Research Article

A Novel S100 Family-Based Signature Associated with Prognosis and Immune Microenvironment in Glioma

Yifang Hu [1] **, Jiahang Song** [2] **, Zhen Wang** [3] **, Jingbao Kan** [1] **, Yaoqi Ge** [1] **, Dan Wang** [1] **, Rihua Zhang** [1] **, Wensong Zhang** [4] **, and Yun Liu** [1,5]

[1] *Department of Geriatrics, The First Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu, China*
[2] *Department of Radiation Oncology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu, China*
[3] *Department of Neurosurgery, The Affiliated Brain Hospital with Nanjing Medical University, Fourth Clinical College of Nanjing Medical University, Nanjing, Jiangsu, China*
[4] *Department of Pharmacy, The First Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu, China*
[5] *Department of Medical Informatics, School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing, Jiangsu, China*

Correspondence should be addressed to Wensong Zhang; zhangwensong91@126.com and Yun Liu; liuyun@njmu.edu.cn

Background. Glioma is the most common central nervous system (CNS) cancer with a short survival period and a poor prognosis. The S100 family gene, comprising 25 members, relates to diverse biological processes of human malignancies. Nonetheless, the significance of S100 genes in predicting the prognosis of glioma remains largely unclear. We aimed to build an S100 family-based signature for glioma prognosis. *Methods.* We downloaded 665 and 313 glioma patients, respectively, from The Cancer Genome Atlas (TCGA) and Chinese Glioma Genome Atlas (CGGA) database with RNAseq data and clinical information. This study established a prognostic signature based on the S100 family genes through multivariate COX and LASSO regression. The Kaplan–Meier curve was plotted to compare overall survival (OS) among groups, whereas Receiver Operating Characteristic (ROC) analysis was performed to evaluate model accuracy. A representative gene S100B was further verified by in vitro experiments. *Results.* An S100 family-based signature comprising 5 genes was constructed to predict the glioma that stratified TCGA-derived cases as a low- or high-risk group, whereas the significance of prognosis was verified based on CGGA-derived cases. Kaplan–Meier analysis revealed that the high-risk group was associated with the dismal prognosis. Furthermore, the S100 family-based signature was proved to be closely related to immune microenvironment. In vitro analysis showed S100B gene in the signature promoted glioblastoma (GBM) cell proliferation and migration. *Conclusions.* We constructed and verified a novel S100 family-based signature associated with tumor immune microenvironment (TIME), which may shed novel light on the glioma diagnosis and treatment.

1. Introduction

Glioma is the most common type of human primary brain cancer, which accounts for approximately 30% of all brain cancer occurrences [1]. Glioma can be divided into low (I-II) or high (III-IV) grades based on the World Health Organization (WHO) criteria. It is difficult to entirely remove tumor tissue during surgery due to the high invasion, infinite proliferation, diffuse infiltration and lack of a clearly boundary of high-grade glioma [2]. Despite advances in

surgery, chemotherapy and radiotherapy, glioma is still associated with a poor prognosis, and its median survival is as short as <15 months [3]. However, these therapeutic strategies are limited by drug resistance and tumor recurrence, which are influenced by a complicated gene regulatory network. Therefore, identification of reliable targets and prognostic biomarkers for glioma is urgently required.

The S100 family is a category of low-molecular-weight (10–14 kda), acidic, calcium-binding protein with an EF-hand motif that was first identified from the bovine brain in

1965 by Blake W. Moore [4]. Currently, 25 family members have been described, 16 are clustered together on chromosome 1q21, a locus susceptible to genomic rearrangements in malignant tumors [5]. S100 proteins participate in regulating some cell processes, like proliferation, differentiation, apoptosis, and immune responses. It has also become evident that many specific S100 genes are abnormally expressed in some human tumors, facilitating cancer genesis and development [6]. S100A4, S100B, and S100P, for example, inhibit the phosphorylation of p53 and subsequently attenuate the tumor-suppressive ability of p53 [7, 8]. S100A8/S100A9 activates the MAPK pathway to promote the proliferation of breast cancer (BC) [9]. Increased S100A11 expression has been observed in lung cancer, which activates Wnt/β-catenin pathways to facilitate the development of drug resistance and cancer metastasis [10]. Additionally, a number of S100 proteins could be identified as molecular biomarkers to diagnose or predict a specific cancer [11].

This study focused on developing a prognostic nomogram based on the S100 family members to explore the clinical significance of this family for glioma prognosis. The prognostic values and expression profiles of the S100 family in glioma samples were comprehensively evaluated using public resources and bioinformatics analysis. We identified five signature-related genes that are associated with the survival of glioma patients; besides, multiple tumor-related pathways were enriched into the high-risk group. Our results indicate that S100 family-based signature may play a critical role in glioma progression and could be considered as prognostic markers and therapeutic targets for glioma in the future.

2. Materials and Methods

2.1. Acquisition of Glioma Datasets. Clinical and FPKM RNAseq data of 703 glioma cases were obtained from TCGA database (https://portal.gdc.cancer.gov/) into the training set. Similarly, we also obtained 325 cases from the CGGA database (http://www.cgga.org.cn) into the validation set. According to patient ID, this study compared clinical features of patients with corresponding transcriptome data. Samples were removed if the data did not match. A total of 665 and 313 patients with complete clinical data were finally selected from TCGA and CGGA database for the next analysis, respectively.

2.2. Construction and Verification of a Risk Score Prognostic Model Based on S100 Gene Family Members. A Cox proportional hazard regression model was constructed to estimate the prognosis of glioma cases in the TCGA training set. Furthermore, the model's prognostic performance was validated in the CGGA validation set. Firstly, the candidate S100 family genes related to prognosis were identified by univariate Cox regression through "survival" package upon the threshold of $P < 0.05$ [12]. Secondly, overfitting genes were removed through LASSO regression via R package "glmnet" function [13]. Thirdly, R package "glmnet" function was also utilized to build a prognosis prediction nomogram by multivariate Cox proportional hazard regression

[14]. The final risk score prognostic model was established with the following formula:

$$\text{risk score} = \sum_{j=1}^{n} \left(\text{Coef}_j \times X_j \right). \tag{1}$$

Here, Coef_j stands for coefficient of multivariate Cox regression for gene j; n represents the overall hub gene number; X_j indicates relative gene j expression within the model.

For exploring the significance of the risk score model in predicting prognosis, glioma cases were classified as a low- or high-risk group according to median risk score value, with high-risk group having a poor prognostic outcome. Then Kaplan–Meier curves were plotted to analyze the OS of two groups of glioma patients through the log-rank test. The sensitivity and specificity of the constructed nomogram were assessed through determining 1-, 3-, and 5-year area under the time-dependent ROC curve (AUC) values using the survival ROC R package, with an AUC > 0.70 denoting a good predictive value.

2.3. Integration of Protein-Protein Interaction (PPI) Network and Identification of Hub Genes. This study constructed the protein-protein interaction (PPI) network using the STRING database (http://www.string-db.org/). Cytoscape (https://cytoscape.org/) is often used for visualizing the complicated network and integrating them with attribute data. In the present work, Cytoscape was utilized for building the PPI network and for analyzing the relationships among S100 family members. Following that, the Maximal Clique Centrality (MCC) algorithm in the Cytoscape software (v 3.7.0) was employed to identify hub genes.

2.4. Construction and Evaluation of the Nomogram. To provide an approach for quantitatively analyzing the OS of glioma, we used the "rms" R packages to construct a nomogram based on clinical variables and the prognosis signature. A calibration curve [15] was plotted to evaluate the nomogram prediction performance by analyzing the consistency of predicted values with actual measurements.

2.5. Gene Set Enrichment Analysis (GSEA). The Hallmark gene set collections (h.all.v7.2) were downloaded in Molecular Signatures Database (MSigDB) v7.1 (https://www.gsea-msigdb.org/gsea/downloads.jsp). The present work carried out GSEA for comparing the biological functions and pathways related to signature-related genes between low- and high-risk groups from both TCGA and CGGA data sets by the use of GSEA v4.0.3 (https://www.gsea-msigdb.org/). In line with GSEA User Guide, the significant gene sets were selected at the thresholds of FDR $q < 0.25$, NOM $p < 0.05$, and |NES| > 1.

2.6. Assessment of Immune Cell Type Fractions. The abundance of immune cell type fraction between low- and high-risk score groups was estimated by CIBERSORT (https://cibersort.stanford.edu/) [16]. CIBERSORT is a new

approach that extensively adopted to characterize cellular components in composite tissues based on gene expression profiling data within cancers, and it can obtain ground truth estimates with high consistency in diverse cancer types. LM22, a white blood cell (WBC) gene signature matrix involving 547 genes, was adopted for distinguishing 22 kinds of tumor-infiltrating immune cells (TIICs), such as regulatory T cells (Tregs), T cells, B cells, natural killer (NK) cells, mast cells, dendritic cells (DCs), monocytes, neutrophils, eosinophils, and macrophages.

2.7. Cell Culture and Transfection. GBM cell lines U251 and T98G were incubated in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum (FBS) at 37°C in a 5% CO_2 incubator. The logarithmic cells were selected for next experiments. S100B knockdown was achieved by transfecting cells with S100B-siRNA using Lipofectamine® 3000 transfection reagent (Invitrogen, USA) according to the manufacturer's protocol. The gene encoding human S100B was inserted into a plasmid vector (Genechem Incorporation, China). The target sequences for siS100B-1 and siS100B-2 were $5'$-GAGACGGUCAUGCAAGAAATT-$3'$ and $5'$-GAGACGGU-CAUGCAAGAAATT-$3'$, respectively.

2.8. RNA Extraction and qRT-PCR. Total RNA was extracted using Trizol reagent (Invitrogen, America), and cDNAs were synthesized using HiScript Synthesis kit (Vazyme, China). Quantitative real-time PCR (qRT-PCR) was then conducted on a StepOnePlus Real-Time PCR system (Applied Biosystems, CA, US) using the Fast SYBR Green Master Mix (Roche, America). Primer sequences in this study are detailed in Table S1.

