Prediction and analysis of hub genes between glioblastoma and low-grade glioma using bioinformatics analysis

Baowei Xu, MD

Abstract
Gliomas are an intractable tumor in the central nervous system. The present study aimed to identify the differentially expressed genes (DEGs) between glioblastoma multiforme (GBM) and low-grade gliomas (LGG) in order to investigate the mechanisms of different grades of gliomas. The Cancer Genome Atlas (TCGA) database was used to identify DEGs between GBM and LGG, and 2641 genes have been found differentially expressed. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were used to determine the related functions and pathways of DEGs. Protein–protein interaction (PPI) network extracted a total of 444 nodes and 1953 interactions, and identified the top 6 hub genes in gliomas. The microarray data of the datasets GSE52009 and GSE4412, which were obtained from Gene Expression Omnibus (GEO) database, were used to externally validate DEGs expression levels. Gene Expression Profiling Interactive Analysis (GEPIA) database which was based on TCGA was used to explore the survival of hub genes in LGG and GBM. Additionally, the Oncomine database and Chinese Glioma Genome Atlas (CGGA) database were used to validate the mRNA expression level and prognostic value of hub genes. Gene Set Enrichment Analysis (GSEA) identified further hub genes-related pathways. In summary, through biological information and survival analysis, 6 hub genes may be new biomarkers for diagnosis and for guiding the choice of treatment strategies for different grades of gliomas.

Abbreviations: BP = biological pathways, CC = cellular component, CGGA = Chinese Glioma Genome Atlas, COL1A1 = collagen type I alpha 1 chain, CXCL8 = C-X-C motif chemokine ligand 8, DEG = differentially expressed gene, FDR = false discovery rate, FN1 = fibronectin 1, GBM = glioblastoma multiforme, GEO = Gene Expression Omnibus, GO = Gene Ontology, GSEA = Gene Set Enrichment Analysis, KEGG = Kyoto Encyclopedia of Genes and Genomes, LGG = low-grade glioma, MF = molecular function, MMP9 = matrix metallopeptidase 9, OS = overall survival, POSTN = periostin, PPI = protein–protein interaction, TCGA = The Cancer Genome Atlas, TIMP1 = TIMP metallopeptidase inhibitor 1.

Keywords: bioinformatics analysis, differentially expressed genes, glioblastoma, survival analysis

1. Introduction
Gliomas are the most common primary tumor of the central nervous system. According to the histopathological characteristics, gliomas can be divided into 4 grades from low to high malignant degree. Low-grade gliomas (LGG) (astrocytomas, oligodendrogliomas, and oligoastrocytomas) are well-differentiated and have a 5-year survival rate of nearly 60%. Despite many treatment strategies, glioblastoma multiforme (GBM) is one of the most lethal brain tumors with a median overall survival (OS) of 14.6 months. Unfortunately, most LGG will progress to GBM within 5 to 10 years. In this process, the expression of many
genes and molecular pathways change. So far, O-6-methylguanine-DNA methyltransferase (MGMT) promoter methylation, EGFR alterations and isocitrate dehydrogenase 1 (IDH1) and isocitrate dehydrogenase 2 (IDH2) mutations have been identified as markers in molecular classification of glioblastomas. Therefore, it is significant to explore new genes and molecular markers for the treatment of different grades of gliomas.

Gene expression profiling analyses have greatly promoted clinical oncology research, mainly in the following aspects: searching for tumor related genes, exploring molecular diagnosis, comparing therapeutic effects and predicting prognosis and recurrence of tumors. In recent years, the application of microarray and high-throughput sequencing technology have enlarged gene expression profiling analysis to an unprecedented scale. At the same time, multiple databases and methods are used to study and verify the differentially expressed genes (DEGs), which makes bioinformatics analysis more accurate and reliable. In addition, through the comparative analysis of DEGs, especially the study of signaling pathways and interaction networks, we can explore the occurrence, development and transformation mechanisms of malignant tumors.

