


DNA Methylation-based Diagnostic and Prognostic Biomarkers for Glioblastoma

Yunliang Tang^{1,2}, Cheng Qing¹, Jiao Wang³,
and Zhenguo Zeng¹ 

Cell Transplantation
Volume 29: 1–14
© The Author(s) 2020
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0963689720933241
journals.sagepub.com/home/ctj


Abstract

Glioblastomas are the most common primary central nervous system malignancy tumor in adults. Glioblastoma patients have poor prognosis, with an average survival period of approximately 14 mo after diagnosis. To date, there are a limited number of effective treatment methods for glioblastoma, and its molecular mechanisms remain elusive. In this article, we analyzed the key biomarkers and pathways in glioblastoma patients based on gene expression and DNA methylation datasets. The 60 hypomethylated/upregulated genes and 110 hypermethylated/downregulated genes were identified in GSE50923, GSE50161, and GSE116520 microarrays. Functional enrichment analyses indicated that these methylated-differentially expressed genes were primarily involved in collagen fibril organization, chemical synaptic transmission, extracellular matrix-receptor interaction, and GABAergic synapse. The hub genes were screened from a protein–protein interaction network; in selected genes, increased NMB mRNA level was associated with favorable overall survival, while elevated CHI3L1, POSTN, S100A4, LOX, S100A11, IGFBP2, SLC12A5, VSNL1, and RGS4 mRNA levels were associated with poor overall survival in glioblastoma patients. Additionally, CHI3L1, S100A4, LOX, and S100A11 expressions were negatively correlated with their corresponding methylation status. Furthermore, the receiver-operator characteristic curve analysis indicated that CHI3L1, S100A4, LOX, and S100A11 can also serve as highly specific and sensitive diagnostic biomarkers for glioblastoma patients. Collectively, our study revealed the possible methylated-differentially expressed genes and associated pathways in glioblastoma and identified four DNA methylation-based biomarkers of glioblastoma. These results may provide insight on diagnostic and prognostic biomarkers, and therapeutic targets in glioblastoma.

Keywords

DNA methylation, glioblastomas, biological markers

Introduction

Primary central nervous system (CNS) tumors include a wide range of tumors, and 80% of CNS tumors are gliomas¹. Glioblastomas (GBM) arise from glial or precursor cells and are the most common primary CNS malignant tumors in adults, with an average survival period of approximately 14 mo after diagnosis^{2,3}. The World Health Organization (WHO) has classified GBM as a grade IV astrocytoma, the highest grade of malignancy, based on the histopathological and clinical features of this tumor⁴. The therapeutic methods commonly used are surgical resection, radiotherapy, and chemotherapy. Unfortunately, there are no standard treatment methods for GBM, and its underlying molecular mechanisms remain unclear⁵. Therefore, it is crucial to explore novel diagnostic and prognostic biomarkers, establish prediction methods, and provide insights into therapeutic targets in GBM.

Epigenetics mechanisms, such as DNA methylation and histone acetylation, depict the changes in gene expression without potential modification of the genome sequence⁶. It is

¹ Department of Critical Care Medicine, The First Affiliated Hospital of Nanchang University, Jiangxi, China

² Department of Rehabilitation Medicine, The First Affiliated Hospital of Nanchang University, Jiangxi, China

³ Department of Endocrinology and Metabolism, The First Affiliated Hospital of Nanchang University, Jiangxi, China

Submitted: January 15, 2020. Revised: March 1, 2020. Accepted: May 19, 2020.

Corresponding Authors:

Zhenguo Zeng, Department of Critical Care Medicine, The First Affiliated Hospital of Nanchang University, 17 Yongwai St, Nanchang, Jiangxi 330006, China.

Email: zengzhenguo@ncu.edu.cn

Jiao Wang, Department of Endocrinology and Metabolism, The First Affiliated Hospital of Nanchang University, 17 Yongwai St., Nanchang, Jiangxi 330006, China.

Email: wangjiao@email.ncu.edu.cn



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

well established that alterations of DNA methylation play a key role in pathogenesis and progression of various tumors, including GBM^{7,8}. In addition, altered methylation patterns in DNA sequences, including hypomethylated oncogenes and hypermethylated suppressors, are significantly correlated with patient survival in various cancers^{9,10}. Thus, the identification of differentially methylated genes (DMGs) combined with differentially expressed genes (DEGs) will significantly benefit the exploration of molecular mechanisms and establishment of effective biomarkers of GBM.

There are evidences that indicate identification of valuable biomarkers using transcriptomics^{11–13}. Although some studies have revealed aberrant DNA methylation in GBM, limited research has been conducted on the comprehensive analysis of gene expression, methylation, and underlying molecular mechanisms^{14,15}. Thus, integrated analyses of methylated-differentially expressed gene (MDEG) profiles are beneficial for the development of therapeutic strategies for GBM.

This study aimed to explore DNA methylation-based diagnostic and prognostic biomarkers in GBM patients. Gene methylation profiling microarrays (GSE50923) and gene expression microarrays (GSE50161 and GSE116520) were analyzed using various bioinformatics tools. These MDEGs may provide insight into therapeutic targets and diagnostic biomarkers for GBM.

Materials and Methods

Microarray Datasets

DNA methylation microarray (GSE50923) and gene expression microarrays (GSE50161 and GSE116520) were obtained from the Gene Expression Omnibus (GEO; National Institutes of Health, Bethesda, MD, USA; <https://www.ncbi.nlm.nih.gov/geo/>). GSE50923 dataset contains 24 control brain tissues and 54 GBM samples obtained using the GPL8490 Platform (HumanMethylation27 BeadChip, Illumina Inc., San Diego, CA, USA). Gene expression profiling of the GSE50161 dataset was performed using the GPL570 platform (Human Genome U133 Plus 2.0 Array, Affymetrix, Inc., Santa Clara, CA, USA) that included 13 nontumor brain tissues and 34 GBM tumor tissues. Gene expression profiling of the GSE116520 dataset was conducted using the GPL10558 platform (HumanHT-12 V4.0 expression Beadchip, Illumina), including 8 nontumor brain tissues and 17 GBM tumor tissues.

Data Processing

We identified DEGs and DMGs using GEO2 R (National Institutes of Health; <http://www.ncbi.nlm.nih.gov/geo/geo2r/>). The adjusted $P < 0.05$ and $|\log\text{fold change}| \geq 2$ were selected for DEGs, and the threshold of adjusted $P < 0.05$ and $|r| > 2$ were selected for DMGs. Subsequently, hypomethylated/upregulated genes and hypermethylated/downregulated genes were intersected by a Venn diagram.

Functional Enrichment Analyses

Functional enrichment analyses were performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID; National Cancer Institute, Frederick, MD, USA; <http://david.abcc.ncifcrf.gov/>) database, including Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses. An enrichment threshold of $P < 0.05$ was considered significant.

Construction of Protein–Protein Interaction Network and Identification of Hub Genes

Protein–Protein Interaction (PPI) network was analyzed using the STRING database (European Molecular Biology Laboratory, Meyerhofstrasse, Heidelberg, Germany; Version 11.0, <http://www.string-db.org/>), and the network was visually constructed using Cytoscape software (National Institute of General Medical Sciences, San Diego, CA, USA; Version 3.7.1, <http://www.cytoscape.org/>). Additionally, CytoHubba, a plugin app of Cytoscape, was used to identify the hub genes.

Survival Analysis of Hub Genes

The prognostic values of hub genes were assessed using gene expression profiling interactive analysis (GEPIA) (Biomedical Pioneering Innovation Center Peking University, Beijing, China; <http://gepia.cancer-pku.cn/>). Samples were divided into two groups (low and high expression) based on mRNA levels using a hazard ratio and log-rank P -value to analyze the overall survival (OS) of GBM patients. Log-rank P -value < 0.05 was considered statistically significant.

Expression and Correlation Analysis of Hub Genes

GEPIA (Biomedical Pioneering Innovation Center Peking University; <http://gepia.cancer-pku.cn/>) is a RNA sequencing database derived from analysis of 9,736 tumors and 8,587 healthy samples from The Cancer Genome Atlas and Genotype-Tissue Expression datasets. The hub gene mRNA levels between tumor and normal samples were analyzed using GEPIA. $P < 0.01$ was considered significant. In addition, the cBioPortal was used for analysis of genetic alterations in hub genes and validation of correlation between gene expression and DNA methylation.

Validation of mRNA Levels In Vitro

Human gliocyte HEB cell lines and human GBM LN229 cells were purchased from Shanghai Biotech Co., Ltd., Shanghai, China. The cells were grown in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum and incubated at 37°C in an atmosphere of 5% carbon dioxide in a humidified incubator. The medium was changed every 2 d.

