

A Six-Immune-Related Genes Prognostic Signature for Glioblastoma

Kate Huang

Wenzhou Medical College First Affiliated Hospital: The First Affiliated Hospital of Wenzhou Medical University

Changjun Rao

Wenzhou Medical University First Affiliated Hospital: The First Affiliated Hospital of Wenzhou Medical University

Jianglong Lu

Wenzhou Medical College First Affiliated Hospital: The First Affiliated Hospital of Wenzhou Medical University

Zhangzhang Zhu

Wenzhou Medical College First Affiliated Hospital: The First Affiliated Hospital of Wenzhou Medical University

Chengde Wang

Wenzhou Medical College First Affiliated Hospital: The First Affiliated Hospital of Wenzhou Medical University

Ming Tu

Wenzhou Medical College First Affiliated Hospital: The First Affiliated Hospital of Wenzhou Medical University

Chaodong Sheng

Wenzhou Medical College First Affiliated Hospital: The First Affiliated Hospital of Wenzhou Medical University

Shuizhi Zheng

Wenzhou Medical College First Affiliated Hospital: The First Affiliated Hospital of Wenzhou Medical University

Xiaofang Chen

Wenzhou Medical College First Affiliated Hospital: The First Affiliated Hospital of Wenzhou Medical University

Qun Li (Sarylina@126.com)

Wenzhou Medical College First Affiliated Hospital: The First Affiliated Hospital of Wenzhou Medical University

Keywords: Glioblastoma, Immune-related genes, Nomogram, Prognostic model

Posted Date: November 1st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-984995/v1

License: © (1) This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Abstract

Background: Glioblastoma (GBM) multiforme is a common malignant brain tumor with high mortality. It is urgently necessary to develop a new treatment because traditional approaches have reached a bottleneck.

Purpose: Here we created an immune-related gene (IRGs)-based prognostic signature to comprehensively define the prognosis of glioblastoma (GBM).

Methods: Glioblastoma samples were abstracted from the Chinese Glioma Genome Atlas (CGGA) and the Gene Expression Omnibus (GEO). We retrieved IRGs from the ImmProt data resource. Univariate Cox analysis was adopted to determine the prognostically remarkable IRGs for individual with GBM. The prognostically optimal IRGs were determined via LASSO regression, and predictive model created. Besides, the association of specific factors with the overall survival (OS) of individuals with GBM was explored via multivariate Cox-regression. Lastly, we constructed a predictive nomogram integrating the independent predictive factors to determine the one-, two-, and three-year OS likelihoods of individuals with GBM. Additionally,gene set enrichment analysis(GSEA) and single sample GSEA(ssGSEA) were performed to understand the correlation between the risk score and immune activity.

Results: Overall, 273 IRGs which exhibited differential expression were identified in GBM tumor in contrast with the non-malignant samples. Of these 273 IRGs, only six were remarkably linked to OS of individuals with GBM, which were employed in constructing the predictive signature. The GBM were categorized into either the high-risk GBM group or the low-risk GBM group. There were remarkable differences between the high-risk GBM and the low-risk GBM groups regarding OS. The AUC for predicting one-,two-, and three-year OS in training set was 0.610,0.698 and 0.694.In line with the AUC of validation set was 0.608,0.692 and 0.678.Besides,the results of ssGSEA showed the score of prognostic signature is closely related to immune activity.

Conclusion: Herein, a robust predictive model based on IRGs was created to estimate the diversity of OS likelihoods in GBM patients, as well as aid future clinical research.

1. Introduction

Glioblastoma multiforme (GBM) is the most frequent malignant brain tumor tied to high mortality along with morbidity. GBM in the USA comprises 14.7%, 47.7%, and 56.6% of all primary brain tumors, malignant brain tumors, and gliomas, respectively.(Preusser M 2011)(Stupp R 2005) At present, treating GBM entails maximal surgical resection and subsequent combination of radiation therapy (RT) with chemotherapeutics. Chemotherapy regimens most often include the alkylating agent, temozolomide (TMZ) according to Stupp's protocol, which has been shown to positively impact long-term outcomes. (Zinn PO 2013)(Darefsky AS 2012) Nonetheless, there are some challenges that need redress, including how to entirely resect the whole tumor based on its location in core or in-operable sites of the brain, as well as its proliferation into neighboring healthy brain tissues. Even with aggressive, as well as

comprehensive treatment, cancer relapse cannot be completely avoided. Patients with GBM exhibit dismal prognosis, with a 5.6% of five-year OS and a median OS of 12-15 months.(Hanif F 2017)(Ostrom QT 2018) Considering the dismal survival of individuals with GBM and the low effectiveness of the current treatment regimes, there is a pivotal need of identifying novel treatment targets, as well as alternative therapeutic approaches.

