Neurology Asia 2020; 25(4) : 527 – 534

Development and validation of a novel survival model for glioblastoma based on glycolysis-related genes

[1]Jing-Bo Gao, [2]Xue Chen, [1]Min Bao

[1]Department of Neurosurgery & [2]Neurology, Shengjing Hospital of China Medical University, Shenyang, Liaoning, China

Abstract

Objective: To construct a glycolysis-related prognostic model to predict individualized survival in patients with glioblastoma (GBM). *Methods:* Clinical data for patients with GBM, including expression levels of glycolysis-related genes (GRGs), were extracted from The Cancer Genome Atlas. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were then carried out and protein-protein interactions were investigated. Univariate and multivariate Cox regression analyses were performed on the GRGs to identify the best prognosis-related genes. We then established and verified a novel prognostic model, based on the expression of differentially expressed GRGs that were significantly associated with overall survival in GBM patients. *Results:* ALDH3B1, CHPF, FBP1, ISG20 and STC1 were chosen to establish the prognostic risk score model. Patients with high risk scores had significantly poorer overall survival than patients with low risk scores.
Conclusion: The glycolysis-related model has significant value in performing individualized survival predictions for GBM patients and could suggest better treatment options for GBM patients. Our results may help to elucidate GBM pathogenesis and contribute to clinical decision-making and individualized treatment.

Keywords: Glioblastoma; differentially expressed genes; prognostic model; glycolysis

INTRODUCTION

Glioblastoma (GBM), a WHO grade IV malignant diffuse glioma with an unpromising overall survival times of about 1.2 years, is a fairly common malignant brain tumor and accounts for 55% of gliomas.[1] Although there have been many molecular and clinical research studies on GBM, a clear set of prognostic biomarkers and predictors of therapeutic responses for GBM patients has not been identified. The diagnosis of GBM relies mainly on histopathological examination, molecular biomarkers of cancer and imaging studies, and estimates of survival rely mainly on histopathological diagnosis and tumor staging. It is, therefore, crucial to identify reliable and accurate prognostic biomarkers to facilitate optimum treatment strategies.

Tumor cells carry out glycolysis, which allows the cells to produce ATP to maintain oxidation-reduction (redox) balance and to carry out macromolecular biosynthesis required for cell growth, proliferation and migration.[2] Compared with normal cell metabolism, glycolysis produces large quantities of lactic acid, even in the presence of oxygen. This phenomenon is called the "Warburg effect" and has been confirmed in many tumors.[3] Since some glycolytic enzymes have been shown to promote the growth of GBM cells, a glycolysis-related gene (GRG) marker may help to determine the prognosis of patients with GBM. In the current investigation, we explored the potential prognostic value GRG markers by integrating the full set of information about expression of GRG and related gene with clinic results obtained from The Cancer Genome Atlas (TCGA).

METHODS

Data pre-processing and identification of differentially expressed GRGs

RNA sequences and clinical data for 169 GBMs and 5 non-tumor tissues were obtained from TCGA. The ensemble set of 733 GRGs was

Address correspondence to: Min Bao, M.D., Ph.D., Department of Neurosurgery, Shengjing Hospital of China Medical University, No. 36, Sao Hao Street, Shenyang, Liaoning 110004, P. R. China. E-mail: 13699250880@163.com

Date of Submission: 15 July 2020; Date of Acceptance: 11 September 2020

gathered using Gene Set Enrichment Analysis (GESA, https://www.gsea-msigdb.org/gsea/index.jsp) and expression information for the GRGs was extracted. Differentially expressed GRGs (DEGRGs) were screened according to a false discovery rate (FDR) < 0.05 and $|\log_2$ fold change (FC)| ≥1.

GO enrichment and KEGG pathway enrichment analyses

DEGRGs were comprehensively tested by analysis of enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) terms. Biological functions in the GO analysis included biological process (BP), molecular function (MF) and cellular component (CC). All enrichment analyses were performed using R 3.6.1. Both FDR and P values <0.05 were regarded as statistically significant.

Identification of the protein-protein interactions network

We submitted the DEGRGs to the STRING database (https://string-db.org/) to characterize protein-protein interactions (PPIs).[4] After construction of the PPI network, we visualized it using Cytoscape (https://cytoscape.org/). We used the molecular complex detection (MCODE) package to select key complexes and genes in the PPI for which MCODE scores and number of nodes were >5.[5] P values <0.05 were regarded as statistically significant.

