Whole-transcriptome sequencing profiling identifies functional and prognostic signatures in patients with PTPRZ1-MET fusion-negative secondary glioblastoma multiforme

BAO-SHI CHEN1*, KUAN-YU WANG2-4*, SHU-QING YU1*, CHUAN-BAO ZHANG1, GUAN-ZHANG LI4, ZHI-LIANG WANG4 and ZHAO-SHI BAO1,3,4

Departments of 1Neurosurgery and 2Gamma Knife Center, Beijing Tiantan Hospital, Capital Medical University, Beijing 100050; 3Center of Brain Tumor, Beijing Institute for Brain Disorders, Beijing 100069; 4Department of Neurosurgery, Beijing Neurosurgical Institute, Capital Medical University, Beijing 100050, P.R. China

Received June 3, 2019; Accepted February 21, 2020

DOI: 10.3892/ol.2020.12049

Correspondence to: Dr Zhao-Shi Bao, Department of Neurosurgery, Beijing Tiantan Hospital, Capital Medical University, 119 South Fourth Ring Road West, Beijing 100050, P.R. China E-mail: baozhaoshi@bjtth.org

*Contributed equally

Abbreviations: ZM-fusion, PTPRZ1-MET fusion; PTPRZ1, protein tyrosine phosphatase receptor type Z1; MET, MET proto-oncogene receptor tyrosine kinase; SIRPD, signal regulatory protein δ; ZNF117, zinc finger protein 117; KIF11, kinesin family member 11; FHDC1, FH2 domain containing 1; DDAH2, dimethylarginine dimethylaminohydrolase 2; C7orf60, base methyltransferase of 25S rRNA 2 homolog; IDH, isocitrate dehydrogenase; sGBM, secondary glioblastoma multiforme; CGGA, Chinese Glioma Genome Atlas; GO, Gene Ontology; GSEA, Gene Set Enrichment Analysis; ROC, receiver operating characteristic; AUC, area under the curve; HR, hazard ratio; PFS, progression-free survival; OS, overall survival

Key words: glioma, secondary glioblastoma multiforme, PTPRZ1-MET fusion, risk score, immune response

Abstract. Gliomas are the most common type of primary brain tumor in adults with a high mortality rate. Low-grade gliomas progress to glioblastoma multiforme (GBM) in the majority of cases, forming secondary GBM (sGBM), followed by rapid fatal clinical outcomes. Protein tyrosine phosphatase receptor type Z1 (PTPRZ1)-MET proto-oncogene receptor tyrosine kinase (MET) (ZM) fusion has been identified as a biomarker for sGBM that is involved in glioma progression, but the mechanism of gliomagenesis and pathology of ZM-negative sGBM has remained to be fully elucidated. A whole-transcriptome signature is thus required to improve the outcome prediction for patients with sGBM without ZM fusion. In the present study, whole-transcriptome sequencing on 42 sGBM samples with or without ZM fusion from the Chinese Glioma Genome Atlas database identified mRNAs with differential expression between patients with and without ZM fusion and the most significant survival-associated genes were identified. A 6-gene signature was identified as a novel prognostic model reflecting survival probability in patients with ZM-negative sGBM. Clinical characteristics in patients with a high or low risk score value were analyzed with the Kaplan-Meier method and a two-sided log-rank test. In addition, ZM-negative sGBM patients with a high risk score exhibited an increase in immune cells, NF-κB-induced pathway activation and a decrease in endothelial cells compared with those with a low risk score. The present study demonstrated the potential use of a next-generation sequencing-based cancer gene signature in patients with ZM-negative sGBM, indicating possible clinical therapeutic strategies for further treatment of such patients.

