



# Bioinformatics analysis of immune infiltration in glioblastoma multiforme based on data using a methylation chip in the GEO database

Song-Zhou Du<sup>1#</sup>, Cheng Chen<sup>2#</sup>, Lu Qin<sup>3</sup>, Xue-Lian Tang<sup>4</sup>

<sup>1</sup>Department of Neurosurgery, Jingzhou Hospital of Traditional Chinese Medicine, The Third Clinical Medical College, Yangtze University, Jingzhou, China; <sup>2</sup>Department of Nuclear Medicine, Jingzhou Central Hospital, The Second Clinical Medical College, Yangtze University, Jingzhou, China; <sup>3</sup>Department of Thyroid Vascular Surgery, Jingzhou Central Hospital, The Second Clinical Medical College, Yangtze University, Jingzhou, China; <sup>4</sup>Department of Respiratory Medicine, Jingzhou Central Hospital, The Second Clinical Medical College, Yangtze University, Jingzhou, China

**Contributions:** (I) Conception and design: SZ Du, XL Tang; (II) Administrative support: L Qin, C Chen, XL Tang; (III) Provision of study materials or patients: SZ Du, C Chen; (IV) Collection and assembly of data: C Chen, L Qin; (V) Data analysis and interpretation: XL Tang, SZ Du; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

<sup>#</sup>These authors contributed equally to this work.

**Correspondence to:** Xue-Lian Tang, Department of Respiratory Medicine, Jingzhou Central Hospital, The Second Clinical Medical College, Yangtze University, Jingzhou 434020, China. Email: XLianTang@outlook.com.

**Background:** Glioblastoma multiforme (GBM) is the most aggressive and malignant tumor of the central nervous system. The study was to obtain the data of immune cell infiltration based on the data of a methylation chip in the GEO, and to clarify its prognostic significance for GBM.

**Methods:** The methylation data of glioblastoma was obtained by using the *Illumina human methylation 450k BeadChip*. The corrected expression was obtained by using *edge R*. *Limma* was used to correct the expression amount of the samples, and *EpiDISH* was used to translate the methylation expression data, so that the expression amount was transformed into the expression matrix of immune cells. The immune cells were then co-expressed, and the proportion and correlation of related immune cells was determined. The results of the cells in each of two groups were analyzed by enrichment and PCA mapping to establish the relevant differences.

**Results:** The data of GBM patients were obtained from the methylation chip of the GEO database. Patients were divided into a long-term (SNU-LTS) (21 cases), and short-term survival group (SNU-STS) (12 cases). There were 73 genes with significant individual differences between the two groups ( $P < 0.05$ ). *EpiDISH* was used to translate the methylation expression data into the expression matrix of immune cells, which showed that the highest proportion of cells in groups were mono cells, while Gran cells and CD8T appeared in a very small number of samples. The positive correlation between mono and B cells was the strongest, while the negative correlation between mono and Gran cells was the strongest. A violin chart shows that there was no significant difference in the infiltration degree of six kinds of immune cells between the two groups. Principal component analysis (PCA) showed that there was individual difference between the two groups, but the overall consistency was high.

**Conclusions:** Data on tumor immune cell infiltration can be obtained by using a methylation chip in the GEO database. This not only extends the application abilities of methylation chips but provides obvious individual differences. The study of tumor immune infiltrating cells may pave the way for targeted therapy in the treatment of GBM.

**Keywords:** Glioblastoma multiforme (GBM); GEO; methylation; immune infiltration

Submitted Oct 12, 2020. Accepted for publication Feb 26, 2021.

doi: 10.21037/tcr-21-74

**View this article at:** <http://dx.doi.org/10.21037/tcr-21-74>

## Introduction

Glioblastoma multiforme (GBM) is the most malignant and invasive central nervous system tumor, and its incidence accounts for about 69% of gliomas (1). The current treatment of GBM is to surgically remove the tumor lesion to the maximum extent combined with chemo- and radiation therapy. However, while the median survival time of low-grade gliomas in China is 78.1 months, for GBM this is only 14.4 months (2). The prognosis for GBM patients is not optimistic, and there is an urgent need for research on targeted therapies to treat this disease.