Selection and validation of prognosis-related GRGs

We carried out univariate Cox regression analysis to further screen major candidate genes. Gene Expression Profiling Interactive Analysis (GEPIA, http://gepia.cancer-pku.cn/) was then used to select and verify mRNA levels of core GRGs in GBM and normal samples.[6] Multivariate Cox regression analysis was performed on screened core GRGs using survival R package. To verify the effects of these core genes on survival, we used R to analyze RNA sequencing expression data and clinical data from GBM patients in the dataset from TCGA. The Human Protein Atlas (https://www.proteinatlas.org/) dataset was used to explore expression of the core GRGs at the translational level.[7]

Prognostic models

Using the preliminary candidate genes selected as described above, we then developed a model to calculate the risk credit, as well as to evaluate prognosis, using multistep Cox regression analysis. The risk score for each patient was defined as follows:

$$Risk\ score = \sum_{i=1}^{n} Expi\beta i,$$

where Exp represents gene expression level and β represents coefficient value. We classified 160 GBM patients into a training group and a test group. Based on the median survival analysis determined by the risk score, we further divided GBM patients in the training group into high- and low-risk subgroups. We compared the overall survival difference between the two subgroups using the log rank test. We also carried out receiver operating characteristic (ROC) curve analysis, using the survival ROC package in R, to validate the prognostic ability of the model.[8] The test group, which comprised the remaining GBM patients with reliable diagnostic information from TCGA, was used to verify the predictive capability of the model. Lastly, we used the rms package in R to construct a nomogram, with calibration plots to predict the overall likelihood of survival. P <0.05 was regarded as statistically significant.

RESULTS

Differentially expressed GRGs in GBM

The GBM data downloaded from TCGA contained expression data from 5 normal and 169 tumor samples. We used R packages to process the data and find DEGRGs. A total of 733 GRGs were used in the analysis and 278 DEGRGs met the screening criteria for the study (FDR <0.05, $|\log_2$FC| >1.0), including 152 with upregulated expression and 126 with downregulated expression (Figure 1A).

GO enrichment and KEGG pathway analyses of differentially expressed GRGs

To study the mechanisms and functions of the identified GRGs, we divided them into two groups: upregulated genes and downregulated genes. The KEGG enrichment analyses verified that up-regulated DEGRGs were significantly enriched in genes associated with glycolysis/gluconeogenesis, carbon metabolism, and serine, glycine and threonine metabolism. The down-regulated DEGRGs were significantly enriched in genes associated with biosynthesis of amino acids, carbon metabolism, pyruvate metabolism and glycolysis/gluconeogenesis. The