In this study, TCGA database was used to identify the DEGs between GBM and LGG. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were used to determine the related functions and pathways of DEGs. Protein–protein interaction (PPI) network extracted important nodes and interactions in gliomas, and identified top 6 hub genes. Subsequently, 2 clinical microarray data (GSE52009 and GSE4412) were obtained from Gene Expression Omnibus (GEO) database. The clinical value of the hub genes was further explored and verified by Gene Expression Profiling Interactive Analysis (GEPIA) and Gene Set Enrichment Analysis (GSEA) based on TCGA database. In order to further improve the reliability of this study, Oncomine database and Chinese Glioma Genome Atlas (CGGA) database were introduced for external validation of the mRNA expression level and prognostic value of hub genes.

2. Material and methods

2.1. Datasets and identification of DEGs

The mRNA expression profiles and clinical data of LGG and GBM patients were obtained from the TCGA data portal, which grades II and III of gliomas were included in the TCGA-LGG project, whereas grade IV was separated in TCGA-GBM project. Moreover, the datasets from the National Center of Biotechnology Information (NCBI) GEO database were divided into LGG and GBM according to the TCGA standard. The gene expression level and clinical data from TCGA were used for identification of DEGs, and Gene expression profile of GSE52009 and GSE4412 from GEO was used to externally validate the hub gene expression level. The GSE52009, which was based on GPL6480 Agilent-014850 Whole Human Genome Microarray was uploaded by Jiang et al, contained 92 LGG samples and 24 GBM samples. The GSE4412, which was based on GPL96 Affymetrix Human Genome U133A Array was uploaded by Nelson SF, included 26 LGG samples and 59 GBM samples. To clarify the DEGs between LGG and GBM, the fold change (FC) >2 and log2FC ≥1 and false discovery rate (FDR) < 0.05 were used as the threshold to identify DEGs. Heat map and volcano plot were generated using the R software package pheatmap and ggplot2, respectively.

2.2. Functional and pathway enrichment analysis

GO analysis was used to determine the functions including biological pathways (BP), cellular component (CC), and molecular function (MF). KEGG pathway enrichment analysis was used to systematically describe relevant pathways for the genes. The clusterProfiler package was used to perform GO analysis and KEGG pathway enrichment analysis. False discovery rate (FDR) of <0.05 was used as the cut-off value.

2.3. PPI network analysis

The online Search Tool the Retrieval of Interacting Genes (STRING) database was used to demonstrate the interaction between various proteins and Cytoscape software was used to reconstruct a PPI network. The mapped network was detected the possible relationship among DEGs. Due to the large number of DEGs while FC >2, we identified PPI network using FC: >4. The cut-off criterion of interaction confidence score was set as >0.4. The important nodes and hub genes in the network were predicted and explored by CytoHubba, a plugin of Cytoscape, and the top 10 genes were selected from the results of each method. The final top 6 hub genes in different methods were selected for further analysis.

2.4. Oncomine database analysis

The expression level of hub genes was identified in the Oncomine database, and analyzed the mRNA expression differences between LGG and GBM. The threshold was determined according to the following values: P value of .05, fold change of 2.0.

2.5. GEPIA database analysis

GEPIA database was used to perform survival analysis which is based on TCGA database. Survival analyses were carried out to achieve Kaplan–Meier plots. The log-rank test was performed during Kaplan–Meier survival analysis. The two-tailed P value was used in this study, and a P value of less than .05 considered statistically significant.

2.6. CGGA database analysis

Chinese Glioma Genome Atlas (CGGA) database which included over 2000 samples from the Chinese cohort, was used for external validation the mRNA expression level and prognostic value of hub genes. The results were downloaded from the CGGA website.

2.7. GSEA analysis

GSEA was conducted to further identify hub genes-related pathways. GSEA version 4.0.3 software was used to analyze data.
level of hub genes was used as the phenotype annotation, and patients in TCGA cohort were divided into low and high categories. The Molecular Signatures Database (MSigDB) of c2 (c2.cp.kegg.v7.1.symbols.gmt) was used to assess the functional differences between the low and high hub gene expression groups. The number of permutations was set to 1000 and the phenotype labels were high and low expression. FDR < 0.25 and P < .05 were set as the cut-off criteria to confirm significant gene sets.

3. Results

3.1. Identification of DEGs between LGG and GBM

Gene expression profiles from the TCGA database were conducted to analyze DEGs. A total of 529 LGG and 169 GBM samples was involved in our further analysis. Differentially expressed genes between GBM and LGG were identified as explained in the methods. The results revealed that a total of 2641 genes were identified to be differentially expressed, including 1428 up-regulated genes and 1213 down-regulated genes (Fig. 1A, B).

