Identification of Biomarkers Related to Prognosis of Glioblastoma and Their Correlation with Immune Infiltration Cells

Wencong Ding

University of South China Affiliated Nanhua Hospital

Guoqiang Jiang

University of South China Affiliated Nanhua Hospital

Songkai Long

University of South China Affiliated Nanhua Hospital

Yongshi Liao

University of South China Affiliated Nanhua Hospital

Jia Liu (✉ liujiajzc@mail.sdu.edu.cn)

University of South China Affiliated Nanhua Hospital

Research

Keywords: Glioblastoma, Bioinformatics, Prognosis, Biomarkers, Immune infiltration

Posted Date: September 15th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-850455/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Glioblastoma (GBM) is the most malignant of all known intracranial tumors, meanwhile most patients have a poor prognosis. In order to improve the poor prognosis of GBM patients as much as possible, it is specifically significant to identify biomarkers related to the gene diagnosis and gene therapy. Herein, a total of 343 GBM specimens and 259 non-tumor specimens were collected from four Gene Expression Omnibus (GEO) datasets and TCGA database, and then analyzed the differentially expressed genes (DEGs) from the above data. Through Venn diagram analysis, 54 common upregulated DEGs and 22 common downregulated DEGs were triumphantly recognized. On account of the degree of formation communication in Protein-protein interaction network (PPIN), the 10 upregulated central genes had been ranked, incorporating LOX, IGFBP3, CD44, TIMP1, FN1, VEGFA, POSTN, COL1A1, COL1A2 and COL3A1. By combining the expression levels and the clinical features of GBM, four hub genes (TIMP1, FN1, POSTN and LOX) were significantly up-regulated and related to poor prognosis. Meanwhile, univariate and multivariate Cox regression analysis results suggested that TIMP1 could be one of the independent prognostic factors for GBM patients. Subsequently, TIMP1 was particularly correlated with the immune marker of macrophage M1, macrophage M2, Neutrophils, tumor associated macrophage and Tregs. In conclusion, these findings investigated that TIMP1 might be a new biomarker to determine prognosis and immune infiltration of GBM patients.

Introduction

Glioma is a malignant intracranial tumor caused by abnormal proliferation of glial cells^[1]. Glioblastoma or glioblastoma multiforme is classified as WHO IV^[2]. GBM is the most common primary intracranial tumor, accounting for 54% of all gliomas. Malignant glioblastoma is characterized by rapid proliferation, rapid infiltration and late diagnosis^[3]. Although many analyses have clarified the potential molecular mechanism of glioma, there is no comprehensive and effective treatment plan, and the long-term survival rate of glioma patients cannot be effectively improved^[4]. In order to improve the prognosis of patients with GBM, the objective of this research is to find new diagnostic and prognostic biomarkers and possible therapeutic targets of GBM, further guide clinical treatment and provide new ideas for improving the outcome of GBM patients.

Bioinformatics is a subject that has developed rapidly in recent years by relying on big data processing and it can extract valuable information for clinical research from massive data generated by multi-source experiments^[5]. Chen et al performed an integrated bioinformatics analysis of the pancreatic cancer data set in TCGA, and established a risk scoring model consisting of DEGs, which provided a more effective method for forecasting survival than the AJCC phase^[6]. Wang et al. used a variety of bioinformatics tools to found that the increased PITX1 expression is correlated with worse relapse-free survival (RFS), disease-specific survival (DSS) and overall survival (OS) in breast cancer patients. Therefore, it may provide some base for oriented drug treatment related to breast cancer^[7]. Although bioinformatics analysis has been widely used in solid tumors, there is little research on bioinformatics analysis of GBM. Therefore, in-depth

studies on the composition, functional changes and immune effects of GBM mutated genes are of positive consequence for forecasting the prognosis of GBM patients and for the exploitation of potential targeted treatment.

Tumor Microenvironment (TME) is the cell environment that tumor cells depend on for survival and growth^[8]. TME consists of peripheral blood vessels, extracellular matrix (ECM), non-tumor cells, and a variety of biological factors that are influenced by malignant and untransformed cells within the tumor^[9]. A number of cell types have been identified in TME using different cell-specific markers, including many different types of cells, such as stromal cells, fibroblasts, and various immune cells. The growth and survival of malignant tumor cells are promoted by growth factors secreted by non-tumor cells in TME. At the same time, these growth factors also act as attractors to stimulate a variety of cells to migrate to TME^[10]. ECM is another key component of TME in addition to special types of cells.^[11] ECM not only participates in the supporting framework of TME cells, but also is an abundant pivotal growth factor^[12]. ECM plays a significant role in the genesis and development of tumors. In the late stage of tumor progression, ECM commonly gets out of control and becomes disorganized. Unnatural ECM can also lead to disordered behavior of TME cells and promote neovascularization and inflammatory response.^[13] TME is distinguished from other normal tissues by these unique properties, and there is new evidence that microenvironment-mediated external stimulation plays a pivotal role in the survival of tumor cells^[14]. Disrupting the tumor microenvironment that surrounds and infiltrates the tumor may provide a proven treatment modality for malignancy^[15]. Therefore, we conducted bioinformatics studies using the Gene Expression Profile Interaction Analysis 2 (GEPIA2) database to analyze the prognostic significance of hub genes. The correlation between core genes expression and GBM immune cell infiltration was detected using Tumor Immune Estimation Resource (TIMER) database.