The major function of human immune system constitutes modulating organ homeostasis, to offer protection against infectious pathogens, as well removal of damaged cells. Research evidence shows that adaptive along with innate immunity play indispensable roles in the onset of cancer, which contributes to the progress and treatment of cancer.(S.R. Woo 2015)(S.K. Biswas 2015) For decades, immunotherapy has been a revolutionary anti-cancer therapy. It has shown considerable benefits, such as enhancing survival in numerous cancers, for instance lung cancer, melanoma, as well as breast cancer. (lams WT 2020)(Emens LA 2018) Through manipulation of the immune system, immunotherapy potentially achieves prolonged cancer remission with minimal complications. Current investigations have documented that anti-cancer responses to immunotherapy might take place in the brain, providing appropriate information for the development of new approaches for treating GBM(Sanmamed, MF 2019). Currently, numerous immunotherapy modalities have been proposed and established for GBM. They include immune checkpoint inhibitors, such as antibodies to cytotoxic T lymphocyte antigen 4(CTLA-4), programmed cell death protein 1(PD-1) along with its ligand programmed death-ligand1 (PD-L1), CAR-T, vaccines and oncolytic viruses. (Dougan M 2009) Generally, a combination strategy involving immunotherapies, surgery, and chemoradiotherapy has been opined as a prospective effective approach for treating GBM. Therefore, the current premise purposed to create an immune-related gene (IRGs)-based prognostic signature to comprehensively define the prognosis of glioblastoma (GBM). Differently expressed immune-linked genes in GBM were obtained. Six immune-related genes (CRH, CRLF1, SERPINA3, SSTR2, TNC and TNFRSF19), drastically linked to the OS of GBM, were determined using univariate Cox along with Lasso regression analyses. A risk score, an independent predictive factor, was defined. Consequently, a nomogram model was constructed using the six immune-linked signatures to prognosticate the prognosis of patients with GBM. The signatures were combined with clinical factors consisting of age, gender, IDH mutation status, radiochemotherapy, as well as MGMT promoter methylation status. The six IRGs signature was found remarkablly associated with clinical characteristics and immune cell population in the tumor microenvironment. These data illustrated that the IRGs signature is a reliable predictive assessment tool for determining high-risk GBM individuals.

2. Materials And Methods

2.1 Study population

The RNA-seq data coupled with related clinical information of individuals with GBM were abstracted from the Chinese Glioma Genome Atlas (CGGA, http://www.cgga.org.cn/). Overall, 139 GBM samples and 249 GBM were included in the mRNAseq_325 and mRNAseq_693 datasets, respectively. Besides, Genotype-Tissue Expression(GTEX) RNA-seq data was abstracted from UCSC Xena (http://xena.ucsc.edu/). This

included a total of 207 samples of non-malignant brain tissues. In addition, mRNAseq data of 158 GBM samples and 8 non-malignant brain tissues were abstracted from the Gene Expression Omnibus (GEO) data resource, with an accession number of GSE16011.

2.2 Immune-related genes

The gene list comprising 1793 IRGs was abstracted from the Immport data resource. There were 1098 IRGs with mRNA expression patterns in both the CGGA and Gtex datasets. About 1244 IRGs was contained in GEO. Finally, tumor-related transcription factors (TFs) were abstracted from the Cistrome Cancer data resource (http://cistrome.org/). All expression data was transformed with log₂ (n+1), which were retained and used for the subsequent analyses.

2.3 Differential expression analysis

The R edge package was adopted to perform differential expression analysis(M.D. Robinson 2010). IRGs which were expressed differentially had a P \boxtimes 0.05 along with an absolute log2 fold change (FC) \boxtimes 1.

2.4 Construction of Risk score

To construct the risk score of IRGs, 237 GBM patients in the mRNAseq693 dataset, with survival information (survival time and death status), were taken as a training set. In addition, 137 individuals with GBM in the mRNAseg325 cohort served as the validation set. The immune genes which were remarkably linked to prognosis were determined via Univariate Cox assessment, with P< 0.05 serving as the cut-off. After that, the training cohort was subject to LASSO regression with the R glmet package to explore the most remarkable prognostic genes. (J. Friedman 2010) Shrinkage of the regression coefficient was done on the genes which were remarkable in the univariate analysis via imposition of the penalty proportional to their size. IRGs that corresponded to the smallest partial probability of deviance were retained finally. The Risk score formula was calculated as follows: -0.0475*Express Value of CRH-0.0260*Express Value of CRLF1+0.0640*Express Value of SERPINA3-0.0162*Express Value of SSTR2+0.0456*Express Value of TNC+0.0272*Express Value of TNFRSF19, for computing the risk of inferior survival likelihood for each GBM sample was created as per the expressions of predictive genes in multivariate Cox-regression assessment. Then, the correlations between IRGs in risk score and remarkably different TFs were analyzed using the pearson method. The regulatory networks were then constructed and visualized using Cytoscape software. The pathway and process enrichment analyses were carried out by using Metascape (http://metascape.org). Finally, the relationship of the risk score with different clinical information (age, IDHmutation, gender, MGMT promoter methylation and chr1p19g codeletion) was determined.