Introduction

Gliomas are the most common type of neuroepithelial tumor of the central nervous system. Glioblastoma multiforme (GBM), referring to grade-IV glioma, is one of the deadliest of human cancer types. Based on the clinical course and molecular characteristics, GBMs may be classified into two subtypes (1-3). Primary GBMs refers to the vast majority of GBMs, which are thought to form de novo in the elderly, with a median overall survival of 15 months after maximal surgical resection, chemotherapy and radiotherapy (RT) (4). Secondary GBMs (sGBMs) typically progress from lower-grade tumors and affect younger patients. Since the Philadelphia chromosome was discovered in chronic myeloid leukemia in 1960, studies performed over the past six decades have identified fusion genes and proteins in a multitude of other types of cancer and through several different approaches (5). Fusion transcripts between fibroblast growth factor receptor 3 and transforming acidic coiled-coil containing protein 3 (6), and between the MYB proto-oncogene and the quaking homolog KH domain RNA binding protein were initially reported as the recurrent fusion transcripts in GBMs and pediatric gliomas, respectively (7). A previous study by our group identified a novel, recurrent fusion rearrangement involving the protein tyrosine phosphatase receptor type Z1 (PTPRZ1) and MET.
the gene signature. The risk score model was developed with survival. The differentially expressed genes associated between ZM-positive and -negative cases with P<0.05 used as a selection threshold. Univariate differentially expressed genes were identified and data were analyzed using a univariate Cox regression model in R language. A total of six mRNAs were selected to develop the risk score with random repeated sampling. ZM-negative patients with high risk scores had a relatively shorter overall survival (OS) time compared with those with low risk scores. The risk score was independent of clinical observations, including sex, age, isocitrate dehydrogenase (IDH) mutation status, O-6-methylguanine-DNA methyltransferase (MGMT) methylation status and therapeutic strategies. The prognostic value of the risk score on patients’ OS was higher compared with the aforementioned clinical information. The immune cell response was enhanced in ZM-negative patients with high risk scores. Increased macrophages rather than endothelial score and NF-kB enrichment were identified in patients with high risk scores. In summary, the risk score signature may be robust in predicting the survival of patients with ZM-negative sGBM and may help identify novel therapeutic targets for further treatment for patients with ZM-negative sGBM.

Materials and methods

Samples. In the present study, 42 tumor samples of confirmed sGBM were selected for the analysis. The samples were collected from January 2005 through to December 2018. The methods of the RNA sequencing were provided in a previous study (9). Our data used in the study were original samples to be used to generate the dataset. The samples were collected at Beijing Tiantan Hospital (Beijing, China) by our group and the RNA sequencing data were uploaded to the Chinese Glioma Genome Atlas (CGGA) RNA sequencing dataset (http://www.cgga.org.cn). For each patient, the following clinical information was collected: Sex, age, IDH status, MGMT promoter methylation status, RT and TMZ chemotherapy. P<0.05 was considered to indicate statistical significance.

Statistical analysis. In the present study, SPSS 16.0 (SPSS, Inc.), R software 3.2.5 (11) and GraphPad Prism 7.0 statistical software (GraphPad Software, Inc.) were used for data analysis and plotting. The R package ‘survival’ (http://www.biocconduct.org/) was used to analyze the most significant survival-associated mRNAs with the highest AUC value in the dataset using the receiver operating characteristic curve analysis. The percentage of macrophages and endothelial cells was calculated by EPIC in R (13). Differentially expressed genes between ZM fusion-positive and -negative cases were determined by Student’s t-test, which was also used to compare the risk scores between two groups. Survival analysis was performed using Kaplan-Meier survival analysis and univariate and multivariate Cox regression analysis. The analyzed factors included risk score, sex, age, IDH status, MGMT promoter methylation status, RT and TMZ chemotherapy. P<0.05 was considered to indicate statistical significance.

