Prognostic Biomarker SYK and its Correlation with Immune Infiltrates in Glioma

Pei Liu (liupei970268117@qq.com)  
Qingdao University  
https://orcid.org/0000-0002-8664-9892

Jiamin Guo  
Qingdao University

Xiaoxiao Xu  
Qingdao University

Haixin Sun  
Qingdao University

Zheng Gong  
Qingdao University

Research

Keywords: TME, TCGA, glioma, SYK, ESTIMATE, CIBERSORT, TICs

Posted Date: August 31st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-839283/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

Background: Tumor microenvironment (TME) has great effects on the development process of glioma, and we sought to identify effective prognostic factors by analyzing data from patients with glioma. In this paper, CIBERSORT and ESTIMATE calculations were employed to figure up the ratio of tumor-infiltrating immune cells (TICs) and the quantity of immune and stromal components in 698 glioma dates from The Cancer Genome Atlas (TCGA) database. In addition, differentially expressed genes (DEGs) were studied by Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis, and single genes associated with prognosis were identified by PPI network and COX combined analysis.

Results: Immune and stromal scores of TME were significantly correlated with glioma patient survival. Through protein–protein interaction (PPI) network and regression analysis of COX, we finally determined that SYK was the best prognostic factor for patients with glioma. Gene Set Enrichment Analysis (GSEA) and CIBERSORT analysis were also employed, with the former showed that high-expression SYK group's genes are principally enriched immune-related activities and the latter revealed that SYK expression was positively associated with T cells CD4 memory resting and Monocytes. All the above experimental analyses provided the theoretical basis for the biological prediction of SYK.

Conclusions: SYK contributes to immune predictors in glioma patients by facilitating the shift of TME from immune dominance to metabolic activity, which provides promising insights into the treatment of glioma.

Background

Of all intracranial tumors, the incidence of glioma is generally high, approximately accounting for 80%[1]. Besides glioma is difficult to be completely removed by surgery, and the prognosis is relatively poor and easy recurrence. As a result of age, gender and other reasons of the difference, the incidence of glioma is different[2]. In spite of the best efforts at multimodal therapy such as surgical excision, complete resection was considered impossible due to the large amount of brain tissue involved. And the recurrence rate of glioma is close to 100%, which is related to the aggressiveness of glioma itself and the special pathophysiological characteristics of the central nervous system[3], and glioma is highly resistant to chemotherapy, so precise gene targeted therapy is an important direction for future treatment.

More and more evidences show the importance of TME in tumor growing trend[4, 5]. The synergistic interaction between cancer cells and their support cells leads to cancer’s immortal proliferation, resistance to apoptosis, and avoidance of immune surveillance, all of which are malignant phenotypes of cancer. Thus, the treatment response and clinical outcome of cancer patients are significantly affected by TME[6, 7]. The structural components of TME mainly consist of two parts: host stromal cells and recruited immune cells. Studying the interaction of TME can combine tumor cells with the microenvironment mechanism, providing biomarkers and potential drug targets for biotherapy[7]. Unlike tumor cells that
constantly mutate, TME mesenchyme is a genetically stable therapeutic target. One particularly interesting stromal element named macrophages, which are tumor-associated monocytes. Infiltrating immune cells, which are mainly composed of macrophages and monocytes, are deemed to have tumor-promoting and immunosuppressive function[8]. Tumor-infiltrating lymphocyte (TIL) have remarkable related to with glioma of the 5-year survival rate, a lot of existing research has indicated that the TICs of TME is a prospective indicator of curative effect[9].

