Research Paper

Integrated clustering signature of genomic heterogeneity, stemness and tumor microenvironment predicts glioma prognosis and immunotherapy response

Yangyang Wu¹, Meng Mao^{1,2,3}, Lin-Jian Wang^{1,3,4}

¹Advanced Medical Research Center of Zhengzhou University, Zhengzhou Central Hospital Affiliated to Zhengzhou University, Zhengzhou 450007, China

²Department of Anesthesiology and Perioperative Medicine, Zhengzhou Central Hospital Affiliated to Zhengzhou University, Zhengzhou 450007, China

³Research of Trauma Center, Zhengzhou Central Hospital Affiliated to Zhengzhou University, Zhengzhou 450007, China

⁴Department of Neurosurgery, Zhengzhou Central Hospital Affiliated to Zhengzhou University, Zhengzhou 450007, China

Correspondence to: Meng Mao, Lin-Jian Wang; **email:** <u>hnmaomeng@zzu.edu.cn</u>, <u>wljgnn@126.com</u>, <u>https://orcid.org/0000-0002-7054-6397</u>

Keywords: glioma, genomic heterogeneity, stemness, tumor microenvironment, risk signatureReceived: May 16, 2023Accepted: August 21, 2023Published: September 11, 2023

Copyright: © 2023 Wu et al. This is an open access article distributed under the terms of the <u>Creative Commons Attribution</u> <u>License</u> (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Background: Glioma is the most frequent primary tumor of the central nervous system. The high heterogeneity of glioma tumors enables them to adapt to challenging environments, leading to resistance to treatment. Therefore, to detect the driving factors and improve the prognosis of glioma, it is essential to have a comprehensive understanding of the genomic heterogeneity, stemness, and immune microenvironment of glioma.

Methods: We classified gliomas into various subtypes based on stemness, genomic heterogeneity, and immune microenvironment consensus clustering analysis. We identified risk hub genes linked to heterogeneous characteristics using WGCNA, LASSO, and multivariate Cox regression analysis and utilized them to create an effective risk model.

Results: We thoroughly investigated the genomic heterogeneity, stemness, and immune microenvironment of glioma and identified the risk hub genes RAB42, SH2D4A, and GDF15 based on the TCGA dataset. We developed a risk model utilizing these genes that can reliably predict the prognosis of glioma patients. The risk signature showed a positive correlation with T cell exhaustion and increased infiltration of immunosuppressive cells, and a negative correlation with the response to immunotherapy. Moreover, we discovered that SH2D4A, one of the risk hub genes, could stimulate the migration and proliferation of glioma cells.

Conclusions: This study identified risk hub genes and established a risk model by analyzing the genomic heterogeneity, stemness, and immune microenvironment of glioma. Our findings will facilitate the diagnosis and prediction of glioma prognosis and may lead to potential treatment strategies for glioma.

INTRODUCTION

Central nervous system (CNS) malignancies are among the cancers with the poorest prognosis [1]. Glioma is the most common primary central nervous tumor in the brain, accounting for approximately 81% of malignant brain tumors. Gliomas usually originate from glial cells or precursor cells and progress to astrocytomas, oligodendrogliomas, ependymomas, or oligoastrocytomas [2, 3]. According to the previous

classification scheme established by the World Health Organization, primary gliomas are classified as grades I to IV [4]. This grading system reflects the malignancy of the tumor, and in general, a higher grade is associated with a worse prognosis. WHO grade I indicates that slow-growing lesions are usually associated with a good prognosis, while WHO grade IV is recognized as highly malignant [5]. WHO grade II and III gliomas, including astrocytomas, oligodendrogliomas, and mixed oligodendrogliomas are defined as low-grade gliomas (LGG). Generally, low-grade gliomas (LGG) show some sensitivity to treatment and have a better prognosis [6, 7]. World Health Organization grade IV gliomas (glioblastomas, GBM) are highly malignant and also the most common gliomas, accounting for approximately 45% of all gliomas [8, 9]. At present, molecular pathological features (such as mutation status of isocitrate dehydrogenase gene IDH1 or IDH2) are used to classify adult glioma. IDH-mutated gliomas usually present with a lower histological grade and have a better prognosis. Median survival is approximately 12 vears [10]. In contrast, IDH wild-type gliomas usually present as glioblastomas (GBM) with a poor prognosis and a median survival of only 12 - 15 months [11–13].

Traditional cancer treatments, such as surgery, chemotherapy, and radiation therapy, have shown limited improvement in the prognosis of patients with glioma. The main reason for the limited therapeutic progress in glioma is the blood-brain barrier consisting of endothelial cells, capillaries, and basement membrane, which prevents most anti-tumor drugs from entering the brain. Thus, although many cancer therapies have been developed, few drugs have been approved for the treatment of glioma [14, 15]. Immunotherapy is currently at the forefront of cancer treatment. Even though immune checkpoint inhibitors (ICI) can penetrate the blood-brain barrier, they still do not improve the prognosis of glioma [16, 17]. Gliomas are characterized by a high degree of heterogeneity and the ability to proliferate aggressively [11]. The heterogeneity of gliomas allows them to adapt to challenging microenvironments, leading to treatment resistance [18, 19]. For patients with glioblastoma, recurrences and resistance to treatment can result in a short average survival time after treatment. To a large extent, this is due to the molecular heterogeneity of gliomas, which affects the overall prognosis and response to treatment [20, 21]. In addition, the tumor microenvironment is another factor in the development and treatment resistance of glioma. The tumor microenvironment (TME) consists of a variety of nontumor cells, such as endothelial cells, stromal cells, mesenchymal cells, and immune cells, which play an extremely important role in tumor progression, recurrence, and drug resistance [22-24]. Communication between glioma cells and adjacent cells and the immune environment accelerates the cancer process and contributes to the formation of glioma stem cells, resulting in treatment resistance [21, 25]. Meanwhile, the interaction between tumor microenvironment (TME) and tumor stem cells significantly affects the aggressive proliferation and molecular heterogeneity of tumors [25].