Results

Identification of survival-associated mRNAs. To characterize the transcriptomic RNA expression in sGBMs, RNA was extracted from 42 sGBM specimens with or without ZM fusion to perform whole-transcriptome sequencing. Student’s t-test was performed to identify differentially expressed genes between samples with and without ZM fusion (Fig. 1A). Univariate Cox regression was performed in the ZM-negative cohort and the genes associated with survival (P<0.05) were selected. The overlapping genes in the two analyses were selected for further analysis. Using the R package ‘survival’, the most significant survival-associated mRNAs with the highest AUC value in the dataset were identified (Fig. 1B), which yielded six genes: Signal regulatory protein δ (SIRPD), kinesin family member 11 (KIF11), zinc finger protein 117 (ZNF117), base methyltransferase of 25S rRNA 2 homolog (C7orf60), FH2 domain containing 1 (FHDC1) and dimethylarginine dimethylaminohydrolase 2 (DDAH2). Each of these genes was differentially expressed between ZM-positive and
negatively expressed genes in the ZM-negative tumor tissues and demonstrated lower expression levels in patients without ZM fusion (Fig. 1C). The HR of the survival analysis of each gene was also determined; the HRs of SIRPD, KIF11, ZNF117 and DDAH2 were positive, which indicated oncogenic characteristics in cell biological processes, whereas the HRs of C7orf60 and FHDC1 were negative, which indicated that these genes may suppress tumor occurrence (Fig. 1D).

**Clinical and genomic characteristics of risk score in patients with ZM-negative sGBM.** To investigate the clinical and genomic characteristics of the six selected genes, different weights were assigned to each gene for an independent risk score. The risk score of each patient was determined as follows: Risk score=(0.5368xSIRPD expression) + (0.6481 x KIF11 expression) + (2.1026 x ZNF117 expression) + (-0.6095 x C7orf60 expression) + (-0.4264 x FHDC1 expression) + (0.4152 x DDAH2 expression). The corresponding coefficients and P-values are presented in Table I. The high-risk score group was distinguished from the low-risk score group by the median score value as the cut-off value. Patients with ZM-negative sGBM with high risk scores exhibited poor prognosis, with significantly shorter OS times compared with those of patients with sGBM with low risk scores (median OS: High risk score, 233 days; low risk score, 572 days; P<0.0001, log-rank test; Fig. 2A). In addition, the progression-free survival (PFS) time in the high- and low-risk groups was evaluated and patients with ZM-negative sGBM in the high-risk group demonstrated
a significantly shorter PFS time compared with patients in the low-risk group, consistent with the trends of the OS time (Fig. 2B). The correlation between the risk score and mRNA expression was analyzed; the results indicated that the risk score may be a potential prognostic indicator (Fig. 2C).

**Clinical indications of risk score as an independent marker for prognosis.** The association between the risk score and clinical indicators, including sex, age, IDH status, MGMT promoter methylation status, RT and TMZ chemotherapy, was analyzed. The results demonstrated that the risk score was independent of sex, age, IDH mutation status, RT and TMZ chemotherapy, but associated with MGMT promoter methylation status (Fig. 3A). Uni- and multivariate Cox regression analysis was performed on these clinical factors. IDH status, risk score and TMZ chemotherapy were identified as potential prognostic markers by univariate Cox analysis (Fig. 3B). In the multivariate analysis, the risk score was the only factor identified as an independent prognostic marker in ZM-negative sGBMs (Table II). Therefore, the risk score was the most significant marker to indicate survival in patients with ZM-negative sGBM.