2.5 Establishment of a nomogram model

After excluding samples with missing clinical information, the final sample sizes for the training and validation sets were 156 patients and 127 patients, respectively. The 156 GBM patients with complete clinical information, consisting of IDH mutation status, age, chemotherapy, MGMT promoter methylation status, gender, 1p19q codeletion status, and radiotherapy in the training set were used to construct the

nomogram model. The 127 individuals with GBM in the verification set were utilized to verify the efficiency of the nomogram model. Consequently, the package "rms" in R, a multivariate Cox assessment was adopted to assess the independent prognostic indicators, consisting of radiotherapy, chemotherapy, and the risk score. Afterwards, these factors were used to create a predictive model, which was adopted to explore the one-year, two-year, as well as three-year OS for GBM. In addition, calibration curves for the test cohort and the verification set were drawn to determine the nomogram's estimation potential regarding the GBM patients' prognosis.

2.6 Immune infiltration analysis of the Risk score

The R "gsva" package was adopted to conduct single-sample gene set enrichment analysis (ssGSEA). The invasion scores of 16 immune cells along with the activity of 13 immune-linked cascades were computed (Chen X 2020). Afterwards, the association of the level of risk score and immune cell invasion with immune-linked cascades was explored.

2.7 Statistical analyses

The remarkable difference in OS between Risk High and Risk Low was estimated using the log-rank test. We expressed the survival result as a Kaplan-Meier (KM) curve, via the survival along with the survminer packages in R. The Chi-square test coupled with the t-test were adopted to compare clinical categorical variables and continuous variables, in that order. The relationship of the IRGs with immune cells was analyzed using Spearman correlation. A P®0.05 signified statistically significant. All the statistical analyses were implemented in SPSS (IBM SPSS 26.0, SPSS INC) and R version 4.0.3.

3. Results

3.1 Differentially expressed Immune-related genes

A total of 273 different expression IRGs, consisting of 202 upregulated genes along with 71 downregulated genes from GSE16011, were identified. Besides, 291 different expression IRGs, consisting of 152 upregulated genes along with 139 downregulated genes, were obtained from the differentially expressed genes in CGGA and GTEX. Subsequently, the intersection set was obtained and used as the final set of different genes, including 89 upregulated genes and 53 downregulated genes. Finally, 16 tumor-related transcription factors (TFs), including 11 upregulated factors and 5 downregulated factors, were obtained from Cistrome.

3.2 Glioblastoma prognostic signature

The 237 patients in the mRNAseq 693 dataset, CGGA, were taken as the test set. The univariate Cox regression data identified 27 genes among the 273 differently expressed IRGs. Then, the multivariate regression was trained using the features selected by LASSO-COX regression analysis. Finally, six genes (CRH, CRLF1, SERPINA3, SSTR2, TNC, and TNFRSF19) were obtained. On calculating the risk score of every patient with the same formula, the patients were stratified into high-risk GMB and low-risk GBM

groups, per the median risk score. The cutoff risk score was 0.700. Expressions of six genes incorporated in the risk score formula were evidently different between the high-risk GBM and low-risk GBM groups. The Kaplan-Meier data exhibited that the OS was considerably different between the high-risk GBM and low-risk GBM groups in the test set (p=1.938e-02). The median OS was 1.093-year (95%CI: 0.758-1.428) for the low-risk GBM group and 0.945-year (95%CI: 0.698-1.193) for the high-risk GBM group. Besides, time-dependent ROC data exhibited that the risk score could efficiently estimate one-year, two-year and three-year OS likelihood. The results for the calculation of one-year AUC= 0.610,2-year AUC=0.698 and 3-year AUC=0.694 are presented in [Figure 1] The Chinese Glioma Genome Atlas was utilized to validate the risk score among the 137 patients in mRNAseq 325. The data exhibited that the expression of six genes was obviously different. The median OS was 1.521-year (95%CI: 0.709-1.115) for high-risk. The resultant KM survival curves for the samples were statistically remarkable (p=8.813e-03). Finally, the ROC curves showed AUC of 0.608 for 1-year, 0.692 for 2-year and 0.678 for 3-year. These values exhibit the robust potential of the prognostic signature to distinguish prognostic different GBM patients [Figure 2]