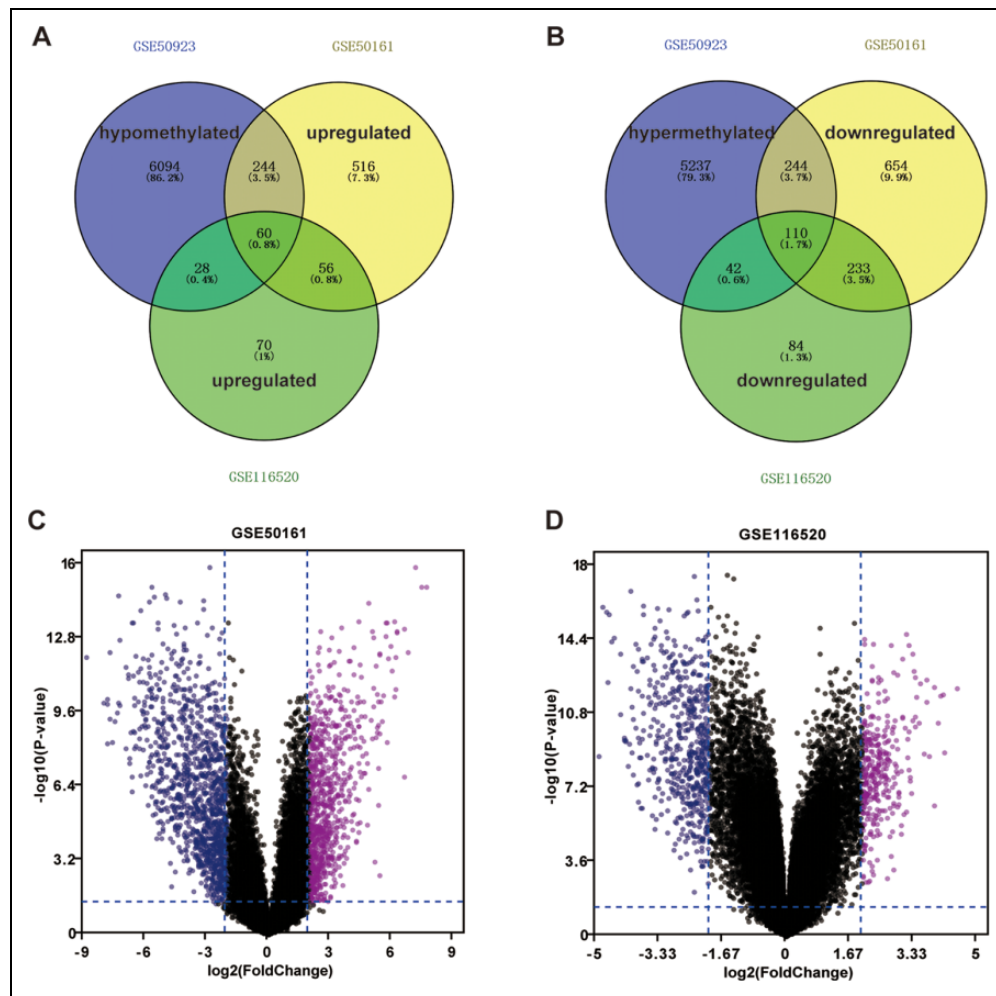


Figure 1. The methylated-differentially expressed genes. (A) Venn diagram of hypomethylation/high-expression genes. (B) Venn diagram of hypermethylation/low-expression genes. (C) Volcano plot for differentially expressed mRNA in GSE50161. (D) Volcano plot for differentially expressed mRNA in GSE116520.

Total RNA was extracted from the cell lines using TRIzol (Invitrogen, Carlsbad, CA, USA) according to manufacturer's protocol. cDNA was synthesized based on standard protocols of EasyScript First-Strand cDNA Synthesis SuperMix (TransGen Biotech, Beijing, China). Quantitative polymerase chain reaction (PCR) analysis was performed using the StepOne Real-Time PCR System (Thermo Fisher Scientific, Rockford, IL, USA). The program was performed as follows: denaturation at 95°C for 15 min, followed by 40 cycles of 10 s at 95°C for annealing, and 32 s at 60°C for extension. The melting curve started at 95°C for 15 s, followed by 60°C for 1 min and ended with 15 s at 95°C. The primers were shown in supplemental Table S1. The $2^{-\Delta\Delta C_t}$ method was used for the relative quantification of gene expression levels following the quantitative real time polymerase chain reaction experiments.

Diagnostic Analysis of Hub Genes

We evaluated the diagnostic effectiveness of hub genes using the receiver-operator characteristic (ROC) curve

analysis with GSE50161 expression profiling. GraphPad prism (GraphPad, San Diego, CA, USA) was used to construct the ROC curve, and the area under the ROC curve exceeding 0.8 was considered significant.

Statistical Analysis

Experimental values are expressed as mean \pm standard deviation of three independent experiments. The statistical significance of differences was obtained by Student's *t*-test in GraphPad Pro (GraphPad). Differences were considered statistically significant when *P*-value was lower than 0.05.

Results

Identification of MDEGs

For the methylation microarray GSE50923, 12,059 DMGs were obtained, comprising 5,633 hypermethylated genes and 6,426 hypomethylated genes (supplemental Tables S2 and S3). In the expression microarrays, 876 upregulated

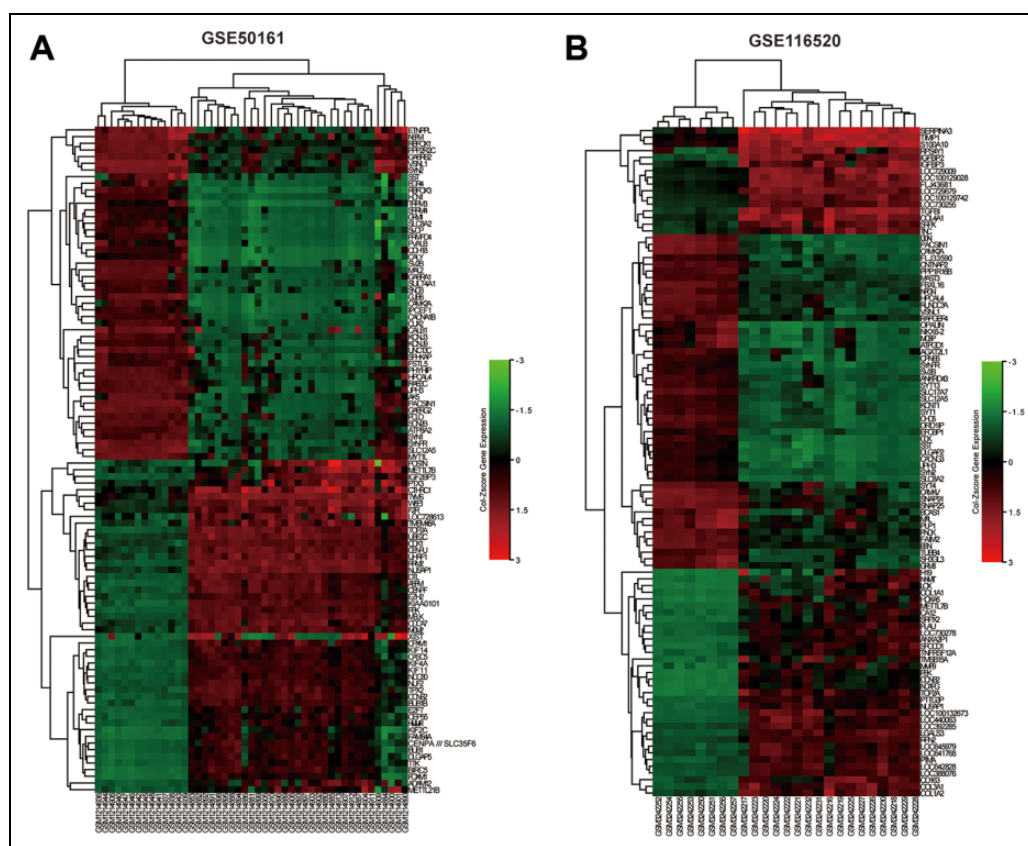


Figure 2. The heatmap of DEGs. (A) Top 100 DEGs in GSE50161 datasets. (B) Top 100 DEGs in GSE116520 datasets. Red: upregulation; Green: downregulation. DEGs: differentially expressed genes.

genes and 1,241 downregulated genes were identified in GSE50161, and 214 upregulated genes and 469 downregulated genes were identified in GSE116520. A total of 60 hypomethylated/upregulated genes were obtained by overlapping the hypomethylation genes and the upregulated genes, and 110 hypermethylated/downregulated genes were obtained by overlapping the hypermethylation genes and the downregulated genes (Fig. 1A, B). The volcano plots of the expression microarray are presented in Fig. 1C, D. Lastly, heat maps of gene expression values were constructed using color patterns indicating the variability in gene expression between normal and GBM groups (Fig. 2A, B).

GO and Pathway Functional Enrichment Analysis

GO and KEGG pathway enrichment analysis were performed using DAVID database to predict the biological function of and pathways followed by the MDEGs. For hypomethylated/upregulated genes, as presented in Table 1 and Fig. 3, the GO was observed to be enriched in collagen fibril organization, extracellular matrix organization, collagen catabolic process, cell division, and cell adhesion in the BP category (Fig. 3A); extracellular space, extracellular

matrix, extracellular region, extracellular exosome, and perinuclear region of cytoplasm in the CC category (Fig. 3B); and extracellular matrix structural constituent, cadherin binding involved in cell–cell adhesion, enzyme binding, protein homodimerization activity, and calcium-dependent protein binding in the MF category (Fig. 3C). Besides, as presented in Table 2 and Fig. 3D, hypomethylated/upregulated genes were enriched in selected cancer-associated pathways, such as PI3K-Akt signaling pathway, microRNAs in cancer, and proteoglycans in cancer. The detailed MDEGs related to the top five enriched pathways are shown in supplemental Table S4.