2.9. Western Blotting. Cell protein was extracted using RIPA lysis buffer (P0013D, Beyotime, China). Equal amounts of protein samples were separated on 12.5% SDS-PAGE gels, which were then electrotransferred onto nitrocellulose (NC) membranes (Pall Corporation, USA). The membranes were blocked for 2 h at room temperature with 5% nonfat milk, then incubated overnight at 4°C with a primary antibody (against S100B, Abcam), followed by the corresponding secondary antibody for 2 h.

2.10. CCK-8 and EdU Proliferation Assay. Cell counting kit-8 (CCK-8) and 5-ethynyl-20-deoxyuridine (EdU) assays were conducted to evaluate the cells' proliferative ability. Cells were seeded in 96-well plates at a density of 2000 cells per well overnight, and cell growth was assayed at different periods utilizing CCK-8 kit (C0038, Beyotime, China) based on the instruction manual. Absorbance at 450 nm was determined by enzyme labeling (Thermo Scientific Multiskan FC, USA). The EdU test was carried out using a Cell-Light EdU Apollo 567 In Vitro Imaging Kit (Ribobio, China) following the manufacturer's protocol. Cells were first stained with 50μM EdU for 2 hours, then fixed with 4% paraformaldehyde and permeabilized with 0.5% Triton X-

100. After three washes with PBS, cells were incubated with $1 \times$ Apollo® for 30 min, followed by DAPI staining. The EdU-positive cells were eventually viewed by fluorescence microscopy (Olympus, Japan).

2.11. Transwell Invasion Assay. The Transwell invasion test was carried out in the 24-well Transwell chambers (Corning, USA) precoated with Matrigel (BD Biosciences, USA). The top chambers were seeded with about 5×10^4 cells in serum-free DMEM media, whereas the bottom chambers were filled with DMEM containing 10% FBS. After 24 h of incubation, the penetrated cells were fixed with 4% methanol, then stained with 0.1% crystal violet. Treated cells in each well were finally photographed at random and counted under an inverted microscope (Nikon, Japan).

2.12. Data Analysis. GraphPad Prism 5.0 (GraphPad Software, Inc., San Diego, CA, USA) or R software (version 4.0.3) was utilized for statistical analysis. Differences between two groups were compared by Mann–Whitney U tests or Students T-tests, whereas those across numerous groups were compared by Kruskal–Wallis H tests or one-way ANOVA. Associations of clinical features with risk scores were analyzed by Fisher's exact test and chi-square test. Survival data were analyzed by Kaplan–Meier analysis. The influence of risk score on OS was evaluated through univariate as well as multivariate Cox regression. The prognosis prediction performance of the risk model was assessed through ROC curve analysis. Each experiment was repeated thrice, and results were expressed as mean ± SD. $^*P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$ were considered significant.

3. Results

3.1. Construction of the S100 Family-Based Signature of Glioma in the TCGA Cohort. The flow chart of this study is shown in Figure 1. A total of 25 genes of S100 family were retrieved from the previous literature and TCGA/CGGA databases [6, 17]. To better explore the interaction among these genes, we established the PPI network comprising 22 nodes and 84 edges (Figure 2(a)). The 10 genes with the highest MCC score were identified by the STRING database and Cytoscape software (Figure 2(b)), suggesting their important role in human cancers.

To develop a specific S100 family-based signature for glioma prognosis, this study carried out univariate, multivariate, and LASSO regression based on the S100 family genes in the TCGA training data set. Finally, five S100 family genes (S100A11, S100A16, S100B, S100PBP, S100A13) were identified as key prognostic genes to establish the risk model (Figures 2(c)–2(e)). Risk score value was determined by this formula: risk score = (S100A11 × 0.431096) + (S100A16 × 0.208028) + (S100B × 0.108247) + (S100PBP × 0.427990) + (S100A13 × 0.107616).

Patients were classified as a low- or high-risk group according to the median risk score value. Figure 3 displays the prediction performance of our constructed nomogram
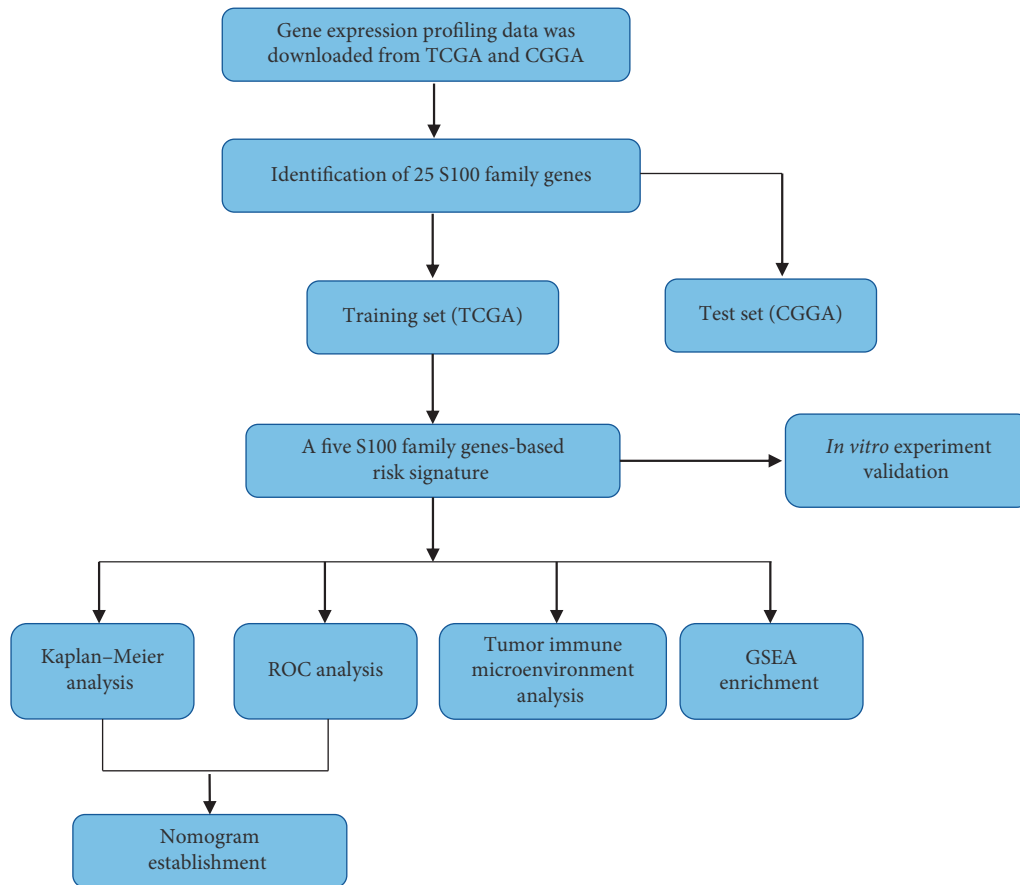
FIGURE 1: Flow chart of study design.

in TCGA and CGGA databases. As revealed by Kaplan-Meier analysis, high-risk group had markedly short OS compared with low-risk group ($P < 0.001$; Figures 3(a) and 3(b)). We then used time-dependent ROC curves to determine the predictive reliability of this model, and the AUCs of ROC analyses are shown in Figures 3(c) and 3(d). Moreover, risk score values of glioma cases and the corresponding distribution are shown in Figures 3(e)–3(f), the survival status of each patient on the dot plot is marked in Figures 3(g)–3(h), and the heatmap of gene expression patterns is described in Figures 3(i)–3(f). The above findings suggested that our constructed prognosis prediction nomogram could be a good prognostic indicator in patients with glioma.

3.2. Independent Prognostic Value of the Five-S100 Family Gene Signature. We performed univariate and multivariate Cox regression to determine whether the five-S100 family gene signature could be independent of other clinical parameters (age, gender, WHO grade, and risk score) as a predictor for patients with glioma. As revealed by univariate analysis, age, risk score value and grade were significantly related to patients' OS in both data sets ($P < 0.001$), only gender was not (Figures 4(a) and 4(c)). Upon multivariate Cox regression, age, risk score value and grade were the independent factors to predict OS for TCGA-derived

patients ($P < 0.001$), whereas only grade and risk score remained statistically significant in the CGGA cohort ($P < 0.001$; Figures 4(b) and 4(d)). The above findings suggested that our constructed model served as an independent predictor for the prognosis of glioma patients.

3.3. Nomogram Analysis. This study constructed the nomogram based on risk score values and clinical features of patients for predicting their risk of survival using "rms" package in R software (Figure 4(e)). The nomogram integrated age, grade, and risk score, and each factor was employed for obtaining relevant score summary, as well as overall score for respective samples. The higher scores indicated a worse prognosis. In the calibration curve, the predictive and actual survival showed a good consistency in 1-, 3-, and 5-year OS (Figure 4(f)). The nomogram model passed the PH assumption and had no statistically significant deviation ($P < 0.05$) (Figure S4). In brief, this nomogram model performed well in predicting glioma survival.