Here, we found that consensus clustering analysis based on tumor stemness, genomic heterogeneity, and immune microenvironment could classify gliomas into distinct subtypes, and the prognosis was significantly different among subtypes. By screening differentially expressed genes, and performing WGCNA, LASSO, and multivariate Cox regression analyses, we identified risk hub genes (RAB42, SH2D4A, and GDF15) associated with glioma stemness, genomic heterogeneity, and the immune microenvironment. The prognostic model constructed using these risk hub genes in this study could predict not only glioma prognosis, but also immunotherapy response in immunotherapy cohort. Although SH2D4A appears to play opposing roles in tumors [26, 27], our experiments demonstrated that knockdown of SH2D4A significantly inhibited the migration and proliferation of glioma cells. In conclusion, we investigated the impact of glioma stemness, genomic heterogeneity, and immune microenvironment-related risk hub genes on glioma and provided new ideas for therapeutic strategies for glioma.

MATERIALS AND METHODS

Datasets and samples

The TCGA dataset (Supplementary Table 1) was downloaded from the Xena Browser at the University of California, Santa Cruz (UCSC, <u>https://xenabrowser.net/datapages/</u>) [28]. The clinical data (Supplementary Table 2) and RNA-seq data of the Chinese Glioma Genome Atlas (CGGA) were obtained from the CGGA data portal (<u>http://www.cgga.org.cn/</u>) [29]. The data for the IMvigor210 cohort were loaded from the R package "IMvigor210CoreBiologies" [30].

Analysis of genomic heterogeneity, stemness, and immune microenvironment

The glioma single nucleotide variant dataset processed by MuTect2 software [31] was downloaded from GDC (https://portal.gdc.cancer.gov/). Tumor mutation burden (TMB) was calculated for each glioma using the tmb function of the R package "maftools", and microsatellite instability (MSI) for each glioma was obtained from previous studies [32, 33]. Glioma stemness scores based on RNA expression and DNA methylation were calculated according to a previous study [34]. Stromal-Score, ImmuneScore, and ESTIMATEScore were calculated for each glioma based on gene expression using the R software package "estimate" [35]. The abundance of tumor-infiltrating immune cells in gliomas was analyzed using the CIBERSOR, MCPCOUNTER and EPIC algorithms on the TIMER2 platform (<u>http://timer.cistrome.org/</u>) [36].

Consensus clustering analysis

Cluster analysis was performed by Consensus-ClusterPlus [37], using agglomerative pam clustering with 1-pearson correlation distances, with 10 repetitions for 80% of the samples. The empirical cumulative distribution function plot was used to determine the optimal number of clusters.

Transfection with siRNA

These siRNAs were synthesized by the Shanghai GenePharma Co. (Supplementary Table 3). U87 MG cells were transfected with negative control siRNA and SH2D4A siRNA according to the manufacturer's instructions of Lipofectamine 3000 (Invitrogen, L3000015). 48 hours after transfection, U87 MG cells were harvested for subsequent western blot and qRT-PCR experiments to verify their knock-down efficiency.

Western blot analysis

Western blot analysis was performed according to our previously reported method with minor modifications [38]. In brief, protein extracts were prepared using RIPA lysis buffer (Epizyme) in the presence of 100x EDTA-free Protease Inhibitor Cocktail (Epizyme). Cell lysates were separated through electrophoresis and electro-transferred to polyvinylidene difluoride (PVDF) membranes (Merck Millipore). The membrane was blocked with 5% non-fat milk in TBS solution for 2 h at room temperature and incubated with anti-SH2D4A (15957-1-AP; Proteintech) or anti-α-Tubulin (11224-1-AP; Proteintech) overnight at 4° C. Blots were then washed thrice with TBST and probed for 2 h at room temperature with HRP- conjugated AffiniPure Goat Anti-Rabbit IgG (SA00001-2; Proteintech). Finally, the labeled proteins were detected using the ECL reagent.

Quantitative real-time PCR

Trizol (Thermo Fisher) was used to extract RNA samples, and then 1 µg of total RNA and NovoScript® Plus All-in-one 1st Strand cDNA Synthesis SuperMix (gDNA Purge) (Novoprotein E047-01B) were used to prepare cDNA according to the manufacturer's instructions. Thereafter, cDNA was then analyzed by NovoStart® SYBR qPCR SuperMix Plus (Novoprotein E096-01A). Relative gene expression was evaluated using the $2^{-\Delta\Delta CT}$ method. The GAPDH was utilized as an internal control, and all primer sequences are compiled in Supplementary Table 4.

Cell proliferation and migration

Cell proliferation was analyzed by Cell Counting Kit-8 (CCK-8) (CellorLab, CX001L) and BeyoClickTM EdU-594 assay kit (Beyotime, C0078S). After 24h of transfection, cell suspensions were planted in 96-well plates. At 24, 48, 72 and 96h, 10ul CCK-8, and 90ul medium were incubated together at 37° C for 2h, then the absorbance at 450 nm was detected by a microplate reader.

For direct observation of the proliferating cells, the 5-Ethynyl-2-deoxyuridine incorporation experiment was also performed according to the specifications. The cells were further incubated with EdU for 2h before fixation, permeabilization, and EdU staining. The cell nuclei were stained with Hoechst33342 for 30min. Finally, the proportion of cells that incorporated EdU was detected using confocal laser microscopy.

To perform migration assays, U87 MG cells were seeded into 6-well plates. Transfection was performed when cell density reached 70-90%. 48 h later the plates were scraped with pipette tips, washed with PBS, and incubated with serum-free medium. The original images and migrated images were obtained using the inversion microscope system. The migrated area was analyzed by Image J software.