For hypermethylated/downregulated genes, as presented in Table 1 and Fig. 4, we observed that the GO was enriched in chemical synaptic transmission, neurotransmitter secretion, long-term synaptic potentiation, axonogenesis, and learning in the BP category (Fig. 4A); cell junction, axon, dendrite, perinuclear region of cytoplasm, and myelin sheath in the CC category (Fig. 4B); and phospholipase binding, GABA-A receptor activity, extracellular ligand-gated ion channel activity; GTPase activator activity, and amino acid transmembrane transporter activity in the MF category (Fig. 4C). Similarly, hypermethylated/downregulated genes were also enriched in some pathways, including GABAergic

Table 1. GO Enrichment Analysis of MDEGs Related With GBM Patients.

Category	Term	Count	%	P-value
<i>Hypomethylation/high-expression genes</i>				
GOTERM_BP_DIRECT	GO:0030199 ~ collagen fibril organization	7	11.66666667	3.49E-09
GOTERM_BP_DIRECT	GO:0030198 ~ extracellular matrix organization	9	15	3.02E-07
GOTERM_BP_DIRECT	GO:0030574 ~ collagen catabolic process	5	8.333333333	6.51E-05
GOTERM_BP_DIRECT	GO:0051301 ~ cell division	8	13.33333333	1.74E-04
GOTERM_BP_DIRECT	GO:0007155 ~ cell adhesion	7	11.66666667	0.004509299
GOTERM_CC_DIRECT	GO:0005615 ~ extracellular space	23	38.33333333	6.88E-11
GOTERM_CC_DIRECT	GO:0031012 ~ extracellular matrix	11	18.33333333	3.42E-08
GOTERM_CC_DIRECT	GO:0005576 ~ extracellular region	18	30	8.52E-06
GOTERM_CC_DIRECT	GO:0070062 ~ extracellular exosome	23	38.33333333	3.41E-05
GOTERM_CC_DIRECT	GO:0048471 ~ perinuclear region of cytoplasm	9	15	8.32E-04
GOTERM_MF_DIRECT	GO:0005201 ~ extracellular matrix structural constituent	7	11.66666667	9.63E-08
GOTERM_MF_DIRECT	GO:0048306 ~ calcium-dependent protein binding	4	6.666666667	0.000987
GOTERM_MF_DIRECT	GO:0098641 ~ cadherin binding involved in cell-cell adhesion	6	10	0.002918456
GOTERM_MF_DIRECT	GO:0019899 ~ enzyme binding	5	8.333333333	0.025911241
GOTERM_MF_DIRECT	GO:0042803 ~ protein homodimerization activity	7	11.66666667	0.035872566
<i>Hypermethylation/low-expression genes</i>				
GOTERM_BP_DIRECT	GO:0007268 ~ chemical synaptic transmission	13	11.81818182	2.99E-08
GOTERM_BP_DIRECT	GO:0007269 ~ neurotransmitter secretion	6	5.454545455	1.55E-05
GOTERM_BP_DIRECT	GO:0060291 ~ long-term synaptic potentiation	5	4.545454545	8.72E-05
GOTERM_BP_DIRECT	GO:0007409 ~ axonogenesis22	6	5.454545455	3.56E-04
GOTERM_BP_DIRECT	GO:0007612 ~ learning	5	4.545454545	4.26E-04
GOTERM_CC_DIRECT	GO:0030054 ~ cell junction	18	16.36363636	1.65E-09
GOTERM_CC_DIRECT	GO:0030424 ~ axon	12	10.90909091	8.38E-08
GOTERM_CC_DIRECT	GO:0030425 ~ dendrite	12	10.90909091	4.85E-06
GOTERM_CC_DIRECT	GO:0048471 ~ perinuclear region of cytoplasm	15	13.63636364	1.80E-05
GOTERM_CC_DIRECT	GO:0043209 ~ myelin sheath	8	7.272727273	3.37E-05
GOTERM_MF_DIRECT	GO:0043274 ~ phospholipase binding	3	2.727272727	0.004900011
GOTERM_MF_DIRECT	GO:0004890 ~ GABA-A receptor activity	3	2.727272727	0.005455649
GOTERM_MF_DIRECT	GO:0005230 ~ extracellular ligand-gated ion channel activity	3	2.727272727	0.01597306
GOTERM_MF_DIRECT	GO:0005096 ~ GTPase activator activity	6	5.454545455	0.024201179
GOTERM_MF_DIRECT	GO:0015171 ~ amino acid transmembrane transporter activity	3	2.727272727	0.03101924

GBM: glioblastomas; GO: Gene Ontology; MDEGs: methylated-differentially expressed genes.

synapse, morphine addiction, retrograde endocannabinoid signaling, nicotine addiction, and synaptic vesicle cycle (Table 2 and Fig. 4D).

PPI Network Construction and Hub Gene Analysis

For hypomethylated/upregulated genes, the PPI network contained 48 nodes and 161 edges (Fig. 5A), while 76 nodes and 282 edges were established in the hypermethylated/downregulated genes networks (Fig. 5C). We next selected hub genes in the PPI networks using CytoHubba. The top 40 hub genes in two datasets were constructed using Cytoscape. The interactions of the top 40 hub hypomethylated/upregulated genes and hypermethylated/downregulated genes are presented in Fig. 5B, D, respectively.

Kaplan–Meier Analysis and Validation of Hub Genes

We investigated the prognostic values of identified hub genes using Kaplan–Meier analysis. For the complete cohort, increase in NMB mRNA was significantly associated

with favorable OS, while elevated CHI3L1, POSTN, S100A4, LOX, S100A11, IGFBP2, SLC12A5, VSNL1, and RGS4 mRNA levels were associated with unfavorable OS in GBM patients (Fig. 6). The remaining hub genes were insignificantly associated with OS (data not shown).

Additionally, we validated the relative mRNA expression of the 10 hub genes in GBM samples and control samples using GEPIA analysis. Compared to healthy brain tissues, CHI3L1, POSTN, S100A4, LOX, S100A11, NMB, and IGFBP2 mRNA levels were significantly higher, while SLC12A5, VSNL1, and RGS4 mRNA levels were significantly lower in GBM tissues (Fig. 7A). To validate the hub genes in cell lines, we also performed real-time quantitative PCR to validate the mRNA level in vitro (Fig. 7B). It was found that most mRNA levels of the survival related hub genes in tumor cells were consistent with the results in GEO and GEPIA database.

We analyzed the genetic alterations present in hub genes in GBM using cBioPortal. Among the datasets analyzed, the frequency of gene alterations, including mutations, fusions, amplifications, deep deletions, and multiple alterations ranged from 0% to 1%, with mutations, amplifications, and