3.2. Functional and pathway enrichment analysis of DEGs

Total, up-regulated and down-regulated DEGs were used for GO and KEGG analysis, respectively. GO analysis showed genes associated with BP were mainly involved in regulation of trans-synaptic signaling in total DEGs, extracellular structure organization in up-regulated DEGs, modulation of chemical synaptic transmission in down-regulated DEGs. Genes associated with CC were mainly involved in synaptic membrane in total DEGs, collagen-containing extracellular matrix in up-regulated DEGs, and synaptic membrane in down-regulated DEGs. And genes associated with MF were mainly involved in channel activity in total DEGs, extracellular matrix structural constituent in up-regulated DEGs, channel activity in down-regulated DEGs (Fig. 2A, C, E). The detailed results were presented in Table 1.

KEGG pathway enrichment analysis was conducted to further investigate the pathways among these DEGs. As shown in Figure 2B, D, F, the significantly enriched key pathways were neuroactive ligand-receptor interaction, cytokine-cytokine receptor interaction and neuroactive ligand-receptor interaction in total, up-regulated and down-regulated DEGs, respectively. The detailed results were presented in Table 2.

3.3. Hub genes screening from PPI network

Protein interaction analysis could clarify the important protein and biological modules. PPI network construction was conducted using STRING tool and Cytoscape software. A total of 444 nodes and 1953 interactions were screened from the PPI network (Fig. 3A). CytoHubba contained several algorithms (Table 3), and the top 10 hub genes using MCC, DEGREE, and EPC algorithm were shown in Figure 3B, D. We chose the top 6 hub nodes as hub genes for further analysis, including TIMP metallopeptidase inhibitor 1 (TIMP1), fibronectin 1 (FN1), collagen type I alpha 1 chain (COL1A1), peristin (POSTN), matrix metallopeptidase 9 (MMP9), C-X-C motif chemokine ligand 8 (CXCL8).

3.4. Gene expression level in TCGA and GEO databases

To demonstrate the mRNA expression level in LGG and GBM, we first visualized the expression level in TCGA. As shown in Figure 4A, hub genes were significantly higher in GBM. We further analyzed 6 hub genes using GSE52009 and GSE4412 datasets. Similarly, hub gene mRNA expression levels were higher in GBM tissues, compared to LGG tissues (Fig. 4B, C).

3.5. The mRNA expression levels of hub genes in Oncomine database

Oncomine, an online microarray database, was used to validate the expression levels of hub genes in LGG and GBM. There we showed six representative datasets for analyzing hub genes. As shown in Figure 5A, TIMP1 in GBM was high expression in Sun, Liang, Bredel, and Freije datasets. FN1 was more highly expressed in GBM from vandenBoom, Nutt, Sun, and Bredel datasets (Fig. 5B). Up-regulation of COL1A1 was found in GBM based on Sun, vandenBoom, Bredel, and Freije datasets (Fig. 5C). Moreover, POSTN expression level was decreased in LGG tissues from Liang, Sun, Nutt, and Freije datasets (Fig. 5D). Compared to GBM, MMP9 was reduced in LGG, demonstrated by Liang, Bredel, Nutt, and Sun datasets (Fig. 5E). As for CXCL8, Liang, vandenBoom, Freije, and Sun datasets showed high expression levels in GBM, compared to LGG (Fig. 5F).

Besides, we used Oncomine dataset to analyze the hub genes expression level differences between tumor and normal tissues. As shown in Figure 5G, the database contained a total of 357, 367, 306, 367, 363, and 370 unique analyses for TIMP1, FN1, COL1A1, POSTN, MMP9, and CXCL8, respectively. In all types of tumors, TIMP1 was ranked with the top 10% of all genes showing statistically significant differences in 121 studies, with 113 studies demonstrated higher expression levels in tumor than normal tissues. Up-regulation of FN1 was found in tumor in 131 studies. 122 significant unique analyses revealed that the mRNA expression level of COL1A1 was higher in tumor. Compared to normal tissues, POSTN was up-regulated in tumors among 95 studies, while reduced expression was found in 18 studies. Higher expression of MMP9 was found in most cancers. As for CXCL8, only 67 studies were listed, and 57 studies showed high level of CXCL8 in tumor tissues. And all hub genes were up-regulated among brain and CNS cancer. Altogether, the transcriptional expression levels of hub genes were significantly up-regulated in GBM, compared to LGG and tumor tissues, compared to normal tissues.