The regulatory network between six genes and TFs was constructed to assess how TFs modulate the clinically relevant IRGs. Triangular nodes represent transcription factors, circular nodes represent prognostic IRGs. Based on the results, higher expression of genes, which were represented by the red node, increased the probability of dismal prognosis. In contrast, elevated expression of genes represented by the green node increased the probability of good prognosis. Red lines designate upregulation, while green lines designate downregulation. The pathway and process enrichment analysis and human disease enrichment were performed in Metascape to determine the function of the IRGs and TFs [Figure 3]The result of the analysis of the risk score proved that patients with different IDH mutation status, different MGMT promoter methylation status, and different 1p19q codeletion status were remarkablly different. Younger patients had a lower risk score. Meanwhile, patients with MGMT promoter methylation, IDH mutant and 1p19q codel had a lower risk score. All these differences in molecular features prove the existence of a strong link between the risk score and the molecular tumor subtype [Figure 4].

3.3 Nomogram and independent validation

The clinical information of 283 patients in two groups is shown in [Table 1]. There was a remarkable difference between the training and verification sets. More patients (85.3%) in the test set had radiotherapy (p=0.04) and more patients (87.8%) in the training set had chemotherapy (p=0.005).

A multivariate COX regression was adopted to explore the independence of the risk score in estimating prognosis. The data illustrated that the risk score could be adopted as an independent variable to estimate the prognosis of individuals with GBM (p=0.005). Moreover, radiotherapy and chemotherapy were also independent prognosis factors. The nomograms were constructed to estimate one-year, two-year and three-year survival probabilities using independent factors (radiotherapy, chemotherapy and risk score). The final nomogram model was well calibrated with a concordance index of 0.63 in the training set. Calibration pots were created for the test set for forecasted 1-year, 2-year and 3-year survival and the for the verification set for visual comparison. The red lines designate the estimated survival rates, while

the gray lines actual the ideal survival rates. All three observed lines are closely aligned, exhibiting good calibration in the test set. Then, the data of the validation set are also acceptable in terms of predictive power [Figure 5].

3.4 Gene set enrichment analysis and assessment of immune infiltration

Gene set enrichment analysis (GSEA) was used to determine the enriched features along with the functional differences between the high risk-GBM and low risk -GBM expression groups. The top 10 entries of the Gene Ontology (GO) term along with the Kyoto Encyclopedia of Genes and Genomes (KEGG) terms were selected [Figure 6]. The high-risk GBM group was primarily abundant in complement activation classical cascade, complement activation, and humoral immune response modulated by circulating immunoglobulin and immunoglobulin complex of GO as well as enriched in allograft rejection. Also, mediation was by Asthma Autoimmune thyroid disease and Grafta Autoimmune thyroid disKEGG. To understand the association of the risk score with immune invasion, the enrichment score of different immune cells and the immune-tied roles and cascades based on the ssGSEA algorithm were calculated. Interestingly, there was a remarkable difference in 13 kinds of immune cells between the high risk GBM group. Besides, the 13 immune cells revealed a remarkable correlation with the risk score. Similarly, there was a remarkable difference in the immune-linked functions between the two groups, which also revealed a remarkable relationship with risk score [Figure 7].

This study further sought to understand whether expressions of core immune checkpoints along with the expressions of HLA family were related to risk score groups. The data illustrated that 6 core checkpoints were expressed differently between the high risk GBM group and the low risk GBM group: PD-L1,B7-H3,CD28,CD40,TIM-3 and PD-1,were positively tied to risk score. The data also illustrated that 19 HLA antigens were expressed differently between the high risk GBM group and the low risk GBM group: HLA-DOA, HLA-DPB2, HLA-A, HLA-DPB1, HLA-H, HLA-B, HLA-DMA, HLA-DOB, HLA-DPA1, HLA-DMB, HLA-DQA1, HLA-DQB1, HLA-DRA, HLA-DRB1, HLA-C, HLA-E, HLA-F, HLA-DQA2, and HLA-J,all these different expressed HLA family genes were also positively correlated with risk score.[Figure 8]

3.5 Hub Gene Drug Sensitivity

The drug sensitivity of the hub genes was explored, based on the Gene Set Cancer Analysis (GSCA) portal ^[32], to give support for drug-targeted therapy. The top targeted therapy drugs were screened using GDSC and CDRP for the six IRGs in the signature: ciclopirox, BRD-K51490254, FK866, Vorinostat, MI-2, NVP-BEZ235, tipifarnib-P1, KU-0063794 and MST-312.