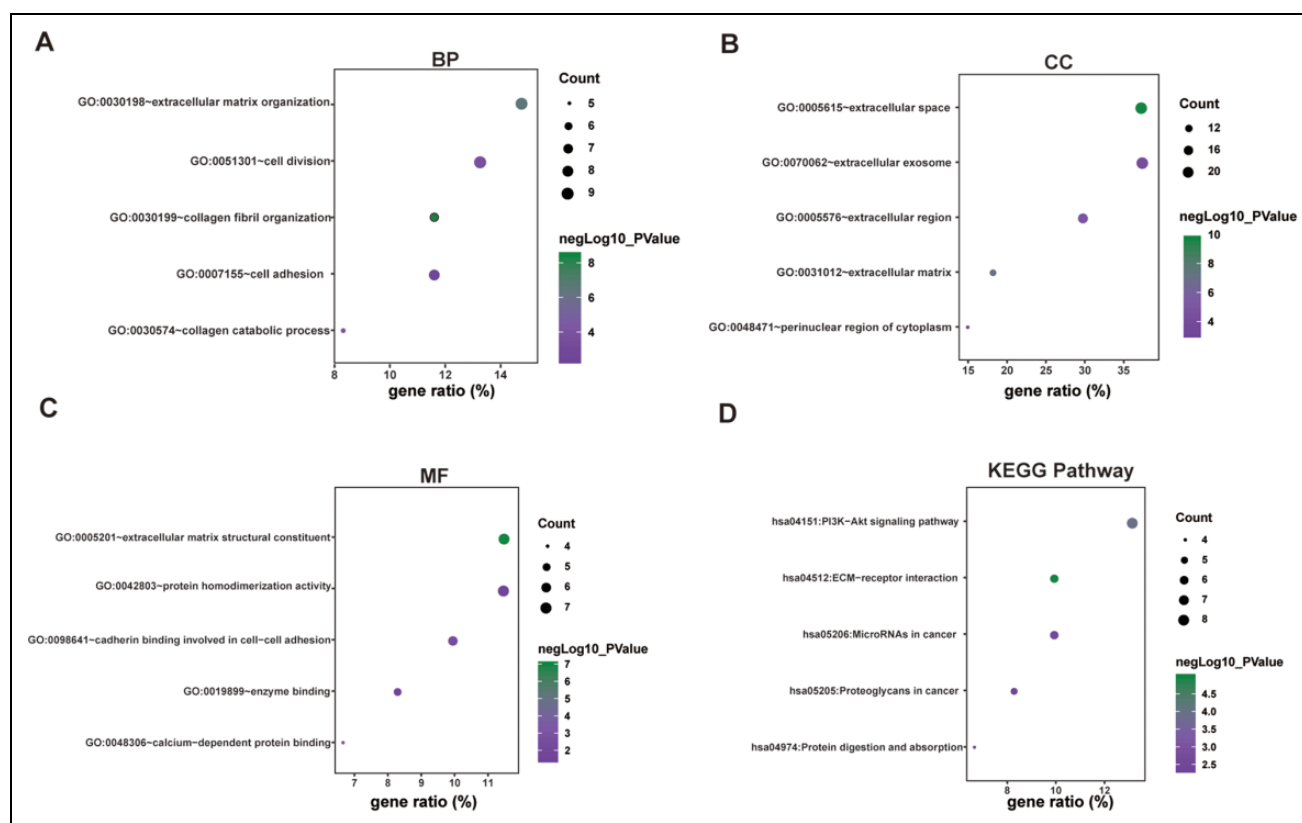


Figure 3. Functional enrichment analyses of hypomethylated/upregulated genes. (A) Significantly enriched GO terms based on biological process. (B) Significantly enriched GO terms based on cellular component. (C) Significantly enriched GO terms based on molecular function. (D) Significantly enriched KEGG pathway. GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes.

Table 2. KEGG Pathway Analysis of MDEGs Related With GBM.

Category	Term	Count	%	P-value
<i>Hypomethylation/high-expression genes</i>				
KEGG_PATHWAY	hsa04512: ECM-receptor interaction	6	10	1.01E-05
KEGG_PATHWAY	hsa04151: PI3K-Akt signaling pathway	8	13.33333333	1.24E-04
KEGG_PATHWAY	hsa05206: MicroRNAs in cancer	6	10	0.002657271
KEGG_PATHWAY	hsa04974: Protein digestion and absorption	4	6.666666667	0.003371441
KEGG_PATHWAY	hsa05205: Proteoglycans in cancer	5	8.333333333	0.004666276
<i>Hypermethylation/low-expression genes</i>				
KEGG_PATHWAY	hsa04727: GABAergic synapse	8	7.272727273	1.24E-06
KEGG_PATHWAY	hsa05032: Morphine addiction	8	7.272727273	1.98E-06
KEGG_PATHWAY	hsa04723: Retrograde endocannabinoid signaling	7	6.363636364	5.02E-05
KEGG_PATHWAY	hsa05033: Nicotine addiction	5	4.545454545	1.34E-04
KEGG_PATHWAY	hsa04721: Synaptic vesicle cycle	4	3.636363636	0.008398367

ECM: extracellular matrix; GBM: glioblastomas; KEGG: Kyoto Encyclopedia of Genes and Genomes; MDEGs: methylated-differentially expressed genes.

deep deletions being the most commonly observed alterations (Fig. 8A).

Following this, we verified the association between the methylation status and the expression of these hub genes (Fig. 8B). It was found that *CHI3L1*, *S100A4*, *LOX*, and *S100A11* mRNA levels were negatively associated with their corresponding methylation status, and were selected the further analysis. *IGFBP2*, *SLC12A5*, *VSNL1*, and *RGS4* mRNA

levels were not significantly associated with the methylation status, and the data of *POSTN* and *NMB* were unavailable in the cBioPortal.

ROC Curve Analysis

We further investigated the diagnostic effectiveness of the four identified hub genes (*CHI3L1*, *S100A4*, *LOX*, and

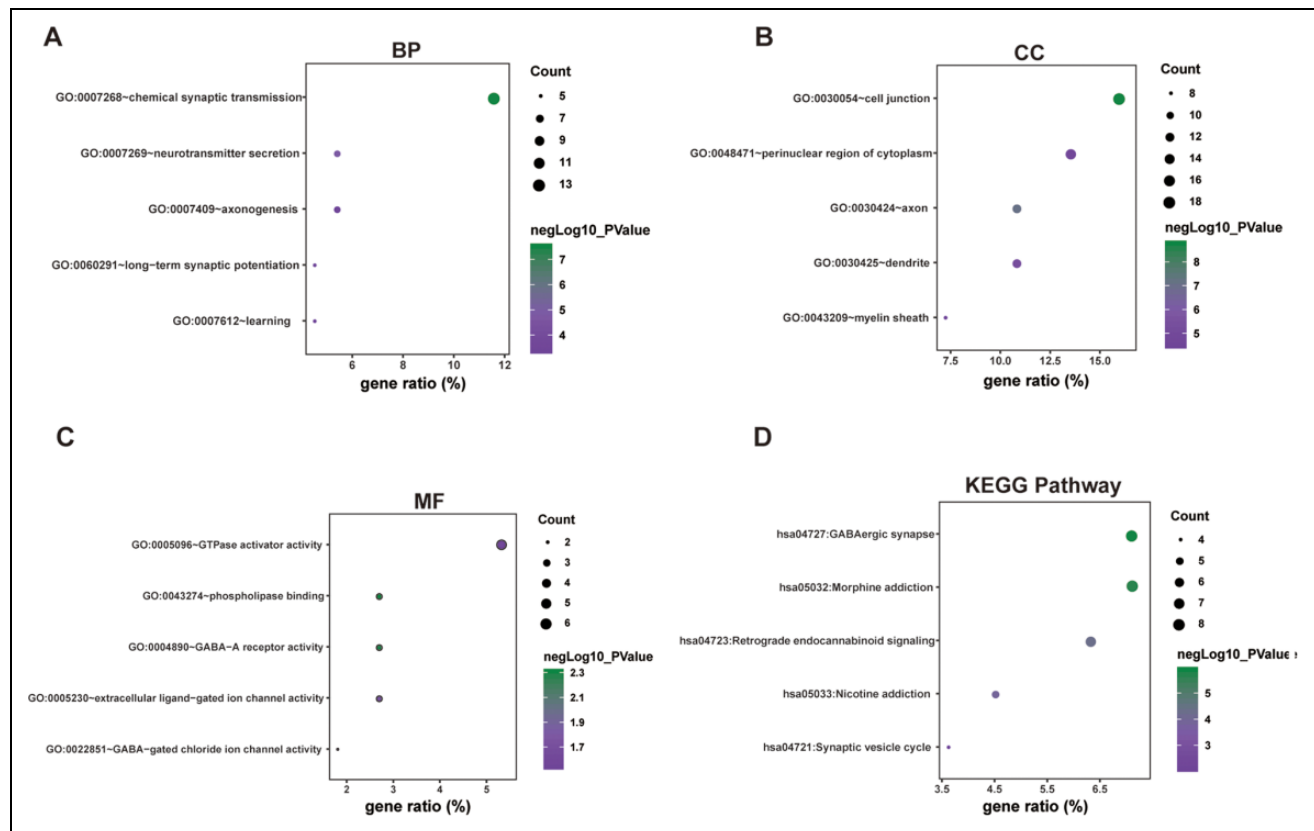


Figure 4. Functional enrichment analyses of hypermethylated/downregulated genes. (A) Significantly enriched GO terms based on biological processes. (B) Significantly enriched GO terms based on cellular components. (C) Significantly enriched GO terms based on molecular function. (D) Significantly enriched KEGG pathway. GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes.

S100A11) using ROC curves. ROC curve analysis revealed that the area under the curve (AUC) was 0.8032 (95% confidence interval [CI], 0.68–0.9264) for *CHI3L1*, 0.8439 (95% CI, 0.734–0.9538) for *S100A4*, 0.9751 (95% CI, 0.9383–1.012) for *LOX*, and 0.8824 (95% CI, 0.7878–0.9769) for *S100A11*. These results suggested that four hypomethylated/upregulated genes (*CHI3L1*, *S100A4*, *LOX*, and *S100A11*) can serve as prognostic biomarkers for distinguishing GBM patients from healthy controls (Fig. 9).

Discussion

DNA methylation can alter gene expression, and subsequently influence the development and progression of carcinomas, including GBM^{16,17}. With advancements in high-throughput technologies, including microarray and ChIP array, it has become feasible to analyze the expression levels. Identifying novel biomarkers of GBM will contribute to improvement in diagnostic, treatment, and prognostic assessment of GBM patients. In the present study, 60 hypomethylated/upregulated and 110 hypermethylated/downregulated overlapping genes, which may be associated with molecular mechanisms underlying important pathways associated with tumor formation, were identified in GBM tissues compared to normal brain

tissues from microarray datasets GSE50923, GSE50161, and GSE116520.