3.6. The Kaplan–Meier plotter of hub genes

The website, http://geopia.cancer-pku.cn/, could provide the prognostic data of the hub genes based on TCGA database. And we found that expression of TIMP1 (HR=3, P=7.8 × 10⁻⁷) was associated with worse OS for LGG patients, as well as FN1 (HR = 1.8, P = .0022), COL1A1 (HR = 2, P = .00028), POSTN (HR = 1.7, P = .0065), not MMP9 (HR = 1.4, P = .093) and CXCL8 (HR = 1.3, P = .13) (Fig. 6A–F). Only FN1 and CXCL8 expression levels were associated with worse OS for GBM patients (HR = 1.5, P = .028, HR = 1.4, P = .049), respectively. Furthermore, POSTN expression had a trend for evaluating the prognosis of GBM patients (Fig. 6G–I).

3.7. Validation in the CGGA database

Next, we used CGGA database, which is based on a Chinese glioma patient information, to validate hub genes expression levels and prognostic value. This database has 3 datasets,
including mRNAseq-325, mRNAseq-693 and mRNA array-301. mRNAseq-325 contained 325 samples which involved 182 LGG, 139 GBM and 4 unclear patients (accession number: PRJCA001746, platform: Illumina HiSeq2000 or 2500), mRNAseq-693 contained 693 samples which involved 443 LGG, 249 GBM, and 1 unclear patients (accession number: PRJCA001747, platform: Illumina HiSeq) and mRNA array-301 contained 301 patients which involved 174 LGG, 124 GBM and
2 unclear patients (platform: Agilent Whole Human Genome). We found that 6 hub genes were associated with tumor grade, except for POSTN in grade II and III (Fig. 7).

As shown in Figure 8, the Kaplan–Meier survival analysis showed that all 6 hub genes were significantly associated with OS, a higher expression level in patients with shorter OS. Furthermore, we evaluated the prognostic value of hub genes in different tumor grade groups. Statistically significant difference was found in patients with grade III and IV group, and no difference was found in tumor grade II group for TIMP1, as well as FN1 and MMP9 (Fig. 8A, B, E). For COL1A1, it only showed the prognostic value in patients with tumor grade II (Fig. 8C).

Figure 2. Functional enrichment analyses of (A) total, (C) up-regulated and (E) down-regulated DEGs (GBM vs LGG). Top 10 biological process, cellular component, and molecular function terms for the DEGs. KEGG pathways enriched for the (B) total, (D) up-regulated and (F) down-regulated DEGs (GBM vs LGG).
POSTN could predict patients with tumor grade II and III, which was similar with TCGA database (Fig. 8D). The expression level of CXCL8 showed the statistically significant difference in patients with tumor grade IV, which met the same result in TCGA database (Fig. 8F).

3.8. Identification of hub genes associated biological pathways

GSEA analysis was performed to determine the biologic characteristics of hub genes using TCGA database. As shown in Figure 9A, TIMP1 regulated biology process mainly associated with sugar metabolism, glutathione metabolism, and leukocyte transendothelial migration, indicating that TIMP1 might regulate metabolism and cell metastasis. Similar results were obtained in COL1A1, POSTN, MMP9, and CXCL8 groups (Fig. 9C–F). It was interesting to find that FN1 might function in cell cycle-related pathways, including glycan biosynthesis, cell cycle, and cell apoptosis (Fig. 9B).

4. Discussion

The formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. The formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations. As a result, the formation of tumor and the increase of malignant degree are caused by the accumulation of many gene alterations.
common malignant tumor in the brain, up to 70% of LGG transform into HGG within 10 years. This transformation process involves changes in many genes and molecular pathways.\[^5\] This study is conducted to determine the DEGs between LGG and GBM through bioinformatics analysis. And further explore and study the molecular functional pathways involved in these DEGs, as well as their impact on tumor prognosis. Therefore, it is possible to use these DEGs as biomarkers and therapeutic targets in clinical practice, which can help the diagnosis and treatment of gliomas, and improve the prognosis of GBM.