Construction of the risk model

In order to construct a scale-free co-expression network, a weighted correlation network analysis (WGCNA) was performed using the R package "WGCNA", and four co-expression modules were finally obtained after merging modules with distances less than 0.25. Genes with high connectivity in the clinically important modules were identified as hub genes. LASSO and multivariate regression analyses were then performed sequentially to screen for positive hub genes that were significantly associated with overall survival (OS). Risk scores were calculated as follows:

$$Risk\ score = \sum_{i=1}^{n} Coef_i * Exp_i$$

Identification of DEGs and enrichment analysis

We identified differentially expressed genes (DEGs) between different clusters by executing the "limma" package in the R software (p< 0.05 and $|FC| \ge 1.5$).

Gene enrichment analysis was conducted using Metascape [39].

Statistical analysis

One-way ANOVA, Wilcox test and t-test were used to analyze the significance of differences in heterogeneity, stemness, gene expression, and infiltration of immune cells in gliomas. LASSO, multivariate Cox regression, and Kaplan-Meier analyses were performed to screen and evaluate the risk signature using the R packages "glmnet" and "survival". Roc curve was drawn using the R package "survivalROC." All statistical analyses were performed using GraphPad Prism and R software, and P values less than 0.05 were considered statistically significant.

Availability of data and material

Bioinformatics datasets presented in this study can be found in online repositories, and the datasets used and/or analyzed during experiments are available from the corresponding author on reasonable request.

Consent for publication

All authors had final approval of the submitted versions and read the journal's authorship statement.

RESULTS

Consensus clustering analysis of stemness, genomic heterogeneity, and the immune microenvironment in glioma

The molecular heterogeneity of gliomas significantly affects the overall prognosis of patients and their response to treatment. Three clusters were identified through unsupervised consensus clustering analysis of genomic heterogeneity of gliomas in the TCGA dataset (Figure 1A). We examined the overall survival of the three clusters and found significant differences in prognosis between them. Cluster 1 had the lowest prognosis, while cluster 3 had the highest prognosis (Figure 1B). Significant differences were observed in TMB and MSI between cluster 1, with the lowest prognosis, and cluster 3, with the highest prognosis (Figure 1C). In addition, gliomas were divided into two subtypes based on the unsupervised consensus clustering analysis of stemness scores in the TCGA dataset (Figure 1D). The prognosis of cluster 2 was more favorable than that of cluster 1 (Figure 1E). Assessment of glioma stemness scores based on RNA expression (RNAss) and DNA methylation (DNAss) revealed that cluster 1 had higher DNAss and lower RNAss than cluster 2 (Figure 1F). Tumor microenvironment is a crucial factor in glioma development and treatment resistance. We conducted another cluster analysis to examine the immune microenvironment of glioma. This analysis revealed that the StromalScore, ImmuneScore, and ESTIMATEScore differed significantly between clusters, with cluster 1 having higher scores than cluster 2. Improved prognosis was observed in cluster 2 (Figure 1G–1I).

Identification of DEGs associated with glioma stemness, genomic heterogeneity, and immune microenvironment

We quantified gene expression among the clusters divided according to genomic heterogeneity, stemness, and tumor microenvironment, and identified 1792 (412 down-regulated, 1380 up-regulated), 1599 (306 downregulated, 1293 up-regulated), and 1806 (709 downregulated, 1097 up-regulated) differentially expressed genes (DEGs), respectively (Figure 2A-2C). And finally, we identified 857 DEGs shared by the three clustering (Figure 2D). Enrichment analysis revealed that differentially expressed genes were involved in immunoregulatory pathways, including Cytokinecytokine receptor interaction, Signaling by Interleukins, regulation of T cell activation, and Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell pathway (Figure 2E). We extracted all proteinprotein interactions among differentially expressed genes from a PPI data source and constructed a PPI network. MCODE algorithm was applied to this network to identify neighborhoods where proteins are densely connected (Figure 2F). GO enrichment analysis was applied to the network to reveal biological meanings, and the results showed that MCODE 3 was enriched in HDACs deacetylate histones; MCODE 4 was enriched in TCR signaling Translocation of ZAP-70 to Immunological synapse, Phosphorylation of CD3 and TCR zeta chains; MCODE 6 was enriched in JAK-STAT signaling pathway, Signaling by Interleukins, and Cytokine Signaling in Immune system (Figure 2F).

Screening of risk hub genes

Through WGCNA analysis of integrated data containing DEGs and clinical characteristics (such as genomic heterogeneity, stemness and tumor immune microenvironment), four co-expression modules were finally determined (Figure 3A–3C). We filtered the most relevant module by evaluating its interrelationships with traits and identified the blue module as the most relevant (Figure 3D, 3E). The hub genes of blue module were extracted. Subsequently, LASSO Cox regression analysis and multivariate Cox regression analysis were performed sequentially to filter variables. Ultimately, we selected RAB42, SH2D4A, and GDF15 as risk hub

genes (Figure 4A–4C). The expression of three risk hub genes in glioma was investigated in the TCGA and CGGA datasets, and RAB42, SH2D4A, and GDF15 were found to be higher expressed in GBM than in LGG (Figure 4D, 4E). In the TCGA and CGGA datasets, the Kaplan-Meier curve showed that RAB42, SH2D4A, and GDF15 significantly affected the prognosis of glioma, with a poor prognosis at high expression (Figure 4F, 4G).