3.4. GSEA Identifies S100 Family-Based Signature-Related Signaling Pathways. The constructed S100 family-based signature had potent stratification ability in the prediction of glioma OS, promoting us to investigate the related signal transduction pathways. GSEA was conducted to compare the
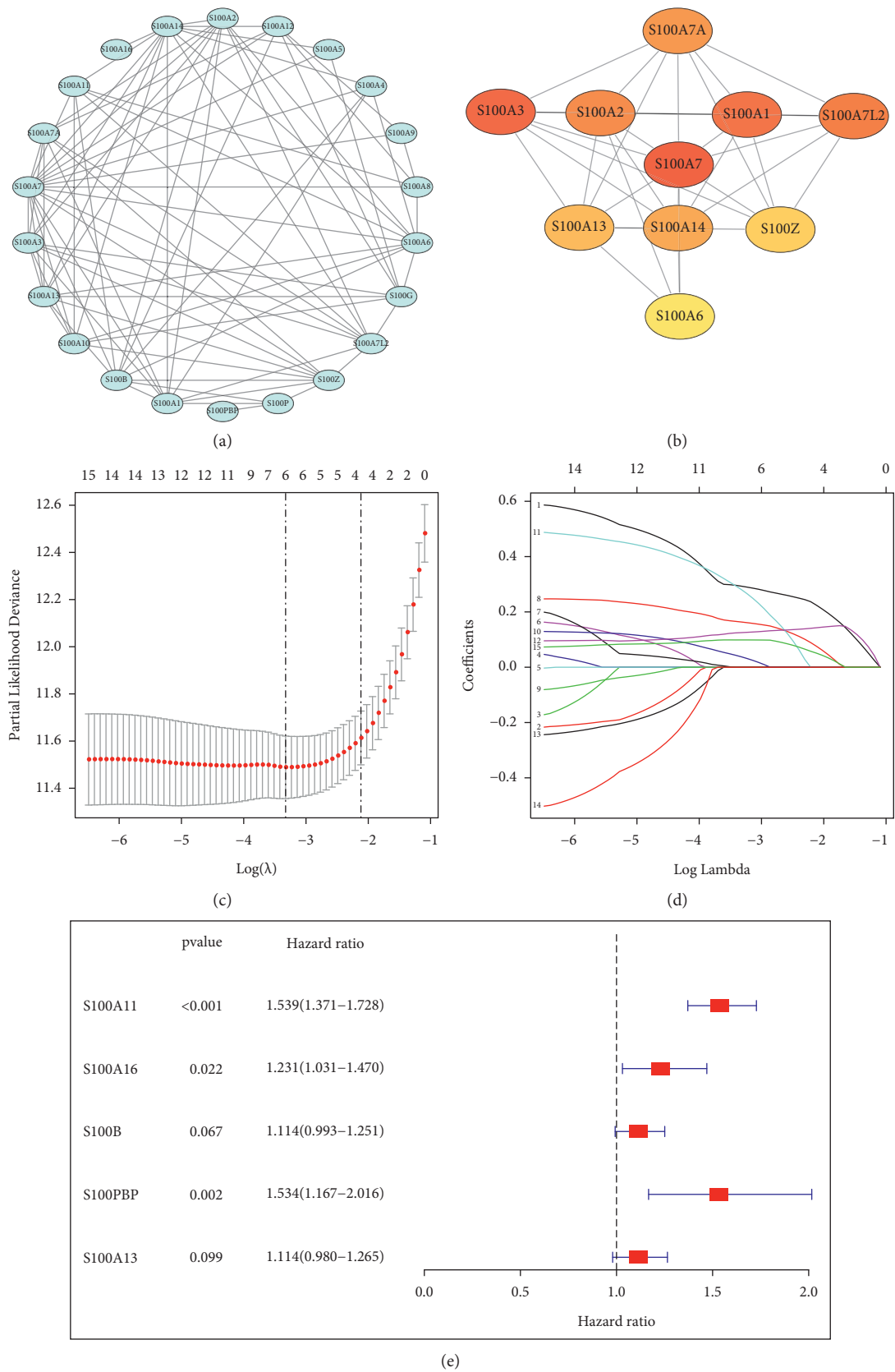
(a)

(b)

(c)

(d)

(e)

FIGURE 2: Establishment of a prognostic signature of glioma on the basis of S100 family genes. (a, b) Protein-protein interaction (PPI) among S100 family genes and the 10 hub genes were identified. (c, d) Multivariate Cox and LASSO regression analyses. (e) Forest plots of risk genes related to the prognosis prediction nomogram.
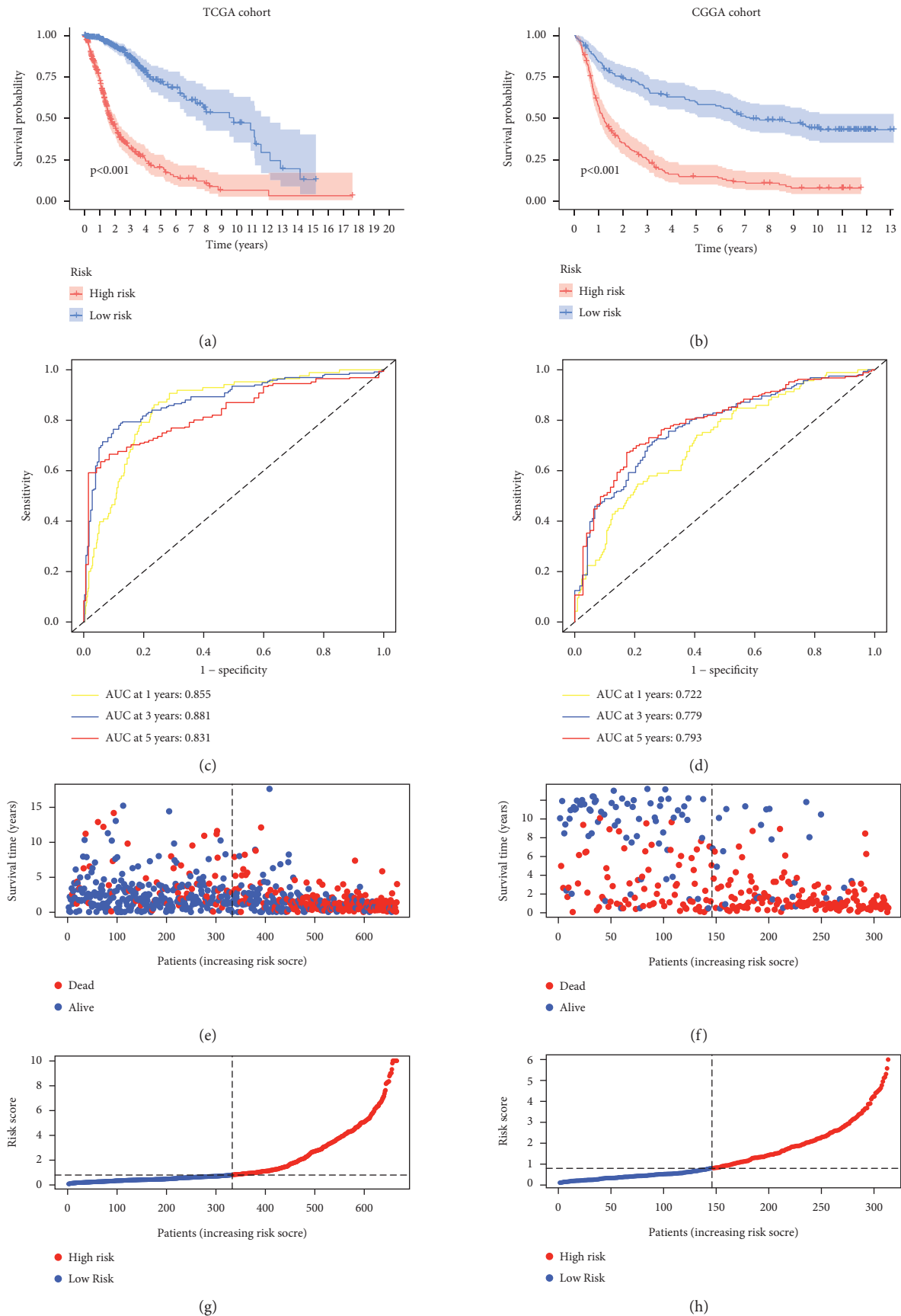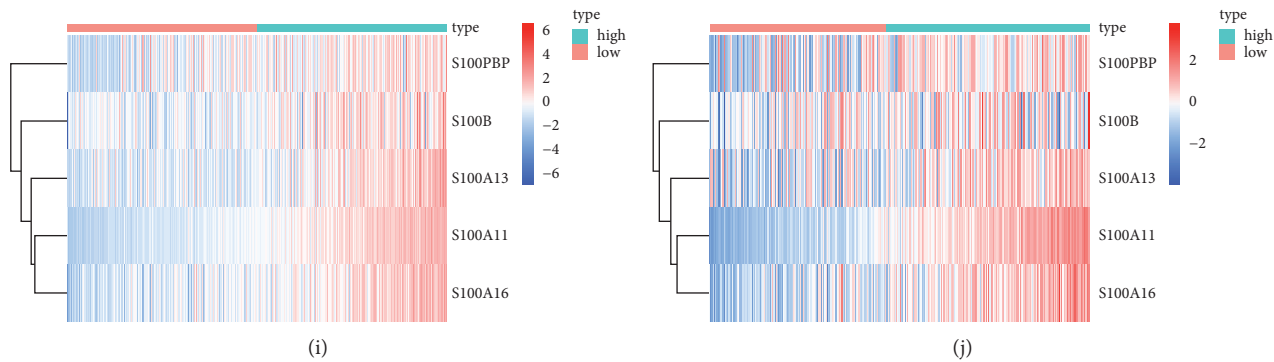
FIGURE 3: Continued.

(i)

(j)

Figure 3: Identification of prognosis prediction performance of S100 family-based model of glioma from TCGA and CGGA databases. (a, b) The Kaplan–Meier (K-M) curve analysis for low- and high-risk patients based on risk score. (c, d) ROC analysis of 1-, 3-, and 5-year survival. (e, f) Patient survival status on the dot pot and survival time with different risk scores. (g, h) Patient distribution of groups and rank of risk score. (i, j) A heatmap of gene levels between high- and low-risk patients.

gene sets enriched in low- and high-risk groups from both data sets. As a result, for TCGA-derived high-risk patients, GSEA pathways were associated with angiogenesis, apoptosis, PI3K-AKT-MTOR signal pathway, epithelial-mesenchymal transition, and glioma stem cell pathway (Figure 5(a)), which were verified using the CGGA database (Figure 5(b)).

3.5. Immune Landscape between Low- and High-Risk Glioma Patients.

More and more studies indicate that tumor development is also affected by the tumor immune microenvironment (TIME). It is necessary to examine the impact of prediction nomogram on TIICs in glioma patients. CIBERSORT with the LM22 signature matrix was adopted for estimating the heterogeneities in 22 kinds of TIIC infiltrating levels in both groups.