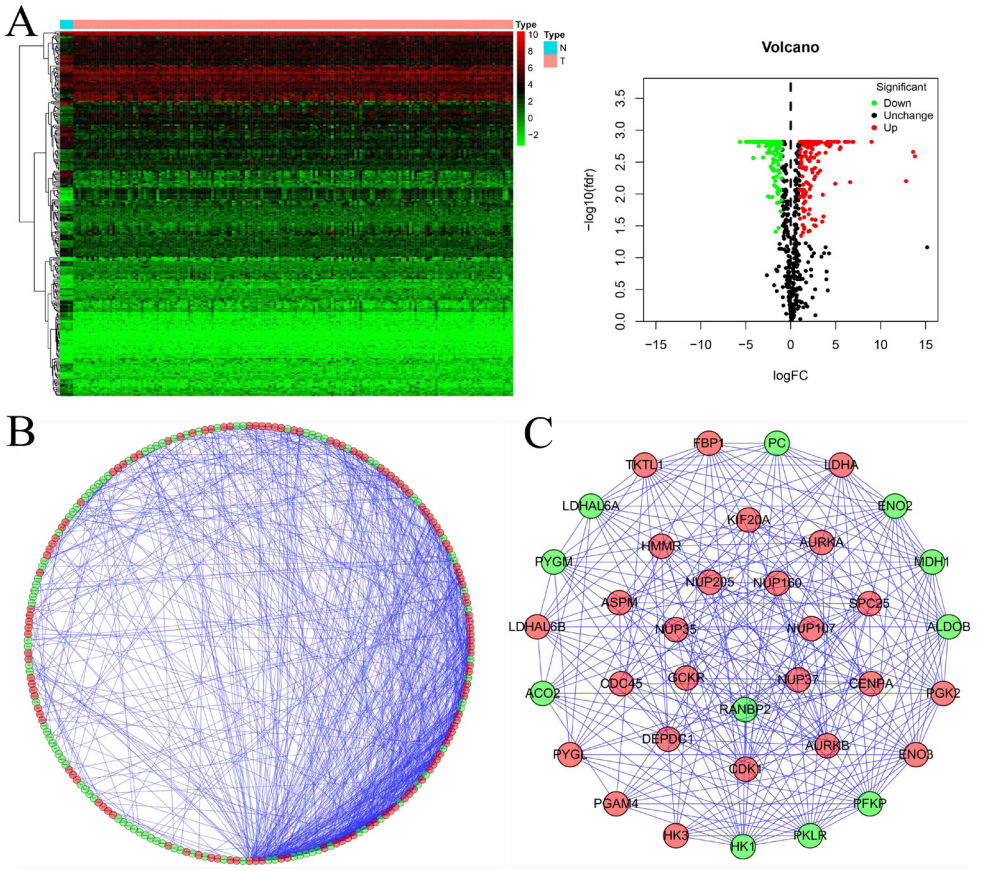

Figure 1. (A) Differentially expressed GRGs in GBM. PPI network and modules analyses. (B) PPI of GRGs; (C) Key module from PPI. Green: down-regulation; Red: up-regulation

GO enrichment analysis illustrated that the up-regulated DEGRGs were significantly enriched in 1) BP: generation of precursor metabolites and energy, coenzyme metabolism, carbohydrate catabolism, nicotinamide nucleotide metabolism, pyridine-containing compound metabolism, oxidoreduction coenzyme metabolism, pyruvate metabolism and glycolysis; 2) CC: cytoplasmic vesicle lumen, vesicle lumen and secretory granule lumen; 3) MF: cofactor binding, carbohydrate binding and carboxylic acid binding (Figure 2A). The down-regulated DEGRGs were significantly enriched in 1) BP: hexose metabolism, glucose metabolism and monosaccharide metabolism; 2) CC: myelin sheath; 3 MF: coenzyme binding (Figure 2B).

PPI networks and hub modules screening

To study the roles of different GRGs in GBM, we used Cytoscape to create a PPI network, which included 218 nodes and 963 edges, using the data obtained from the STRING database (Figure 1B). We used the mode tool to process the co-expression network, and determined possible key modules and the first important module. The network consisted of 36 nodes and 230 edges (Figure 1C). The key module contained an abundance of glycolysis/gluconeogenesis- and carbon metabolism-related genes, based on KEGG enrichment analysis (Count ≥10, P <0.05, FDR <0.05).

Selection and validation of prognosis-related GRGs

In total, we identified 218 key DEGRGs from the PPIs. To explore the prognostic significance of these genes, we carried out univariate Cox regression analysis and identified 16 candidate hub-GRGs that were associated with prognosis (Figure 3A). To verify mRNA expression levels of the five key DEGRGs in GBM, we compared expression levels of the five genes in 163 GBM tissues and 207 normal tissues. We found that the 13 genes were differentially expressed in