GO analysis indicated that the hypomethylation/high-expression genes were primarily enriched in functions associated with collagen fibril organization, extracellular space, and extracellular matrix structural constituent, indicating that hypermethylation primarily affected genes associated with the extracellular matrix. This finding is consistent with the fact that the extracellular matrix participates in pathophysiological process of GBM^{18,19}. The hypermethylation/low-expression genes were primarily enriched in functions associated with chemical synaptic transmission, cell junction, and phospholipase binding, suggesting that GBM tumorigenesis may be associated with these biological functions. The KEGG pathway analysis of hypomethylation/high-expression genes indicated that these were enriched in functions associated with extracellular matrix (ECM)-receptor interaction, PI3K-Akt signaling pathway, microRNAs in cancer, protein digestion and absorption, and proteoglycans in cancer. The hypermethylation/low-expression genes were primarily enriched in functions associated with GABAergic synapse, morphine addiction, retrograde endocannabinoid signaling, nicotine addiction, and synaptic vesicle cycle. Among the enriched KEGG pathways, several terms play

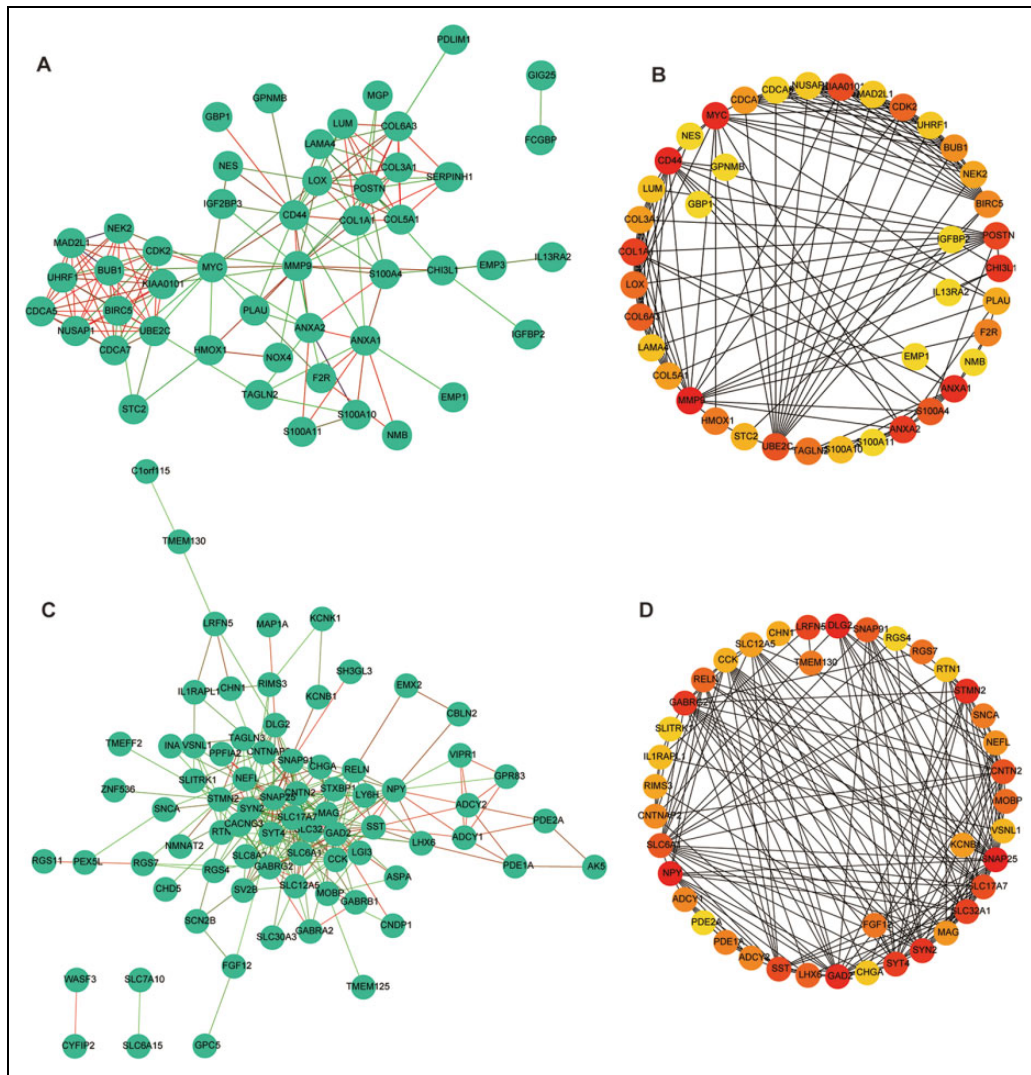


Figure 5. Top 40 hub genes identified in PPI networks. (A) PPI network of hypermethylation/low-expression genes. (B) Top 40 hub genes of the significant hypermethylation/low-expression genes. (C) PPI network of hypermethylation/low-expression genes. (D) Top 40 hub genes of the significantly hypermethylated/low-expression genes. PPI: protein-protein interaction.

crucial roles in tumorigenesis. For instance, ECM-receptor interaction pathway is involved in the proliferation and invasion of cancer cell²⁰. PI3K-Akt signaling pathway is one of the major cellular signaling pathways that play a significant role in basic intracellular functions and is a therapeutic target in human cancer²¹.

The PPI network constructed using identified MDEGs led to the identification of high-degree hub genes in the local network. Lastly, CHI3L1, POSTN, S100A4, LOX, S100A11, NMB, IGFBP2, SLC12A5, VSNL1, and RGS4 were identified as the key genes high connectivity degree through survival analysis. Increase in NMB mRNA was significantly associated with unfavorable survival, while elevated CHI3L1, POSTN, S100A4, LOX, S100A11, IGFBP2, SLC12A5, VSNL1, and RGS4 mRNA levels were associated with unfavorable OS in GBM patients. Moreover,

we validated the expression of these hub genes in GBM and observed that the results were consistent with our previous microarray data. We also validated the mRNA levels of the hub genes using the GEPIA database and real-time quantitative PCR, and confirmed that the expression levels of most of the genes matched the results from GEO database. However, the expression of IGFBP2 was not completely consistent with GEO data. This may be due to the difference between in vivo and in vitro studies. Among the hub genes, CHI3L1, S100A4, LOX, and S100A11 mRNA levels correlated negatively with their corresponding methylation status and were selected for further analysis. However, the negatively association between methylation and matched mRNA levels of IGFBP2, SLC12A5, VSNL1, and RGS4 were not be observed. This may be due to DNA methylation is not the main mechanism of regulation of these gene's expression. It