Construction of the risk signature

A risk model was then constructed in the TCGA dataset using the risk hub genes RAB42, SH2D4A, and GDF15. Kaplan-Meier curve showed that high-risk scores were associated with poor prognosis. The risk score had high sensitivity and specificity in predicting 1-, 3-, and 5-year survival rates in glioma patients (Figure 5A–5C). The results of the risk model were also confirmed in the CGGA dataset (Figure 5D–5F). In addition, we also validated the risk model in GBM and LGG, respectively (Supplementary Figure 1). The results of the correlation analysis showed that risk scores and risk hub genes RAB42, SH2D4A and GDF15 were highly correlated with the clinical traits such as TMB, MSI, RNAss, DNAss, StromalScore, ImmuneScore and ESTIMATEScore (Figure 6A). To better understand the relationship between the risk signature and clinical characteristics, we analyzed the distribution of survival status, WHO classification, risk scores, MGMT, and IDH



Figure 1. Consensus clustering analysis in glioma. (A) Consensus clustering was performed based on the genomic heterogeneity of gliomas. (B) Kaplan-Meier curves displaying prognostic differences between different clusters. (C) The differences in TMB and MSI between clusters. (D) Consensus clustering was performed based on the stemness of gliomas. (E) Kaplan-Meier curves displaying prognostic differences in DNAss and RNAss between clusters. (G) Consensus clustering is performed based on the microenvironment of gliomas. (H) Kaplan-Meier curves displaying prognostic differences in the microenvironment scores between clusters.

status of glioma patients. We identified significant differences in clinical characteristics between patients with low-risk scores and those with high-risk scores. Patients in the low-risk score group exhibited an IDH-Mutant, Methylated phenotype and had a better prognosis, and GBM had a significantly higher risk score than the corresponding LGG subtype (Figure 6B–6E).

The risk score is an independent risk factor for glioma prognosis

We performed univariate Cox regression analysis to investigate the independent prognostic factors for glioma.

And the analysis showed that risk score, age, MGMT promoter status, WHO classification, and IDH status were significantly associated with prognosis (Figure 7A). Multivariate Cox regression analysis showed that risk score and age were significantly associated with prognosis, suggesting that risk score is an independent prognostic factor for glioma (Figure 7B). Following that, we developed a nomogram survival prediction model for glioma patients based on independent prognostic parameters, and plotted calibration curves. The results showed good agreement between the predicted outcome and the 1-, 3- and 5-year overall survival of patients (Figure 7C, 7D).



Figure 2. DEGs screening and enrichment analysis. (**A**) Differentially expressed genes (Fold change >1.5 and P < 0.05) were screened between different clusters distinguished in genomic heterogeneity, (**B**) stemness, and (**C**) tumor microenvironment, respectively. (**D**) Venn diagram showing the overlap of identified differentially expressed genes. (**E**) Enrichment analysis of the differentially expressed genes. (**F**) Protein-Protein Interaction Networks, PPI.

Risk signature is highly associated with tumor immune cells

The above analysis mentioned that DEGs were enriched in immunoregulatory process, and risk signature was highly correlated with the immune microenvironment, so we further analyzed the relationship between risk signature and infiltrating tumor immune cells. Infiltrating analysis showed elevated infiltration of T cell regulatory (Tregs), tumor associated macrophages,



Figure 3. Weighted correlation network analysis. (A, B) Soft-thresholding powers. (C) Clustering of module genes in the TCGA cohort. (D) Module-trait relationships. (E) Scatter plot of correlation between GS and MM. *, P < 0.05; **, P < 0.01; ***, P < 0.001; ****, P < 0.0001.

tumor associated fibroblasts, neutrophils and endothelial cells in patients in the high-risk score group compared to patients in the low-risk score group (Figure 8A–8C). Moreover, the risk signature and risk hub genes RAB42, SH2D4A, and GDF15 were highly positively correlated with the increase of these immunosuppressive tumor infiltrating cells (Figure 8D). Although CD4+ and CD8+ T cells increased in high-risk group, they might be exhausted T cells, characterized by progressive loss of T cell function and ultimately loss of cascade response. Therefore, we analyzed the relationship between risk signature and T cell exhaustion. The

results showed that the risk signature was positively correlated with exhausted T cell signature (Figure 8E, 8F), and the expression of exhausted markers HAVCR2, TIGIT, LAG3, PDCD1, and LAYN was upregulated in high-risk group (Figure 8G).

Risk signature is associated with immunotherapy outcome

In order to analyze the predictive effect of risk signature on the efficacy of immune checkpoint inhibitors (ICI) therapy, we downloaded gene expression profiles and



Figure 4. Identification of risk hub genes related to the overall survival of glioma. (**A**, **B**) LASSO regression analysis of hub genes in the TCGA cohort. (**C**) Multivariate Cox regression analysis of hub genes in the TCGA cohort. (**D**, **E**) Violin plot showing the expression levels of the 3 risk hub genes between LGG and GBM in the TCGA and CGGA datasets, respectively. (**F**, **G**) Kaplan-Meier curves displayed 3 risk hub genes were significantly related to poor prognosis in the TCGA and CGGA datasets, respectively.

clinical data of the IMvigor210 cohort. All samples were divided into high and low-risk groups according to the risk signature. Patients with treatment response [complete remission (CR) or partial remission (PR)] had significantly lower risk scores than those without response [stable (SD) or progressive (PD)] (Figure 9A). By evaluating the distribution of CR/PR and SD/PD in the high-risk and low-risk groups, we found that the low-risk group responded better to ICI treatment than the high-risk group (Figure 9B), and had a significantly better prognosis than the high-risk group (Figure 9C). We then analyzed the prognostic value of risk signature and risk hub genes in pan-cancer (Figure 9D, 9E), and the results showed that the low-risk group had a better prognosis in LIHC\PAAD\MESO\UVM (Figure 9E).

SH2D4A affects the migration and proliferation of glioma cells

To validate the expression of the risk hub genes in glioma, we analyzed the immunohistochemistry pathological specimen data. The results showed that the expression level of SH2D4A was increased in GBM relative to LGG (Figure 10A). However, the function of SH2D4A in glioma has not been reported in the literature so far. To investigate the effect of SH2D4A on glioma cells, we knocked down SH2D4A in U87 MG cells by transfecting specific siRNAs (Figure 10B, 10C). CCK-8 and Edu assays suggested that SH2D4A significantly affected the proliferation of U87 MG cells (Figure 10D–10F). Cell migration assays confirmed that





Figure 6. The relationship between risk score and clinical phenotype. (A) Correlations between 3 risk hub genes and tumor stemness, genomic heterogeneity, and immune microenvironment. (B) Sankey Diagram displayed the distribution of the survival status, IDH status, WHO grade, risk score and MGMT promoter status. (C) Analysis of the risk scores in different IDH status, (D) MGMT promoter status, and (E) WHO grades, respectively. *, P < 0.05; **, P < 0.01; ***, P < 0.001; ****, P < 0.001.