As shown in Figure 6, high-risk patients with glioma showed markedly increased M2 macrophage and Treg proportions, whereas apparently decreased activated Mast cell proportion in both CGGA and TCGA cohorts. Additionally, high-risk group had a markedly increased proportion of resting Mast cells compared with low-risk group in TCGA, but there was no significant difference in CGGA. Based on the Human Protein Atlas (HPA) database (http://www.proteinatlas.org), we also found that M2-related molecular (CD163) and Treg-related makers (CD25, STAT5B, and IL-10) were highly expressed in GBM tissues (Figure S1). The qRT-PCR results revealed that two important immune checkpoints, TGF-β and IL-10, were significantly higher in GBM cells than in NHA control (Figure S2).

3.6. Immune Checkpoint Inhibitors (ICIs) of S100 Family-Based Signature.

ICIs have vital functions in maintaining immune homeostasis, which can be utilized by cancer cells to escape the immune response. Here, this study analyzed ICIs levels in both groups. As a result, high-risk patients had markedly increased levels of checkpoints like PD-1, PD-L1, PD-L2, TIM-3, LAG3, and CTLA-4 compared with low-group patients (Figures 7(a)–7(h)). Additionally, the levels

of immunosuppressive cytokines (ARG1, TGF-β, IL-10, IDO1, MOS2, MOS3) increased among high-risk cases (Figures 7(i)–7(j)).

The above findings suggested that high-risk patients were more likely to develop the immunosuppressive microenvironment by upregulating immune checkpoints and immunosuppressive cytokines.

3.7. Verification of the Target Genes.

Furthermore, the databases CGGA and GEPIA were adopted to verify the relationship between the expression of those 5 signature-related genes and patient survival. For the GEPIA-derived cohort, the expression levels of S100A11, S100A16, and S100B in low-grade glioma (LGG) and GBM samples increased compared with those in noncarcinoma samples, while S10013 in LGG tissue was upregulated in normal tissues (Figure 8). In TCGA and CGGA database, we performed a series of survival analyses to reveal the prognostic value of target genes of the signature in glioma patients. According to Figure 9, the group with high expression of S100A11, S100A13, S100A16, S100B, and S100PBP showed shorter OS relative to the group with low expression for all patients with glioma in the TCGA database ($P < 0.001$). In the CCGA database, the OS rate in patients with high levels of S10011, S100A16, and S100B was markedly worse than those with low expression ($P < 0.001$). Then, a subgroup survival analysis was also performed for patients with LGG and HGG (Figure S3). In general, upregulated target gene mRNA expression of the S100 family-based signature predicted dismal prognostic outcomes.

3.8. S100B Mediates GBM Cell Growth and Migration.

To further validate this prognostic model, S100B gene was selected as a representative to carry out the functional experiments for the following reasons. Firstly, the S100A11, S100A16, and S100B expression levels were significantly upregulated in glioma tissues than in normal controls from GEPIA database. Secondly, the survival analysis further
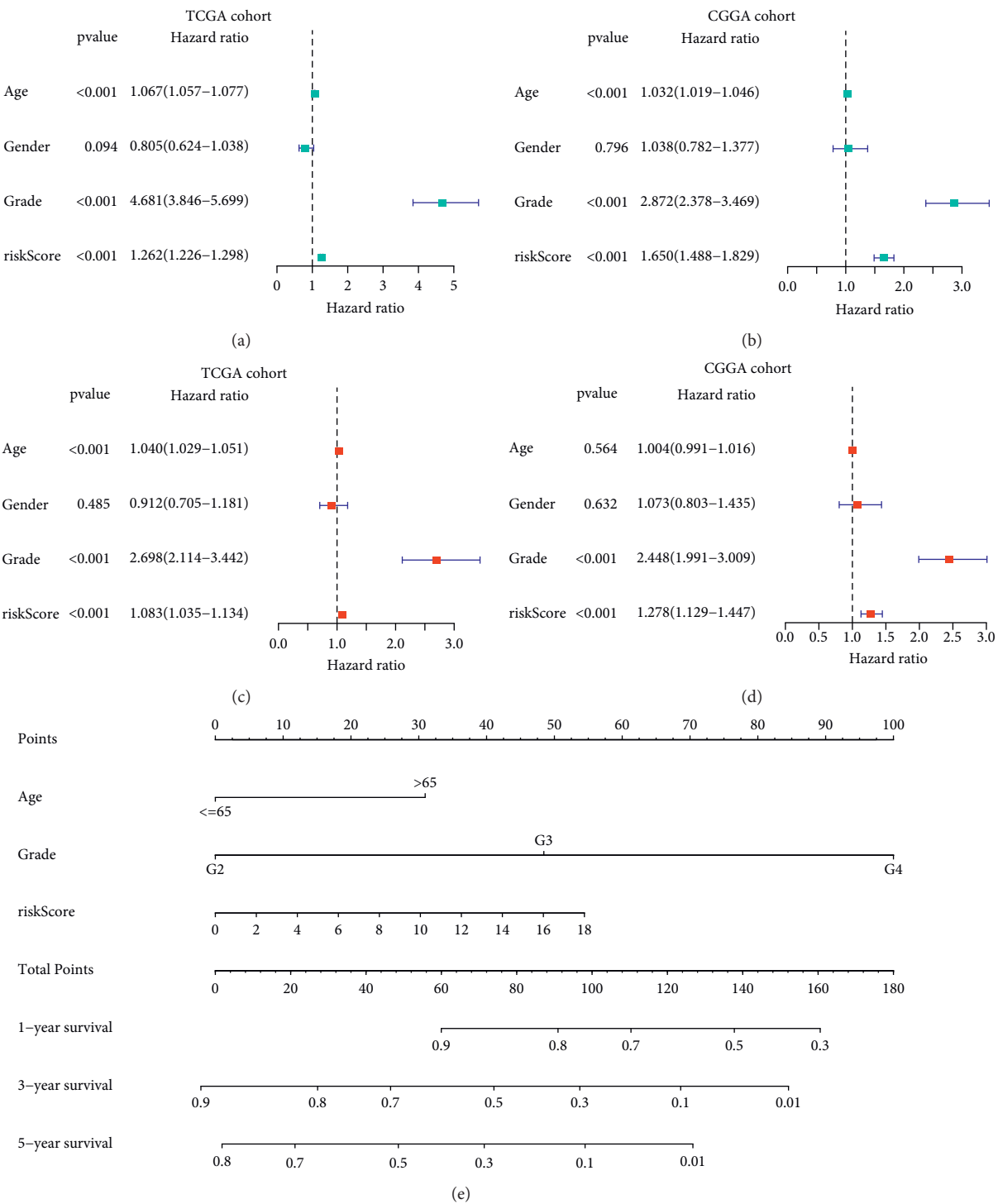
(a)

(b)

(c)

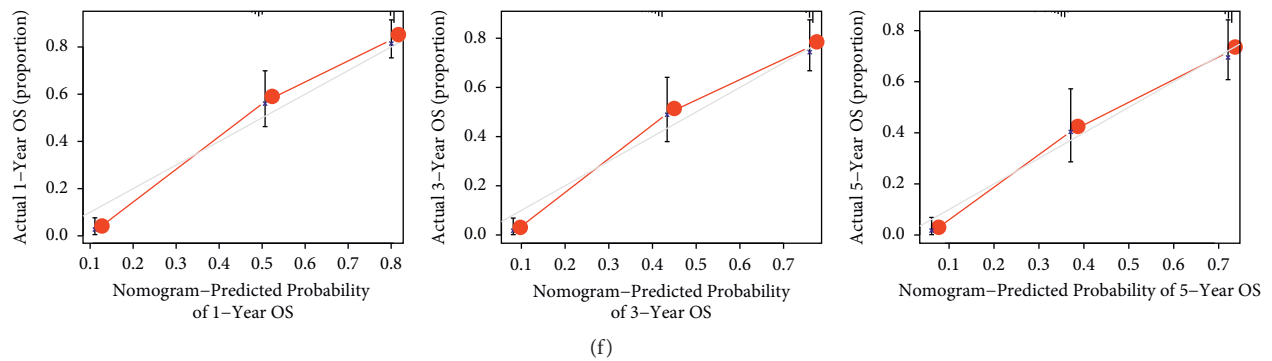(d)

(e)

Figure 4: Continued.

(f)

FIGURE 4: Independent prognosis significance of S100 family-based signature. (a–d) Univariate as well as multivariate Cox regression for OS in glioma cases from TCGA and CGGA databases. (e) A model constructed by incorporating clinical features and signature. (f) Calibration curve showing the nomogram prediction performance in 1-, 3-, and 5-year OS.

demonstrated the significant prognostic power of these 3 signature-related genes in TCGA and CGGA cohorts. Then, the qRT-PCR analysis showed that S100B expression was increased more markedly in GBM cell lines (Figure 10(a)), indicating that S100B may serve as an important prognostic biomarker.

S100B expression level was relatively higher in U251 cells than in T98G cells. Therefore, we knocked down the S100B expression in U251 cells by si-S100B transfection, and upregulated S100B in T98G cells (Figure 10(b)). This study conducted CCK-8 and EdU assays for detecting S100B's effect on the proliferation of GBM cells. As revealed by CCK-8 assay, downregulated S100B expression markedly inhibited U251 cell viability, with its overexpression in T98G cells and yielded the opposite effect (Figure 10(c)). According to EdU assay, inhibiting S100B dramatically decreased the percentage of EdU-positive U251 cells, and overexpressing S100B increased EdU-positive T98G cells (Figure 10(d)). Also, this study conducted a transwell assay for investigating S100B's function in GBM migration and invasion. Silencing S100B expression by siRNA obviously decrease the number of invaded cells, whereas upregulating S100B resulted in more invaded cells (Figure 10(e)). The findings suggested that S100B promoted GBM cell growth and migration.
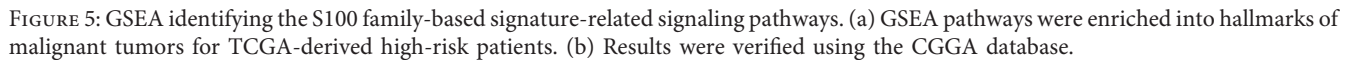
4. Discussion

Glioma is the most common type of brain tumor originating from neuroglial progenitor cells. Typically, the blood-brain barrier (BBB), comprising capillaries, basilar membranes, and endothelial cells, is a major reason for limiting the progress of antitumor drugs. With the recent advances in high-throughput technology, identifying novel prognostic markers and therapeutic targets may help improve glioma survival. Many S100 family proteins showed high expression levels within the nervous system [18]. We speculated that S100 family genes might exhibit a potent prognostic value for glioma patients.