Results

Identification of DEGs

The DEGs in the GSE4290, GSE7696, GSE13276, GSE29796 and TCGA datasets were identified by GEO2R. Add up to 1756 DEGs were recognized in the GSE4290, add up to 938 DEGs were identified in the GSE7696, add up to 546 DEGs were recognized in the GSE13276, add up to 3873 DEGs were recognized in the GSE29796, and add up to 2283 DEGs were recognized in TCGA(GBM). Through visual hierarchical clustering analysis, the volcano map of these DEGs is clearly shown (Figure 1A), where the scarlet and blue dots respectively represent upregulated and downregulated genes. Then, the Venn plot of the co-expression genes containing overlapping DEGs in the GSE4290, GSE7696, GSE13276, GSE29796 and TCGA datasets were constructed (Figure 1B, C). These 54 upregulated DEGs and 22 downregulated DEGs were plotted in Table 2. Add up to 54 overlapping upregulated genes were acquired as core genes for further analysis.

GO and KEGG Enrichment Analyses

We conducted GO and KEGG enrichment analysis of 54 upregulated DEGs and 22 downregulated DEGs to systematically understand the biological role of these DEGs. Figure 2 respectively lists the top enriched GO terms and Figure 3A lists the top 30 KEGG pathways.

GO BP displayed that 54 upregulated DEGs were obviously enriched in the lectin pathway of extracellular matrix organization, extracellular structure organization, cell-substrate adhesion and urogenital system development. GO CC analysis shown that, the most enriched terms were collagen-containing extracellular matrix, basement membrane, endoplasmic reticulum lumen and complex of collagen trimers. The top four significantly enriched MF terms included extracellular matrix structural constituent, extracellular matrix structural constituent conferring tensile strength, platelet-derived growth factor binding and collagen binding (Figure 2A).

GO BP displayed that 22 downregulated DEGs were obviously enriched in the protein localization to axon, regulation of SNARE complex assembly, response to magnesium ion and SNARE complex assembly. GO CC analysis shown that, the top obviously enriched terms were main axon, juxtaparanode region of axon, voltage-gated potassium channel complex and potassium channel complex. The top four significantly enriched MF terms included structural constituent of myelin sheath, magnesium ion binding, nucleoside monophosphate kinase activity and phosphotransferase activity (Figure 2B).

In addition, the four obvious enrichment pathways of these 54 upregulated DEGs attracted our attention were ECM-receptor interaction, Focal adhesion, Pathways in cancer, and the p53 signaling pathway (Figure 3A). Pathway diagram shows one of the top enrichment pathways, ECM-receptor interaction (Figure 3B).

Module screening from the PPI network

The PPI pairs in 54 upregulated DEGs were determined by using the STRING. The Cytoscape identified the top 10 central genes for connectivity of the degree score (Table 3). Cytoscape identifies the most tightly connected modules (Figure 4B). KEGG analysis revealed that the 10 hub genes were significantly enriched in the pathways of ECM-receptor interaction, Focal adhesion, Amoebiasis, Protein digestion and absorption, Bladder cancer, Pathways in cancer, mTOR signaling pathway, Shigellosis and the p53 signaling pathway (Figure 4C).

Validation of mRNA Expression in GBM.

The results of GEPIA database displayed that the mRNA expression levels of 10 genes in GBM tissues were expressively higher than that in non-malignant cerebral cortex specimens ($P < 0.01$) (Figure 5A). Notably, the protein levels of CD44, COL1A1, COL1A2, COL3A1, FN1, POSTN, TIMP1 and VEGFA were not voiced in normal cerebral cortex tissues, nevertheless, moderately and highly expressed of these genes were took stock of GBM tissues (Figure 5B). As a whole, the consequences of this study manifested that the hub genes were overexpressed at both transcriptional and translational expression levels in GBM patients.

Changes in Hub gene frequency and prognostic value in GBM.

The CBioPortal database was used to assess the frequency of genetic changes in these selected central genes from GBM. Approximately 43.45% of GBM clinical patients displayed vital changes in these central genes (Figure 6A). The mRNA alteration was the noticeably vital factors for the altered hub genes in 36 patients (24.83%) of GBM. The consequences showed that the percentage variation in the mRNA expression of LOX, IGFBP3, CD44, TIMP1, FN1, VEGFA, POSTN, COL1A1, COL1A2 and COL3A1 in GBM were 2.8, 18, 3, 4, 3, 8, 5, 10, 16 and 5%, separately (Figure 6B).

Kaplan-Meier diagrams were used to contradistinguish DSS and progression free survival (PFS) in GBM patients with or without changes in the mRNA expression levels of central genes. As uncovered in Figure 6C, GBM sufferers with changed central gene expression represented notably worse DSS compared to those with unchanged central gene expression ($P=0.0117$). Similarly, GBM sufferers with changed hub gene expression showed observably poor PFS ($P=0.0251$) confronted with those with unchanged central gene expression (Figure 6D).

Survival Curve of Hub Genes

Corresponding AUCs for these hub genes were also obtained, including CD44 (AUC=0.981), COL1A1 (AUC=0.970), COL1A2 (AUC=0.993), COL3A1 (AUC=0.996), FN1 (AUC=0.997), IGFBP3 (AUC=0.972), LOX (AUC=0.975), POSTN (AUC =0.961), TIMP1 (AUC=0.981), VEGFA (AUC=0.924) (all $p < 0.0001$) (Figure 7). From the above results, it can be concluded that these variables have high accuracy in predicting the outcomes of non-tumor patients and GBM patients.