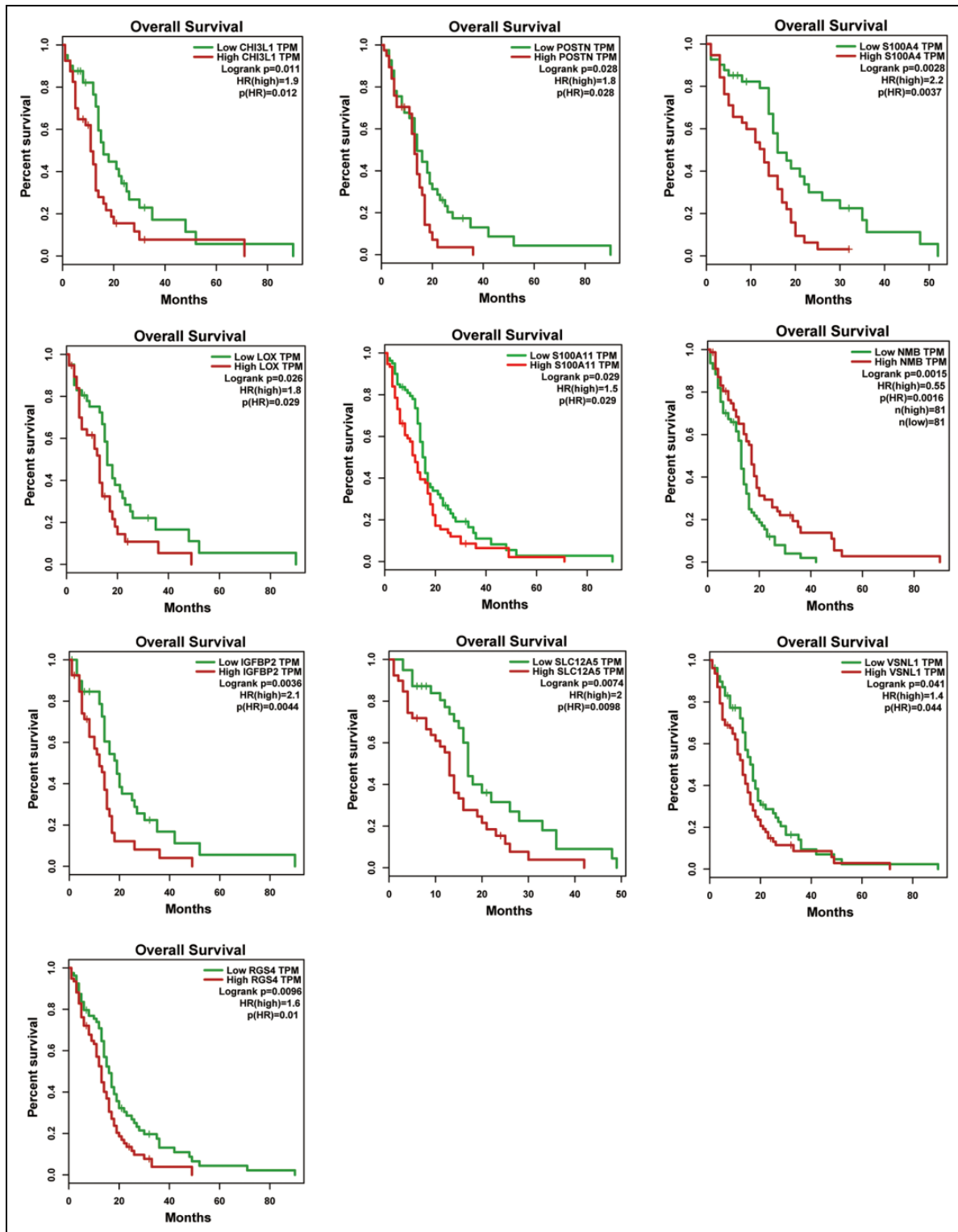


Figure 6. Identification of hub genes by prognosis analyses. The association between CHI3L1, POSTN, S100A4, LOX, S100A11, NMB, IGFBP2, SLC12A5, VSNL1, and RGS4 mRNA levels and overall survival in GBM patients. GBM: glioblastomas.

is widely acknowledged that epigenetic regulation is involved in tumor-related gene expression including transcriptional regulation (DNA methylation, histone

modification, chromatin remodeling, and X chromosome inactivation) and post-transcriptional regulation (microRNA or long noncoding RNA). These genes may be regulated by

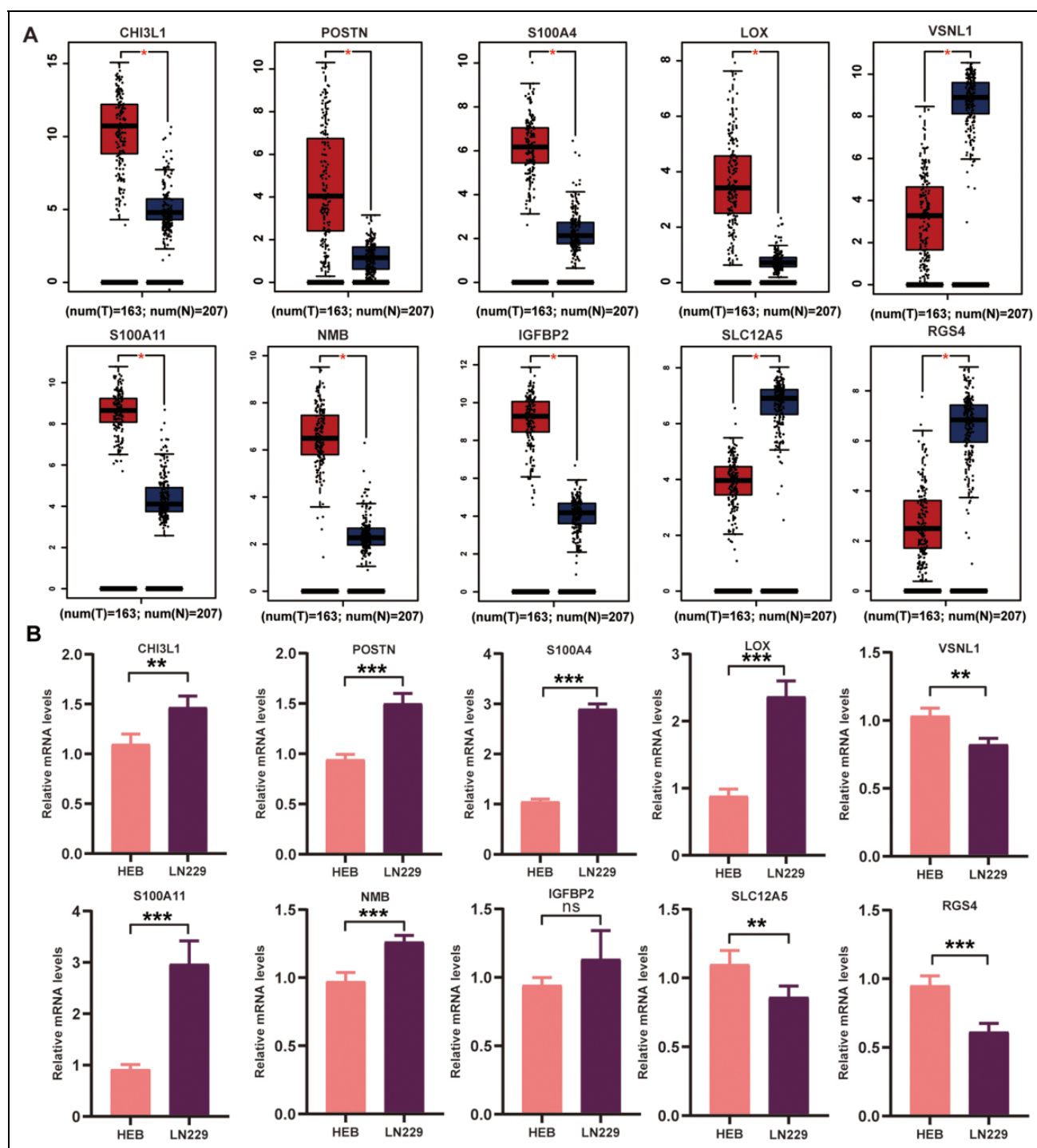


Figure 7. Validation of the mRNA levels of selected hub genes. CHI3L1, POSTN, S100A4, LOX, S100A11, NMB, IGFBP2, SLC12A5, VSNL1, and RGS4 mRNA expressions in GBM patients. (A) Validation in GEPIA database. (B) Validation in HEB cell lines (control) and LN229 cell lines (tumor). The data are expressed as the mean \pm SD of three independent experiments (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). GBM: glioblastomas; SD: standard deviation.

other epigenetic regulations. Furthermore, we also observed that CHI3L1, S100A4, LOX, and S100A11 methylation are potential diagnostic probes for GBM detection that are highly sensitive and specific.

CHI3L1, also known as YKL-40 or HC-gp39, plays a role in cell proliferation, differentiation, apoptosis, angiogenesis, inflammation, and extracellular tissue remodeling²². Numerous inflammatory and malignant diseases possibly