4. Discussion

Glioblastoma multiforme (GBM), an aggressive primary malignant brain tumor, is common in adults. Currently, treatment strategies for GBM consist of surgery alone which is adopted for an early-stage disease, whilst adjuvant radio/chemotherapy integrated with surgical resection is adopted for advanced stage. Nevertheless, the outcome of most GBM patients remains poor. For instance, surgical resection does not yield a satisfactory outcome since cancer cells may have developed metastasis.(Cunha ML 2019) In addition, there are still controversies as to whether systemic adjuvant treatment can be administered after surgery considering potential adverse effects or tumor heterogeneity. (Abdul KU 2018) Therefore, it is essential to identify critical biomarkers to estimate GBM prognosis. In the current premise, an immune-related gene (IRGs)-based prognostic signature was explored as a comprehensive method to define the prognosis of glioblastoma (GBM) and provided importance in most analyses.

Much recent investigations have focused on the association of IRG expression with the onset and progress of diverse (Wilson CL 2005) Comprehensive research evidence has documented that IRGs harbor stable capacity to estimate the prognosis of patients, numerous IRGs with robust estimation roles have been (Jia D 2018) Hitherto, some existing nomograms have employed IRGs as predictive factors of individuals with glioma. A recent premise established an immune-linked gene pairs nomogram for estimating the survival of individuals with GBM (Wang S 2021), as well as a risk model on the basis of 20 differentially expressed IRGs was demonstrated to exhibit efficient OS estimation potential for LGG.(Song LR 2020) Nevertheless, numerous prognosis-linked nomograms are limited by variables, for instance sample size coupled with inadequate verification.

After an array of analyses on the basis of the CGGA dataset a prognostic signature consisting of six IRGs (CRH, CRLF1, SERPINA3, SSTR2, TNC, and TNFRSF19) were constructed. Herein, all patients were stratified into low-risk GBM and high-risk GBM cohorts, per the median risk score in two datasets. The OS of the low-risk cohort was better in contrast with that of the high-risk GBM cohort in both the test and verification sets. Besides, the ROC curves showed the efficiency of the prognostic signature for forecasting one-year, two-year and three-year OS of GBM. To further explore its clinical application, the relationship of the risk score with IDH1 mutation, 1p19q codeletion and MGMT promoter status were investigated. In additon, we created nomograms, on the basis of the IRGs signature risk score, radiotherapy and chemotherapy for the training set. In the verification process, we established that the model had robust potential to estimate the prognosis of GBM patients. Besides, we determined the connection of the risk score with invasion of several immune cells, HLA family and core checkpoints, the data exhibited that high risk patients had high immune infiltration and were more likely to benefit from immunotherapy.Finally,some hub genes targeted agents had also been found for the subsequent experiment.

The six IRGs needs further investigation. In our risk score signature, CRH, CRLF1 and SSTR2 were considered beneficial to OS, while SERPINA3, TNC and TNFRSF19 were considered deleterious. The results of single-cell RNA sequencing exhibited that high expression of CRH leads to the downregulation of invasion (r=-0.44), DNA repair (r=-0.37) and Epithelial-Mesenchymal Transitio (EMT) (r=-0.37) (Panossian A 2018) (Cancer SEA 2019). It was previously reported that CRLF1 cross talks with MYH9, triggering PTC cell growth along with metastasis via the ERK/ETV4 cascade, in vitro along with in vivo. (Yu ST 2020) Besides, CRLF1 may contribute to neuroprotection because of its activity in enhancing neuronal cell survival and in modulating neuronal apoptosis. (Niada S 2018) High expression of CRLF1 was negatively associated with invasion (r=-0.44), DNA repair (r=-0.39) and DNA damage (r=-0.38). In

addition,high expressin of SSTR2 was negatively associated with invasion (r=-0.38) and EMT (r=-0.32). All three genes were negative for invasion, considered inhibiting tumor enlargement and recurrence, which were remarkablly beneficial to patients' prognosis.

Montserrat Lara-Velazquez et al. evaluated the impacts of silencing and over-expression (OE) of SERPINA3 on cell migration, viability, and cell proliferation.(Appay R 2018) The authors reported that SERPINA3 KD caused a reduction in cell growth, migration, infiltration, as well as stem cell characteristics, whilst SERPINA3 OE caused elevated cell migration. The data of single-cell RNA sequencing showed that high expression of SERPINA3 leads to the upregulation of hypoxia (r=0.32) and inflammation (r=0.31) ^[24].Tenascin-c (TNC) participates in Vasculogenic mimicry (VM) formation. Vasculogenic mimicry (VM) is the generation of vessel-like structures via highly infiltrative tumor cells. The VM has been regarded one of the numerous mechanisms which account for the failure of anti-angiogenesis treatment in individuals with glioma(Cai HP 2019). The last gene, TNFRSF19, a member of the TNF receptor superfamily is negatively linked to patient survival. It triggers glioblastoma cell migration along with infiltration *in vitro* by Pyk-Rac 1 signaling, JAK1-STAT3 and PDZ-RhoGF.(Liu CJ 2018) Hence,we propose that the six IRGs in our signature might be promising molecular targets for GBM treatment.