A growing amount of open-sourced online platforms and genomic data have made it possible to explore family gene expression levels in glioma as well as the corresponding

clinical significance. The present work analyzed S100 family genes and constructed a robust nomogram on this basis to predict glioma OS. Using Cox hazards and LASSO regression, five S100 family genes were identified for the prognostic model. The nomogram reliability, prediction performance, and stability were next analyzed and validated. As a result, the constructed signature could discriminate glioma prognosis with high accuracy. In addition, we created a nomogram consisting of clinical features and risk scores to present a personalized survival prediction for each patient with glioma. The calibration curves showed that the predicted patient survival was close to the actual measurement, indicating good predictive effects of the nomogram for survival time. Our signature, therefore, has great potential to be a clinical prognostic and predictive biomarker of glioma.

By focusing on the five signature-related genes of S100 family in this study, most of which have important functions in cancer genesis and development. S100A11, is also called calgizzarin or S100C, is localized both in the nucleus and in the cytoplasm. S100A11 binds to RAGE receptor thereby increasing epidermal growth factor (EGF) protein expression and stimulating cell growth [19]. It has been reported that S100A11 shows overexpression in many cancer types, including glioblastoma (GBM) [20–23]. S100A13 is identified to be involved in the nonclassical protein export, containing fibroblast growth factor (FGF), interleukin-1α (IL-1α), and synaptotagmins [24]. Growing evidence showed that S100A13 has a strong relationship with tumorigenesis [25–27], and it has been proved as a novel biomarker for papillary thyroid carcinomas (PTC) [28, 29]. Interestingly, S100A13 shows differential expression in brain development process, which suggests that it is of great importance to maintain the function of nervous system [30]. S100A16 is a recently discovered member of the S100 family obtained from astrocytoma, which is structurally more stable than other S100 genes [31]. S100A16 overexpression is detected in different cancers, including pancreas cancer, lung cancer, ovarian cancer, bladder cancer, and thyroid gland cancers [32]. S100B is a nervous system-specific protein that is mainly secreted from astrocytes. S100B is widely involved in the regulation of phosphorylation, protein degradation, cellular proliferation, and differentiation.

FIGURE 5: GSEA identifying the S100 family-based signature-related signaling pathways. (a) GSEA pathways were enriched into hallmarks of malignant tumors for TCGA-derived high-risk patients. (b) Results were verified using the CGGA database.
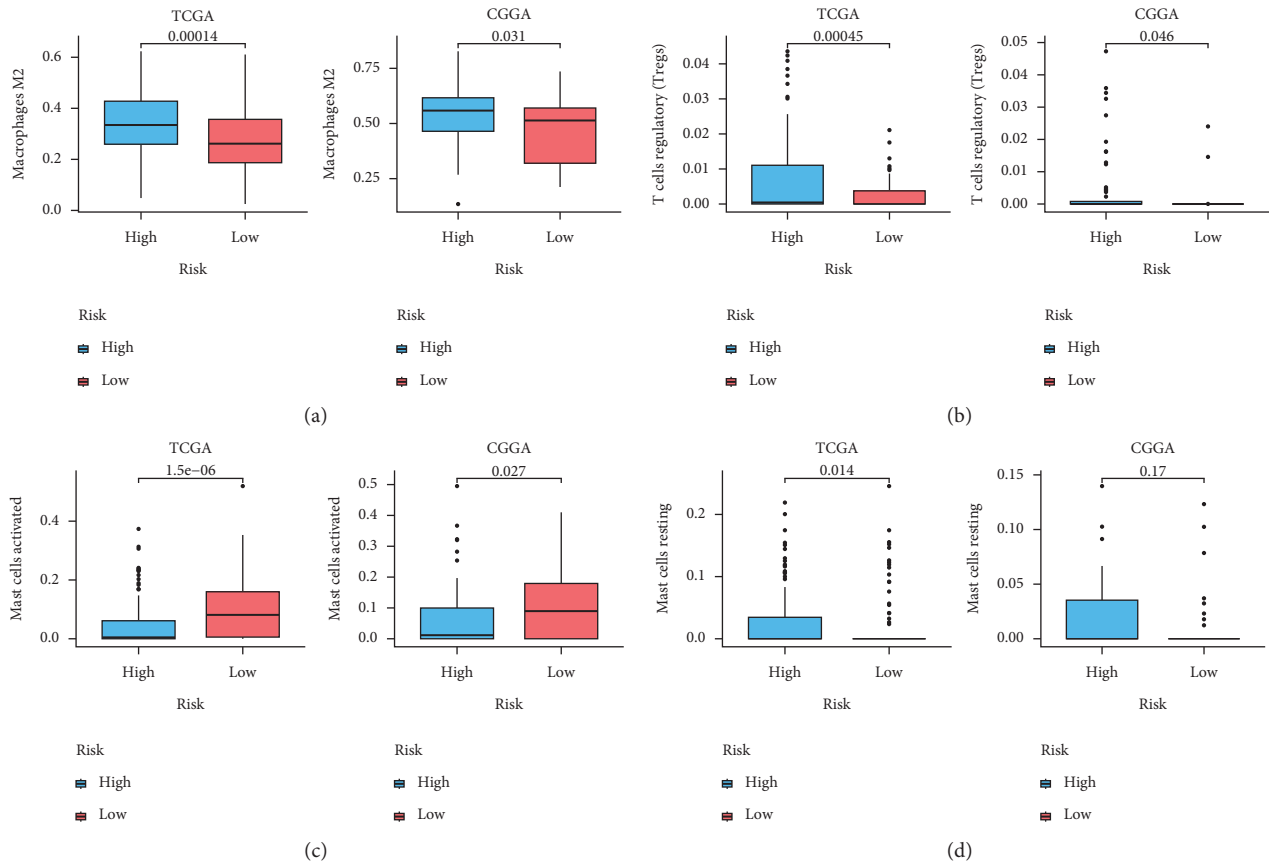
FIGURE 6: Association of risk score with TIIC infiltration. Box plots showing the TIICs with significant difference: (a) M2 macrophages, (b) Regulatory *T* cells, (c) activated Mast cells, and (d) resting Mast cells between low- and high-risk patients with glioma in TCGA and CGGA databases.

Additionally, serum of S100B is used as the diagnostic biomarker for melanoma for a long time, which has also been adopted to be the candidate predicting factor for lung cancer brain metastasis recently [33, 34]. S100PBP is differentially expressed in various organs and disease states, which is dependent on tissue and cancer type. In breast cancer, S100PBP expression was markedly related to patient prognosis and different metastatic sites [35]. S100PBP level is suggested to be related to pancreatic ductal adenocarcinoma [36]. The biological roles of these five genes in cancer have partially provided clues for understanding the diagnostic and prognostic significance of the risk model in glioma. In our study, we demonstrated that most of these genes show high expression within GBM, and glioma patients with high expression levels have a shorter survival time than those with low expression. Moreover, we chose S100B as representative in subsequent functional analyses. The results showed that S100B expression markedly increased within GBM cells, and S100B promoted GBM cells growth, invasion, and migration. More investigations are needed to explore the molecular mechanisms underlying S100B and roles of other markers in the model in GBM.

Functional annotation of the S100 family-based signature via GSEA showed that there are a series of biological functions, such as PI3K-AKT-MTOR signaling,

angiogenesis, apoptosis, epithelial-mesenchymal transition, and glioma stem cell pathways. It is worth mentioning here that apart from cancer-associated pathway, cancer stem cells (CSCs) are highly enriched in the high-risk group. CSCs are known as a rare population of self-renewing tumor cells, which contribute mainly to tumor recurrence and resistant to therapy [37, 38]. These findings indicated that high-risk patients based on the prognostic signature are more predisposed to tumorigenesis, recurrence, and resistance. In the future, more functional experiments are expected to explore the role of these five signature-related genes in glioma stem cells.