In the article, We analyzed all samples of glioma in the TCGA database by utilizing the ESTIMATE and CIBERSORT algorithms, the ratio of TIC as well as the ratio of immune and stromal components were acquired. Next, we used PPI network build and Cox regression analysis to analyze the differentially expressed genes (DEGs) in normal samples and tumor samples, and the underlying association between DEGs and immune-related activities was investigated by GO and KEGG enrichment analysis. With these data, we developed an optimal biomarker for biological prediction as a prognostic indicator called spleen tyrosine kinase (SYK). SYK is a non-receptor type of PTKs in the Src family, which is a non-receptor tyrosine kinase. SYK has been widely reported in many hematopoietic malignancies and some epithelial primary tumors as a pro-survival factor and it has been detected in many cells, such as melanocytes, human nasal fibroblasts, and liver cells, and so on[10–12]. Consequently, SYK is likely to be involved in the pathogenesis of glioma, and the DEGs between immune components and stromal components in glioma samples are carefully studied.

Result

1. There Was an Association in Survival and Score in glioma Patients

To ascertain the association between immune and stromal estimates ratio and survival rate, survival analysis was demonstrated for ImmuneScore, StromalScore, as well as ESTIMATEScore, using the Kaplan–Meier survival analysis. The count of immune or stromal components of TME were expressed as an estimate in ImmuneScore or StromalScore. In comparison with the median, glioma patients were segmented into high and low groups. The results in Figure 1 showed, the scores of immune cells and stromal cells content were markedly connected with the survival in glioma patients. In other words, these revealed that the immune, stromal and estimate components of TME were more appropriate indicators of prognosis in patients with glioma(p<0.01).

2. The Score Was Correlated with Clinicopathologic Stage in glioma Patients

Clinical information on glioma patients in TCGA database was obtained to ascertain the association in the ratio of immune and stromal components with clinicopathological trait. The figure2 showed that immune score and stromal score were markedly positively association in tumor age and grade (Figure2A,
Figure2C, Figure2D, and Figure2F, p< 0.001). From the results, we can see that the quantity of immune and stromal components was relevant to glioma evolution, such as age and grade, but there was little correlation with gender.

3. The DEG Shared by The ImmuneScore and StromalScore Were Principally Enriched in Immune-related Genes

We performed comparative analyses of high-score and low-score samples to ensure the definite variation in the gene profiles of immune and stromal components of TME. Contrasted to the median, we got 1655 DEGs in the high-score and low-score samples of Immunescore. There has 1026 up-regulated genes and 629 down-regulated genes in all genes (Figure3A, Figure3C, and Figure3D). Analogously, 1813 differential genes were acquired in StromalScore, with 1204 up-regulated genes and 609 down-regulated genes (Figure3B, Figure3C, and Figure3D). The intersecting point analysis of the Venn plot displayed that 950 up-regulated genes overlapped in the aggregate in ImmuneScore and StromalScore analysis, with 488 down-regulated genes overlapped in ImmuneScore and StromalScore. These genes (total 1438 genes) may be influencing the status of TME. At the GO enrichment analysis, we can conclude the DEGs was majority consistent with GO term in relation to immunity. For instance, neutrophil activation, neutrophil mediated immunity and leukocyte migproportionn (Figure 3E). Similarly, Neuroactive ligand-receptor interaction, Cytokine-cytokine receptor interaction and Tuberculosis enrichment were shown at the Kyoto Encyclopedia of KEGG enrichment analysis in the same way (Figure 3F). Hence, this entire function in DEGs seemingly to be connected with immune-related activities, which manifests that the exist of immune factors has an influence for certain in TME of glioma.

4. Intersection Analysis of PPI Network and Univariate COX Regression

According to STRING database, the PPI network was built by using Cytoscape software (3.7.2). Figure 4A displayed the interaction between genes (confidence 0.95), and the bar chart shown the previous 30 genes permutated by the count of nodes (Figure 4B). We intersected the 30 genes in PPI and the genes with a pvalue of <0.05 by the univariate cox regression, and screened out thirty genes (Figure 4C). To determine the value of risk among the 30 factors, we performed univariate Cox regression analysis on glioma patients. Figure 4D shows that only a few genes have low risk values.