Primary GBM is usually divided into four types according to the genetic changes involved. These are, a neuron type, pre-neuronal type, classic type and interstitial type (3). With the continuous development of biological technology, especially the application of high-throughput sequencing and genome sequencing, it has become possible to evaluate the genetic and epigenetic changes of various types of tumors at the genome-wide level. Cell and molecular genetics related studies have confirmed that there are a large number of chromosomal abnormalities and gene mutations in GBM (4). However, in addition to genetic changes, epigenetic changes may also be involved in the development and evolution of each subtype (5). As one of the important mechanisms of epigenetic regulation, DNA methylation plays an important role in the occurrence and evolution of GBM (6).

Research shows that the infiltration of immune cells in tumors is closely related to clinical outcomes. Immune cells infiltrated in tumors are most likely to serve as drug targets to improve patient survival (7) as they can be isolated after migrating from blood to tissue (8). However, it is difficult to analyze tumor prognosis with gene methylation combined with tumor immune infiltration, and few studies have reported it.

Fortunately, with the development of bioinformatics, gene methylation chip data has gradually become a powerful tool to study tumor targeted therapy (9). The *Illumina Human Methylation 450K BeadChip* can detect nearly 450,000 methylation sites in the human genome with single-base resolution. This technology completely covers 96% of CpG islands, in addition to CpG sites other than CpG islands, the non-CpG methylation sites of human stem cells, and the differential methylation sites of normal tissues and tumor (multiple cancers) tissues, encoding as required. The sites of CpG island, miRNA promoter region and GWAS disease-related regions