Figure 7. Risk score is an independent prognostic factor for glioma. (A) Univariate Cox regression analyses showed that clinical features such as the risk score, age, MGMT promoter status, WHO grade, and IDH status were significantly correlated with prognosis. (B) Multivariate Cox analysis showed the risk score remained associated with the prognosis. (C) The Nomogram was constructed to predict prognosis in patients at 1-, 3-, and 5 years in the TCGA dataset. (D) The calibration curve of the nomogram.

knockdown of SH2D4A significantly inhibited the migration of U87 MG cells (Figure 10G–10H).

DISCUSSION

Primary central nervous system tumors arise from heterogeneous cells within the central nervous system (CNS) [40]. Although traditional treatments for glioma, such as surgical resection, temozolomide (TMZ), and radiation therapy, have been used, they offer limited progress in combating the tumor's progression. Immunotherapy is a recent breakthrough in cancer treatment and is currently at the forefront of cancer therapy [41]. However, it still requires further

Figure 8. Risk signature is associated with infiltrating immune cells. (A) The level of immune cell infiltration between different risk subgroups evaluated by CIBERSORT, (B) MCPCOUNTER and (C) EPIC algorithm, respectively. (D) Correlation of infiltrating immune cells with risk scores and 3 risk hub genes. (E, F) Correlation of exhausted T cell signature with risk scores and 3 risk hub genes. (G) Violin plot showing the expression levels of the exhausted markers between low- and high-risk groups in the TCGA dataset. *, P < 0.05; **, P < 0.01; ****, P < 0.001; ****, P < 0.001.

improvement for gliomas [16, 17]. Previous studies have shown that there is significant genetic, epigenetic, and microenvironmental heterogeneity in each glioma. Intratumoral heterogeneity has a significant influence on tumor recurrence in gliomas, and this presents various challenges in developing targeted therapies [42]. Thus, improvement of patient prognosis requires a better understanding of the genomic heterogeneity, stemness, and immune microenvironment of gliomas.

Figure 9. Analysis of risk signature in immunotherapy cohort and pan cancer. (A) Violin plot depicting the risk scores between SD/PD and CR/PR groups in the IMVIgor210 cohort. (B) The response rate of the high-risk score group and low-risk score group to immune therapy in the IMVIgor210 cohort. (C) Survival analysis of patients with low-risk scores and high-risk scores in the IMVIgor210 cohort. (D) Survival analysis of 3 risk hub genes in pan cancer. (E) Kaplan-Meier survival curve of patients in high- and low-risk score groups in pan cancer. **, P < 0.01.

To address this issue, we conducted a consensus clustering analysis of gliomas' genomic heterogeneity, stemness, and microenvironment. The results showed that all samples can be classified into subgroups, and the prognostic differences between them were significant (Figure 1). Therefore, identification of driver hub genes affecting these three aspects may help to address treatment failure. Interestingly, genes differentially expressed in different groups in all three consensus clustering analyses were enriched in immunomodulatory interactions (Figure 2E), suggesting that genomic heterogeneity, stemness, and the tumor microenvironment may all contribute to reprogramming the immune status of gliomas, which may have implications for therapy.

Through WGCNA, LASSO and multivariate Cox regression analyses, we identified RAB42, SH2D4A and GDF15 as high-risk hub genes for glioma, and we

employed these genes to create a risk model that showed high sensitivity and specificity in prognosticating patients' 1-, 3-, and 5-year survival (Figure 5). Undoubtedly, the risk model showed an extremely strong correlation with genomic heterogeneity, stemness, and immune microenvironment scores (Figure 6). In addition, we found that the risk signature was closely associated with T cell exhaustion (Figure 8E-8G), which could theoretically be partially reversed by inhibiting the PD-1 pathway. However, we also found that immunosuppressive cells, such as tumor-associated macrophages, tumor-associated fibroblasts, and T cell regulatory (Tregs), were elevated in the high-risk group (Figure 8A–8C). The presence of these immunosuppressive cells may be an important contributor to T-cell exhaustion and to the low response rate to immune checkpoint inhibitor therapy. And these results further corroborate the influence of genomic

heterogeneity, stemness and tumor microenvironment on the immune status of glioma (Figures 2, 8).

Studies have shown that RAB42 can promote the proliferation, migration and invasion of glioma cells [43, 44]. And in addition to these functions, GDF15 can regulate immune infiltration of glioma and promote stem cell-like phenotype [45, 46]. However, the role of SH2D4A in glioma remains unclear. SH2D4A is localized to human chromosome 8p21.3 and encodes SH(2)A. Earlier studies indicate that chromosome regions harboring SH2D4A are frequently deleted in various cancer types, and the deletion or downregulation of this gene is related to poor survival and hepatocarcinogenesis [26, 47]. In addition, SH2D4A was found to promote the oncogenic progression of HCT15 and LoVo CRC cells [16]. Given the conflicting roles of SH2D4A in different tumor types, we validated its effects on glioma cells. Our results suggest that SH2D4A contributes to the migration and proliferation of glioma cells (Figure 10). However, the effects and mechanisms of SH2D4A on glioma stemness and microenvironment needs to be further investigated.

CONCLUSIONS

In this study, we utilized consensus clustering analysis of genome heterogeneity, stemness, and tumor microenvironment to identify hub genes associated with risk in glioma. We then constructed a risk model capable of accurately predicting glioma prognosis and treatment response to immunotherapy. Our findings could pave the way for new strategies to treat glioma.

