Wu MC, Cong WM. Novel microvascular invasion-based prognostic nomograms to predict survival outcomes in patients after R0 resection for hepatocellular carcinoma. J Cancer Res Clin Oncol. 2017;143:293–303.
14. Reya T, Morrison SJ, Clarke MF, Weissman IL. Stem cells, cancer, and cancer stem cells. Nature. 2001;414:105–11.
15. Zhou S, Schuetz JD, Bunting KD, Colapietro AM, Sampath J, Morris JJ, Lagutina I, Grosveld GC, Osawa M, Nakauchi H, Sorrentino BP. The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of stem cells and is a molecular determinant of the side-population phenotype. Nat Med. 2001; 7:1028–34.
16. Fan K, Fan Z, Cheng H, Huang Q, Yang C, Jin K, Luo G, Yu X, Liu C. Hexokinase 2 dimerization and interaction with voltage-dependent anion channel promoted resistance to cell apoptosis induced by gemcitabine in pancreatic cancer. Cancer Med. 2019;8:5903–15.
17. Liu K, Du S, Gao P, Zheng J. Verteporfin suppresses the proliferation, epithelial-mesenchymal transition and stemness of head and neck squamous carcinoma cells via inhibiting YAP1. J Cancer. 2019;10:4196–207.
18. Sundar SJ, Hsieh JK, Manjila S, Lathia JD, Sloan A. The role of cancer stem cells in glioblastoma. Neurosurg Focus. 2014;37:E6.
19. Loria R, Laquintana V, Bon G, Trisciuoglio D, Frapolli R, Covello R, Amoreo CA, Ferraresi V, Zoccali C, Novello M, et al. HMGA1/E2F1 axis and NFkB pathways regulate LPS progression and trabectedin resistance. Oncogene. 2018;37:5926–38.
20. Lee HJ, Zhuang G, Cao Y, Du P, Kim HJ, Settleman J. Drug resistance via feedback activation of Stat3 in oncogene-addicted cancer cells. Cancer Cell. 2014;26:207–21.
21. Aung PP, Parra ER, Barua S, Sui D, Ning J, Mino B, Ledesma DA, Curry JL, Nagarajan P, Torres-Cabala CA, et al. B7-H3 expression in Merkel cell carcinoma-associated endothelial cells correlates with locally aggressive primary tumor features and increased vascular density. Clin Cancer Res. 2019;25:3455–67.
22. Romano S, Tufano M, D'Arrigo P, Vigorito V, Russo S, Romano MF. Cell Stemness, epithelial-to-Mesenchymal transition, and Immunoevasion: intertwined aspects in Cancer metastasis. Semin Cancer Biol. 2019.
23. Romero D. Immunotherapy: PD-1 says goodbye, TIM-3 says hello. Nat Rev Clin Oncol. 2016;13:202–3.
24. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer. 2012;12:252–64.
25. Gabrusiewicz K, Rodriguez B, Wei J, Hashimoto Y, Healy LM, Maiti SN, Thomas G, Zhou S, Wang Q, Elakkad A, et al. Glioblastoma-infiltrated innate immune cells resemble M0 macrophage phenotype. JCI Insight. 2016;1(2).
26. Bradshaw A, Wickremesekera A, Tan ST, Peng L, Davis PF, Itinteang T. Cancer stem cell hierarchy in Glioblastoma Multiforme. Front Surg. 2016;3:21.
27. Denzinger M, Link A, Kurz J, Krauss S, Thoma R, Schlensak C, Wendel HP, Krajewski S. Keratinocyte growth factor modified messenger RNA accelerating cell proliferation and migration of keratinocytes. Nucleic Acid Ther. 2018;28:335–47.
28. Mao XG, Xue XY, Wang L, Zhang X, Yan M, Tu YY, Lin W, Jiang XF, Ren HG, Zhang W, Song SJ. CDH5 is specifically activated in glioblastoma stemlike cells and contributes to vasculogenic mimicry induced by hypoxia. Neuro-Oncology. 2013;15:865–79.

29. Thomas SJ, Snowden JA, Zeidler MP, Danson SJ. The role of JAK/STAT signalling in the pathogenesis, prognosis and treatment of solid tumours. Br J Cancer. 2015;113:365–71.
30. Messina JL, Yu H, Riker AI, Munster PN, Jove RL, Daud AI. Activated stat-3 in melanoma. Cancer Control. 2008;15:196–201.
31. Bharti R, Dey G, Mandal M. Cancer development, chemoresistance, epithelial to mesenchymal transition and stem cells: a snapshot of IL-6 mediated involvement. Cancer Lett. 2016;375:51–61.
32. Zhang K, Che S, Pan C, Su Z, Zheng S, Yang S, Zhang H, Li W, Wang W, Liu J. The SHH/Gli axis regulates CD90-mediated liver cancer stem cell function by activating the IL6/JAK2 pathway. J Cell Mol Med. 2018;22:3679–90.