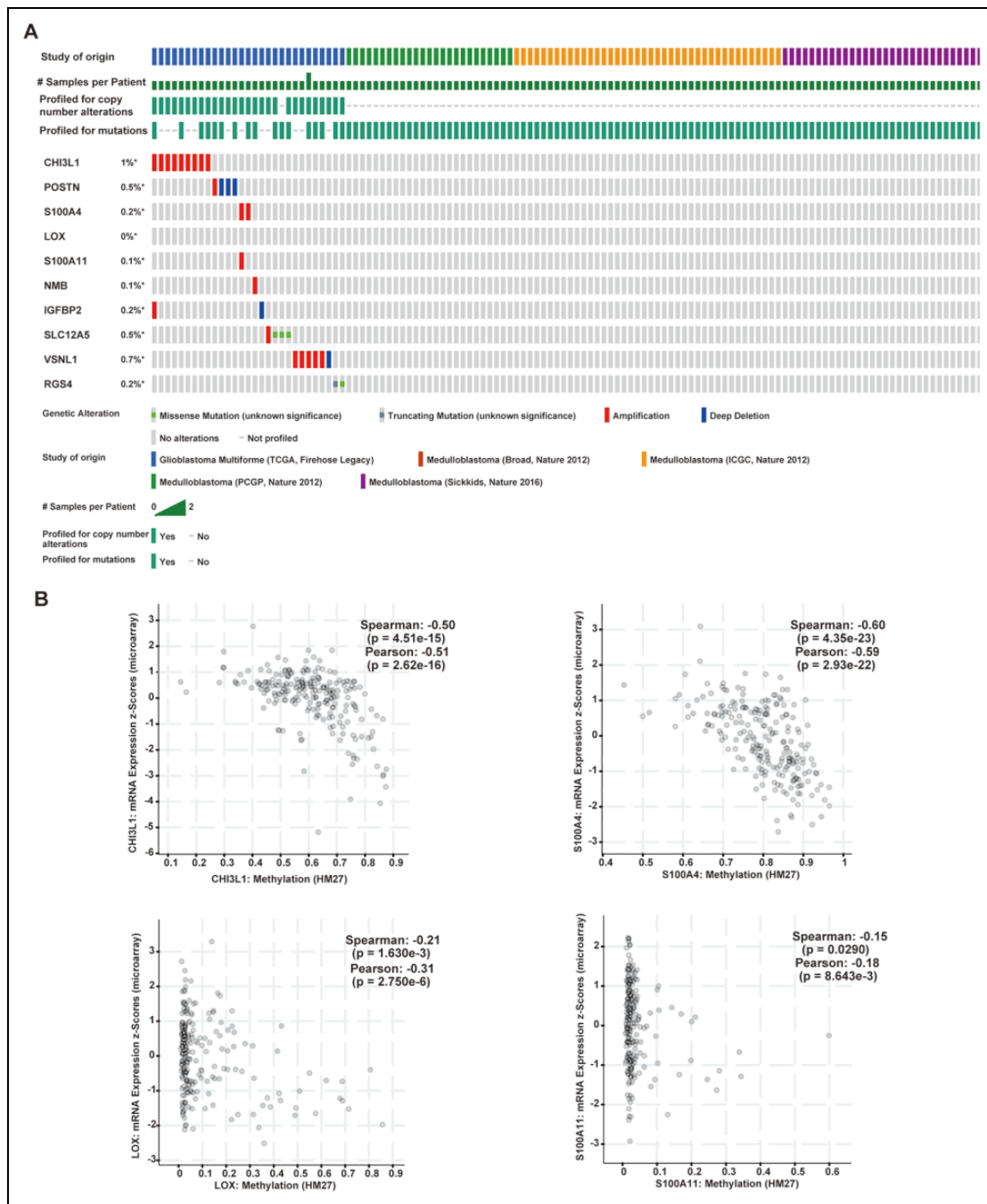


Figure 8. Genetic alteration of hub genes and the association between mRNA expression and DNA methylation. (A) A visual summary of alteration based on a query of the 10 hub genes. (B) Correlation between mRNA and methylation of CHI3L1, S100A4, LOX, and S100A11 in GBM patients. GBM: glioblastomas.

contribute to the serum levels of CHI3L1. It is also one of the significantly induced genes in GBM^{23,24}. However, the exact biological role of CHI3L1 in the GBM pathogenesis has not been elucidated yet. Hormigo's study demonstrated that it can be used as a predictor of survival of GBM patients²⁵. However, there are data that indicate that CHI3L1 has no significant impact on survival of GBM patients^{26,27}. However, majority of these studies had a small sample size. In our

study, we observed that mRNA expression of CHI3L1 was significantly higher in GBM tissues, and reduced CHI3L1 mRNA levels were associated with favorable OS. Moreover, ROC curve analysis indicated that it could serve as a valuable biomarker for distinguishing patients with GBM from healthy controls.

S100A4 is a gene encoding small calcium-binding proteins that interact with other proteins²⁸. A growing body of

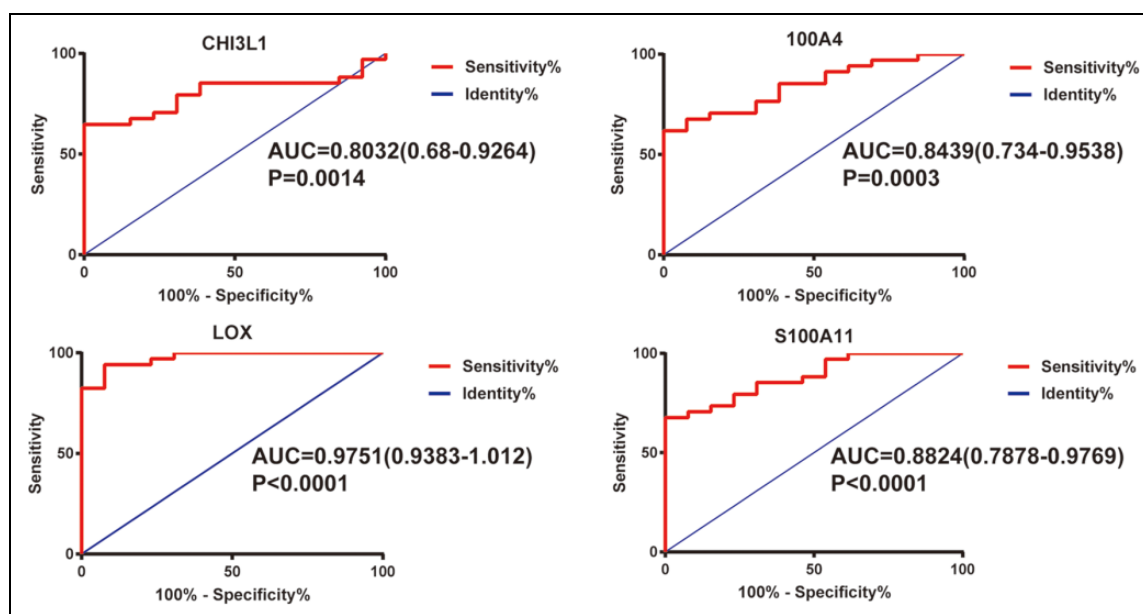


Figure 9. ROC curves of hub genes for GBM diagnosis. The X-axis represents false-positive rate, presented as “1-Specificity.” The Y-axis represents true positive rate, presented as “Sensitivity”. GBM: glioblastomas; ROC: receiver-operator characteristic.

evidence indicates that 100A4 plays a significant role in the progression and metastasis of various types of tumors, including breast cancer, colorectal cancer, cervical cancer, and lung cancer^{29–32}. S100A4 downregulation can suppress *in vitro* and *in vivo* GBM tumor progression, suggesting that S100A4 may be an effective therapeutic target³³. S100A4 has been regarded as a biomarker and regulator of glioma stem cells that is critical for mesenchymal transition in GBM³⁴. However, there are limited studies that have assessed the diagnostic and prognostic value of S100A4. In the current study, we observed that 100A4 could function as a highly specific and sensitive biomarker for GBM diagnosis and predict poor survival.

LOX is associated with crucial biological processes, including cell motility, cell signaling, and gene regulation. It is a secretory amine oxidase that plays a vital role in modifying the primary tumor microenvironment by cross-linking collagen and elastin in the ECM³⁵. Previous studies have revealed that the mRNA expression of LOX was significantly higher in GBM patients, and increased expression and activity of LOX had positive correlation with the malignant grade of astrocytomas³⁶. Consistent with the findings of this study, we observed that the mRNA expression of LOX was significantly higher in GBM tissues. In addition, we also observed that the increased LOX mRNA levels associated with unfavorable OS and LOX have diagnostic value in GBM treatment.

S100A11, a lesser-known protein of the S100 family, is upregulated in various cancers, including lung cancer, bladder cancer, and liver cancer^{37–39}. Previous studies have demonstrated that S100A11 may enhance proliferation and invasion of gastric cancer⁴⁰. To date, the significance of the

role of S100A11 in GBM remains undetermined. To further investigate the DNA methylation-based diagnosis and prognosis of S100A11, we conducted ROC curve analysis and survival analysis. The results indicated that S100A11 could function as a highly specific and sensitive biomarker for GBM diagnosis, and lower mRNA expression of S100A11 was associated with better survival, suggesting it could be used as a biomarker for GBM patients in the future.

In conclusion, we have developed 10 MDEGs expression that accurately predict survival and prognosis of GBM patients. Furthermore, CHI3L1, S100A4, LOX, and S100A11 provide helpful information for the selection of therapeutic strategies.

Data Availability Statement

The microarray data (GSE50923, GSE116520, and GSE50161) used to support the findings of this study are available in the Gene Expression Omnibus (GEO) repository (NIH, <http://www.ncbi.nlm.nih.gov/geo>). Also, the data used to support the findings of this study are included within the article.

Ethical Approval

Ethical approval is not applicable for this article.

Statement of Human and Animal Rights

This article does not contain any studies with human or animal subjects.

Statement of Informed Consent

There are no human subjects in this article and informed consent is not applicable.


Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Natural Science Foundation of China (Nos. 81760351 and 81460015).

ORCID iD

Zhenguo Zeng  <https://orcid.org/0000-0002-9864-7923>

Supplemental Material

Supplemental Material for this article is available online.