5. Correlation Between Survival Analysis and SYK Expression in Patients with glioma

An interesting gene, SYK, is thought to regulate the growth of epithelial cells in human breast cancer. Research found out that SKY has a great relationship with the occurrence and development of tumors [12,
We analyzed the data obtained from TCGA database that compared with normal tissues, SKY expression in the tumor group was significantly increased (p<0.001) (Figure 5A), and survival analysis of SKY displayed that the survival rate of glioma patients with low SKY expression was observably higher compared to the glioma patients with high SKY expression, the entire survival of tumor patients was observably reduced (Figure 5B). Multivariate Cox survival analysis confirmed that high SKY expression level was an autocephalous predictor of undesirable prognosis in glioma patients (p <0.001). The results show in figure 5C that SKY is a high-risk gene in glioma. These results indicated that SKY significant high expression in glioma samples than in normal samples, and the expression of SKY was an autocephalous predictor of undesirable prognosis in glioma patients.

6. SYK May Be an Indicator of TME regulation

Since SYK levels were inversely associated with the survival in glioma patients, we divided the data of glioma patients into high group of expression and low group of expression and severally contrasted them to the median level of SYK expression in GSEA. As we can see from Figure 6A, SYK high expression group principally enriched immune-related activities. For example, B cell receptor signaling pathway, hematopoietic cell lineage as well as autoimmune thyroid disease. In contrast, there was almost no gene sets enrichment in the low expression group of SYK. The afore-mentioned consequences revealed that SYK can be a potential indicator possibly of TME status.

7. Correlation Between SYK Expression and the Ratio of TICs

To further verify the association in the expression of SYK and the immune microenvironment, we analyzed the ratio of tumor infiltrating immune subsets by utilizing CIBERSORT algorithm and established a spectrum of 22 immune cells in glioma samples (Figure 7AB). Immune cells difference analysis results show that there are 6 different immune cells were related to the expression of SYK (Figure 7C), Immune cell correlation analysis results show that there are 9 different immune cells were related to the expression of SYK (Figure 7D). The difference analysis and correlation analysis show that there are 4 different immune cells that were associated with the SYK expression. Among them, two kinds of immune cells were positively correlation to the SYK expression, comprising CD4 memory resting T cell and Monocytes and two different types of immune cells were inversely associated with the SYK expression, including T cells follicular helper and Macrophages M0. All of these results were reflected ulteriorly that the part of SYK in modulating immune response and playing a key part in TME.

Discussion

Through this article, we used data derives from the TCGA database to determine TME-related genes which were related to survival in patients with glioma. Through Cox regression analysis, SYK was finally
confirmed in DEGs to be closely associated with survival of glioma patients. SYK was determined to participate in immunization activities. In the end, we discovered that SYK may be a preponderant target of TME status for glioma patients by a series of bioinformatics analyses.

In the process of tumor occurrence and development, TME is always in effect from beginning to end. Therefore, exploring the latent remedial targets of TME refactoring and further promoting the change of TME from tumor-friendly to tumor-suppressive is of value for carefully research. From previous studies, we can know that the occurrence of tumors is strongly associated with the immune microenvironment. In the transcriptional analysis of glioma data by downloaded from the TCGA database, we found that the immune component of TME has a certain effect on the postoperative predictive outcomes of patients. In particular, the evolve in glioma (such as intrusion and transfer) is related to the quantity of immune and stromal components in TME. This above-mentioned analysis results, which indicate the importance of investigating the mutual effect in tumor cells and immune cells, offer new horizons to the development of more promising treatment options.