33. Curtarelli RB, Goncalves JM, Dos Santos LGP, Savi MG, Nor JE, Mezzomo LAM, Rodriguez Cordeiro MM. Expression of Cancer stem cell biomarkers in human head and neck carcinomas: a systematic review. Stem Cell Rev. 2018;14:769–84.
34. Asleh K, Lyck Carstensen S, Tykjaer Jorgensen CL, Burugu S, Gao D, Won JR, Jensen MB, Balslev E, Laenkholm AV, Nielsen DL, et al. Basal biomarkers nestin and INPP4B predict gemcitabine benefit in metastatic breast cancer: samples from the phase III SBG0102 clinical trial. Int J Cancer. 2019;144:2578–86.
35. Liu Z, Han H, He X, Li S, Wu C, Yu C, Wang S. Expression of the galectin-9-Tim-3 pathway in glioma tissues is associated with the clinical manifestations of glioma. Oncol Lett. 2016;11:1829–34.
36. Morantz RA, Wood GW, Foster M, Clark M, Gollahon K. Macrophages in experimental and human brain tumors. Part 2: studies of the macrophage content of human brain tumors. J Neurosurg. 1979;50:305–11.
37. Poon CC, Sarkar S, Yong VW, Kelly JJP. Glioblastoma-associated microglia and macrophages: targets for therapies to improve prognosis. Brain. 2017;140:1548–60.
38. Segerman A, Niklasson M, Haglund C, Bergstrom T, Jarvius M, Xie Y, Westermark A, Sonmez D, Hermansson A, Kastemar M, et al. Clonal variation in drug and radiation response among Glioma-initiating cells is linked to proneural-Mesenchymal transition. Cell Rep. 2016;17:2994–3009.
39. Wang Q, Hu B, Hu X, Kim H, Squatrito M. Scarpace L, deCarvalho AC, Lyu S, Li P, Li Y, et al: tumor evolution of Glioma-intrinsic gene expression subtypes associates with immunological changes in the microenvironment. Cancer Cell. 2017;32:42–56 e46.
40. Binder H, Willscher E, Loeffler-Wirth H, Hopp L, Jones DTW, Pfister SM, Kreuz M, Gramatzki D, Fortenbacher E, Hentschel B, et al. DNA methylation, transcriptome and genetic copy number signatures of diffuse cerebral WHO grade II/III gliomas resolve cancer heterogeneity and development. Acta Neuropathol Commun. 2019;7:59.
41. Klughammer J, Kiesel B, Roetzer T, Fortelny N, Nemc A, Nenning KH, Furtner J, Sheffield NC, Datlinger P, Peter N, et al. The DNA methylation landscape of glioblastoma disease progression shows extensive heterogeneity in time and space. Nat Med. 2018;24:1611–24.
42. Saddawi-Konefka R, Seelige R, Gross ET, Levy E, Searles SC, Washington A Jr, Santosa EK, Liu B, O'Sullivan TE, Harismendy O, Bui JD. Nrf2 induces IL-17D to mediate tumor and virus surveillance. Cell Rep. 2016;16:2348–58.
43. Iamjan SA, Thanoi S, Watiktinkorn P, Reynolds GP, Nudmamud-Thanoi S. Genetic variation of GRIA3 gene is associated with vulnerability to methamphetamine dependence and its associated psychosis. J Psychopharmacol. 2018;32:309–15.
44. Hatakeyama S. TRIM proteins and cancer. Nat Rev Cancer. 2011;11:792–804.
45. Yaguchi H, Okumura F, Takahashi H, Kano T, Kameda H, Uchigashima M, Tanaka S, Watanabe M, Sasaki H, Hatakeyama S. TRIM67 protein negatively regulates Ras activity through degradation of 80K-H and induces neuritogenesis. J Biol Chem. 2012;287:12050–9.
46. Boyer NP, Monkiewicz C, Menon S, Moy SS, Gupton SL. Mammalian TRIM67 functions in brain development and behavior. eNeuro. 2018;5(3).
47. von der Heyde S, Wagner S, Czerny A, Nietert M, Ludewig F, Salinas-Riester G, Arlt D, Beissbarth T. mRNA profiling reveals determinants of trastuzumab efficiency in HER2-positive breast cancer. PLoS One. 2015;10:e0117818.
48. Kumar S, Bandyopadhyay U. Free heme toxicity and its detoxification systems in human. Toxicol Lett. 2005;157:175–88.
49. Cine N, Baykal AT, Sunnetci D, Canturk Z, Serhatli M, Savli H. Identification of ApoA1, HPX and POTEE genes by omic analysis in breast cancer. Oncol Rep. 2014;32:1078–86.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.