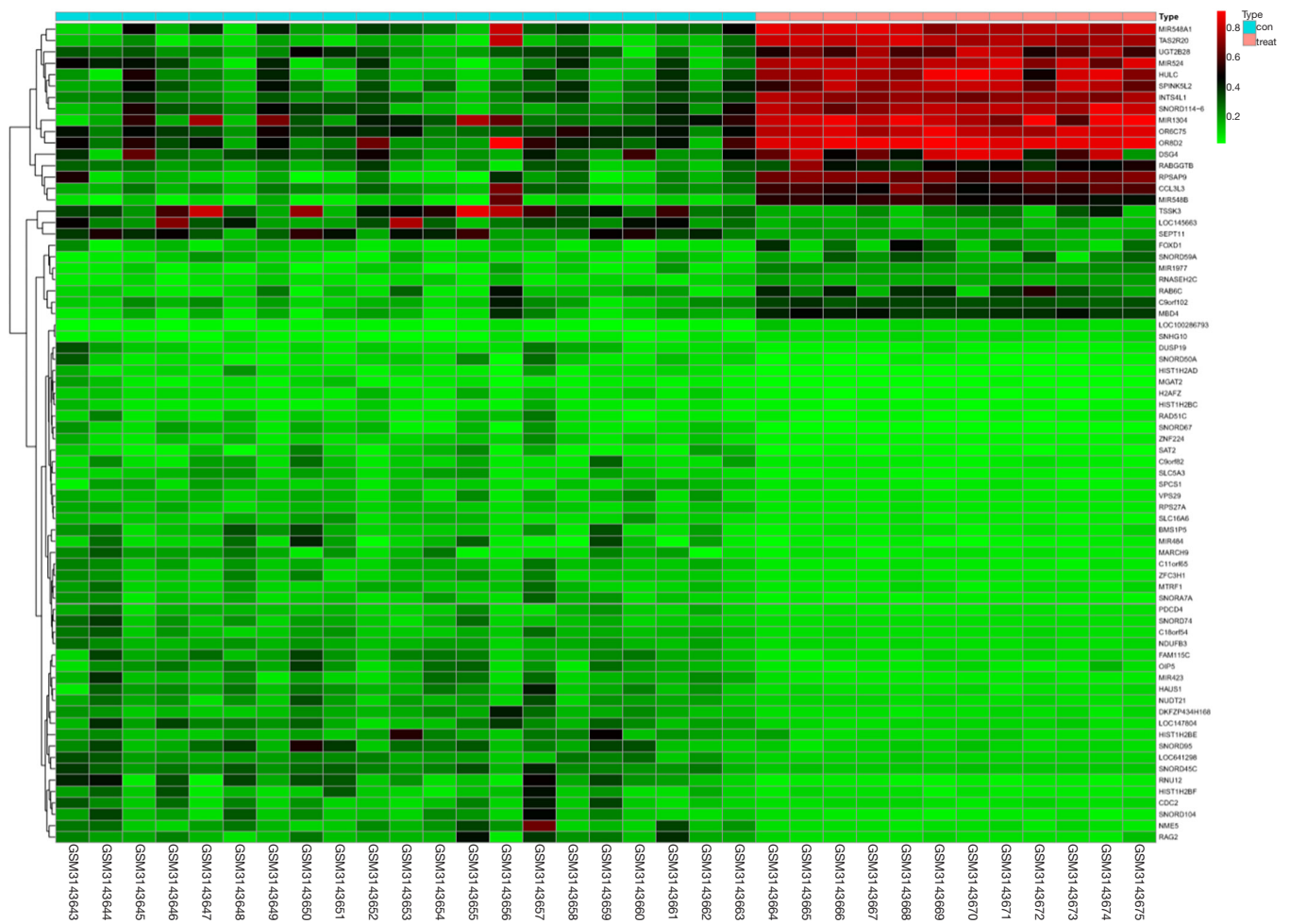
outside the region also cover 90% of the sites of the *Human Methylation 27 BeadChip* (10). Moreover, *EpiDISH* is an R software package used to infer the proportion of a priori known cell subtypes present in a mixture of cell types. The package contains references for eight cell subtypes (B cells, CD4T cells, CD8T cells, NK cells, monocytes, neutrophils, eosinophils, and granulocytes) (11) and provides a function that allows the identification of differentially methylated cell-types and their directionality of change in Epigenome-Wide Association Studies. Importantly, *EpiDISH* can be used directly on 450k DNA array and EPIC array data (12); therefore methylation chip data can also be used for tumor immune infiltration pattern analysis.

In this study, we obtained the data of the *450k BeadChip* methylation chip in the GEO database and that of immune cell infiltration through the *EpiDISH* software package transformation. The proportion, correlation and differential expression of related immune cells was then analyzed, and its prognostic significance for GBM clarified. We contend the results of this study will lay the foundation for the development of targeted therapy for this aggressive malignancy.

## Methods

### Data acquisition

The gene methylation chip data of GBM with survival data were downloaded from the GEO database with *RTGCA toolkit R* (version 3.5.1). The samples were divided into two groups: long-term survival (SNU-LTS: GSM3143643-GSM3143663) and short-term survival (SNU-STs: GSM3143664-GSM3143675) based of patient survival. The related data were downloaded, sorted, and annotated. We then used glioma data and the ID conversion of relevant GEO data to convert methylation sites to gene sites. Using *edge R* for differential analysis, the corrected expression and gene differential expression were obtained, and the Heatmap file was used to draw a heat map. *Limma* was used to correct the expression of samples. The corrected expression file was used for the software transformation of immune cells. *EpiDISH* was used to translate the expression amount of methylation data, so that the expression amount was transformed into the expression matrix of immune cells. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).



**Figure 1** The heatmaps of genomic methylation in SNU-LTS (con) and SNU-STs (treat). SNU-LTS, long-term survival tumor samples; SNU-STs, short-term survival tumor samples.