In addition, accumulating evidence suggests that TIME exerts an important part in glioma progress and development [39]. As a result, this study also examined the association of TIIC infiltration levels with risk score value in the prognosis mode. The high-risk group showed increased fractions of Treg cells and M2 macrophages phenotype. Macrophages can be divided into classically activated M1 macrophages and alternatively activated M2 macrophages. It is clear that M1 macrophages are involved in the antitumor immune response, and M2 macrophages are mainly responsible for tumor initiation, growth, and metastasis. It is also noted that Treg could promote tumor progression by specifically inhibiting tumor-reactive *T* cells [40]. Consistent
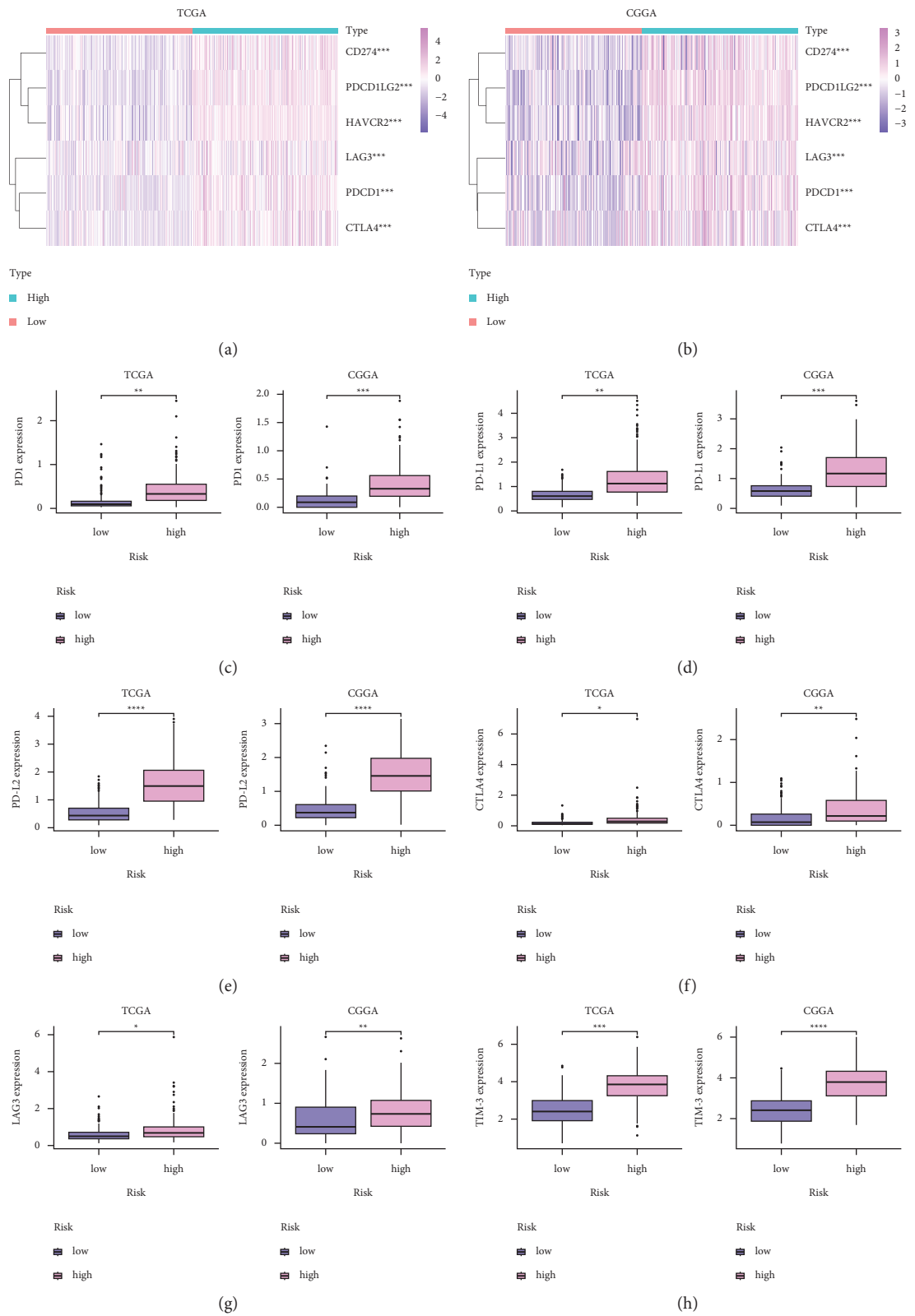
(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

FIGURE 7: Continued.
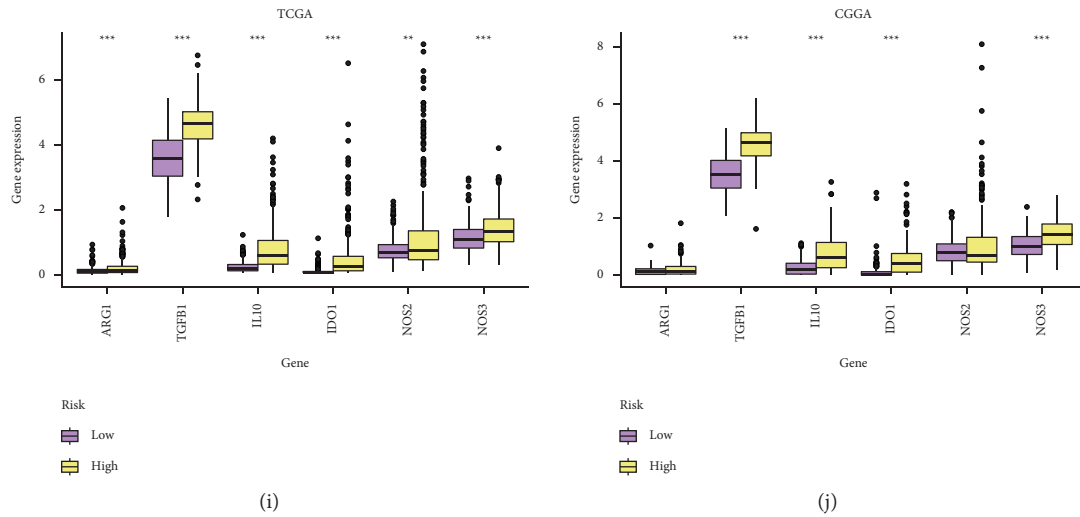
(i)   (j)

FIGURE 7: The estimation of two prognostic subtypes in immunosuppressive microenvironment. (a, b) Heatmap showing expressions of immune checkpoints between high- and low-risk patients derived from TCGA and CGGA databases. (c) PD-1 levels between high- and low-risk patients; (d) PD-L1 levels within high- and low-risk patients; (e) PD-L2 levels within low- and high-risk groups; (f) CTLA-4 expression in low- and high-risk patients; (g) LAG3 levels within low- and high-risk patients; (h) TIM-3 levels within low- and high-risk patients. (i, j) Cytokine levels in tumor immunosuppressive microenvironment between low- and high-risk patients. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, and $^{****}P < 0.0001$.
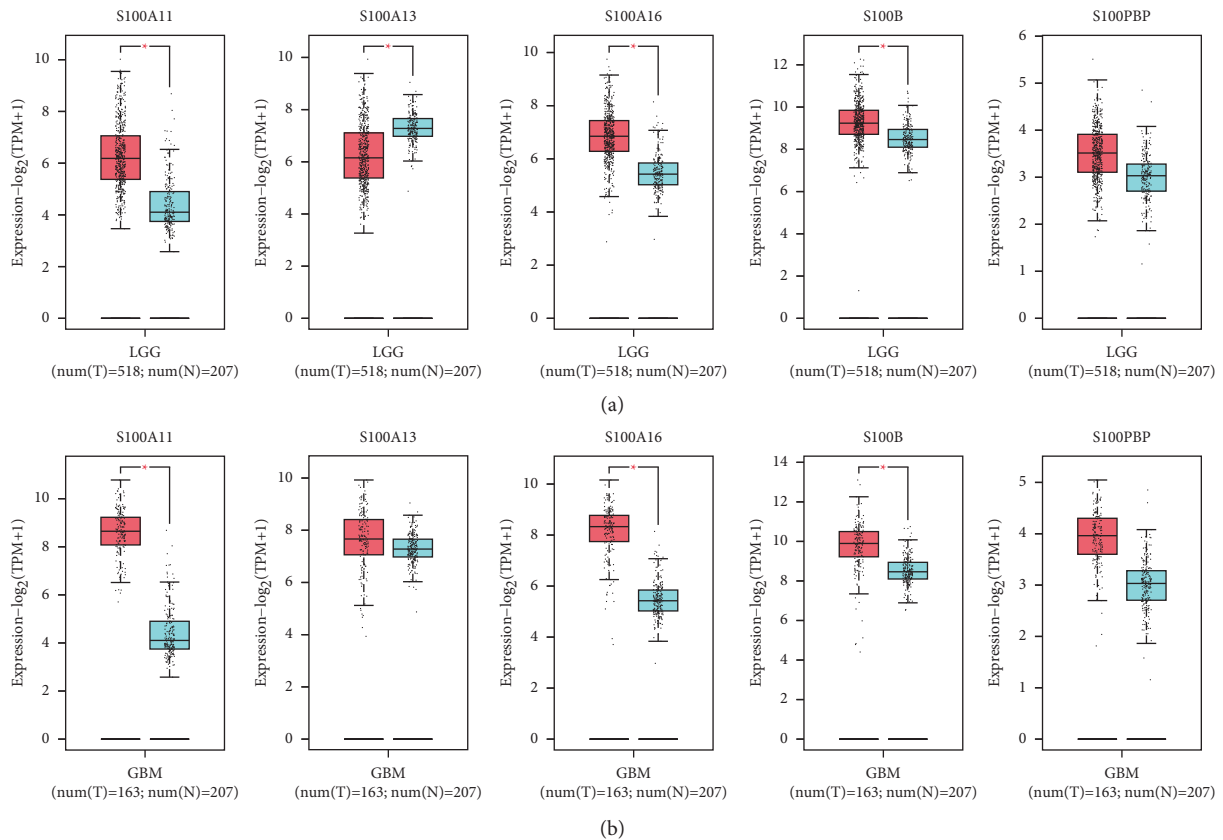

(a)

(b)