AUTHOR CONTRIBUTIONS

Yangyang Wu, Meng Mao, and Lin-jian Wang conceived and designed the experiments. Yangyang Wu performed the experiments and drafted the manuscript. Meng Mao and Lin-jian Wang approved the final version of the manuscript.

ACKNOWLEDGMENTS

We would like to thank the researchers for their contributions.

CONFLICTS OF INTEREST

The authors declare no competing financial interests and no conflicts of interest for this work.

FUNDING

This work was supported by the National Natural Science Foundation of China (82101401).

REFERENCES

- Rouse C, Gittleman H, Ostrom QT, Kruchko C, Barnholtz-Sloan JS. Years of potential life lost for brain and CNS tumors relative to other cancers in adults in the United States, 2010. Neuro Oncol. 2016; 18:70–7. <u>https://doi.org/10.1093/neuonc/nov249</u> PMID:26459813
- Zhang N, Zhang L, Qiu B, Meng L, Wang X, Hou BL. Correlation of volume transfer coefficient Ktrans with histopathologic grades of gliomas. J Magn Reson Imaging. 2012; 36:355–63. https://doi.org/10.1002/jmri.23675 PMID:22581762
- Ostrom QT, Bauchet L, Davis FG, Deltour I, Fisher JL, Langer CE, Pekmezci M, Schwartzbaum JA, Turner MC, Walsh KM, Wrensch MR, Barnholtz-Sloan JS. The epidemiology of glioma in adults: a "state of the science" review. Neuro Oncol. 2014; 16:896–913. <u>https://doi.org/10.1093/neuonc/nou087</u> PMID:<u>24842956</u>
- Ostrom QT, Gittleman H, Truitt G, Boscia A, Kruchko C, Barnholtz-Sloan JS. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2011-2015. Neuro Oncol. 2018; 20:iv1–v86. <u>https://doi.org/10.1093/neuonc/noy131</u> PMID:30445539
- van den Bent MJ. Interobserver variation of the histopathological diagnosis in clinical trials on glioma: a clinician's perspective. Acta Neuropathol. 2010; 120:297–304. <u>https://doi.org/10.1007/s00401-010-0725-7</u> PMID:20644945
- Suzuki H, Aoki K, Chiba K, Sato Y, Shiozawa Y, Shiraishi Y, Shimamura T, Niida A, Motomura K, Ohka F, Yamamoto T, Tanahashi K, Ranjit M, et al. Mutational landscape and clonal architecture in grade II and III gliomas. Nat Genet. 2015; 47:458–68. https://doi.org/10.1038/ng.3273 PMID:25848751
- Baumert BG, Hegi ME, van den Bent MJ, von Deimling A, Gorlia T, Hoang-Xuan K, Brandes AA, Kantor G, Taphoorn MJ, Hassel MB, Hartmann C, Ryan G, Capper D, et al. Temozolomide chemotherapy versus radiotherapy in high-risk low-grade glioma (EORTC 22033-26033): a randomised, open-label, phase 3 intergroup study. Lancet Oncol. 2016; 17:1521–32. <u>https://doi.org/10.1016/S1470-2045(16)30313-8</u> PMID:<u>27686946</u>
- Ostrom QT, Cioffi G, Gittleman H, Patil N, Waite K, Kruchko C, Barnholtz-Sloan JS. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2012-2016. Neuro Oncol. 2019 (Suppl 5); 21:v1–100.

https://doi.org/10.1093/neuonc/noz150 PMID:<u>31675094</u>

- Ostrom QT, Gittleman H, Xu J, Kromer C, Wolinsky Y, Kruchko C, Barnholtz-Sloan JS. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2009-2013. Neuro Oncol. 2016; 18:v1–75. <u>https://doi.org/10.1093/neuonc/now207</u> PMID:<u>28475809</u>
- Franceschi E, Tosoni A, Bartolini S, Minichillo S, Mura A, Asioli S, Bartolini D, Gardiman M, Gessi M, Ghimenton C, Giangaspero F, Lanza G, Marucci G, et al. Histopathological grading affects survival in patients with IDH-mutant grade II and grade III diffuse gliomas. Eur J Cancer. 2020; 137:10–7. <u>https://doi.org/10.1016/j.ejca.2020.06.018</u> PMID:<u>32721633</u>
- Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJ, Belanger K, Brandes AA, Marosi C, Bogdahn U, Curschmann J, Janzer RC, Ludwin SK, et al, European Organisation for Research and Treatment of Cancer Brain Tumor and Radiotherapy Groups, and National Cancer Institute of Canada Clinical Trials Group. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N Engl J Med. 2005; 352:987–96.

https://doi.org/10.1056/NEJMoa043330 PMID:<u>15758009</u>

- Molenaar RJ, Radivoyevitch T, Maciejewski JP, van Noorden CJ, Bleeker FE. The driver and passenger effects of isocitrate dehydrogenase 1 and 2 mutations in oncogenesis and survival prolongation. Biochim Biophys Acta. 2014; 1846:326–41. <u>https://doi.org/10.1016/j.bbcan.2014.05.004</u> PMID:<u>24880135</u>
- Wang Z, Bao Z, Yan W, You G, Wang Y, Li X, Zhang W. Isocitrate dehydrogenase 1 (IDH1) mutation-specific microRNA signature predicts favorable prognosis in glioblastoma patients with IDH1 wild type. J Exp Clin Cancer Res. 2013; 32:59. <u>https://doi.org/10.1186/1756-9966-32-59</u> PMID:<u>23988086</u>
- Ballabh P, Braun A, Nedergaard M. The blood-brain barrier: an overview: structure, regulation, and clinical implications. Neurobiol Dis. 2004; 16:1–13. <u>https://doi.org/10.1016/j.nbd.2003.12.016</u> PMID:<u>15207256</u>
- Oberoi RK, Parrish KE, Sio TT, Mittapalli RK, Elmquist WF, Sarkaria JN. Strategies to improve delivery of anticancer drugs across the blood-brain barrier to treat glioblastoma. Neuro Oncol. 2016; 18:27–36. <u>https://doi.org/10.1093/neuonc/nov164</u> PMID:<u>26359209</u>