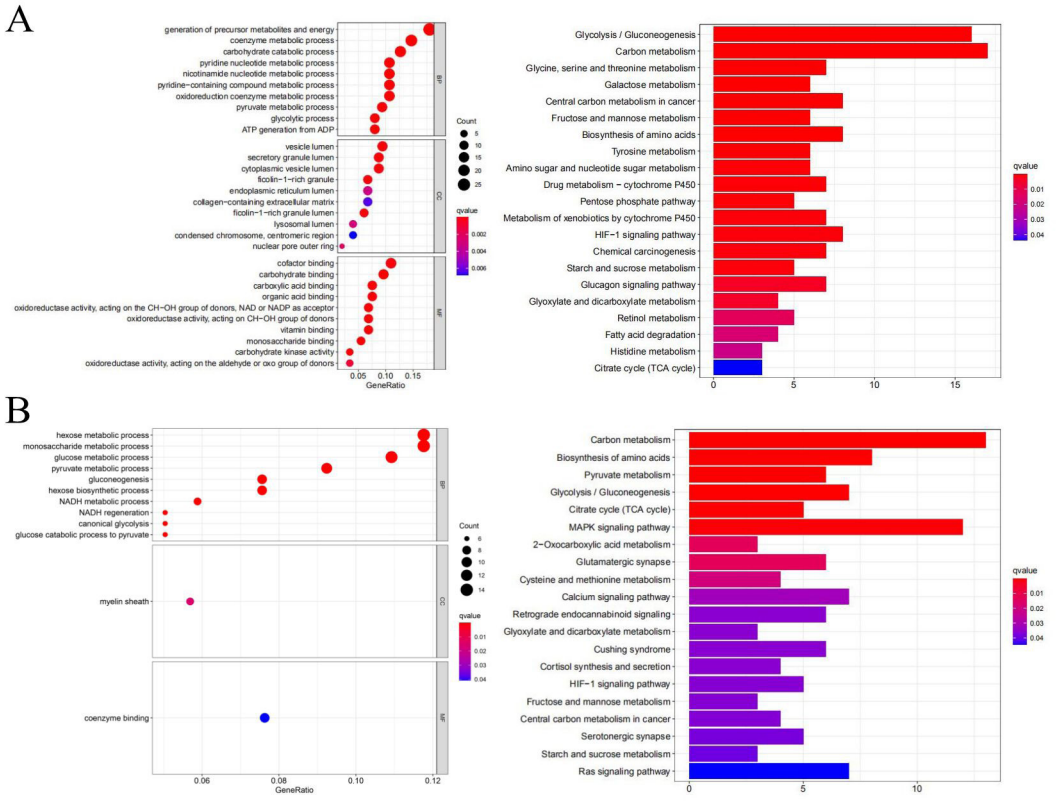

Figure 2. GO and KEGG analyses regarding DEGRGs

GBM tissues and normal tissues (Figure 3B). Multivariate Cox regression analysis was then carried out to determine the effect of these 13 prognostically relevant candidate hub-GRGs on clinical outcomes and patient survival time. A total of five hub-GRGs (aldehyde dehydrogenase 3 family member B1 (ALDH3B1), chondroitin polymerizing factor (CHPF), fructose-bisphosphatase 1 (FBP1), interferon-stimulated gene 20 (ISG20) and stanniocalcin 1 (STC1)) were found to be independent predictors of GBM (Figure 4A). To further test the prognostic value of the five key DEGRGs for GBM, we used the Kaplan-Meier plotter to discover correlations between core DEGRGs and overall survival. The Kaplan-Meier plot identified five core DEGRGs (ALDH3B1, CHPF, FBP1, ISG20, and STC1). The results of the log rank test showed that five

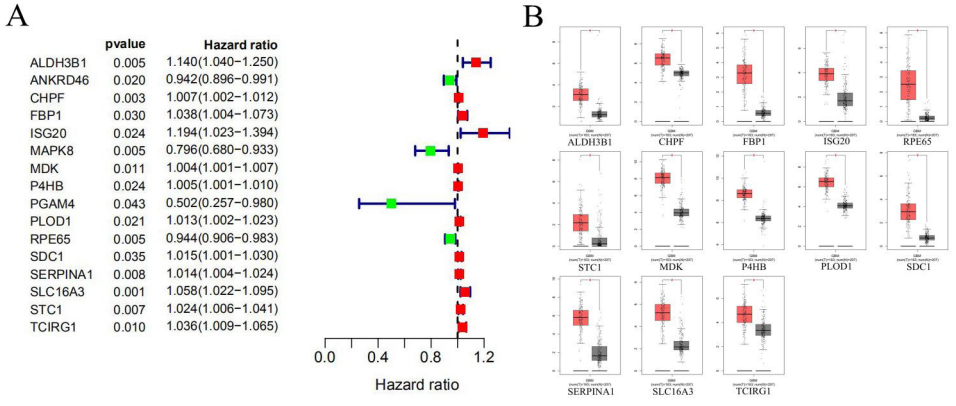

Figure 3. (A) Analysis through univariate Cox regression for core GRG identifications. (B) Validation of core GRG expression by GEPIA