Our prognostic index was based on gene expression data provided by CCGA, however some limitations remain. First, our data were all abstracted from publicly accessible datasets. Even though all data were analyzed after normalization, due to differences of microarray with sequencing technology, some systematic errors likely remained.Second,some of the clinical data were missed,decreasing our sample size to a large extent.Moreover,the conclusions made from a series of limited bioinformatics analyses are inadequate and require further validation via comprehensive experiments, as well as clinical studies.

5 Conclusion

In this study, an immunogenomic landscape analysis was performed and an IRG-related prognostic signature for GBM was constructed. The results of this premise provide a more comprehensive understanding of the immune response in the TME and prospective immune treatment targets for clinical practice.

Declarations

Data Availability

The dataset used and/or analyzed during the current study are available from the corresponding author on a reasonable request.

ETHICAL STATEMENT:

Ethics approval and consent to participate

The research didn't involve animal experiments and human specimens, no ethics related issues.

Consent for Publication

Allowed for publication.

Availability of data and materials

The dataset used and/or analyzed during the current study are available from the corresponding author on a reasonable request.

Competing Interests

The authors declare that they have no confict of interest.

Funding

None

Authors' Contributions

(I) Conception and design: Kate Huang, Changjun Rao, (II) Administrative support: Qun Li, Xiaofang Chen, (III) Provision of study materials or patients: Jianglong Lu, Zhangzhang Zhu, Chengde Wang, (IV) Collection and assembly of data: Chaodong Sheng, Shuizhi Zheng, (V) Data analysis and interpretation: Kate Huang Changjun Rao, (VI) Manuscript writing: All authors, (VII) Final approval of manuscript: All authors.

Acknowledgments

None

Authors' information

Katre Huang and Changjun Rao contributed equally to this work.

Department of Pathology, First Affiliated Hospital of Wenzhou Medical University, Wenzhou 325000, China

Kate Huang

Department of Neurosurgery, First Affiliated Hospital of Wenzhou Medical University, Wenzhou 325000, China

Changjun Rao,,Jianglong Lu,Zhangzhang Zhu,Chengde Wang,Ming Tu,Chaodong Sheng,Shuizhi Zheng,Xiaofang Chen&Qun Li

Corresponding author

Correspondence to Qun Li & Xiaofang Chen.

References

- 1. Preusser M(2011), de Ribaupierre S, Wöhrer A, et al. Current concepts and management of glioblastoma. Ann Neurol. 2011,70(1):9–21.
- 2. Stupp R(2005), 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, Gorlia T, Allgeier A, Lacombe D, Cairncross JG, Eisenhauer E, Mirimanoff RO. European Organisation for Research and Treatment of Cancer Brain Tumor and Radiotherapy Groups, National Cancer Institute of Canada Clinical Trials Group. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N Engl J Med. 2005 Mar 10,352(10):987-96.
- Zinn PO(2013), Colen RR, Kasper EM, Burkhardt JK. Extent of resection and radiotherapy in GBM: A 1973 to 2007 surveillance, epidemiology and end results analysis of 21,783 patients. Int J Oncol. 2013, 42(3):929–934.
- Darefsky AS(2012), King JT Jr, Dubrow R. Adult glioblastoma multiforme survival in the temozolomide era: a population-based analysis of surveillance, epidemiology, and end results registries. Cancer. 2012, 118(8):2163–2172.
- Hanif F(2017), Muzaffar K, Perveen K, Malhi SM, SimjeeShU. Glioblastoma multiforme: a review of its epidemiology and pathogenesis through clinical presentation and treatment. Asian Pac J Cancer Prev. 2017, 18(1):3–9.
- Ostrom QT(2018), 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 Oct 1,20.
- 7. S.R. Woo(2015), L. Corrales, T.F. Gajewski. Innate immune recognition of cancer, Annu. Rev. Immunol. 33 (2015) 445–474
- 8. S.K. Biswas, (2015) Metabolic reprogramming of immune cells in cancer progression, Immunity 43
 (3) (2015) 435-449
- 9. lams WT(2020), Porter J, Horn L. Immunotherapeutic approaches for small-cell lung cancer.Nat Rev Clin Oncol(2020) 17:300–12.
- Emens LA(2018). Breast Cancer Immunotherapy: Facts and Hopes.Clin Cancer Res (2018) 24:511– 20.
- 11. Sanmamed MF(2018), Chen L. A Paradigm Shift in Cancer Immunotherapy: From Enhancement to Normalization. Cell. 2018 Oct 4,175(2):313-326.
- 12. Dougan M(2009), Dranoff G. Immune therapy for cancer. Annu Rev Immunol. 2009,27:83-117.