Survival Analysis and Cox Regression Analysis in GBM

OS and DFS analyses of the central genes were farther conducted by GEPIA database. As shown in Figure 8A, high expressions of TIMP1, FN1, LOX, and POSTN in GBM patients were significantly correlated with worse OS. Adverse DFS were also significantly scanned in GBM patients with elevated TIMP1, FN1, LOX, and POSTN expression levels (Figure 8B). The results of Log-rank P test showed that only the OS and DFS of TIMP1, FN1, LOX and POSTN had statistical significance simultaneously.

Furthermore, univariate cox regression analysis revealed that FN1 was remarkably correlated with the OS (HR 1.66401, 95% CI=1.54102,1.79681, $p<0.0001$), LOX was remarkably correlated with the OS (HR 1.66184, 95% CI=1.55921,1.77123, $p<0.0001$), POSTN was observably correlated with the OS (HR 1.37421, 95% CI=1.32001,1.43063, $p<0.0001$) and TIMP1 was observably correlated with the OS (HR 1.56064, 95% CI=1.48016,1.64549, $p<0.0001$). Moreover, multivariate cox regression analysis revealed that TIMP1 was isolated risk factor for OS (HR 1.44208, 95% CI=1.16537,1.7845, $p= 0.00076$). The results are summarized in Figure 8C.

Immune Infiltration Analysis of GBM

The degree of immune cell infiltration in tumor tissue is an isolated prognostic factor for different tumor survival. Therefore, we focused on analyzing the association between TIMP1 expression and infiltrating immune cells in GBM. TIMER2.0 was revealed that the expression of TIMP1 remarkably associated with the infiltration of CD4+ T cell ($\text{Rho}=0.335$, $P=6.28\text{e-}05$), B cell ($\text{Rho}=0.408$, $P=7.63\text{e-}07$), T cell regulatory (Tregs) ($\text{Rho}=0.514$, $P=1.29\text{e-}10$), Neutrophils ($\text{Rho}=0.312$, $P=2.10\text{e-}04$), Cancer associated fibroblast ($\text{Rho}=0.653$, $P=5.46\text{e-}18$), Macrophage M1 ($\text{Rho}=0.492$, $P=1.04\text{e-}09$) and Myeloid dendritic cell ($\text{Rho}=0.338$, $P=5.40\text{e-}05$) in GBM. By comparison, it negatively correlated with cancer purity ($\text{Rho}=-0.41$, $P=5.81\text{e-}07$) (Figure 9A).

To further explore the connection with GBM and the level of immune cell infiltration, we used TIMER2.0 to evaluate the connection with TIMP1 expression of various immune cells and immune marker genes in GBM. As the results revealed that, the expression level of TIMP1 was remarkably positively associated with most immune cell markers in GBM. Whereas, the result displayed that TAM marker (CCL2), Macrophage M1 marker (PTGS2), Macrophage M2 markers (CD163, VSIG4, MS4A4A), Neutrophil marker (ITGAM), Treg marker (TGFB1) were observably positively associated with TIMP1 expression level in GBM (Table 4). This suggests that TIMP1 plays a dynamic role in the glioma immune microenvironment. Finally, higher Tregs infiltration was related with worse prognosis for TIMP1 in GBM (Figure 9B, $\text{HR}=2.05$, $P=0.0395$). Equally, higher Cancer associated fibroblast and Mast cell infiltration also associated with poor outcome in GBM (Figure 9C, D, $\text{HR}=1.81$, $P=0.0497$, $\text{HR}=1.98$, $P=0.0341$).

Discussion

GBM is the most invasive and worst prognosis of high-grade gliomas (WHOIII and WHOIV gliomas). Although surgical resection, radiotherapy and chemotherapy have achieved tremendous advance in modern medicine, the median survival time of patients after initial diagnosis is only 12–15 months. difficult radical resection of gliomas and resistance to radiotherapy and chemotherapy are the main causes of recurrence and treatment failure^[2, 3]. As a result of, early diagnoses would obviously enhance the clinical prognosis of GBM. Bioinformatics analysis can quickly and accurately identify biomarkers related to GBM development, and has important research value in the study of diagnostic markers and prognostic genes.

Bioinformatics is a subject that has developed rapidly in recent years by relying on big data processing and it can extract valuable information for clinical research from massive data generated by multi-source experiments^[5, 16]. Chen et al performed a extensive bioinformatics analysis of the pancreatic cancer data set in TCGA, and established a risk scoring model consisting of DEGs, which provided a better approach for forecasting survival than the AJCC phase^[6]. Using a variety of bioinformatics tools, Wang et al. found that increased PITX1 expression is correlated with worse RFS, DSS and OS in breast cancer sufferers. Meanwhile, PITX1 was positively correlated with metastatic RFS and DFS. Therefore, it may provide some basis for targeted drug therapy related to breast cancer^[7]. Zhou et al used software R to normalize the data of native GBM samples and nontumorous samples in the databases. The data were then analyzed jointly. A total of 7 genes were discriminated as great prognostic biomarkers. The discriminated DEGS

were used to the grasping of the molecular mechanisms of GBM, and those DEGs hold promise as prognostic biomarkers for future transactions of GBM sufferers^[17]. Although bioinformatics analysis has been widely used in solid tumors, there is little research on bioinformatics analysis of GBM. In this study, 5 gene expression profiles of GBM were retrieved from GEO and TCGA databases, including 343 glioblastoma specimens and 259 normal specimens. 54 common upregulated DEGs expression profiles and 22 common downregulated DEGs expression profiles were favorably identified, severally. The 10 central genes were selected according to the connectivity of PPIN, incorporating LOX, IGFBP3, CD44, TIMP1, FN1, VEGFA, POSTN, COL1A1, COL1A2 and COL3A1. Subsequently, we performed mRNA expression verification, gene change frequency, survival analysis, COX analysis and tumor immune infiltration analysis of these 10 hub genes. Finally, four hub genes had been screened out including TIMP1, FN1, POSTN and LOX. These identified central genes may play a crucial role in GBM.