- 16. Xu S, Tang L, Li X, Fan F, Liu Z. Immunotherapy for glioma: Current management and future application. Cancer Lett. 2020; 476:1–12. <u>https://doi.org/10.1016/j.canlet.2020.02.002</u> PMID:<u>32044356</u>
- Reardon DA, Brandes AA, Omuro A, Mulholland P, Lim M, Wick A, Baehring J, Ahluwalia MS, Roth P, Bähr O, Phuphanich S, Sepulveda JM, De Souza P, et al. Effect of Nivolumab vs Bevacizumab in Patients With Recurrent Glioblastoma: The CheckMate 143 Phase 3 Randomized Clinical Trial. JAMA Oncol. 2020; 6:1003–10.

https://doi.org/10.1001/jamaoncol.2020.1024 PMID:<u>32437507</u>

- Johnson KC, Anderson KJ, Courtois ET, Gujar AD, Barthel FP, Varn FS, Luo D, Seignon M, Yi E, Kim H, Estecio MR, Zhao D, Tang M, et al. Single-cell multimodal glioma analyses identify epigenetic regulators of cellular plasticity and environmental stress response. Nat Genet. 2021; 53:1456–68. <u>https://doi.org/10.1038/s41588-021-00926-8</u> PMID:<u>34594038</u>
- Weller M, van den Bent M, Tonn JC, Stupp R, Preusser M, Cohen-Jonathan-Moyal E, Henriksson R, Le Rhun E, Balana C, Chinot O, Bendszus M, Reijneveld JC, Dhermain F, et al, and European Association for Neuro-Oncology (EANO) Task Force on Gliomas. European Association for Neuro-Oncology (EANO) guideline on the diagnosis and treatment of adult astrocytic and oligodendroglial gliomas. Lancet Oncol. 2017; 18:e315–29.

https://doi.org/10.1016/S1470-2045(17)30194-8 PMID:28483413

 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.

https://doi.org/10.1007/s00401-016-1545-1 PMID:27157931

- Barthel L, Hadamitzky M, Dammann P, Schedlowski M, Sure U, Thakur BK, Hetze S. Glioma: molecular signature and crossroads with tumor microenvironment. Cancer Metastasis Rev. 2022; 41:53–75. <u>https://doi.org/10.1007/s10555-021-09997-9</u> PMID:34687436
- Ribeiro Franco PI, Rodrigues AP, de Menezes LB, Pacheco Miguel M. Tumor microenvironment components: Allies of cancer progression. Pathol Res Pract. 2020; 216:152729. <u>https://doi.org/10.1016/j.prp.2019.152729</u>

https://doi.org/10.1016/j.prp.2019.15272 PMID:<u>31735322</u>

- Poon CC, Sarkar S, Yong VW, Kelly JJ. Glioblastomaassociated microglia and macrophages: targets for therapies to improve prognosis. Brain. 2017; 140:1548–60. <u>https://doi.org/10.1093/brain/aww355</u> PMID:28334886
- 24. Arneth B. Tumor Microenvironment. Medicina (Kaunas). 2019; 56:15. <u>https://doi.org/10.3390/medicina56010015</u> PMID:<u>31906017</u>
- 25. Cole AP, Hoffmeyer E, Chetty SL, Cruz-Cruz J, Hamrick F, Youssef O, Cheshier S, Mitra SS. Microglia in the Brain Tumor Microenvironment. Adv Exp Med Biol. 2020; 1273:197–208. https://doi.org/10.1007/978-3-030-49270-0 11 PMID:33119883
- 26. Quagliata L, Andreozzi M, Kovac M, Tornillo L, Makowska Z, Moretti F, Heim MH, Heinimann K, Piscuoglio S, Terracciano LM. SH2D4A is frequently downregulated in hepatocellular carcinoma and cirrhotic nodules. Eur J Cancer. 2014; 50:731–8. <u>https://doi.org/10.1016/j.ejca.2013.11.018</u> PMID:<u>24315626</u>
- Kim JC, Kim JH, Ha YJ, Kim CW, Tak KH, Yoon YS, Kwon YH, Roh SA, Cho DH, Kim SK, Kim SY, Kim YS. Analysis of genomic pathogenesis according to the revised Bethesda guidelines and additional criteria. J Cancer Res Clin Oncol. 2021; 147:117–28. <u>https://doi.org/10.1007/s00432-020-03391-8</u> PMID:<u>32960359</u>
- Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012; 2:401–4. <u>https://doi.org/10.1158/2159-8290.CD-12-0095</u> PMID:<u>22588877</u>
- 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. https://doi.org/10.1038/sdata.2017.24

PMID:28291232

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

https://doi.org/10.1038/nature25501 PMID:29443960

31. Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, Barretina J, Boehm JS,

Dobson J, Urashima M, Mc Henry KT, Pinchback RM, Ligon AH, et al. The landscape of somatic copy-number alteration across human cancers. Nature. 2010; 463:899–905.

https://doi.org/10.1038/nature08822 PMID:20164920

- Bonneville R, Krook MA, Kautto EA, Miya J, Wing MR, Chen HZ, Reeser JW, Yu L, Roychowdhury S. Landscape of Microsatellite Instability Across 39 Cancer Types. JCO Precis Oncol. 2017; 2017. <u>https://doi.org/10.1200/PO.17.00073</u> PMID:<u>29850653</u>
- Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, Porta-Pardo E, Gao GF, Plaisier CL, Eddy JA, Ziv E, Culhane AC, Paull EO, et al, and Cancer Genome Atlas Research Network. The Immune Landscape of Cancer. Immunity. 2018; 48:812–30.e14. <u>https://doi.org/10.1016/j.immuni.2018.03.023</u> PMID:<u>29628290</u>
- 34. Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, Kamińska B, Huelsken J, Omberg L, Gevaert O, Colaprico A, Czerwińska P, Mazurek S, et al, and Cancer Genome Atlas Research Network. Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. Cell. 2018; 173:338–54.e15. https://doi.org/10.1016/j.cell.2018.03.034

https://doi.org/10.1016/j.cell.2018.03 PMID:29625051

35. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, Carter SL, Getz G, Stemke-Hale K, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013; 4:2612.