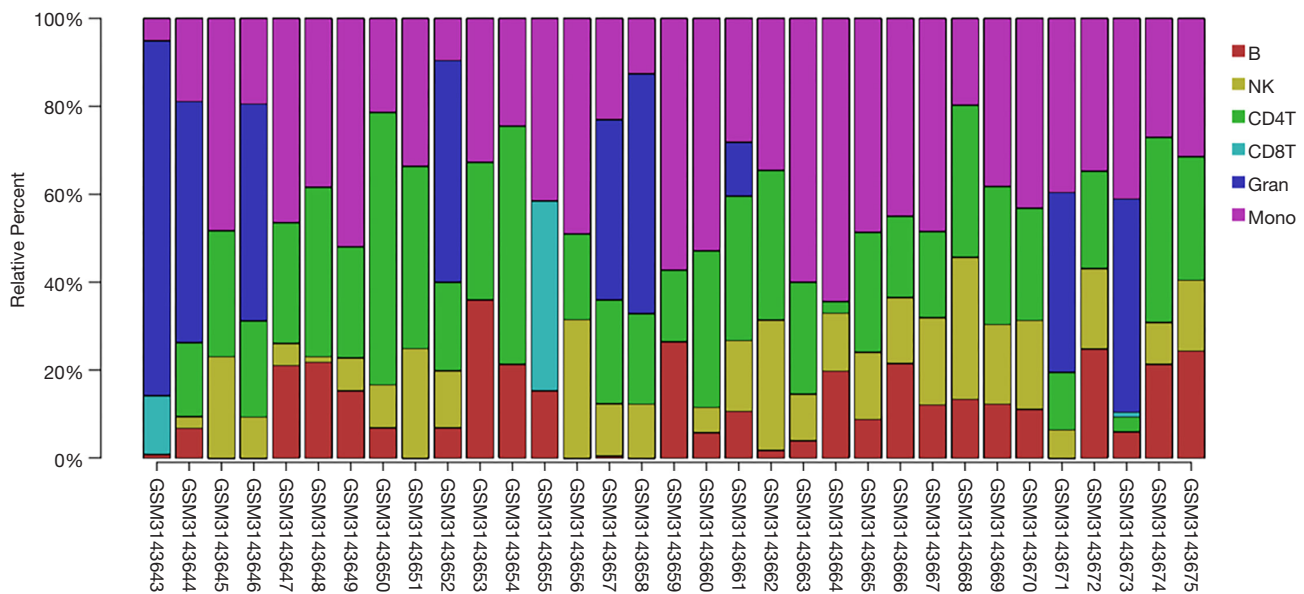
### Statistical analysis

SPSS software 22.0 (IBM Corporation, Armonk, NY, USA) was used for data analysis. All statistical tests were bilateral, and a *p* value less than 0.05 had statistical significance. When the continuous variable conformed to normal distribution, the independent sample *t*-test was used for comparison, and when the continuous variable deviated from the normal distribution, the *U*-test was used for comparison. Pearson correlation analysis and Spearman correlation analysis were then used.

### Results

#### Differential expression of methylation gene

There were 21 samples in the SNU-LTS group and 12 in the SNU-STs group. The difference analysis was carried out by edgeR package and the corrected expression amount and gene differential expression were obtained. The behavior gene name was listed as the sample name and the Heatmap file used to draw the heat map (Figure 1). The results showed that 73 genes (DKFZP434H168, INTS4L1, MIR524, etc.) in the two groups were different, and the



**Figure 2** The proportion of six immune cells in SNU-LTS and SNU-STs. SNU-LTS, long-term survival tumor samples; SNU-STs, short-term survival tumor samples.

difference was statistically significant.