FN1 encodes fibronectin, a glycoprotein that exists in the form of dimers or polymers on the cell surface and in the extracellular matrix.^[18] POSTN encodes a protein that displays as a significant role in tissue development and regeneration. These proteins are secreted to the extracellular matrix for tissue repair, such as granulation hyperplasia, mucosal repair and ventricular remodeling.^[19] LOX is a protein-coding gene that mainly encodes the lysyl-oxidase family and leads to a variety of transcriptional variants, one of the precoding proteins was hydrolyzed to produce a regulatory propeptide and a ripening enzyme^[20].

TIMP1, also known as Fibroblast Collagenase Inhibitor, is a Protein Coding gene^[21]. This gene was primarily known as an endogenous inhibitor of matrix metalloproteinases (MMPs)^[22]. Among its related pathways are HIF-1 signaling pathway, IL6-mediated signaling events and ECM organization. This gene family encodes proteins that are inhibitors of MMPs involved in extracellular matrix degradation^[23]. With the exception of its suppressive role against most of the known MMPs, as a cytokine and a key regulator of ECM degradation, TIMP1 has a variety of functions related to tumor microenvironment and progression^[24], Tumor Cells Proliferation^[25], anti-apoptotic activity in cancer^[26–28]. Intriguingly, MMPs have been found to be synthesized mainly by adjacent and intervening stromal cells—as same as TIMP1 is secreted in the tumor microenvironment^[29, 30]. Similarly, we discovered that TIMP1 was negatively in connection with tumor purity through the TIMER2.0 database, suggesting that TIMP1 was mainly derived from stromal cells rather than GBM cells. A large number of research results suggest that TIMP1 is often highly expressed in several types of human cancer cells, containing prostate cancer^[31], lung cancer^[32], melanoma^[33], breast cancer^[30] and glioblastoma^[34]. Based on the above conclusions, we suggest that FN1, POSTN, LOX and TIMP1 can be used in cancer diagnosis and as prognostic indicators.

Infiltration of immune cells into tumors is an essential factor affecting tumor occurrence and progression^[35, 36]. For instance, Chen et al proposed that the expression of CXCL10 is related to the infiltration of miscellaneous immune cells. Tumor-associated macrophages express CXCL10 in pancreatic cancer, and its acceptor CXCR3 is extremely expressed in T cells. Their study showed that the expression of CXCL10 was actively in connection with multiple immune core proteins, suggesting that CXCL10 can be used as one of the available targets of immunotherapy.^[6] Therefore, we conducted

Immune correlation analysis of the four core genes (FN1, POSTN, LOX, TIMP1). Notably, TIMP1 of these four genes is expressively involved in and affects tumor immunity by result comparison. The high expression of TIMP1 combined with the high proportion of Tregs, Cancer associated fibroblast and MAST cell suggested a poor prognosis in GBM patients. On the side, we also detected that TIMP1 was particularly connected with the immune marker of tumor associated macrophage, macrophage M1, macrophage M2, Neutrophils and Tregs. Hence, the results confirm the findings that the value of TIMP1 as a gene diagnostic for GBM, as well as the prognostic value of GBM patients with high TIMP1 expression, and are cheek by jowl connected with the immune microenvironment of GBM.

In addition, the main limitations of this study are as follows. First of all, the expression levels of hub genes, such as TIMP1, POSTN, LOX, and FN1, has not been confirmed by immunohistochemistry in clinical specimens. Secondly, the role of these hub genes in the appearance and evolution of GBM is still vague and should to be verified by in vitro and in vivo functional researches.

Conclusion

In this study, we report the results of a comprehensive bioinformatics analysis of biomarkers associated with glioblastoma prognosis and their association with immune-infiltrating cells. The wholesale data analysis in this research afforded a comprehensive bioinformatics analysis of DEGS that may be concerned in the progress of GBM. The shortcoming of insufficient samples is overcome by using five open access databases for new joint analysis. However, the molecular mechanism of TIMP1 in GBM still needs to be further studied. In general, our experiments recommended the requirement to more discreetly score TIMP1 as a biomarker and therapeutic target, in order to develop the new antibodies for GBM exploration and prevention, with the hope of providing new ideas for clinical diagnosis and treatment of GBM patients and effectively improving sufferer prognosis.

Materials And Methods

GEO Gene Expression Data

Keywords such as "Glioblastoma", "Homo sapiens" and "Array expression profile" were searched in GEO database and TCGA database, and GBM mRNA expression datasets retrieved by the above keywords were screened. After a systematic review, four GSE profiles (GSE4290, GSE7696, GSE13276 and GSE29796) and one TCGA profile (Glioblastoma multiforme) were selected and downloaded. The detail showed in Table 1. All data are freely available and no human or animal experiments were involved in this study.

DEGs Filtering

GEO data obtained DEGs between GBM and nontumorous brain tissue by using the GEO2R tool to standardize and convert original data from previously selected datasets into log2 format, setting screening criteria for log-fold change (FC) ≥ 2 and P value < 0.05 in each files. Then, the volcano diagram

of DEGs was revealed by visual hierarchical clustering analysis, and the overlapping genes were inspected in Venn diagram.