The 5-year survival rate and overall survival time of patients with LGG and GBM are significantly different.\[^26\] GBM grows rapidly, has high invasiveness, is easy to cause neurological symptoms, not sensitive to chemotherapy, and easy to relapse after treatment.\[^4\] GBM can be divided into primary and secondary GBM, which are difficult to be differentiated histologically. Secondary GBM develops from LGG. Existing studies have shown that IDH mutation is a clear molecular marker for the diagnosis of secondary GBM, and it is more accurate than clinical and pathological diagnostic criteria.\[^17\] In addition, TP53 mutation and 19q deletion are more common gene changes in secondary GBM.\[^8,28\] Although there are some differences between primary and secondary GBM, they still have similar gene expression profiles.\[^27\] However, there are great differences between LGG and GBM in terms of biological behavior, histology, and gene expression profiles.\[^28\] Therefore, it is necessary to use the existing biological information analysis to further explore the differences between LGG and GBM, and help clinical diagnosis and treatment.

We used GO analysis to further clarify the functions of DEGs in biological pathways, cellular component and molecular function. Down-regulated DEGs are mainly related to the synaptic membrane structure and the activity of membrane channel, and participate in the regulation of trans-synaptic signaling. This result revealed that the down-regulated DEGs between LGG and GBM control the biological behavior of gliomas by regulating signals among different neurons. However, the up-regulated DEGs increased the malignant degree of glioblastoma mainly by changing the extracellular matrix, cytokines, and microenvironment. The above results showed that up-regulated and down-regulated DEGs changed the molecular functions, signal transduction and biological behavior of gliomas by regulating signals among different neurons. However, the up-regulated DEGs increased the malignant degree of glioblastoma mainly by changing the extracellular matrix, cytokines, and microenvironment. The above results showed that up-regulated and down-regulated DEGs changed the molecular functions, signal transduction and biological behavior of gliomas by regulating signals among different neurons. However, the up-regulated DEGs increased the malignant degree of glioblastoma mainly by changing the extracellular matrix, cytokines, and microenvironment.
PPI network analysis was used to reveal the interactions between proteins encoded by DEGs. We identified PPI network using FC>4 to focus on DEGs with more clinical significance. Similarly, we set the cut-off criterion of interaction confidence to be greater than 0.4. The important nodes and hub DEGs in the PPI network were predicted and explored by CytoHubba. Finally, we screened out the 6 most important DEGs from PPI network analysis, and made a detailed and in-depth study on them. The 6 hub genes we focused on have been described above, as shown below: TIMP1, FN1, COL1A1, POSTN, MMP9, and CXCL8.

The expression level of TIMP1 increased significantly in malignant tumors, which can inhibit the apoptosis of tumor cells.
The high expression of this gene correlates with a poor prognosis of the patients.\cite{29} In malignant gliomas, TIMP1 is related to the decomposition of extracellular matrix and can promote the invasion and movement of tumor cells.\cite{30} Fibronectin-1 encoded by FN1 is an important cell adhesion molecule in the extracellular matrix. Fibronectin-1 has the functions of regulating cell adhesion, growth and differentiation, promoting cell migration and proliferation, as well as ion exchange and information transfer.\cite{31} FN1 plays an important role in the invasion and metastasis of malignant tumors, and it is also one of the current research hotspots. At present, there is a few studies focusing on FN1 and gliomas.\cite{32} The COL1A1 gene provides instructions for making part of a large molecule called type I collagen.\cite{33} A recent study has shown that this gene is related to the poor prognosis of malignant gliomas.\cite{34} By knocking down of COL1A1 in invasive gliomas, the progression of tumor can be reduced and the survival rate of experimental animals can be significantly improved.\cite{35} POSTN encodes an unstructured extracellular matrix protein called peristin. The protein can interact with multiple integrins and participate in cell proliferation, cell migration, and epithelial-mesenchymal transition.\cite{35} This gene is expressed in the process of inflammation and many kinds of cancers, and is related to the resistance to antiangiogenic therapy and poor prognosis of gliomas.\cite{36,37} The overexpression of MMP9 can also promote the invasion and progression of gliomas, and is related to poor prognosis.\cite{38-40} High expression of CXCL8 in gliomas with release more cytokines, stimulate inflammation, promote cell proliferation, and lead to tumor progression.\cite{41} The mechanism