### *The landscape of immune infiltration*

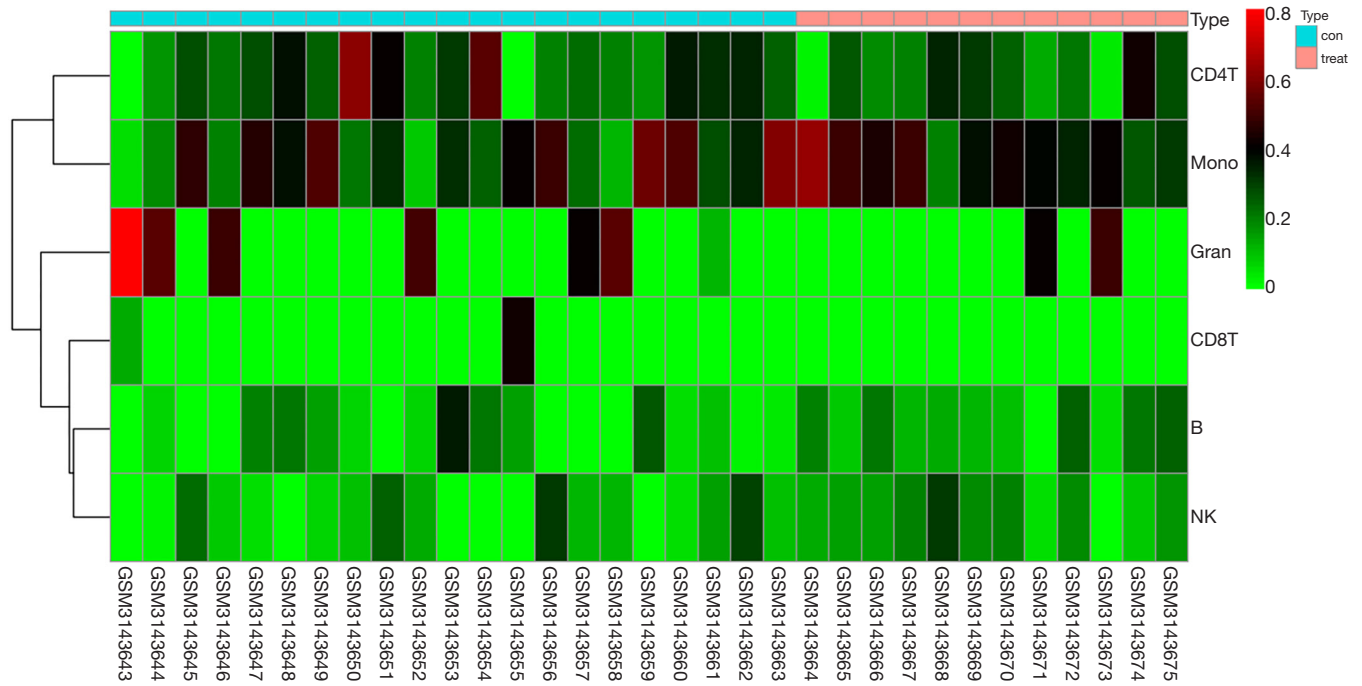
Methylation expression data was translated into the expression matrix of immune cells by *EpiDISH* software. A histogram shows the proportion of six immune cells in each sample (*Figure 2*). The abscissa of the column chart was the sample name, the ordinate was the proportion of relevant immune cells, and the legend was the name of the immune cells. The results show that the highest proportion of immune cells in the two groups were Mono cells, whilst B cells, NK cells and CD4T cells accounted for a certain proportion in the two groups. Very few samples from either group, however, contained Gran cells and CD8T. The results also showed that the proportion of immune cells in tumor tissue is different among six groups. We then drew a heatmap of six immune cells, shown as *Figure 3*. These results also show that the proportion of immune cells in six samples is different between the two groups. We then continued to explore the correlation between immune cells, as shown in *Figure 4*, and found the correlation coefficient between six immune cells. This revealed the strongest positive correlation was between Mono cells and B cells ( $r=0.22$ ), and the strongest negative correlation was between Mono cells and Gran cells ( $r=-0.64$ ).