Through systematic analysis, we finally screened the most significant gene SYK for the survival of glioma. SYK is a pivotal section in immune cell signaling pathways as it modulating proliferation, differentiation, and cell survival by activating triggers a series of signal transduction pathways as an element of immune receptor signal transduction[16, 17]. Also, SYK is a crucial part of the B lymphocyte signal receptor[18], it can regulate a variety of biological functions of B lymphocytes, and it is closely connected with the activation and maturation of B cells[11]. Be short of spleen tyrosine kinase, an expression product of SYK involved in TME’s immune SYK, leads to impaired development and maturation of immune cells, and in severe cases, to severe combined immunodeficiency disease (SCID). Thus to the mutation, abnormal proliferation of cells lose immunity, leading to the occurrence of tumor[19]. Up to now, there have been a number of studies on SYK inhibitors[20, 21]. Entospletinib, which is an inhibitor of SYK, has shown a towardly results in clinical trials in the treatment of B-cell malignancies, also several oral SYK inhibitors has already being evaluated in clinical trials, containing fostamatinib (R788), entospletinib (GS-9973), and TAK-659[22–24]. It is revealed that SYK inhibition can block the propagation and migration in glioma cells in vitro[25]. Furthermore, as report goes, SYK may be referred to the regulation of macrophage polarization in TME[26]. Hence, we made a deeper analysis of the relationship in SYK expression and TME. Immune-related signaling pathways were identified by GSEA analysis. For instance, FC gamma R mediated phagocytosis, B cell receptor signaling pathway, and hematopoietic cell lineage, were both benefications markedly in the SYK high-expression group. Violin plot showed that T cell CD4 memory resting, Monocytes and Macrophages M2 in SYK high-expression group were higher than SYK low-expression group, which revealed the expression of SYK was closely associated with immune cells in the TME. Macrophages are roughly segmented into M1 and M2 genres according to their functions. The M1 macrophage was participated in the inflammatory response, pathogen clearance, and antitumor immunity. Nevertheless, the M2 macrophage was different from M1, it influenced an anti-inflammatory response, wound healing, and pro-tumorigenic properties[27, 28]. As can be seen from the violin diagram in this article that M1 was lower and M2 was higher in the high-expression group of SYK, further supporting the chance that SYK take part in tumor-promoting of glioma.
Conclusions

Current studies suggest that SYK may affect the proliferation and migration of glioma cells by affecting B cell receptor signaling pathways and hematopoietic cell lineage. In summary, SYK is a potential cancer-promoting gene, and it is expected to be a new target for the cure of glioma and afford new ideas for the therapy of glioma. Therefore, we need to further investigate the accuracy of the combined analysis to define SYK expression, tumor-infiltrating B-cell subtypes, and mutagen-driven patterns before glioma patients are treated with SYK inhibitors.

Methods

1. Data Preparation

The data of transcriptome RNA-seq were available through the GDC portal of TCGA GDC official website(https://portal.gdc.cancer.gov/). There has an aggregate number of 698 glioma patients with mRNA expression information.

2. Acquisition of ImmuneScore, StromalScore, and ESTIMATEScore

Use the estimate package[13] in the R language to compute the ratio of immune-stromal component of each sample in TME. There was expressed in three different forms, including ImmuneScore, StromalScore and ESTIMATEScore, which were positively related to the proportion of immune, stromal, and the sum of the two, respectively, indicating that as the score for each component increased, their corresponding component accounted for a larger proportion of the TME.

3. Survival Analysis

Survival analysis was performed using R language (survival and survminer packages had been loaded), use the ggsurvplot function to plot the survival curve. Of a total of 698 tumor samples, 510 had detailed survival records ranging from 0 to 12 years can be chosen to conduct survival analysis. Besides, the Kaplan-Meier method was used to depict the survival spline, and the logarithmic rank was utilized to test for statistical significance; p < 0.05 was considered meaningful.

4. Discrepancy Analysis of Scores with Clinic–Pathological Stages

The getatable clinical data of the glioma samples can be obtained from TCGA database. Data is analyzed by R of version 4.0.1. Besides, Wilcoxon rank sum or Kruskal–Wallis rank sum test was
regarded as the importance test according to the count of clinical stages for comparison.