### Table 3

| Rank methods in cytoHubba | Catalog | MCC | MNC | Degree | EPC | BottleNeck | EcCentricity | Closeness | Radiality | Betweenness | Stress |
|---------------------------|---------|-----|-----|--------|-----|-------------|--------------|-----------|-----------|-------------|--------|
| Hub gene top 10           | CXCL8   | FN1 | FN1 | FN1    | FN1 | CXCL8       | CXCL8        | CXCL8     | CXCL8     | CXCL8       | CXCL8  |
|                           | FN1     | FN1 | FN1 | FN1    | FN1 | CXCL8       | CXCL8        | CXCL8     | CXCL8     | CXCL8       | CXCL8  |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| CCP1        | CCP1         | CCP1      | CCP1      | CCP1        | CCP1   |
|                           | CCP1    | CCP1| CCP1| CCP1   | CCP1| CCP1        | CCP1         | CCP1      | CCP1      | CCP1        | CCP1   |
|                           | COL1A2  | COL1A2| COL1A2| COL1A2| COL1A2| COL1A2        | COL1A2      | COL1A2  | COL1A2  | COL1A2      | COL1A2 |
|                           | POSTN   | POSTN| POSTN| POSTN   | POSTN| POSTN        | POSTN        | POSTN    | POSTN    | POSTN        | POSTN  |
|                           | ANXA1   | ANXA1| ANXA1| ANXA1   | ANXA1| ANXA1        | ANXA1        | ANXA1   | ANXA1   | ANXA1        | ANXA1  |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |
|                           | MMP9    | MMP9| MMP9| MMP9   | MMP9| MMP9         | MMP9         | MMP9     | MMP9     | MMP9        | MMP9   |

Figure 4. Top 6 hub genes mRNA expression levels in TCGA and GEO database between LGG and GBM. (A) TCGA; (B) GSE52009; (C) GSE4412. Differences between groups were analyzed by the Student t test, *P < 0.05; **P < 0.01; ***P < 0.001.
Figure 5. Top 6 hub gene mRNA expression levels in Oncomine database. (A) TIMP1; (B) FN1; (C) COL1A1; (D) POSTN; (E) MMP9; (F) CXCL8; (G) The hub genes in human cancers, the number in the colored cell represents the number of studies meeting thresholds. The more stressed red (over-expression) or blue (under-expression) indicates a more highly significant relationship. Differences between groups were analyzed by the Student t test or one way ANOVA.
Figure 6. Kaplan-Meier plots and log-rank tests for the prognostic value of the hub genes in LGG (A-F, A, TIMP1; B, FN1; C, COL1A1; D, POSTN; E, MMP9; F, CXCL8) and GBM (G-L, G, TIMP1; H, FN1; I, COL1A1; J, POSTN; K, MMP9; L, CXCL8).
by which CXCL8 works is related to inflammatory stimulation, tumor angiogenesis, and JAK/STAT1/HIF-1α/Snail pathway activation.[42,43]

Oncomine database is one of the largest cancer gene chip databases in the world, and is committed to the data standardization and analysis of gene expression profile data of tumor samples.[17] Oncomine database can be used to analyze gene expression differences, predict co-expressed genes, and classify according to the clinical information such as tumor staging and tissue types.[44] CGGA database is the largest functional genomics database of gliomas in Asia. Its information covers matched samples of different pathological types and malignant degrees. In this study, the expression level and prognostic value of hub genes were validated by the above 2

Figure 7. The hub genes mRNA expression levels in CGGA database. (A) TIMP1; (B) FN1; (C) COL1A1; (D) POSTN; (E) MMP9; (F) CXCL8. Differences between groups were analyzed by the Student t test, *P < 0.05; **P < 0.01, ***P < 0.001, ns, not significant.
Through multiple databases and multiple methods of interactive verification, we can improve the reliability and accuracy of bioinformatics analysis in this study, and reduce the research deviation caused by single database analysis.

However, there are several limitations in this study. The results above, identified in TCGA database and validated in GEO, Oncomine database, and CGGA datasets, were not validated using new tissue samples. Second, the functions of hub genes were not functionally tested, which would be conducted in our further studies.