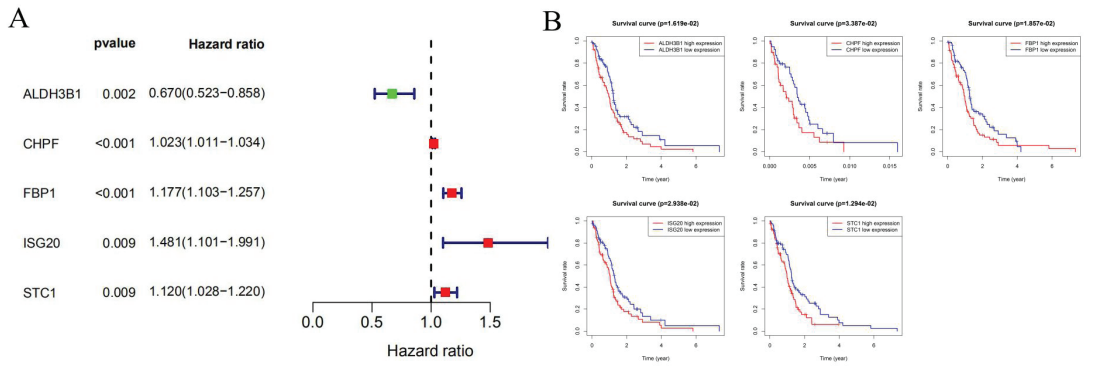

Figure 4. (A) Analysis through multivariate Cox regression for core GRG identifications. (B) Validation of core GRG survival

DEGRGs in GBM patients were significantly related to overall survival ((Figure 4B). We then used immunohistochemistry to measure protein expression of the five DEGRGs in GBM tissues. The four core DEGRGs (ALDH3B1, CHPF, FBP1 and STC1) identified from the Human Protein Atlas database were significantly increased in GBM tissues compared with normal tissues (Figure 5). There was no protein expression information for ISG20 in the Human Protein Atlas database.

Prognosis-related genetic risk score model

We used the five key GRGs identified by multivariate Cox regression analysis to build a prediction model. We calculated a risk score for each patient using the formula: *Risk score* = (–0.400649198*ExpALDH3B1) + (0.022555795*ExpCHPF)+(0.163230086* ExpFBP1) + (0.392493898*ExpISG20) + (0.113233127*ExpSTC1) and then performed survival analysis to assess predictive ability.