**Functional pathway annotation in patients with different risk scores.** To investigate the potential functional pathways activated in patients with high or low risk scores, GO analysis was performed on genes positively or negatively associated with the risk score. Genes enriched in ‘immune response’, ‘inflammatory response’, ‘positive regulation of T-cell proliferation’ and ‘cell division’ were significantly positively associated with the risk score in patients (Fig. 4A), whereas those enriched in relatively normal GO terms were negatively associated with the risk score in patients (Fig. 4B). To study the potential effect of immune cell differences in the present cohort, reference gene expression profiles of a major tumor-infiltrating immune cell type (macrophages) were established and expression profiles were derived from endothelial cells. The reference profiles were obtained as cell type averages from the single-cell RNA sequencing data of melanoma samples from Tirosh and colleagues (14). The reference cohort contained only samples from primary tumors and non-lymphoid metastasis. The reference gene expression profiles from each of the immune and endothelial cells were used to predict the expression profile of bulk tumor and to evaluate their association with the risk score of patients with ZM-negative sGBM. The results demonstrated that patients with sGBM without ZM fusion and a high risk score exhibited a significantly higher macrophage signature expression compared with that of patients with a low risk score (Fig. 5B). By contrast, patients with a high risk score exhibited a lower endothelial cell signature expression compared with that in patients with a low risk score (Fig. 5A), indicating that these pathways may induce immune cells and suppress endothelial cells in patients with a high risk score. GSEA was performed with the risk score ranging from low to high; ZM-negative patients with higher risk scores tended to exhibit a higher inflammatory response and tumor necrosis factor (TNF)-α signaling activation by NF-κB (Fig. 5C and D). These results may provide further therapeutic strategies for patients with ZM-negative sGBM.

**Discussion**

Based on whole-transcriptome sequencing on 42 sGBM samples from our own internal data, six genes were identified as an independent signature with the most significant survival-associated mRNAs and the best AUC in sGBMs without ZM fusion.

**SIRPD** is a member of a signal regulatory protein family involved in signal transduction and cell adhesion (15). No evidence has been confirmed in association with cancer initiation or progression.

The ZNF117 gene encodes a protein containing multiple C2H2-type zinc finger motifs. It has been demonstrated to participate in the maturation and differentiation of adipocytes and may control fat accumulation (16). The association between ZNF117 and human cancer requires further investigation.

**KIF11**, which is located on chromosome 22, encodes a kinesin-like motor protein that participates in chromosome positioning, centrosome separation and bipolar spindle formation during cell mitosis (17). Interactions between KIF11, MCF7 and phosphatase and tensin homolog regulate chromosome stability (18). KIF11 has been identified as a novel prognostic biomarker and a treatment target for various types of cancer, including lung squamous cell carcinoma, prostate cancer and chronic myeloid leukemia (19-22). In addition, previous studies demonstrated that KIF11 was an important regulator of activity in tumor stem cells of esophageal and colorectal cancer (23) and that upregulation of KIF11 was associated with high-grade astrocytoma (24). Targeting KIF11 also altered the cell fate and reduced glioma cell invasion (25). Thus, KIF11 may be a potential therapeutic target in glioma treatment.

The FHDC1 gene is located on human chromosome 4q31.3 (26) and encodes a protein product involved in the developmental stages of the mouse brain and the actin- and microtubule-dependent regulation of Golgi morphology of human cells (27). However, a limited number of studies have demonstrated an association between the FHDC1 gene and cancer development and progression. Further clinical studies on FHDC1-induced glioma initiation and progression require to be performed.

---

Table I. HR (95% CI) and P-values generated from the univariate Cox regression analysis in the dataset for each gene.

<table>
<thead>
<tr>
<th>Gene</th>
<th>HR</th>
<th>95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIRPD</td>
<td>1.711</td>
<td>1.025-2.855</td>
<td>0.040</td>
</tr>
<tr>
<td>KIF11</td>
<td>1.912</td>
<td>1.125-3.248</td>
<td>0.017</td>
</tr>
<tr>
<td>ZNF117</td>
<td>8.187</td>
<td>1.306-51.334</td>
<td>0.025</td>
</tr>
<tr>
<td>C7orf60</td>
<td>0.544</td>
<td>0.339-0.871</td>
<td>0.011</td>
</tr>
<tr>
<td>FHDC1</td>
<td>0.653</td>
<td>0.441-0.966</td>
<td>0.033</td>
</tr>
<tr>
<td>DDAH2</td>
<td>1.515</td>
<td>1.079-2.126</td>
<td>0.016</td>
</tr>
</tbody>
</table>