- 13. Robinson MD(2010), McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010 Jan 1,26(1):139-40.
- 14. Friedman J(2010), Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw. 2010,33(1):1-22.
- 15. Chen X(2020), Fan X, Zhao C, Zhao Z, Hu L, Wang D, Wang R, Fang Z. Molecular subtyping of glioblastoma based on immune-related genes for prognosis. Sci Rep. 2020 Sep 23,10(1):15495.
- 16. Cunha ML(2019), Maldaun MV. Metastasis from glioblastoma multiforme: a meta-analysis. Revista da Associação Médica Brasileira. 2019 Apr 11,65:424-33.
- 17. Abdul KU(2018), Houweling M, Svensson F, Narayan RS, Cornelissen FM, Küçükosmanoglu A, Metzakopian E, Watts C, Bailey D, Wurdinger T, Westerman BA. WINDOW consortium: a path towards increased therapy efficacy against glioblastoma. Drug Resistance Updates. 2018 Sep 1,40:17-24.
- 18. Wilson CL(2005), Miller CJ. Simpleaffy: A BioConductor Package for Affymetrix Quality Control and Data Analysis. Bioinformatics (2005) 21:3683–5.
- 19. Jia D(2018), Li S, Li D, Xue H, Yang D, Liu Y. Mining TCGA Database for Genes of Prognostic Value in Glioblastoma Microenvironment. Aging (Albany NY) (2018) 10:592–605.
- 20. Wang S(2021), Xu X. An Immune-Related Gene Pairs Signature for Predicting Survival in Glioblastoma. Front Oncol. 2021 Mar 30,11:564960.
- 21. Song LR(2020), Weng JC, Li CB, Huo XL, Li H, Hao SY, et al. Prognostic and Predictive Value of an Immune Infiltration Signature in Diffuse Lower-Grade Gliomas. JCI Insight (2020) 5(8):e133811.
- Panossian A(2018), Seo EJ, Efferth T. Novel molecular mechanisms for the adaptogenic effects of herbal extracts on isolated brain cells using systems biology. Phytomedicine. 2018 Nov 15,50:257-284.
- 23. Cancer SEA: a cancer single-cell state atlas. Nucleic Acids Res. 2019.
- 24. Yu ST(2020), Sun BH, Ge JN, Shi JL, Zhu MS, Wei ZG, Li TT, Zhang ZC, Chen WS, Lei ST. CRLF1-MYH9 Interaction Regulates Proliferation and Metastasis of Papillary Thyroid Carcinoma Through the ERK/ETV4 Axis. Front Endocrinol (Lausanne). 2020 Aug 25,11:535.
- 25. Niada S(2018), Giannasi C, Gualerzi A, Banfi G, Brini AT. Differential Proteomic Analysis Predicts Appropriate Applications for the Secretome of Adipose-Derived Mesenchymal Stem/Stromal Cells and Dermal Fibroblasts. Stem Cells Int. 2018 Aug 6,2018:7309031.
- 26. Appay R(2018), Tabouret E, Touat M, Carpentier C, Colin C, Ducray F, Idbaih A, Mokhtari K, Uro-Coste E, Dehais C, Figarella-Branger D, POLA network. Somatostatin receptor 2A protein expression characterizes anaplastic oligodendrogliomas with favorable outcome. Acta Neuropathol Commun. 2018 Sep 7,6(1):89.
- 27. Cai HP(2019), Wang J, Xi SY, Ni XR, Chen YS, Yu YJ, Cen ZW, Yu ZH, Chen FR, Guo CC, Zhang J, Ke C, Wang J, Chen ZP. Tenascin-cmediated vasculogenic mimicry formation via regulation of MMP2/MMP9 in glioma. Cell Death Dis. 2019 Nov 21,10(12):879.

28. Liu C. J.(2018), Hu F. F., Xia M. X., Han L., Zhang Q., Guo A. Y. GSCALite: a web server for gene set cancer analysis. Bioinformatics 34 3771–3772. 10.1093/bioinformatics/bty411.

Tables

Table 1: Characters of GBM patients in the test set and the verification set.