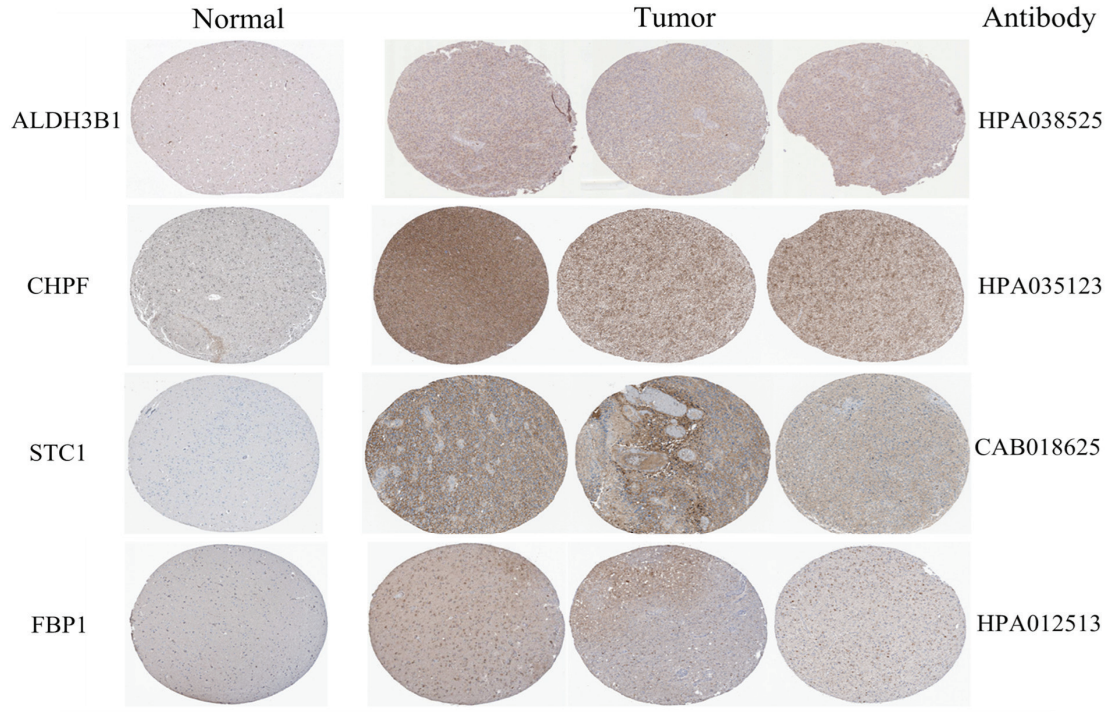

Figure 5. Validation of core GRG expression by the Human Protein Atlas database

We classified 160 GBM patients into a training group and a test group. We classified the 82 GBM patients in the training group into high- and low-risk subgroups based on median risk score. The high-risk group was found to have worse survival than the low-risk group (Figure 6A). The expression heatmaps of the signatures consisted of five GRGs in the high- and low-risk groups, patient survival status and risk score (Figure 6A). The same formula was used for the test group of 78 GBM patients, and the results indicated that the prognostic model has good specificity and sensitivity (Figure 6B). To assess the prognostic ability of the five GRG biomarkers, we performed time-dependent ROC analysis. Using the GRG risk scoring model, we found that the areas under the ROC curves in the training group and the test group were 0.850 and 0.643, respectively (Figure 6).

Nomogram Construction based on core GRGs

To implement a quantitative method for GBM prognosis, we integrated the 5 GRG signatures and established a nomogram. According to multivariate Cox analysis, the points in a nomogram are used to assign points to each variable. We drew a horizontal line to determine the points for each variable, calculated the total points for each patient by counting all variable points, and then normalized them to a distribution within [0, 100]. We drew a line between total points and the prognosis axis to calculate the estimated survival rate of GBM patients in one, three and five years, which may help doctors to make clinical decisions for GBM patients (Figure 7).

DISCUSSION

Based on the GBM data from TCGA, 215 GRGs were found to be differentially expressed between tumor tissues and normal tissues. We analyzed the biological pathways associated with the GRGs and constructed co-expression and PPI networks. Functional pathway enrichment analysis showed that the DEGRGs were highly enriched in glycolysis. It is worth noting that the hub PPI network module analysis showed that GBM is correlated with glycolysis/gluconeogenesis and carbon metabolism. The switch from oxidative phosphorylation to glycolysis is a major response of cancer cells to hypoxia. Hypoxia is a ubiquitous source of cell stress in the microenvironment of the tumor and plays an important role in tissue inflammation and malignancy.[9] Rapidly