Figure 8. Kaplan-Meier plots and log-rank tests for the prognostic value of hub genes in CGGA database. (A) TIMP1; (B) FN1; (C) COL1A1; (D) POSTN; (E) MMP9; (F) CXCL8.
5. Conclusion
This study carried out a systematic bioinformatics analysis of DEGs between LGG and GBM. The results provide potential biomarkers and therapeutic targets for gliomas. However, further experimental verification is needed to directly determine the role of these DEGs in glioma.

Author contributions
BX designed the study, collected the data, performed the bioinformatics analysis, analyzed the data and wrote the manuscript.
Writing – original draft: Baowei Xu.
References

[1] Quinn , Ostrom , Haley , et al. CBTRUS statistical report: primary brain and other central nervous system tumours diagnosed in the United States in 2011-2015. Neuro-oncology 2018;20(1):1–86.

[2] Venneti S, Huse JT. The evolving molecular genetics of low-grade glioma. Adv Anat Pathol 2018;25(2):94–101.

[3] Claus EB, Black PM. Survival rates and patterns of care for patients diagnosed with supratentorial low-grade gliomas: data from the SEER program, 1973-2001. Cancer 2006;106:1358–63.

[4] Perry JR, Lapenerie N, O'Callaghan CJ, et al. Short-course radiation plus temozolomide in elderly patients with glioblastoma. N Engl J Med 2017;376:1027–37.

[5] Gustavo DL, Loraine C, Cassia R. Gliomas and the vascular fragility of the blood brain barrier. Front Cell Neurosci 2014;8:418.

[6] RH D, JD, SF, et al. Prognostic value of O-6-methylguanine-DNA methyltransferase (MGMT) protein expression in glioblastoma excluding nontumour cells from the analysis. Neuropathol Appl Neurobiol 2018;44:172–84.

[7] Eckel-Passow JE, Lachance DH, Molinaro AM, et al. Glioma groups based on 1p/19q, IDH, and TERT promoter mutations in tumors. N Engl J Med 2015;372(26):2606–15.

[8] Kim YW, Liu TJ, Koul D, et al. Identification of novel synergistic targets for rational drug combinations with PI3 kinase inhibitors using siRNA synthetic lethality screening against GBM. Neuro Oncology 2011;13:367–75.

[9] García-Escudero R, Paramio JM. Gene expression profiling of a tool for basic analysis and clinical application of human cancer. Mol Carcinog 2008;47:573–9.

[10] Weeraratna AT. Discovering causes and cures for cancer from gene expression analysis. Ageing Res Rev 2005;4:50–63.

[11] Li J, Lu Y, Akhani R, et al. TCPA: a resource for cancer functional proteomics data. Nat Methods 2013;10:1046–7.

[12] Cheng W, Ren X, Zhang C, et al. Bioinformatic profiling identifies an immune-related risk signature for glioblastoma. Neurology 2016;10.1212/WNL.0000000000002770.

[13] Jiang CM, Wang XH, Shu J, et al. Analysis of differentially expressed genes based on microarray data of glioma. Int J Clin Exp Med 2015;8:17321–32.

[14] Subramanian A, Kuehn H, Gould J, et al. GSEA-P: a desktop application for Gene Set Enrichment Analysis. Bioinformatics 2007;23:3251–3.

[15] Tang Z, Li C, Kang B, et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 2017;51(1):W1–W11.

[16] Liu M, Xu Z, Du Z, et al. The Identification of key genes and pathways in glioma by bioinformatics analysis. J Immunol Res 2017;2017:1–9.

[17] Rhodes DR, Kalyana-Sundaram S, Mahavisno V, et al. Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. Neoplasia 2007;9:166–80.

[18] Yang P, Yan W, Zhang W, et al. Whole-genome messenger RNA profiling reveals genes involved in malignant progression of glioma. Zhonghua Yi Xue Za Zhi 2013;93:595–5.

[19] Freij WA, Castro-Vargas FF, Tang Z, et al. Gene expression profiling of gliomas strongly predicts survival. Cancer Res 2004;64:6503–10.

[20] Ashburner MM, Ball CA C, Blake JA J, et al. Gene ontology: tool for the unification of biology. The gene ontology consortium. Nat Genet 2000;25:25–9.

[21] Minoura K, Susumu G. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000;28.