### *Differential analysis of tumor infiltration of six immune cells*

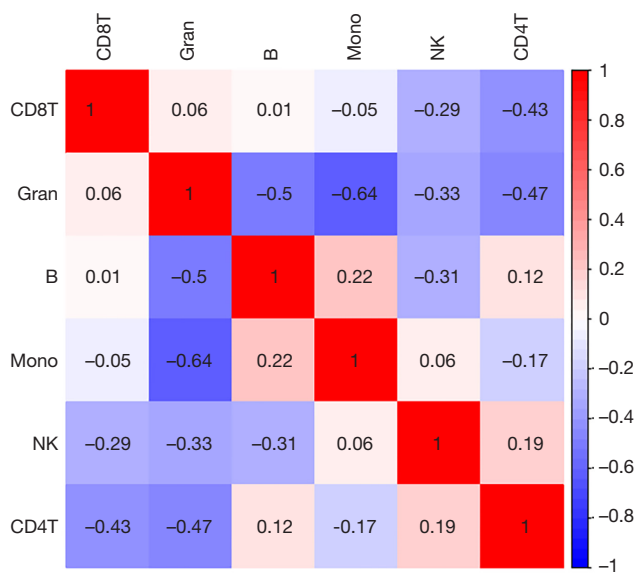
The different infiltrates of six immune cells in the SNU-LTS and SNU-STs groups were compared. The results show that there is no significant difference in the infiltration degree between the two groups ( $P>0.05$ , *Figure 5*). The results of principal component analysis (PCA) show that while there are individual differences between the SNU-LTS and SNU-STs groups, there is also a certain degree of consistency, suggesting PCA analysis had no obvious significance (*Figure 6*).

### **Discussion**

The importance of DNA methylation in the occurrence and development of GBM continues to be of interest to researchers. Epigenetic regulation has the characteristics of complexity, variability, and space-time specificity, which makes the study of epigenetic regulation very difficult (13). At present, most research on DNA methylation is limited to the use of drug treatment or knockout of DNA methyltransferase so that the effect of gene changes on downstream targets can be observed. However, the epigenetic changes of GBM are across a wide range, and such research often has little effect on its treatment. Finding a correlation between the regulatory epigenetic changes and



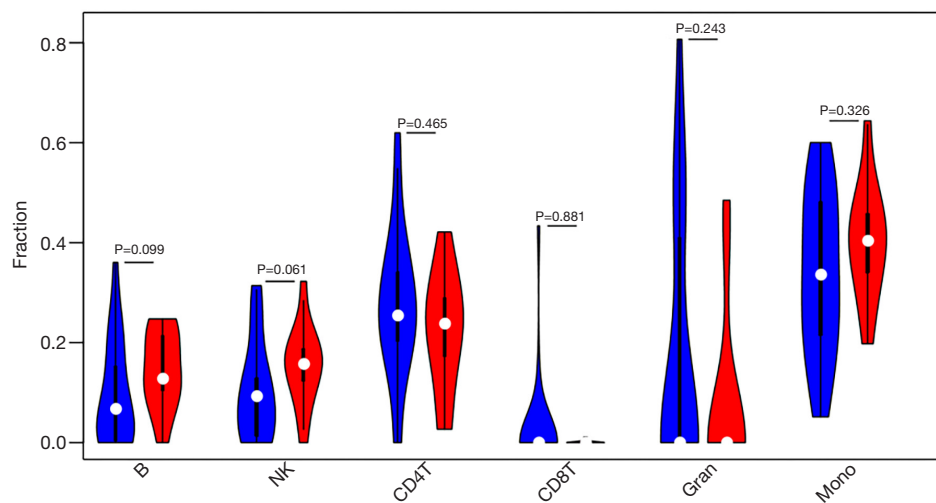
**Figure 3** The heatmaps of six immune cells in SNU-LTS (con) and SNU-STs (treat). The horizontal axis shows the clustering information of samples which were divided into two major clusters. SNU-LTS, long-term survival tumor samples; SNU-STs, short-term survival tumor samples.