References

- Weller M, Wick W, Aldape K, Brada M, Berger M, Pfister SM, Nishikawa R, Rosenthal M, Wen PY, Stupp R, Reifenberger G. Glioma. *Nat Rev Dis Primers*. 2015;1:15017.
- Ostrom QT, Gittleman H, Farah P, Ondracek A, Chen Y, Wolinsky Y, Stroup NE, Kruchko C, Barnholtz-Sloan JS. CBTRUS statistical report: Primary brain and central nervous system tumors diagnosed in the United States in 2006-2010. *Neuro Oncol*. 2013;15(Suppl 2):ii1-ii56.
- Smoll NR, Schaller K, Gautschi OP. Long-term survival of patients with glioblastoma multiforme (GBM). *J Clin Neurosci*. 2013;20(5):670-675.
- Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, Ohgaki H, Wiestler OD, Kleihues P, Ellison DW. The 2016 world health organization classification of tumors of the central nervous system: a summary. *Acta Neuropathol*. 2016;131(6):803-820.
- Urbanska K, Sokolowska J, Szmidt M, Sysa P. Glioblastoma multiforme - an overview. *Contemp Oncol (Pozn)*. 2014;18(5):307-312.
- Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, Xie M, Zhang Q, McMichael JF, Wyczalkowski MA, Leiserson MDM, et al. Mutational landscape and significance across 12 major cancer types. *Nature*. 2013;502(7471):333-339.
- Dupont C, Armant DR, Brenner CA. Epigenetics: definition, mechanisms and clinical perspective. *Semin Reprod Med*. 2009;27(5):351-357.
- Dahlrot RH, Larsen P, Boldt HB, Kreutzfeldt MS, Hansen S, Hjelmberg JB, Kristensen BW. Posttreatment effect of MGMT methylation level on glioblastoma survival. *J Neuropathol Exp Neurol*. 2019;78(7):633-640.
- Liu J, Li H, Sun L, Wang Z, Xing C, Yuan Y. Aberrantly methylated-differentially expressed genes and pathways in colorectal cancer. *Cancer Cell Int*. 2017;17:75.
- Yang X, Gao L, Zhang S. Comparative pan-cancer DNA methylation analysis reveals cancer common and specific patterns. *Brief Bioinform*. 2017;18(5):761-773.
- Wang Y, Ruan Z, Yu S, Tian T, Liang X, Jing L, Li W, Wang X, Xiang L, Claret FX, Nan K, et al. A four-methylated mRNA signature-based risk score system predicts survival in patients with hepatocellular carcinoma. *Aging (Albany NY)*. 2019;11(1):160-173.
- Li H, Liu JW, Liu S, Yuan Y, Sun LP. Bioinformatics-based identification of methylated-differentially expressed genes and related pathways in gastric cancer. *Dig Dis Sci*. 2017;62(11):3029-3039.
- Yi L, Luo P, Zhang J. Identification of aberrantly methylated differentially expressed genes in breast cancer by integrated bioinformatics analysis. *J Cell Biochem*. 2019;120(9):16229-16243.
- Ma X, , Shang F, , Zhu W, , Lin Q. CXCR4 expression varies significantly among different subtypes of glioblastoma multiforme (GBM) and its low expression or hypermethylation might predict favorable overall survival. *Expert Rev Neurother*. 2017;17(9):941-946.
- Hegi ME, Genbrugge E, Gorlia T, Stupp R, Gilbert MR, Chinot OL, Nabors LB, Jones G, Van Criekinge W, Straub J, Weller M. MGMT promoter methylation Cutoff with safety margin for selecting glioblastoma patients into trials omitting temozolomide: a pooled analysis of four clinical trials. *Clin Cancer Res*. 2019;25(6):1809-1816.
- Gyorffy B, Bottai G, Fleischer T, Munkacsy G, Budczies J, Paladini L, Borresen-Dale AL, Kristensen VN, Santarpia L. Aberrant DNA methylation impacts gene expression and prognosis in breast cancer subtypes. *Int J Cancer*. 2016;138(1):87-97.
- Meng W, Jiang Y, Ma J. Is the prognostic significance of O6-methylguanine- DNA methyltransferase promoter methylation equally important in glioblastomas of patients from different continents? A systematic review with meta-analysis. *Cancer Manag Res*. 2017;9:411-425.
- Jiang X, Zhou T, Wang Z, Qi B, Xia H. HSP47 promotes glioblastoma stemlike cell survival by modulating tumor microenvironment extracellular matrix through TGF-beta pathway. *ACS Chem Neurosci*. 2017;8(1):128-134.
- Herrera-Perez M, Voytik-Harbin SL, Rickus JL. Extracellular matrix properties regulate the migratory response of glioblastoma stem cells in three-dimensional culture. *Tissue Eng Part A*. 2015;21(19-20):2572-2582.
- Tang X, Hou Y, Yang G, Wang X, Tang S, Du YE, Yang L, Yu T, Zhang H, Zhou M, Wen S, et al. Stromal miR-200 s contribute to breast cancer cell invasion through CAF activation and ECM remodeling. *Cell Death Differ*. 2016;23(1):132-145.
- Alzahrani AS. PI3K/Akt/mTOR inhibitors in cancer: at the bench and bedside. *Semin Cancer Biol*. 2019;59:125-132.
- Johansen JS, Schultz NA, Jensen BV. Plasma YKL-40: a potential new cancer biomarker? *Future Oncol*. 2009;5(7):1065-1082.
- Nigro JM, Misra A, Zhang L, Smirnov I, Colman H, Griffin C, Ozburn N, Chen M, Pan E, Koul D, Yung WK, et al. Integrated array-comparative genomic hybridization and expression array profiles identify clinically relevant molecular subtypes of glioblastoma. *Cancer Res*. 2005;65(5):1678-1686.
- Tanwar MK, Gilbert MR, Holland EC. Gene expression microarray analysis reveals YKL-40 to be a potential serum marker

- for malignant character in human glioma. *Cancer Res.* 2002; 62(15):4364–4368.
25. Hormigo A, Gu B, Karimi S, Riedel E, Panageas KS, Edgar MA, Tanwar MK, Rao JS, Fleisher M, DeAngelis LM, Holland EC. YKL-40 and matrix metalloproteinase-9 as potential serum biomarkers for patients with high-grade gliomas. *Clin Cancer Res.* 2006;12(19):5698–5704.
 26. Seroo NV, Delfino KR, Southey BR, Beever JE, Rodriguez-Zas SL. Cell cycle and aging, morphogenesis, and response to stimuli genes are individualized biomarkers of glioblastoma progression and survival. *BMC Med Genomics.* 2011;4:49.
 27. Boisselier B, Marie Y, El Hallani S, Kaloshi G, Iershov A, Kavsan V, Psimaras D, Thillet J, Hoang-Xuan K, Delattre JY, Sanson M. No association of (-131C→G) variant of CH13L1 gene with risk of glioblastoma and prognosis. *J Neurooncol.* 2009;94(2):169–172.
 28. Mishra SK, Siddique HR, Saleem M. S100A4 calcium-binding protein is key player in tumor progression and metastasis: pre-clinical and clinical evidence. *Cancer Metastasis Rev.* 2012; 31(1-2):163–172.
 29. Chen N, Sato D, Saiki Y, Sunamura M, Fukushige S, Horii A. S100A4 is frequently overexpressed in lung cancer cells and promotes cell growth and cell motility. *Biochem Biophys Res Commun.* 2014;447(3):459–464.
 30. Niu Y, Wang L, Cheng C, Du C, Lu X, Wang G, Liu J. Increased expressions of SATB1 and S100A4 are associated with poor prognosis in human colorectal carcinoma. *Apmis.* 2015;123(2):93–101.
 31. Liu M, Liu J, Yang B, Gao X, Gao LL, Kong QY, Zhang P, Li H. Inversed expression patterns of S100A4 and E-cadherin in cervical cancers: implication in epithelial-mesenchymal transition. *Anat Rec (Hoboken).* 2017;300(12):2184–2191.
 32. Ismail TM, Bennett D, Platt-Higgins AM, Al-Medhity M, Barraclough R, Rudland PS. S100A4 elevation empowers expression of metastasis effector molecules in human breast cancer. *Cancer Res.* 2017;77(3):780–789.
 33. Liang J, Piao Y, Holmes L, Fuller GN, Henry V, Tiao N, de Groot JF. Neutrophils promote the malignant glioma phenotype through S100A4. *Clin Cancer Res.* 2014;20(1): 187–198.
 34. Chow KH, Park HJ, George J, Yamamoto K, Gallup AD, Graber JH, Chen Y, Jiang W, Steindler DA, Neilson EG, Kim BYS, et al. S100A4 Is a biomarker and regulator of glioma stem cells that is critical for mesenchymal transition in glioblastoma. *Cancer Res.* 2017;77(19):5360–5373.
 35. Erler JT, Bennewith KL, Nicolau M, Dornhofer N, Kong C, Le QT, Chi JT, Jeffrey SS, Giaccia AJ. Lysyl oxidase is essential for hypoxia-induced metastasis. *Nature.* 2006;440(7088): 1222–1226.
 36. da Silva R, Uno M, Marie SK, Oba-Shinjo SM. LOX expression and functional analysis in astrocytomas and impact of IDH1 mutation. *PLoS One.* 2015;10(3):e0119781.
 37. Miyazaki M, Sakaguchi M, Akiyama I, Sakaguchi Y, Nagamori S, Huh NH. Involvement of interferon regulatory factor 1 and S100C/A11 in growth inhibition by transforming growth factor beta 1 in human hepatocellular carcinoma cells. *Cancer Res.* 2004;64(12):4155–4161.
 38. Memon AA, Sorensen BS, Meldgaard P, Fokdal L, Thykjaer T, Nexø E. Down-regulation of S100C is associated with bladder cancer progression and poor survival. *Clin Cancer Res.* 2005; 11(2 Pt 1):606–611.
 39. Woo T, Okudela K, Mitsui H, Tajiri M, Rino Y, Ohashi K, Masuda M. Up-regulation of S100A11 in lung adenocarcinoma - its potential relationship with cancer progression. *PLoS One.* 2015;10(11):e0142642.
 40. Koh SA, Lee KH. HGF-mediated S100A11 overexpression enhances proliferation and invasion of gastric cancer. *Am J Transl Res.* 2018;10(11):3385–3394.