https://doi.org/10.1038/ncomms3612 PMID:24113773

- 36. Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, Li B, Liu XS. TIMER2.0 for analysis of tumor-infiltrating immune cells. Nucleic Acids Res. 2020; 48:W509–14. <u>https://doi.org/10.1093/nar/gkaa407</u> PMID:<u>32442275</u>
- Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics. 2010; 26:1572–3. <u>https://doi.org/10.1093/bioinformatics/btq170</u> PMID:20427518
- 38. Shen C, Yu J, Zhang X, Liu CC, Guo YS, Zhu JW, Zhang K, Yu Y, Gao TT, Yang SM, Li H, Zheng B, Huang XY. Strawberry Notch 1 (SBNO1) promotes proliferation of spermatogonial stem cells via the noncanonical Wnt pathway in mice. Asian J Androl. 2019; 21:345–50. <u>https://doi.org/10.4103/aja.aja_65_18</u> PMID:<u>30198493</u>

- 39. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun. 2019; 10:1523. <u>https://doi.org/10.1038/s41467-019-09234-6</u> PMID:30944313
- 40. Lapointe S, Perry A, Butowski NA. Primary brain tumours in adults. Lancet. 2018; 392:432–46. <u>https://doi.org/10.1016/S0140-6736(18)30990-5</u> PMID:<u>30060998</u>
- Nabors LB, Portnow J, Ammirati M, Baehring J, Brem H, Butowski N, Fenstermaker RA, Forsyth P, Hattangadi-Gluth J, Holdhoff M, Howard S, Junck L, Kaley T, et al. NCCN Guidelines Insights: Central Nervous System Cancers, Version 1.2017. J Natl Compr Canc Netw. 2017; 15:1331–45. <u>https://doi.org/10.6004/jnccn.2017.0166</u> PMID:<u>29118226</u>
- 42. Nicholson JG, Fine HA. Diffuse Glioma Heterogeneity and Its Therapeutic Implications. Cancer Discov. 2021; 11:575–90. <u>https://doi.org/10.1158/2159-8290.CD-20-1474</u> PMID:<u>33558264</u>
- 43. Sun L, Yan T, Yang B. The Progression Related Gene *RAB42* Affects the Prognosis of Glioblastoma Patients. Brain Sci. 2022; 12:767. <u>https://doi.org/10.3390/brainsci12060767</u> PMID:<u>35741652</u>

- 44. Liu B, Su Q, Xiao B, Zheng G, Zhang L, Yin J, Wang L, Che F, Heng X. RAB42 Promotes Glioma Pathogenesis via the VEGF Signaling Pathway. Front Oncol. 2021; 11:657029. https://doi.org/10.3389/fonc.2021.657029 PMID:<u>34912698</u> Erratum in: Front Oncol. 2022; 12:1034167. https://doi.org/10.3389/fonc.2022.1034167 PMID:<u>36276111</u>
- Zhu S, Yang N, Guan Y, Wang X, Zang G, Lv X, Deng S, Wang W, Li T, Chen J. GDF15 promotes glioma stem cell-like phenotype via regulation of ERK1/2-c-Fos-LIF signaling. Cell Death Discov. 2021; 7:3. <u>https://doi.org/10.1038/s41420-020-00395-8</u> PMID:<u>33431816</u>
- Hasanpour Segherlou Z, Nouri-Vaskeh M, Noroozi Guilandehi S, Baghbanzadeh A, Zand R, Baradaran B, Zarei M. GDF-15: Diagnostic, prognostic, and therapeutic significance in glioblastoma multiforme. J Cell Physiol. 2021; 236:5564–81. https://doi.org/10.1002/jcp.30289 PMID:33580506
- Ploeger C, Huth T, Sugiyanto RN, Pusch S, Goeppert B, Singer S, Tabti R, Hausser I, Schirmacher P, Désaubry L, Roessler S. Prohibitin, STAT3 and SH2D4A physically and functionally interact in tumor cell mitochondria. Cell Death Dis. 2020; 11:1023. <u>https://doi.org/10.1038/s41419-020-03220-3</u>

PMID:<u>33257655</u>

SUPPLEMENTARY MATERIALS

Supplementary Figure

Supplementary Figure 1. Validation of risk model in GBM and LGG. (A) Kaplan-Meier curves displaying prognostic differences between high- and low-risk groups in GBM and (B) LGG cohorts, respectively.

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 1, 2.

Supplementary Table 1. Clinical information for the TCGA glioma dataset.

Supplementary Table 2. Clinical information for the CGGA glioma dataset.

Supplementary Table 3. siRNAs used in this study.

Name	Sequence
siNC-sense	5'-UUCUCCGAACGUGUCACGUTT-3'
siNC-antisense	5'-ACGUGACACGUUCGGAGAATT-3'
siSH2D4A#1-sense	5'-GAUCCGACGAUGGAAAGAATT-3'
siSH2D4A#1-antisense	5'-UUCUUUCCAUCGUCGGAUCTT-3'
siSH2D4A#2-sense	5'-GCGAACACCAUCUAGAUAATT-3'
siSH2D4A#2-antisense	5'-UUAUCUAGAUGGUGUUCGCTT-3'

Supplementary Table 4. qPCR primer list.

Name	Sequence
SH2D4A-F	5'-CTGGAGCAAGGATCGAGGC-3'
SH2D4A -R	5'-CAGCTCTTACAAATCTGCTTCGT-3'