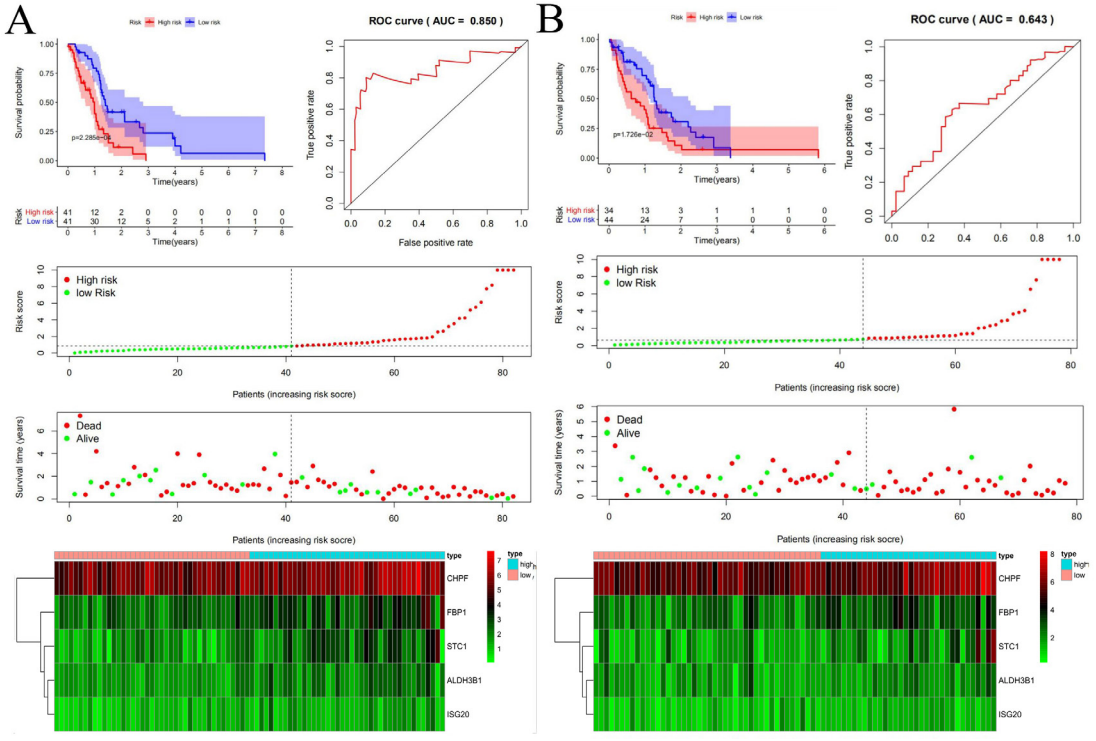

Figure 6. Risk score analysis of five-genes prognostic model in TCGA cohort

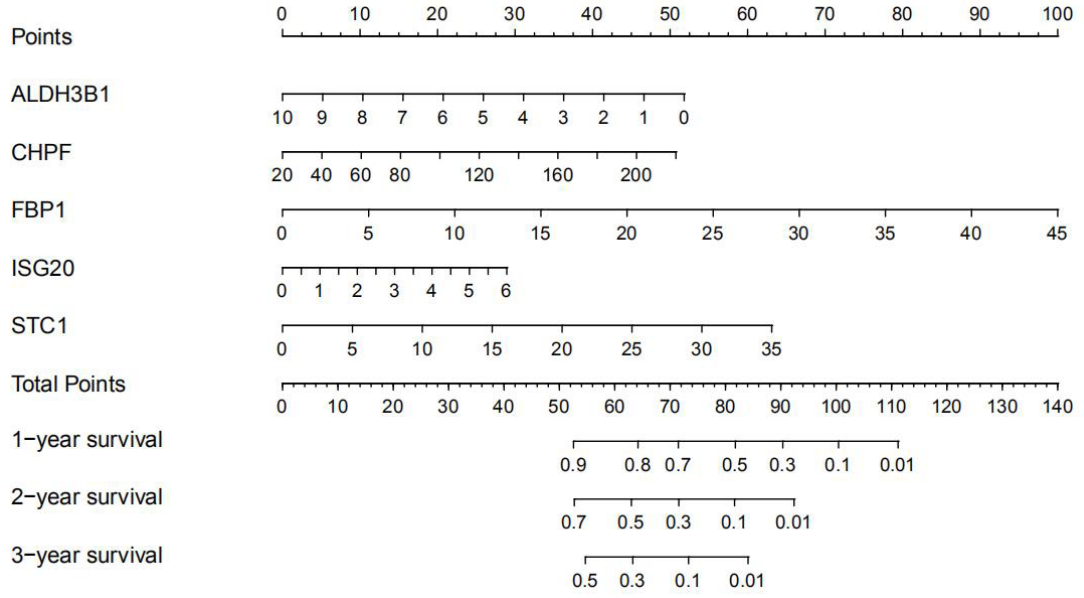

Figure 7. Nomogram for predicting GBM patient OS in identification TCGA cohort

proliferating tumors are exposed to a hypoxic microenvironment because of their high density, high metabolic consumption and interruption of blood flow due to immature angiogenesis.[10] The cellular response to hypoxia promotes highly malignant and metastatic behavior, as well as resistance to chemotherapy. Hypoxia promotes the invasive and mesenchymal characteristics of transformed cells, which plays an important role in the pathogenesis of cancer, including GBM.[11] The spread of GBM is closely associated with proliferation of microvessels in hypoxic brain regions. Glycolysis and carbon cycling are sequential metabolic adaptations of GBM during hypoxia.[12,13] We also carried out univariate and multiple Cox regression analyses on survival, and ROC on GRGs, to further explore their clinical value. Five GRGs were classified as prognosis-related. We built a survival model to estimate GBM prognosis, based on five prognosis-related GRG genes. Our findings may lead to the development of new biomarkers for the diagnosis and prognosis of GBM patients. STC1 is a novel atypical Notch ligand, which is regulated by several microRNAs in GBM, and is an important regulator of GBM stem cells.[14,15] STC1 regulates the migration and invasion of GBM through the TGF-β / Smad4 signaling pathway.[16] ISG20 promotes local tumor immunity and leads to low survival rate for human glioma patients.[17] Changes in FBP1 expression can alter metabolic process that affect the aggressive of GBM.[18] The Warburg effect is the basis of tumorigenesis, proliferation, migration and metastasis, and contributes to a unique tumor microenvironment. Recently, anti-Warburg therapies, which focus not only on the characteristics of the tumor microenvironment or direct inhibition of glycolysis but also on inhibition of glycolysis by increasing gluconeogenesis, have been developed. Screening for key glycolysis genes and regulating the expression of key glycolysis transporters and enzymes to change the tumor microenvironment may provide potential therapeutic strategies.