[22] Damian S, Andrea F, Stefan W, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 2015;43:447–52.

[23] Shen F, Guo Q, Hu Q, et al. ReBi, a good prognosis predictor, links cell cycle and migration to glioma tumorigenesis. Oncol Lett 2018;15:4404–10.

[24] Vecchi M, Nuciforo P, Romagnoli S, et al. Gene expression analysis of early and advanced gastric cancers. Oncogene 2007;26:4284–94.

[25] Carter SL, Eklund AC, Kohane IS, et al. A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. Nat Genet 2006;38:1043–8.

[26] Werner JM, Kuhl S, Stavrinou P, et al. Expression of FAS-L differs from primary to relapsed low-grade gliomas and predicts progression-free survival. Anticancer Res 2017;37:6639–48.

[27] Ohgaki H, Kleihues P. The definition of primary and secondary glioblastoma. Clin Cancer Res 2013;19:764–72.

[28] Reon BJ, Jordan A, Ying Z, et al. Expression of IncRNAs in low-grade gliomas and glioblastoma multiforme: an in silico analysis. PLoS Med 2016;13:e1002192.

[29] Li Y, Jiang-fei W, Guang-zu G, et al. Plasma levels of tissue inhibitor of matrix metalloproteinase-1 correlate with diagnosis and prognosis of glioma patients. Chin J Med 2013;126:4295–300.

[30] Singh MK, Bhattacharya D, Chaudhuri S, et al. T11TS inhibits glioma angiogenesis by modulation of MMPs, TIMPs, with related integrins av and TGF-b1 expressions. Tumor Biol 2013;35:

[31] Ohnishi T, Hiraga S, Iizumoto S, et al. Role of fibronectin-stimulated tumor cell migration in glioma invasion in vivo: clinical significance of fibronectin and fibronectin receptor expressed in human glioma tissues. Clin Exp Metastasis 1998;16:729–41.

[32] Jian , Gong , Zhao-xia , et al. miRNA1271 inhibits cell proliferation in neuroglioma by targeting fibronectin I. Mol Med Rep 2017;16:143–50.

[33] Lee EX, Lam DH, Wu C, et al. Glioma gene therapy using induced pluripotent stem cell derived neural stem cells. Mol Pharmaceutics 2011;8:1515–24.

[34] Koj Y, Hideki M, Rysuske H, et al. Quantification of phosphoryl gene-expression signature of gliomas and glioblastoma-derived spheres. Neuro Onco 2014;(suppl_3):sul_3.

[35] Andrea C, Patrick JD, Anna EA, et al. TMIC-58. the cellular and molecular basis for mesenchymal transformation in gliomas. Neuro Onco 2019;(Supplement_6):Sulement_6.

[36] Young PS, Yuj P, John dG. Al23Peristin regulates tumor resistance to antiangiogenic therapy through emt and angiogenesis-related mechanisms in glioma stem cell models. Neuro Onco 2014;(suppl_5):sul_5.

[37] Park SY, Piao Y, Jeong KJ, et al. Peristin (POSTN) regulates tumor resistance to antiangiogenic therapy in glioma models. Mol Cancer Therapeut 2016;15:1877–97.

[38] Cai H, Wang J, Xi S, et al. Tenascin-c-mediated vasculogenic mimicry formation via regulation of MMP2/9MP9 in glioma. Cell Death Dis 2019;10:8879.

[39] Erneline T, Françoise B, Patrizia F, et al. MMP2 and MMP9 as candidate biomarkers to monitor bevacizumab therapy in high-grade glioma. Neuro Onco 2015;8.

[40] Feng S, Miao G, Han Y, et al. CARMA3 and CARM433 is overexpressed in human glioma cells. Biochim Biophys Acta 2014;1842:30.

[41] Chen Z, Mou L, Pan Y, et al. CXCL8 promotes glioma progression by activating the JAK/STAT1/HIF-1alpha signaling axis. OncoTargets Ther 2019;12:8125–38.

[42] Brat DJ, Bellai AC, Meir EGV. The role of interleukin-8 and its receptors in gliomagenesis and tumoral angiogenesis. Neuro Onco 2005;7:122–33.

[43] Chimnayian A. Oncomine and caBIG advance cancer bioinformatics - the scientist - magazine of the life sciences. Scientist 2010;19.