Functional Annotation and Signaling Pathway Analysis of the Hub Genes

To reveal the functions of DEGs, an online platform for data analysis and visualization (<http://www.bioinformatics.com.cn>) was used to dispose the GO function and KEGG paths of common upregulated and downregulated DEGs. GO term and KEGG pathways with $P < 0.05$ was considered statistically significant, meanwhile Adj. $P < 0.05$ was regarded as statistically significant.

PPI Network Analysis

We used the Search Tool for the Retrieval of Interacting Genes (STRING)^[37] to process and analyze the obtained DEGs. The filtrated 54 common upregulated DEGs had beforehand been handed over to the STRING. High-degree nodes are considered to be the most significant in this network. In this experiment, we selected the 10 genes with the highest connectivity as the central genes after the statistics of Cytoscape (V3.7.1) plugin cytoHubba.

Detection of Hub Genes in GBM

The GEPIA2 database^[38], was used to visualize the mRNA expression of screened out genes in GBM and non-malignant brain specimens. Human Protein Atlas (HPA) was used to determine the protein expression levels of 10 HUB genes in non-malignant brain specimens and GBM organizations.^[39]

Extract RNAseq Data and Establish Survival and Risk Curves

Use the UCSC XENA (<https://xenabrowser.net/datapages/>)^[40] after the unified processing of RNAseq data of TCGA and GTEX, the GBM data of TCGA and corresponding non-malignant specimen data in GTEX were extracted. After log2 transformation of RNASeq data in TPM format, expression comparison among samples was conducted, and the overall capability was assessed using receiver-operating characteristic (ROC) curve and area under ROC curve (AUC). For each hub gene, we used the pROC package to institute the ROC curve, calculated the AUC value and visualized the ROC curve with the ggplot2 package. The closer the AUC value is to 1, indicating better overall performance.

Survival Analyses, Univariate and Multivariate COX Analysis for Hub Genes

The relationship between overall survival (OS), disease-free survival (DFS) and core genes expressed in GBM sufferers was assessed by GEPIA2. Log-rank test results with $P < 0.01$ were statistically significant. The hub genes with independent prognostic factors were further screened by univariate and multivariate Cox regression analysis.

Immune Infiltration Analysis

The infiltration of multiple immune cells in GBM specimens was analyzed by tumor immune estimation resources (TIMER) ^[41]. TIMER2.0 was used to analyze the correlation between the expression of TIMP1 and the degree of infiltration of six kinds of immune cells, including regulatory cells and cancer-related fibroblasts etc.

Declarations

Data Availability Statement

All the experimental data involved in this study are obtained from open source. Please refer to the materials and methods in this paper for specific access.

Author Contribution

WD, SL, GJ and JL performed bioinformatics and biostatistical data analysis. WD was responsible for reviewing relevant literature, writing manuscripts, and producing figures. JL edited and modified the manuscript. JL and YL were involved in study design and project management. YL will bear all the expenses of this study. All authors contributed to this paper, which was submitted and published.

Conflict of Interest

There is no objection to the authorship and order, and no intellectual property dispute. In addition, the research was conducted without any commercial or financial relationship and there was no potential conflict of interest.

Funding

This study was supported by Natural Science Foundation of Hunan Province (2021JJ70044).

Ethics Statement

The author promises that all the results of this paper are original research, and all non-original research results are canonically cited as far as possible. In addition, there is no academic misconduct in this study, such as plagiarizing other people's articles, submitting more than one manuscript, modifying data results and so on.

References

1. Dong, Z., G. Zhang, M. Qu, et al. (2019). Targeting Glioblastoma Stem Cells through Disruption of the Circadian Clock. *Cancer Discov*, 9(11), 1556-1573. DOI: 10.1158/2159-8290.CD-19-0215.
2. Noushmehr, H., D.J. Weisenberger, K. Diefes, et al. (2010). Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell*, 17(5), 510-522. DOI: 10.1016/j.ccr.2010.03.017.