HR, hazard ratio; SIRPD, signal regulatory protein δ; KIF11, kinesin family member 11; ZNF117, zinc finger protein 117; C7orf60, base methyltransferase of 25S rRNA 2 homolog; FHDC1, FH2 domain containing 1; DDAH2, dimethylarginine dimethylaminohydrolase 2.
Figure 2. (A) Survival analysis of patients with ZM-negative sGBM in CGGA dataset. (B) PFS analysis in patients with ZM-negative sGBM in CGGA dataset. (C) The correlation between the risk score and the six gene expression. ZM, protein tyrosine phosphatase receptor type Z1-MET proto-oncogene receptor tyrosine kinase fusion; sGBM, secondary glioblastoma multiforme; CGGA, Chinese Glioma Genome Atlas; OS, overall survival; PFS, progression-free survival; SIRPD, signal regulatory protein δ; KIF11, kinesin family member 11; ZNF117, zinc finger protein 117; C7orf60, base methyltransferase of 25S rRNA 2 homolog; FHDC1, FH2 domain containing 1; DDAH2, dimethylarginine dimethylaminohydrolase 2.

Figure 3. (A) Risk score associated with different clinical characteristics of patients. The horizontal lines indicate the median risk score. The bars indicate the maximum and minimum risk score. The boxes indicated the 25-75% risk score. (B) Forest plot of univariate Cox analysis. HR, hazard ratio; IDH, isocitrate dehydrogenase; MGMT, O-6-methylguanine-DNA methyltransferase; NS, no significance; MUT, mutated; WT, wild-type; unmeth, unmethylated MGMT promoter; CGGA, Chinese Glioma Genome Atlas; sGBM, secondary glioblastoma multiforme; RT, radiotherapy; TMZ, temozolomide.
DDAH2 encodes a dimethylarginine dimethylaminohydro-
lase, which is important in nitric oxide generation by regulating
the cellular concentrations of methylarginines. DDAH2 has
been demonstrated to be associated with various types of
disease, including hypertension (28), type 2 diabetes (29) and
cancer. Upregulation of DDAH2 is associated with the inva-
siveness of lung adenocarcinoma by inducing proliferation and
capillary-like tube formation of vascular endothelial cells (30).
Further studies may focus on the therapeutic applications and
the tumorigenic mechanisms of DDAH2 in other types of
human cancer.

C7orf60, also known as the base methyltransferase of 25S
rRNA 2 homolog gene, was identified as a genomic event in
human complementary DNA (31,32) and associated with
smoking cessation (33) and Saccharomyces cerevisiae infec-
tion (34). However, C7orf60 has not been demonstrated to be
associated with cancer initiation or development.

In the present study, these six genes were assigned as an
independent signature with the most significant survival-asso-
cialed with secondary glioblastoma multiforme in the Chinese Glioma
Genome Atlas RNA sequencing database.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Univariate analysis</th>
<th>Multivariate analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR</td>
<td>95% CI</td>
</tr>
<tr>
<td>Age at diagnosis</td>
<td>1.006</td>
<td>0.966-1.047</td>
</tr>
<tr>
<td>Sex</td>
<td>1.203</td>
<td>0.563-2.569</td>
</tr>
<tr>
<td>MGMT</td>
<td>0.468</td>
<td>0.176-1.243</td>
</tr>
<tr>
<td>IDH</td>
<td>0.369</td>
<td>0.163-0.836</td>
</tr>
<tr>
<td>Radiotherapy</td>
<td>0.503</td>
<td>0.223-1.133</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>0.364</td>
<td>0.144-0.924</td>
</tr>
<tr>
<td>Risk score</td>
<td>3.116</td>
<td>2.032-4.777</td>
</tr>
</tbody>
</table>

All variables were considered as continuous variables. HR>1 indicates an increased risk, while HR<1 indicates a protective effect. HR, hazard ratio; IDH, isocitrate dehydrogenase; MGMT, O-6-methylguanine-DNA methyltransferase.