5. Heatmaps

Generate heatmaps of DEGS by using the R language with package pheatmap. In which way, the results can be divided into high expression group and low expression group.

6. Analysis of DEGs in ImmuneScore and StromalScore Between High-Score Group and Low-Score Group

The ImmuneScore and StromalScore were compared with median scores, severally, and we marked 698 tumor samples with high score or low score. R language with Limma package was applied to analyze the differentiation of gene expression, and DEGs were obtained by comparing the high score samples with the low score samples. We converted the log2 values obtained from the high-score vs the low-score groups of DEGs into decimals, and fold change (FC) of the result larger than 1 and false discovery rate (FDR)<0.05 were seen as meaningful.

7. Analysis of GO and KEGG Enrichment Analysis

The DAVID database(http://david.ncifcrf.gov) merges biological data and analysis tools that can be used for genetic difference analysis as well as pathway enrichment. A total of 1438 DEGs were outlined with GO and KEGG enrichment analysis by using clusterProfiler, enrichplot, and ggplot2 package in R language. Only the results with both p- and q-value of <0.05 were supposed dramatically enriched.

8. PPI Network and COX Regression Analysis

Used the STRING database to build PPI network, and then the Cytoscape version 3.7.2 was used to rebuild the network. The network was constructed by using nodes with confidence of interaction higher than 0.95. Besides, the intersection of the previous 30 genes in the PPI network was taken as an example, and univariate Cox regression study was analyzed by using the survival packet R language.

9. Gene Set Enrichment Analysis

GSEA is a reliable and also practical gene analysis method[14], which confirms the correlation between the target gene we need and the signal transduction pathway by ranking the high and low expression status of the target gene. GSEA analysis included the whole transcriptome of entire tumor samples, and just gene sets with NOM p < 0.05 and FDR q < 0.06 were regarded as dramatically.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumor microenvironment</td>
<td>TME</td>
</tr>
<tr>
<td>tumor-infiltrating immune cells</td>
<td>TICs</td>
</tr>
<tr>
<td>The Cancer Genome Atlas</td>
<td>TCGA</td>
</tr>
<tr>
<td>differentially expressed genes</td>
<td>DEGs</td>
</tr>
<tr>
<td>Gene Ontology</td>
<td>GO</td>
</tr>
<tr>
<td>Kyoto Encyclopedia of Genes and Genomes</td>
<td>KEGG</td>
</tr>
<tr>
<td>protein–protein interaction</td>
<td>PPI</td>
</tr>
<tr>
<td>Gene Set Enrichment Analysis</td>
<td>GSEA</td>
</tr>
<tr>
<td>Tumor-infiltrating lymphocyte</td>
<td>TIL</td>
</tr>
<tr>
<td>differentially expressed genes</td>
<td>DEGs</td>
</tr>
<tr>
<td>spleen tyrosine kinase</td>
<td>SYK</td>
</tr>
<tr>
<td>fold change</td>
<td>FC</td>
</tr>
<tr>
<td>false discovery rate</td>
<td>FDR</td>
</tr>
</tbody>
</table>

### Declarations

**Ethics approval and consent to participate**

Not applicable

**Consent for publication**

Not applicable

**Availability of data and materials**

The datasets generated during and/or analysed during the current study are available in the TCGA database (https://portal.gdc.cancer.gov/)

**Competing interests**

The authors declare that they have no conflicts of interests
Funding

Not applicable

Authors' contributions

PL designed and performed the experiments. XX and JG analyzed and interpreted the data. PL searched the literature and wrote the manuscript. HS and ZG revised the manuscript. All the authors reviewed and approved the final manuscript.

Acknowledgements

Not applicable

References


Determine Lung Cancer Cell Fate and Shape the Tumor Immune Microenvironment. Immunity. 2018;49:764–79.