3. Purow, B. and D. Schiff. (2009). Advances in the genetics of glioblastoma: are we reaching critical mass? *Nat Rev Neurol*, 5(8), 419-426. DOI: 10.1038/nrneurol.2009.96.
4. Zhou, X., S. Xie, S. Wu, et al. (2017). Golgi phosphoprotein 3 promotes glioma progression via inhibiting Rab5-mediated endocytosis and degradation of epidermal growth factor receptor. *Neuro Oncol*, 19(12), 1628-1639. DOI: 10.1093/neuonc/nox104.
5. Kanehisa, M., S. Goto, Y. Sato, et al. (2014). Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res*, 42(Database issue), D199-205. DOI: 10.1093/nar/gkt1076.
6. Chen, S., F. Huang, S. Chen, et al. (2021). Bioinformatics-Based Identification of Tumor Microenvironment-Related Prognostic Genes in Pancreatic Cancer. *Front Genet*, 12(1171), 632803. DOI: 10.3389/fgene.2021.632803.
7. Wang, Q., S. Zhao, L. Gan, et al. (2020). Bioinformatics analysis of prognostic value of PITX1 gene in breast cancer. *Biosci Rep*, 40(9). DOI: 10.1042/BSR20202537.
8. Zhang, Z., Q. Li, F. Wang, et al. (2021). Identifying Hypoxia Characteristics to Stratify Prognosis and Assess the Tumor Immune Microenvironment in Renal Cell Carcinoma. *Front Genet*, 12, 606816. DOI: 10.3389/fgene.2021.606816.
9. Gilbert, L.A. and M.T. Hemann. (2010). DNA damage-mediated induction of a chemoresistant niche. *Cell*, 143(3), 355-366. DOI: 10.1016/j.cell.2010.09.043.
10. Hanahan, D. and L.M. Coussens. (2012). Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell*, 21(3), 309-322. DOI: 10.1016/j.ccr.2012.02.022.
11. Ye, Z., M. Zheng, Y. Zeng, et al. (2020). Bioinformatics Analysis Reveals an Association Between Cancer Cell Stemness, Gene Mutations, and the Immune Microenvironment in Stomach Adenocarcinoma. *Front Genet*, 11, 595477. DOI: 10.3389/fgene.2020.595477.
12. Wishart, A.L., S.J. Conner, J.R. Guarin, et al. (2020). Decellularized extracellular matrix scaffolds identify full-length collagen VI as a driver of breast cancer cell invasion in obesity and metastasis. *Sci Adv*, 6(43). DOI: 10.1126/sciadv.abc3175.
13. Lu, P., V.M. Weaver and Z. Werb. (2012). The extracellular matrix: a dynamic niche in cancer progression. *J Cell Biol*, 196(4), 395-406. DOI: 10.1083/jcb.201102147.
14. Cairns, R., I. Papandreou and N. Denko. (2006). Overcoming physiologic barriers to cancer treatment by molecularly targeting the tumor microenvironment. *Mol Cancer Res*, 4(2), 61-70. DOI: 10.1158/1541-7786.MCR-06-0002.
15. Hui, L. and Y. Chen. (2015). Tumor microenvironment: Sanctuary of the devil. *Cancer Lett*, 368(1), 7-13. DOI: 10.1016/j.canlet.2015.07.039.
16. Coussens, P.M. and W. Nobis. (2002). Bioinformatics and high throughput approach to create genomic resources for the study of bovine immunobiology. *Vet Immunol Immunopathol*, 86(3-4), 229-244. DOI: 10.1016/s0165-2427(02)00005-3.
17. Zhou, L., H. Tang, F. Wang, et al. (2018). Bioinformatics analyses of significant genes, related pathways and candidate prognostic biomarkers in glioblastoma. *Mol Med Rep*, 18(5), 4185-4196. DOI: 10.3892/mmr.2018.9411.

18. Sun, W., Y. Qin, Z. Wang, et al. (2021). The NEAT1_2/miR-491 Axis Modulates Papillary Thyroid Cancer Invasion and Metastasis Through TGM2/NFkappab/FN1 Signaling. *Front Oncol*, 11, 610547. DOI: 10.3389/fonc.2021.610547.
19. Cheng, N., M.Y. Wang, Y.B. Wu, et al. (2020). Circular RNA POSTN Promotes Myocardial Infarction-Induced Myocardial Injury and Cardiac Remodeling by Regulating miR-96-5p/BNIP3 Axis. *Front Cell Dev Biol*, 8, 618574. DOI: 10.3389/fcell.2020.618574.
20. Li, Q., J. Liu, W. Liu, et al. (2020). LOX-1 Regulates P. gingivalis-Induced Monocyte Migration and Adhesion to Human Umbilical Vein Endothelial Cells. *Front Cell Dev Biol*, 8, 596. DOI: 10.3389/fcell.2020.00596.
21. Yang, Y., Q. Feng, K. Hu, et al. (2021). Using CRISPRa and CRISPRi Technologies to Study the Biological Functions of ITGB5, TIMP1, and TMEM176B in Prostate Cancer Cells. *Front Mol Biosci*, 8, 676021. DOI: 10.3389/fmolb.2021.676021.
22. Docherty, A.J., A. Lyons, B.J. Smith, et al. (1985). Sequence of human tissue inhibitor of metalloproteinases and its identity to erythroid-potentiating activity. *Nature*, 318(6041), 66-69. DOI: 10.1038/318066a0.
23. Nosoudi, N., P. Nahar-Gohad, A. Sinha, et al. (2015). Prevention of abdominal aortic aneurysm progression by targeted inhibition of matrix metalloproteinase activity with batimastat-loaded nanoparticles. *Circ Res*, 117(11), e80-9. DOI: 10.1161/CIRCRESAHA.115.307207.
24. Ries, C. (2014). Cytokine functions of TIMP-1. *Cell Mol Life Sci*, 71(4), 659-672. DOI: 10.1007/s00018-013-1457-3.
25. Hayakawa, T., K. Yamashita, K. Tanzawa, et al. (1992). Growth-promoting activity of tissue inhibitor of metalloproteinases-1 (TIMP-1) for a wide range of cells A possible new growth factor in serum. *FEBS Letters*, 298(1), 29-32. DOI: 10.1016/0014-5793(92)80015-9.
26. Guedez, L., W.G. Stetler-Stevenson, L. Wolff, et al. (1998). In vitro suppression of programmed cell death of B cells by tissue inhibitor of metalloproteinases-1. *J Clin Invest*, 102(11), 2002-2010. DOI: 10.1172/JCI2881.
27. Vorotnikova, E., M. Triesand S. Braunhut. (2004). Retinoids and TIMP1 prevent radiation-induced apoptosis of capillary endothelial cells. *Radiat Res*, 161(2), 174-184. DOI: 10.1667/rr3107.
28. Chromek, M., K. Tullus, J. Lundahl, et al. (2004). Tissue inhibitor of metalloproteinase 1 activates normal human granulocytes, protects them from apoptosis, and blocks their transmigration during inflammation. *Infect Immun*, 72(1), 82-88. DOI: 10.1128/IAI.72.1.82-88.2004.
29. Coussens, L.M. and Z. Werb. (1996). Matrix metalloproteinases and the development of cancer. *Chem Biol*, 3(11), 895-904. DOI: 10.1016/s1074-5521(96)90178-7.
30. Cheng, G., X. Fan, M. Hao, et al. (2016). Higher levels of TIMP-1 expression are associated with a poor prognosis in triple-negative breast cancer. *Mol Cancer*, 15(1), 30. DOI: 10.1186/s12943-016-0515-5.
31. Oh, W.K., R. Vargas, S. Jacobus, et al. (2011). Elevated plasma tissue inhibitor of metalloproteinase-1 levels predict decreased survival in castration-resistant prostate cancer patients. *Cancer*, 117(3), 517-