FIGURE 8: Expression of five signature-related genes based GEPIA database. (a) S100A11, S100A13, S100A16, S100B, and S100PBP levels between low-grade glioma (LGG) and normal tissues. (b) Gene levels within glioblastoma (GBM) and noncarcinoma samples.
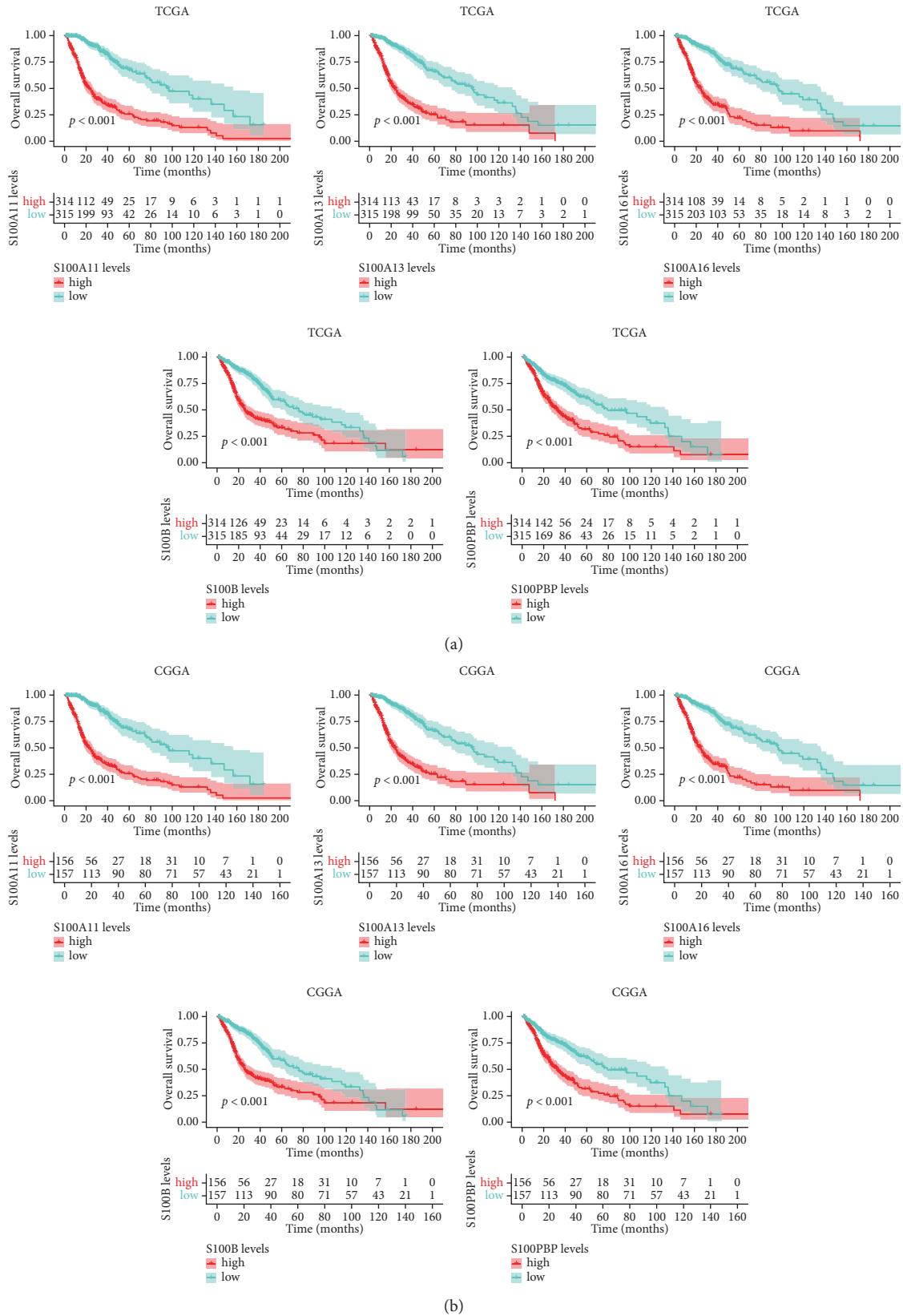
Figure 9: Prognosis of five signature-related genes based on TCGA and CGGA database. (a) Prognostic value of S100A11, S100A13, S100A16, S100B, and S100PBP in glioma from TCGA database. (b) Prognostic value of these genes in glioma from CGGA database.
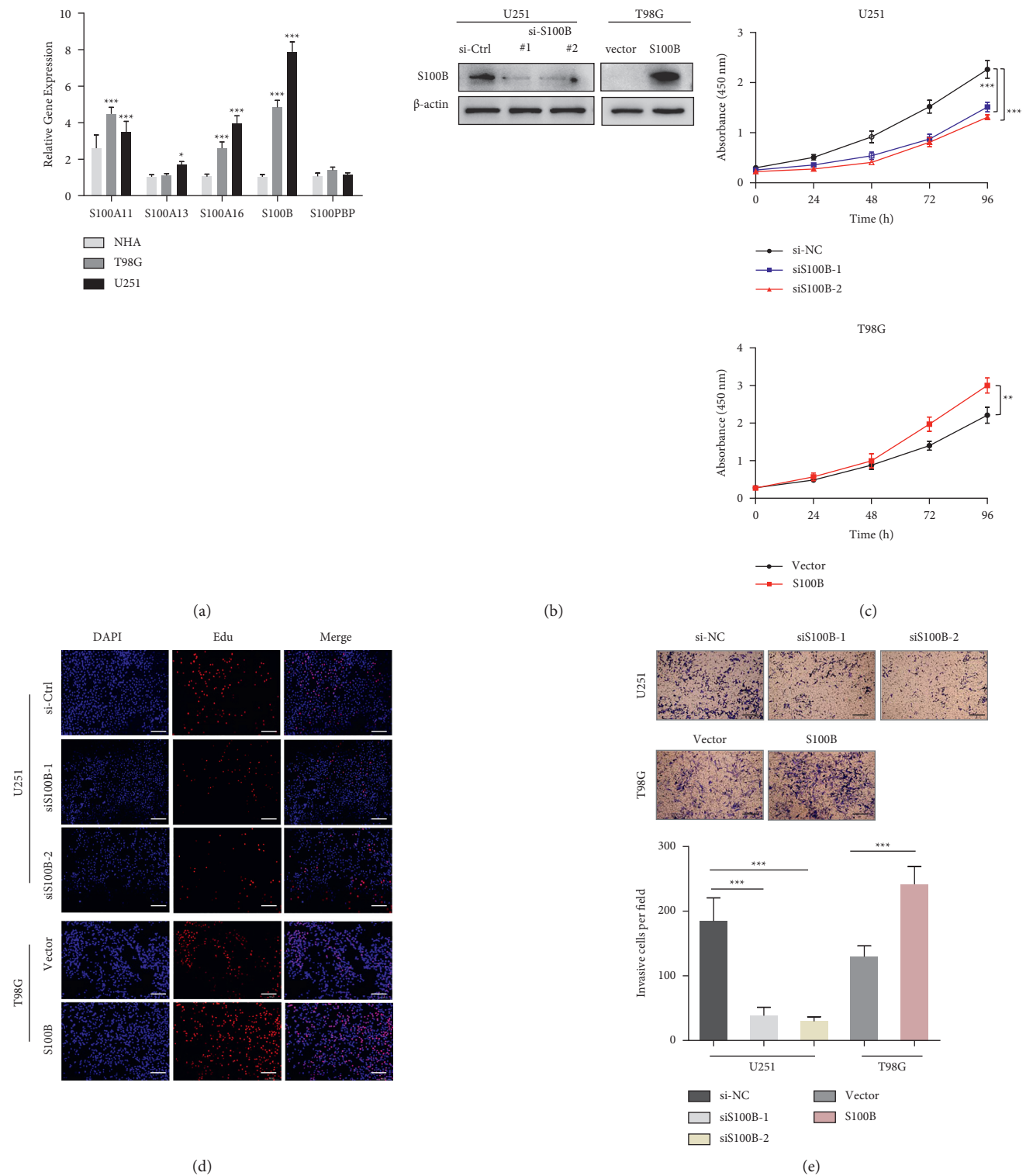
(a)

(b)

(c)

(d)

(e)

Figure 10: S100B mediates GBM cell proliferation and migration. (a) qPCR results showing 5 signature-related genes levels within GBM cells and NHA controls. (b) WB results revealing S100B was efficiently knocked down within U251 cells and overexpressed within T98G cells. (c) CCK-8 assay of U251 cells subjected to specific siRNA transfection and T98G cells after S100B plasmid or vector transfection. (d) EdU assay of U251 cells subjected to specific siRNA transfection and T98G cells S100B plasmid transfection. (e) Transwell assay of U251 cells subjected to specific siRNA transfection and T98G cells S100B plasmid transfection. $^*P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$.

with these, our study suggested that the increased tumor-associated M2 macrophages and Tregs are related to poor prognosis of glioma, probably due to their involvement in immune invasion.

Tumor immune cytokines and checkpoints are considered important factors to determine glioma prognosis and efficacy [41]. Interleukin-10 (IL-10) and transforming growth factor-β (TGF-β) represent two typical immunosuppressive cytokines within TIME. TGF-β is known to inhibit immune responses through suppressing the activity of NK-cells, regulating the generation of proinflammatory cytokines, and changing the differentiation of *T* cells [42]. IL-10 is an anti-inflammatory cytokine that is broadly expressed by various immune cells, including M2 macrophages, myeloid dendritic cells (DCs), Th1, Th2, and Treg cells. Especially, Treg-derived IL-10 can enhance Treg function and involve in Treg-induced immune regulation [43]. In our study, immunosuppressive cytokines, TGF-β and IL-10, were upregulated in the high-risk group. In addition, as the immune checkpoints are often used to escape immune surveillance by cancer cells, we also explore the response of checkpoint inhibitors (e.g., PD-1, PD-L1, PD-L2, CTLA-4, LAG3, and TIM-3) and discovered that many of these genes significantly increased in the high-risk group. Based on our results, high-risk glioma patients may have a better response for immunotherapy.

However, there are some limitations in the present study. Firstly, the data downloaded from public sources was restricted and incomplete, as well as no clinical samples were used for validation. Secondly, the five genes in the signature required more in vitro and in vivo experiments to verify their function in glioma. Finally, further researches in multicenter, large-scale, and prospective clinical trials are needed to confirm the risk model's predictive efficacy.

5. Conclusion

This work first constructed and validated an S100 family-based signature for the prognosis of glioma. This risk signature can be used as a factor to independently predict the glioma patients' OS. We also proved an important value of this model in glioma immune microenvironment. Moreover, we identified that S100B, as an important biomarker, could promote GBM cell growth and invasion in vitro. Our study provided a prognostic model and promising biomarkers for glioma diagnosis and treatment.

Data Availability

All data sets used in the present work are included within this manuscript. These data are available in TCGA (https://portal.gdc.cancer.gov/) and CGGA (http://www.cgga.org.cn) databases.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

YH, JS and ZW contributed equally to this work. YH and JS conceived, designed, and wrote the manuscript. ZW analyzed the data. JK, YG, and RZ performed the experiments. WZ and YL revised the manuscript. All authors have contributed to the article and approved the final manuscript.

Acknowledgments

Supplementary Materials

Figure S1: representative IHC staining for makers relating to M2 macrophages (CD163) and Treg cells (CD25, STAT5B, and IL-10) in normal and tumor tissues from the Human Protein Atlas. Figure S2: qPCR analysis of TGF-β and IL-10 expression in NHA and glioma cell lines. Figure S3: subgroup analysis for prognosis of five genes incorporated into the signature. Figure S4: plot of the *p* values of the global PH assumption tests. Table S1: primer sequences used in this study (). (*Supplementary Materials*)

References

[1] Q. T. Ostrom, H. Gittleman, L. Stetson, S. Virk, and J. S. Barnholtz-Sloan, "Epidemiology of intracranial gliomas," *Progress in Neurological Surgery*, vol. 30, pp. 1–11, 2018.