Using bioinformatics analyses, we identified expression of GRGs in GBM tissues and determined their prognostic value in GBM patients. The GRGs may be involved in the progression, tumorigenesis, invasion and metastasis of GBM. We constructed a GRG coding gene prognostic model, which may provide a useful independent prognostic index for GBM. Compared with a single biomarker, the predictive effect of all participating genes provides a more accurate projection. Further validation of our findings in a laboratory setting, and in an independent cohort of GBM patients with similar characteristics, is essential before using the GRGs as a prognostic tool in the clinic. Glycolysis-related risk signals can effectively predict which patients are at high risk and which patients have a poor prognosis. More accurate individualized treatment strategies could, therefore, be developed for GBM patients with high risk scores, who need more aggressive

treatment strategies and closer follow-up. Because treatment, sampling location, and evolution of heterogeneous tumor cells may all change gene expression, the risk score can only reflect the survival prospects of patients according to gene expression at the time of testing. It may thus be necessary to measure gene expression again for some patients with a long course of disease. To our knowledge, this is the first description of a GRG-related GBM prognosis model. Our results may help to elucidate GBM pathogenesis and contribute to clinical decision-making and individualized treatment.

DISCLOSURE

Financial support: None

Conflict of interest: None

REFERENCES

1. Wesseling P, Capper D, WHO 2016 classification of gliomas. *Neuropathol Appl Neurobiol* 2018; 44: 139-50.
2. Akins NS, Nielson TC, Le HV. Inhibition of glycolysis and glutaminolysis: An emerging drug discovery approach to combat cancer. *Curr Top Med Chem* 2018;18(6):494-504.
3. Spencer NY, Stanton RC. The Warburg effect, lactate, and nearly a century of trying to cure cancer. *Semin Nephrol* 2019;39(4):380-93.
4. Szklarczyk D, Gable AL, Lyon D, *et al*. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2019;47(D1):D607-D613.
5. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 2003;4:2.
6. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res* 2017;45(W1):W98-W102.
7. Thul PJ, Åkesson L, Wiking M, *et al*. A subcellular map of the human proteome. *Science* 2017;356(6340):eaal3321.
8. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics* 2000;56(2):337-44.
9. Vaupel P, Mayer A. Hypoxia in cancer: significance and impact on clinical outcome. *Cancer Metastasis Rev* 2007;26(2):225-39.
10. Marie-Egyptienne DT, Lohse I, Hill RP. Cancer stem cells, the epithelial to mesenchymal transition (EMT) and radioresistance: potential role of hypoxia. *Cancer Lett* 2013;341(1):63-72.
11. Ortensi B, Setti M, Osti D, Pelicci G. Cancer stem cell contribution to glioblastoma invasiveness. *Stem Cell Res Ther* 2013;4(1):18.
12. Zhang K, Xu P, Sowers JL, *et al*. Proteome analysis of hypoxic glioblastoma cells reveals sequential metabolic adaptation of one-carbon metabolic pathways. *Mol Cell Proteomics* 2017;16(11):1906-21.
13. Rosenzweig A, Blenis J, Gomes AP. Beyond the Warburg effect: How do cancer cells regulate one-carbon metabolism?. *Front Cell Dev Biol* 2018;6:90.
14. Li Y, He ZC, Zhang XN, *et al*. Stanniocalcin-1 augments stem-like traits of glioblastoma cells through binding and activating NOTCH1. *Cancer Lett* 2018;416:66-74.
15. Sakata J, Sasayama T, Tanaka K, *et al*. MicroRNA regulating stanniocalcin-1 is a metastasis and dissemination promoting factor in glioblastoma. *J Neurooncol* 2019;142(2):241-51.
16. Xiong Y, Wang Q. STC1 regulates glioblastoma migration and invasion via the TGF-β/SMAD4 signaling pathway. *Mol Med Rep* 2019;20(4):3055-64.
17. Gao M, Lin Y, Liu X, *et al*. ISG20 promotes local tumor immunity and contributes to poor survival in human glioma. *Oncoimmunology* 2018;8(2):e1534038.
18. Son B, Lee S, Kim H, *et al*. Decreased FBP1 expression rewires metabolic processes affecting aggressiveness of glioblastoma. *Oncogene* 2020;39(1):36-49.