525. DOI: 10.1002/cncr.25394.
32. Gouyer, V., M. Conti, P. Devos, et al. (2005). Tissue inhibitor of metalloproteinase 1 is an independent predictor of prognosis in patients with nonsmall cell lung carcinoma who undergo resection with curative intent. *Cancer*, 103(8), 1676-1684. DOI: 10.1002/cncr.20965.
 33. Kluger, H.M., K. Hoyt, A. Bacchiocchi, et al. (2011). Plasma markers for identifying patients with metastatic melanoma. *Clin Cancer Res*, 17(8), 2417-2425. DOI: 10.1158/1078-0432.CCR-10-2402.
 34. Aaberg-Jessen, C., K. Christensen, H. Offenberg, et al. (2009). Low expression of tissue inhibitor of metalloproteinases-1 (TIMP-1) in glioblastoma predicts longer patient survival. *J Neurooncol*, 95(1), 117-128. DOI: 10.1007/s11060-009-9910-8.
 35. Chen, P., D. Zhao, J. Li, et al. (2019). Symbiotic Macrophage-Glioma Cell Interactions Reveal Synthetic Lethality in PTEN-Null Glioma. *Cancer Cell*, 35(6), 868-884 e6. DOI: 10.1016/j.ccell.2019.05.003.
 36. Wang, X., X. Luo, C. Chen, et al. (2020). The Ap-2alpha/Elk-1 axis regulates Sirpalph-dependent tumor phagocytosis by tumor-associated macrophages in colorectal cancer. *Signal Transduct Target Ther*, 5(1), 35. DOI: 10.1038/s41392-020-0124-z.
 37. Szklarczyk, D., J.H. Morris, H. Cook, et al. (2017). The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res*, 45(D1), D362-D368. DOI: 10.1093/nar/gkw937.
 38. Tang, Z., C. Li, B. Kang, et al. (2017). GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*, 45(W1), W98-W102. DOI: 10.1093/nar/gkx247.
 39. Asplund, A., P.H. Edqvist, J.M. Schwenk, et al. (2012). Antibodies for profiling the human proteome- The Human Protein Atlas as a resource for cancer research. *Proteomics*, 12(13), 2067-2077. DOI: 10.1002/pmic.201100504.
 40. Vivian, J., A.A. Rao, F.A. Nothaft, et al. (2017). Toil enables reproducible, open source, big biomedical data analyses. *Nat Biotechnol*, 35(4), 314-316. DOI: 10.1038/nbt.3772.
 41. Li, T., J. Fan, B. Wang, et al. (2017). TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Res*, 77(21), e108-e110. DOI: 10.1158/0008-5472.CAN-17-0307.

Tables

Table 1. The GBM patients vs normal samples capacity of GEO datasets

datasets	sample capacity	
	Tumor tissues	Normal tissues
GSE4290	81	23
GSE7696	80	4
GSE13276	5	5
GSE29796	14	20
TCGA	163	207
ALL	343	259

Table 2. The common DEGs of four gene expression profiles (adj. P-val. <0.05, |logFC|>2.0).

Common DEGs	Gene symbol
Upregulated DEGs	TOP2A; ASPM; NDC80; IGFBP2; RRM2; LTF; COL4A1; IGF2BP3; PDPN; MEOX2; SHOX2; COL3A1; SOX4; COL1A1; CDK1; NAMPT; COL1A2; CHI3L1; VEGFA; LOX; IGFBP3; CD44; COL4A2; NES; POSTN; EZH2; TNC; ANXA1; TGFB1; PTX3; SOX11; CFI; BTN3A2; NID1; KIAA0101; ADAM12; HAS2; KLHDC8A; ANXA2; FN1; EMP1; VIM; MGP; TIMP1; PHLDA1; CENPU; CKS2; ENPEP; COL6A2; LAMC1; SMC4; BUB1B; C1R TGIF1
Downregulated DEGs	AK5; KCNK1; SNCA; GRM3; PEX5L; MOBP; STXBP6; RAB40B; SPOCK3; DLG2; MAL; CNTNAP2; REPS2; SH3GL3; TPPP; ANK3; MBP; ZNF365; ATP8A1; PAIP2B; SEC14L5; SEPTIN4
DEGs, differentially expressed genes	

Table 3. Top ten hub genes with higher degree of connectivity.

Gene symbol	Gene description	Degree
FN1	Fibronectin 1	25
TIMP1	TIMP Metalloproteinase Inhibitor 1	21
CD44	CD44 Molecule (Indian Blood Group)	21
COL1A1	Collagen Type I Alpha 1 Chain	20
VEGFA	Vascular Endothelial Growth Factor A	20
COL1A2	Collagen Type I Alpha 2 Chain	18
POSTN	Periostin	16
COL3A1	Collagen Type III Alpha 1 Chain	15
LOX	Lysyl Oxidase	14
IGFBP3	Insulin Like Growth Factor Binding Protein 3	14

Table 4. Correlation analysis between TIMP1 and relate genes and markers of immune cells in TIMER.