Figure 4. (A) Genes positively-associated with the risk score analyzed by Gene Ontology analysis in protein tyrosine phosphatase receptor type Z1-MET proto-oncogene receptor tyrosine kinase fusion-negative secondary glioblastoma multiforme. (B) Genes negatively-associated with the risk score analyzed by Gene Ontology analysis in ZM fusion-negative secondary glioblastoma multiforme.
nostic value independent of other clinical features, including age, gender, IDH mutation and MGMT promoter methylation status.

Over the past decades, the tumor microenvironment and a number of tumor-associated cell types, including endothelial cells, pericytes, immune inflammatory cells, cancer-associated fibroblasts and stem and progenitor cells of the tumor stroma, have increasingly been demonstrated to contribute to the biology of numerous tumors and to regulate signaling that controls their development and progression. Immune cells differ in their gene expression profiles depending on their state and site of origin (for example blood or tumors) (35). Tumor-associated endothelial cell phenotypes have been implicated in cancer development and tumor-associated angiogenesis (36-38). These endothelial cells participate in establishing lymphatic vessels (39) that serve as channels for the seeding of metastases commonly observed in a number of cancer types. Tumor-promoting inflammatory cells include various macrophage subtypes that serve diverse and crucial roles in giving rise to tumorigenesis and malignancy (40). Qian and Pollard (41) demonstrated that tumor-associated macrophages suppressed cytotoxic T-cell and natural killer cell activity, which have been independently identified as myeloid-derived suppressor cells.

The results of the present study demonstrated that ZM-negative sGBM patients with a high risk score exhibited significantly higher macrophage expression compared with that of patients with a low risk score, whereas patients with a high risk score exhibited a lower endothelial cell signature expression compared with those with a low risk score, indicating increased immune cells and suppressed endothelial cells in samples from patients with a high risk score. In addition, NF-κB-induced TNF-α signaling activation tended to be enriched in patients with a high risk score, implying potential therapeutic targets for treatment in these patients.

The major advantage of the present study was the use of next-generation sequencing and systematic analysis under multidimensional conditions. The results identified a risk score that was associated with immune response in patients with ZM-negative sGBM and may guide further clinical management. However, the extent of the immune response and the type of macrophage activation that may promote glioma progression in ZM-negative sGBM still requires to be clarified. The cellular components within the tumor microenvironment requires to be studied further to reveal their clinical implications. In addition to the requirement for exploration of the immunoregulatory role of glioma cells, the results of the present study require validation in independent cohorts. Another limitation of the
present study is the relatively low number of sGBM samples and future studies may be performed in other, larger datasets, including The Cancer Genome Atlas.

In conclusion, the present study demonstrated the clinical utility of next-generation sequencing-based cancer gene profiling in sGBM. The results revealed the diagnostic value of a six gene-based risk score signature and clinical features of patients with ZM-negative sGBM. An increase in immune cells and a decrease in endothelial cells were identified in patients with a high risk score; potential therapeutic strategies were identified as NF-κB-induced TNF-α signaling activation in patients with ZM-negative sGBM with a high risk score.

Acknowledgements

Not applicable.

Funding

The present study was funded by the National Natural Science Foundation of China (grant no. 81502606), Beijing Municipal Administration of Hospitals’ Youth Program (grant no. QML20160502), General Technological Program of Peking Education Commission (grant no. KM201810025023) and Beijing Nova Program (grant no. Z17110000117022).

Authors’ contributions

BC and KW designed the study and drafted the manuscript. SY revised the manuscript. SY, CZ, GL and ZW performed the data analyses. ZB contributed to the conception of the study and drafted the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

The present study was approved by the Beijing Tiantan Hospital institutional review board (Beijing, China) and informed consent was obtained from all individual participants included in this study.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References