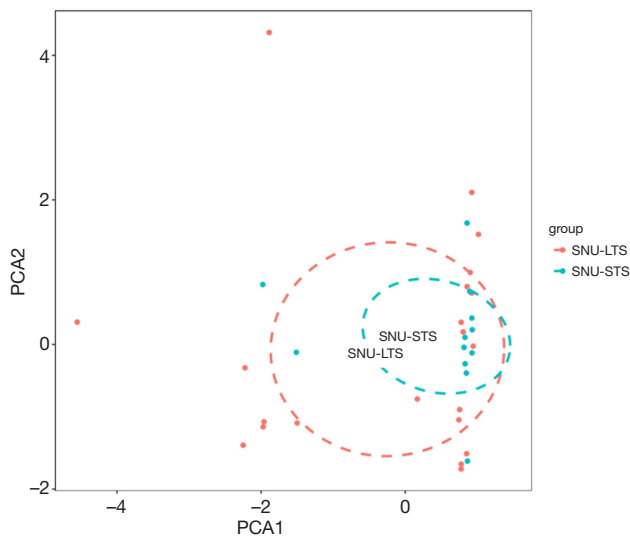
**Figure 4** Correlation matrix between six immune cells. Red means positive correlation, blue means negative correlation, and the darker the color, the stronger the correlation.

other tumor pathogeneses to indirectly affect the overall epigenetic changes of GBM will have a profound impact on its treatment. In addition, it is difficult to analyze the whole genome methylation group in a large number of GBM patients and the procedure is expensive (14). Therefore, it is important to reduce the cost of detection or find a substitute detection method. Developments in gene editing methods and human genome information mining suggest there is a close relationship between methylation and immune cell infiltration, and the study of their correlation is an important topic for future research (4). We obtained the data of the *450k BeadChip* methylation chip in the GEO database and that of immune cell infiltration through the *EpiDISH* software package transformation. The proportion, correlation and differential expression of related immune cells was then analyzed, which is of great significance in determining the prognosis of GBM.

We explored the differences in methylation gene expression among 33 GBM patients with different life cycles, and then acquired the proportion, difference, and



**Figure 5** Variance analysis of six immune cells in SNU-LTS (blue) and SNU-STs (red). SNU-LTS, long-term survival tumor samples; SNU-STs, short-term survival tumor samples.



**Figure 6** Principal component analysis (PCA) of samples in SNU-LTS and SNU-STs. SNU-LTS, long-term survival tumor samples; SNU-STs, short-term survival tumor samples.

correlation of tumor immune infiltrating cells through data conversion. Statistical analysis of the methylation chip data showed 73 genes with significant differences in expression among patients with different life cycles, and individual gene methylation showed considerable differences. Whilst methylation genes were up regulated to some extent, some genes were down regulated. Following this, we determined

the proportion of immune cells infiltrated by tumor tissue. Mono cells were found to be infiltrated more commonly in tumor tissue than CD8T and Gran cells, and the individual differences were large. There was no significant difference between the two groups, indicating that correlation between the survival cycle and the infiltration of immune cells was not high.

These results are consistent with those reported in other studies which have shown that extensive DNA hypomethylation is a common event in GBM. In animal experiments, induction of extensive DNA hypomethylation in mice led to genomic instability and tumor formation. At least 10 million CpG dinucleotides in each haploid genome of GBM are affected by DNA hypomethylation. The hypomethylation sites include satellite 2 (SAT2) around the centromere of chromosomes 1 and 16, D4Z4 near the telomeres of chromosomes 4q and 10q, and the Alu element with dot. However, severe demethylation of the SAT2 sequence were often accompanied by change of the copy number of nearby euchromatin, which suggested that hypomethylation may be a precondition for GBM gene change (15). At least 1,548 abnormal methylation CpG loci (1,307 genes) have been identified when studying the genome-wide methylation differences between GBM and normal brain tissue. Among these, hypomethylation genes numbered twice as many as hypermethylation genes (16). The analysis of low methylation genes in the top ten showed that they were related to innate immunity and acquired

immunity. In view of whether DNA hypomethylation can judge the prognosis of GBM and serve as a potential diagnostic index.