Description	Gene markers	GBM			
		None		Purity	
		Cor	P	Cor	P
CD8+ T Cell	CD8A	0.093	0.251	-0.009	0.915
	CD8B	0.168	0.038	0.069	0.420
T Cell (general)	CD3D	0.275	**	0.088	0.308
	CD3E	0.266	**	0.121	0.160
	CD2	0.273	**	0.092	0.285
B Cell	CD19	0.043	0.600	-0.026	0.760
	CD79A	-0.037	0.654	-0.087	0.313
Monocyte	CD86	0.313	***	0.111	0.198
	CSF1R	0.359	***	0.166	0.053
TAM	CCL2	0.579	***	0.490	***
	CD68	0.406	***	0.210	0.014
	IL10	0.397	***	0.214	0.012
Macrophage M1	NOS2	0.185	0.022	0.266	*
	IRF5	0.232	*	-0.005	0.954
	PTGS2	0.421	***	0.330	***
Macrophage M2	CD163	0.639	***	0.556	***
	VSIG4	0.470	***	0.316	**
	MS4A4A	0.498	***	0.380	***
Neutrophils	CEACAM8	-0.049	0.549	-0.086	0.315
	ITGAM	0.447	***	0.279	**
	CCR7	0.275	**	0.149	0.081
Natural killer cell	KIR2DL1	-0.007	0.929	-0.037	0.668
	KIR2DL3	-0.020	0.803	-0.083	0.335
	KIR2DL4	0.140	0.083	0.072	0.401
	KIR3DL1	0.036	0.656	0.023	0.793
	KIR3DL2	0.115	0.158	0.115	0.180

	KIR3DL3	0.034	0.675	-0.007	0.935
	KIR2DS4	0.174	0.032	0.136	0.112
Dendritic cell	HLA-DPB1	0.316	***	0.175	0.041
	HLA-DQB1	0.241	*	0.102	0.234
	HLA-DRA	0.339	***	0.167	0.051
	HLA-DPA1	0.223	*	0.100	0.245
Treg	STAT5B	-0.226	*	-0.132	0.124
	TGFB1	0.411	***		*

* P<0.001; ** P<0.001; *** P<0.0001

Figures

Figure 1

DEGs identification in different datasets of GBM. (A) Volcano diagrams of DEGs from the GBM vs. normal tissues in different datasets. $|\text{Log}_2\text{FC}| > 2.0$, $p\text{-value} < 0.05$. (B) The common upregulated genes in both GSE4290, GSE7696, GSE13276 and GSE29796 and GEPIA datasets were determined using Venn diagram. (C) The common downregulated genes in both GSE4290, GSE7696, GSE13276 and GSE29796 and GEPIA datasets were determined using Venn diagram.

Figure 2

Bioinformatics analyses of DEGs in GBM. (A) GO functions of the common upregulated DEGs in GBM. (B) GO functions of the common downregulated DEGs in GBM.

Figure 3

KEGG Pathway Enrichment Analyses of DEGs in GBM. (A) 54 upregulated DEGS enrichment pathways. (B) The top enrichment pathways, ECM-receptor interaction.

Figure 4

PPI network and hub genes identification. (A) PPI network was structured by the 54 upregulated DEGs using STRING database. (B) The top 10 hub genes in the PPI network were screened ground on their connectivity degree. The genes such as CD44, COL1A1, COL1A2, COL3A1, FN1, IGFBP3, LOX, POSTN, TIMP1 and VEGFA are represented from red (high degree) to yellow (low degree). (C) Gene Ontology (GO) chord diagram of the top genes in GBM.

Figure 5

Validation of mRNA expression of the top 10 hub genes in GBM. (A) The top genes expression levels in GBM (from TCGA database) and non-malignant pancreas (from GTEx database) were analyzed by GEPIA. ($p < 0.01$) (B) Representative immunohistochemistry images of the genes in GBM and non-malignant cerebral cortex specimens from the HPA.

Figure 6

Changes in Hub gene frequency and prognostic value in GBM. (A) The cBioPortal was used to calculate the percent- ages of GBM sufferers with the 10 altered hub genes. (B) mRNA expression alterations of the 10 hub genes in GBM sufferers. (C) DFS and (D) PFS of GBM patients with altered (red) and unaltered (blue) mRNA expression of 10 central genes.

Figure 7

The AUC analysis of hub genes. AUC analysis of CD44, COL1A1, COL1A2, COL3A1, FN1, IGFBP3, LOX, POSTN, TIMP1 and VEGFA for distinguishing GBM samples from normal tissues.

Figure 8

OS, DFS and Univariate, Multivariate Cox analysis of the hub genes overexpressed in GBM patients. (A) OS of the hub genes overexpressed in GBM sufferers. (B) DFS of the hub genes overexpressed in GBM sufferers. (C) Univariate and Multivariate Cox analysis of TIMP1 expression and other clinical pathological factors for OS. HR, hazard ratio.

Figure 9

Correlation between immune infiltrates and TIMP1 expression in GBM. (A) TIMP1 expression was definitely connected with CD4+ T cell, B cell, T cell regulatory (Tregs), Neutrophils, Cancer associated fibroblast, Macrophage M1 and Myeloid dendritic cell. (B) Higher infiltration of Tregs, Cancer associated fibroblast and Mast cell correlated with worse prognosis.