**Figures**

**Figure 1**

Association between score and survival in glioma patients. (A) Kaplan–Meier survival spline of ImmuneScores with \( p<0.01 \). (B) Kaplan–Meier survival spline of StromalScore with \( p<0.01 \). (C) Kaplan–Meier survival spline of ESTIMATEScore with \( p<0.01 \).
Figure 2

Association between ImmuneScore and StromalScore with clinical stage. (A-C) Relationship of ImmuneScore and patient age, gender, and grade. The P-value was estimated by Wilcoxon rank sum test. (D-F) Relationship of StromalScore and patient age, gender, and grade. The P-value was estimated by Wilcoxon rank sum test.
Figure 3

Heatmaps, Venn plots, and enrichment analysis of GO and KEGG of DEGs. (A) Heatmap of DEGs obtained by comparing the high and low score group in ImmuneScore. The abscissa is the name with 100 genes and the vertical is the sample ID (not shown in the figure). DEGs were obtained by Wilcoxon rank sum test ($q = 0.05$ and FC>1) after log2 conversion as the importance threshold. (B) Heatmap of DEGs obtained by comparing the high and low score group in StromalScore. (C) Venn plots for up-regulated DEGs both to
ImmuneScore and StromalScore with q< 0.05 and FC>1 after log2 conversion as the DEGs importance filtering threshold. (D)Venn plots for down-regulated DEGs both to ImmuneScore and StromalScore with q< 0.05 and FC>1 after log2 conversion as the DEGs importance filtering threshold. (E)The barplot made by GO enrichment analysis, incorporated Biological Process(BP), Cellular Component(CC)and Molecular Function(MF).The abscissa represents the number of gene enrichment, Color represents the significance of gene enrichment. (F)The barplot made by KEGG enrichment analysis.

Figure 4
PPI network and univariate COX analysis (A) Interaction network built by Cytoscape. (B) The previous 30 genes sorted by the count of nodes. (C) Venn plot showed the corporate factors intersected the previous 30 nodes in PPI and the most prominently factors in univariate COX. (D) Univariate COX regression analysis was performed with the common DEGs.

**Figure 5**

The differentiated expression of glioma samples and association with survival and univariate COX with glioma patients (A) Differentiated expression of SYK in the normal and tumor samples. (B) Survival analysis of glioma patients which the SYK expression was different. Patients were divided into high expression or low expression based on the comparison with the median expression level with p < 0.001 by log-rank test. (C) Multivariate Cox regression analysis for SYK.

**Figure 6**

Samples with high and low expression of SYK were analyzed by GSEA (A) The enriched gene sets in KEGG of SYK samples with high expression. Every row on behalf of a pathway and has a different color. The left side neighbour the center axis is the up-regulated genes, correspondingly, the down-regulated genes are located to the other side of the x-axis. Just genomes with NOM p < 0.05 and FDR q < 0.05 were deemed to markedly. Just a few dominant gene sets were unfolded in this figure. (B) The enriched genomes of SYK samples with low expression.
Figure 7

TIC profile and association analysis of tumor samples. (A) The Barplot shows the ratio of 22 TICs of glioma samples. The x-coordinate is the sample name, and the y-coordinate is the amount of immune cells. Each color represents one type of immune cell. (B) Heatmap displayed the relevant with 22 TICs types. The numbers in each square manifest the association between the two types of cells. The shadow of each small box on behalf of the association value between the two cells. (C) Take the expression of
SYK as the median value, Violin plot displayed the proportion of 22 immune cells in tumor samples of glioma with high and low expression. The abscissa is the name of immune cells, and the number of immune cells is shown on the vertical axis. The samples were separated as the high-expression and low-expression group, which the red on behalf of the high-expression group and the green on behalf of the low-expression group. P<0.05 was regarded as significant. (D) Scatter plot displayed the association in 9 varieties of TICs ratio with the SYK expression (p < 0.001). (E) The 4 kinds of TIC related to SYK expression determined by difference are shown in the Venn plot.