[2] H. Malkki, "Novel biomarkers for high-grade astrocytoma," *Nature Reviews Neurology*, vol. 10, no. 3, p. 121, 2014.

[3] M. A. Meyer, "Malignant gliomas in adults," *New England Journal of Medicine*, vol. 359, no. 17, p. 1850, 2008.

[4] B. W. Moore, "A soluble protein characteristic of the nervous system," *Biochemical and Biophysical Research Communications*, vol. 19, no. 6, pp. 739–744, 1965.

[5] N. Kozlyuk, A. J. Monteith, V. Garcia, S. M. Damo, E. P. Skaar, and W. J. Chazin, "S100 proteins in the innate immune response to pathogens," *Methods in Molecular Biology*, vol. 1929, pp. 275–290, 2019.

[6] X. Hua, H. Zhang, J. Jia, S. Chen, Y. Sun, and X. Zhu, "Roles of S100 family members in drug resistance in tumors: status and prospects," *Biomedicine & Pharmacotherapy*, vol. 127, Article ID 110156, 2020.

[7] A. Gibadulinova, M. Pastorek, P. Filipcik et al., "Cancer-associated S100P protein binds and inactivates p53, permits therapy-induced senescence and supports chemoresistance," *Oncotarget*, vol. 7, no. 16, pp. 22508–22522, 2016.

[8] T. Yang, J. Cheng, Y. Yang et al., "S100B mediates stemness of ovarian cancer stem-like cells through inhibiting p53," *Stem Cells*, vol. 35, no. 2, pp. 325–336, 2017.

[9] C. Yin, H. Li, B. Zhang et al., "RAGE-binding S100A8/A9 promotes the migration and invasion of human breast cancer cells through actin polymerization and epithelial-mesenchymal transition," *Breast Cancer Research and Treatment*, vol. 142, no. 2, pp. 297–309, 2013.

[10] M. Meng, L. Sang, and X. Wang, "S100 calcium binding protein A11 (S100A11) promotes the proliferation, migration and invasion of cervical cancer cells, and activates wnt/β-catenin signaling," *OncoTargets and Therapy*, vol. 12, pp. 8675–8685, 2019.

[11] C. Allgöwer, A. L. Kretz, S. von Karstedt, M. Wittau, D. Henne-Bruns, and J. Lemke, "Friend or foe: S100 proteins in cancer," *Cancers*, vol. 12, no. 8, 2020.

[12] A. Linden and P. R. Yarnold, "Modeling time-to-event (survival) data using classification tree analysis," *Journal of Evaluation in Clinical Practice*, vol. 23, no. 6, pp. 1299–1308, 2017.

[13] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *Journal of Statistical Software*, vol. 33, no. 1, pp. 1–22, 2010.

[14] N. Simon, J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for cox's proportional hazards model via coordinate descent," *Journal of Statistical Software*, vol. 39, no. 5, pp. 1–13, 2011.

[15] A. Iasonos, D. Schrag, G. V. Raj, and K. S. Panageas, "How to build and interpret a nomogram for cancer prognosis," *Journal of Clinical Oncology*, vol. 26, no. 8, pp. 1364–1370, 2008.

[16] A. M. Newman, C. L. Liu, M. R. Green et al., "Robust enumeration of cell subsets from tissue expression profiles," *Nature Methods*, vol. 12, no. 5, pp. 453–457, 2015.

[17] L. L. Gonzalez, K. Garrie, and M. D. Turner, "Role of S100 proteins in health and disease," *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, vol. 1867, no. 6, Article ID 118677, 2020.

[18] C. W. Heizmann, "Ca2+-Binding proteins of the EF-hand superfamily: diagnostic and prognostic biomarkers and novel therapeutic targets," *Methods in Molecular Biology*, vol. 1929, pp. 157–186, 2019.

[19] A. Safronova, A. Araujo, E. T. Camanzo et al., "Alarmin S100A11 initiates a chemokine response to the human pathogen Toxoplasma gondii," *Nature Immunology*, vol. 20, no. 1, pp. 64–72, 2019.

[20] K. Ohuchida, K. Mizumoto, S. Ohhashi et al., "S100A11, a putative tumor suppressor gene, is overexpressed in pancreatic carcinogenesis," *Clinical Cancer Research*, vol. 12, no. 18, pp. 5417–5422, 2006.

[21] M. C. Anania, C. Miranda, M. G. Vizioli et al., "S100A11 overexpression contributes to the malignant phenotype of papillary thyroid carcinoma," *Journal of Clinical Endocrinology & Metabolism*, vol. 98, no. 10, pp. E1591–E1600, 2013.

[22] Y. Liu, X. Han, and B. Gao, "Knockdown of S100A11 expression suppresses ovarian cancer cell growth and invasion," *Experimental and Therapeutic Medicine*, vol. 9, no. 4, pp. 1460–1464, 2015.

[23] Y. Tu, P. Xie, X. Du et al., "S100A11 functions as novel oncogene in glioblastoma via S100A11/ANXA2/NF-κB positive feedback loop," *Journal of Cellular and Molecular Medicine*, vol. 23, no. 10, pp. 6907–6918, 2019.

[24] H. Matsunaga and H. Ueda, "Stress-induced non-vesicular release of prothymosin-α initiated by an interaction with S100A13, and its blockade by caspase-3 cleavage," *Cell Death & Differentiation*, vol. 17, no. 11, pp. 1760–1772, 2010.

[25] A. Pierce, N. Barron, R. Linehan et al., "Identification of a novel, functional role for S100A13 in invasive lung cancer cell lines," *European Journal of Cancer*, vol. 44, no. 1, pp. 151–159, 2008.

[26] D. Massi, M. Landriscina, A. Piscazzi et al., "S100A13 is a new angiogenic marker in human melanoma," *Modern Pathology*, vol. 23, no. 6, pp. 804–813, 2010.

[27] S. Miao, T. Qiu, Y. Zhao et al., "Overexpression of S100A13 protein is associated with tumor angiogenesis and poor survival in patients with early-stage non-small cell lung cancer," *Thoracic Cancer*, vol. 9, no. 9, pp. 1136–1144, 2018.

[28] J. Martínez-Aguilar, R. Clifton-Bligh, and M. P. Molloy, "A multiplexed, targeted mass spectrometry assay of the S100 protein family uncovers the isoform-specific expression in thyroid tumours," *BMC Cancer*, vol. 15, no. 1, p. 199, 2015.

[29] J. Zhong, C. Liu, Y.-j. Chen et al., "The association between S100A13 and HMGA1 in the modulation of thyroid cancer proliferation and invasion," *Journal of Translational Medicine*, vol. 14, no. 1, p. 80, 2016.

[30] S. G. Rani, S. K. Mohan, and C. Yu, "Molecular level interactions of S100A13 with amlexanox: inhibitor for formation of the multiprotein complex in the nonclassical pathway of acidic fibroblast growth factor," *Biochemistry*, vol. 49, no. 11, pp. 2585–2592, 2010.

[31] E. Babini, I. Bertini, V. Borsi et al., "Structural characterization of human S100A16, a low-affinity calcium binder," *Journal of Biological Inorganic Chemistry*, vol. 16, no. 2, pp. 243–256, 2011.

[32] I. Marenholz and C. W. Heizmann, "S100A16, a ubiquitously expressed EF-hand protein which is up-regulated in tumors," *Biochemical and Biophysical Research Communications*, vol. 313, no. 2, pp. 237–244, 2004.

[33] R. Gaynor, R. Irie, D. Morton, and H. R. Herschman, "S100 protein is present in cultured human malignant melanomas," *Nature*, vol. 286, no. 5771, pp. 400-401, 1980.

[34] H. Choi, V. Puvenna, C. Brennan et al., "S100B and S100B autoantibody as biomarkers for early detection of brain metastases in lung cancer," *Translational Lung Cancer Research*, vol. 5, no. 4, pp. 413–419, 2016.

[35] W. S. Al Amri, D. E. Baxter, A. M. Hanby et al., "Identification of candidate mediators of chemo response in breast cancer through therapy-driven selection of somatic variants," *Breast Cancer Research and Treatment*, vol. 183, no. 3, pp. 607–616, 2020.

[36] P. S. Yang, H. H. Hsu, T. C. Hsu et al., "Genome-wide scan for copy number alteration association with relapse-free survival in colorectal cancer with liver metastasis patients," *Journal of Clinical Medicine*, vol. 7, no. 11, 2018.

[37] D. Zhang, D. G. Tang, and K. Rycaj, "Cancer stem cells: regulation programs, immunological properties and immunotherapy," *Seminars in Cancer Biology*, vol. 52, no. Pt 2, pp. 94–106, 2018.

[38] F. Sharifzad, S. Ghavami, J. Verdi et al., "Glioblastoma cancer stem cell biology: potential theranostic targets," *Drug Resistance Updates*, vol. 42, pp. 35–45, 2019.

[39] S. Roesch, C. Rapp, S. Dettling, and C. Herold-Mende, "When immune cells turn bad-tumor-associated microglia/macrophages in glioma," *International Journal of Molecular Sciences*, vol. 19, no. 2, 2018.

[40] Y. Ohue and H. Nishikawa, "Regulatory T(Treg) cells in cancer: can Treg cells be a new therapeutic target?" *Cancer Science*, vol. 110, no. 7, pp. 2080–2089, 2019.

[41] T. F. Cloughesy, A. Y. Mochizuki, J. R. Orpilla et al., "Neo-adjuvant anti-PD-1 immunotherapy promotes a survival benefit with intratumoral and systemic immune responses in recurrent glioblastoma," *Nature Medicine*, vol. 25, no. 3, pp. 477–486, 2019.

[42] C. Larson, B. Oronsky, C. A. Carter et al., "TGF-beta: a master immune regulator," *Expert Opinion on Therapeutic Targets*, vol. 24, no. 5, pp. 427–438, 2020.

[43] M. Saraiva and A. O'Garra, "The regulation of IL-10 production by immune cells," *Nature Reviews Immunology*, vol. 10, no. 3, pp. 170–181, 2010.