Variables	Training set(N=156)	Validation set(N=127)	P-Value
Age(mean(SD))[Range]	47.93(13.631)[14-73]	46.48(12.648)[8-79]	0.919
Gender (N(%))			0.141
Male	86(55.1)	81(63.8)	
Female	70(44.9)	46(36.2)	
IDHmutation (N (%))			0.070
Mutant	31(19.9)	37(29.1)	
Wildtype	125(80.1)	90(70.9)	
MGMTp_methylation(N(%))			0.563
Methylated	84(53.8)	64(50.4)	
Un-methylated	72(46.2)	63(49.6)	
1p19q_codeletion (N(%))			0.598
Non-codel	145(92.9)	120(94.5)	
Codel	11(7.1)	7(5.5)	
Radiotherapy (N (%))			0.04*
Treated	133(85.3)	96(75.6)	
Untreated	23(14.7)	31(24.4)	
Chemotherapy (N (%))			0.005*
Treated	137(87.8)	95(74.8)	
Untreated	19(12.2)	32(25.2)	
Death status (N (%))			0.201
Alive	24(15.4)	13(10.2)	
Dead	132(84.6)	114(89.8)	

Figures



Construction of IRGs prognostic signature.(a) Coefficients of determined characteristics are exhibited via lambda parameter Partial probability deviance relative to log (λ)was generated via LASSO Cox regression approach. (b) Prognostic assessment of the gene signature in the CGGAmRNAseq693 cohort. Top:The dotted line designates the median risk score and stratified the patients into low-risk GBM and high-risk GBM groups.Middle: Survival status of the patients. More dead patients matching to the higher risk

score.Bottom: Heatmap illustrating the expression patterns of the prognostic genes in low risk GBM and high-risk GBM groups.(c)Kaplan-Meier survival assessment of the gene signature. Time-dependent ROC assessment the of the gene signature.



Figure 2

Patients in the verification set (CGGAmRNAseq325) were used to verify the risk score.(a) Top:The dotted line designates the median risk score and stratified the patients into low-risk GBM and high-risk GBM groups.Middle: Survival status of the patients.Bottom: Heatmap of the prognostic genes in low risk GBM and high-risk GBM groups.(b)Kaplan-Meier survival curve and Time-dependent ROC curve of the validation set.



PPI of IRGs and TFs.(a) The results of PPI network analysis of six IRGs and fifteen TFs. (Triangular nodes represent transcription factors, circular nodes represent prognostic IRGs, and the higher the expression of genes represented by the red node, the higher the probability of poor prognosis. Similarly, the higher the expression of the genes represented by the green node, the higher the probability of good prognosis. Red lines designate upregulation, while green lines designate downregulation). (b) Pathway and Process enrichment analysis of PPI networks in Metascape. (c) Diseases Enrichment Analysis of PPI networks in DisGeNET.



Boxplots showing the distribution of risk scores in GBM samples categorized by different factors, consisting of (a)age, (b) gender,(c)IDH mutation status,(d) MGMT promoter methylation status, (E)Chr1p19q codelerion status.



The construction of nomogram.(a) Multivariate COX regression was adopted to select the independent variables, including radiotherapy, chemotherapy, and risk score. (b) The nomogram using radiotherapy,chemotherapy and risk score. For each patient, three lines are drawn upward to verify the points received from the three predictors of the nomogram. The sum of these points situates on the Total Points'axis. Then a line is drawn downward to assess the one-year, two-year, and three-year overall survival of GBM.(c) The calibration curve for the evaluation of the nomogram. The Y-axis designates the actual survival, while the X-axis designates nomogram estimated one-year, two-year and three-year OS of patients in the training set. (d) The predicted one-year, two-year along with three-year OS in the verification set.



GSEA analysis between the high risk GBM group and the low risk GBM group.(a)GO enrichment analysis. (b)KEGG enrichment analysis.



The correlation between riskscore and immune microenvironment (a) Heatmap indicating the relationship between the high-risk GBM and low-risk GBM groups with the expression of immune cells and immunelinked functions. (b) Boxplots illustrating the level of 16 immune cells in the high-risk GBM and low-risk GBM groups(c) Radar plot illustrating the association of the risk score with 16 immune cells. (d) Boxplots illustrating the level of 13 immune-linked functions in the high-risk GBM and low-risk GBM groups.(e) Radar plot illustrating the correlation of the risk scores with the 13 immune-related functions. The p-values were uniformly designated using the following symbols: *p < 0.05, **p < 0.01, ***p < 0.001



Figure 8

The predicting efficacy of immunetherapy in different risk groups.(a) Boxplot illustrating the difference in immune checkpoints between the high risk GBM group and the low risk GBM group.(b)Radar plot illustrating the correlation of the risk scores with the immune checkpoints. (c) Boxplots illustrating the differences between the high-risk GBM group and the low-risk GBM group in the HLA family.(d) Radar plot illustrating the correlation of the risk scores with the expression of HLA family.The p-values were uniformly designated using the following symbols: *p < 0.05, **p < 0.01, ***p < 0.001.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplement.docx
- uniCox.xls