Exploration of the correlation between DNA methylation and tumor infiltrating immune cells is in its infancy, and many factors have not been taken into account, such as the age of patients, the course and stage of the disease, and the treatment received. The sample size obtained in this study is small, and it is difficult to obtain a clear correlation trend. More research is required to verify these results.

## Conclusions

This study verified that the data of tumor infiltration of immune cells can be obtained by gene methylation chip. This method is simple, convenient, and has broad implications for tumor related research. The results of this study showed that the individual differences of gene methylation expression and infiltration ratio of immune cells are significant in GBM patients. However, the analysis also showed that there is no significant correlation between the survival period and the infiltration degree of immune cells. Despite this we conclude that cells with a high proportion of infiltration in GBM tumor tissue are potential targets for its treatment.

## Acknowledgments

*Funding:* None.

## Footnote

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/tcr-21-74>). The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-

commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

1. Neamati F, Asemi Z. The effects of melatonin on signaling pathways and molecules involved in glioma. *Fundam Clin Pharmacol* 2020;34:192-9.
2. Yang Y, Sui Y, Xie B, et al. GliomaDB: A Web Server for Integrating Glioma Omics Data and Interactive Analysis. *Genomics Proteomics Bioinformatics* 2019;17:465-471.
3. Hwang I, Choi SH, Park CK, et al. Dynamic Contrast-Enhanced MR Imaging of Nonenhancing T2 High-Signal-Intensity Lesions in Baseline and Posttreatment Glioblastoma: Temporal Change and Prognostic Value. *Am J Neuroradiol* 2020;41:49-56.
4. Jang BS, Kim IA. A Radiosensitivity Gene Signature and PD-L1 Status Predict Clinical Outcome of Patients with Glioblastoma Multiforme in The Cancer Genome Atlas Dataset. *Cancer Res Treat* 2020;52:530.
5. Braczynski AK, Capper D, Jones DTW, et al. High density DNA methylation array is a reliable alternative for PCR-based analysis of the MGMT promoter methylation status in glioblastoma. *Pathol Res Pract* 2020;216:152728.
6. Feldheim J, Kessler AF, Monoranu CM, et al. Changes of O-Methylguanine DNA Methyltransferase (MGMT) Promoter Methylation in Glioblastoma Relapse-A Meta-Analysis Type Literature Review. *Cancers (Basel)* 2019;11:1837.
7. Petty AJ, Yang Y. Tumor-Associated Macrophages in Hematologic Malignancies: New Insights and Targeted Therapies. *Cells* 2019;8:1526.
8. Weiss JM. The promise and peril of targeting cell metabolism for cancer therapy. *Cancer Immunol Immunother* 2020;69:255-61.
9. Kazmi N, Sharp GC, Reese SE, et al. Hypertensive Disorders of Pregnancy and DNA Methylation in Newborns. *Hypertension* 2019;74:375-83.
10. Braun PR, Han S, Hing B, et al. Genome-wide DNA methylation comparison between live human brain and peripheral tissues within individuals. *Transl Psychiatry* 2019;9:47.
11. Zheng SC, Breeze CE, Beck S, et al. EpiDISH web server: Epigenetic Dissection of Intra-Sample-Heterogeneity with online GUI. *Bioinformatics (Oxford, England)*

- 2019;36:1950-1.
12. Du R, Carey V, Weiss ST. deconvSeq: Deconvolution of Cell Mixture Distribution in Sequencing Data. *Bioinformatics* (Oxford, England) 2019;35:5095-102.
  13. Yu J, Loh XJ, Luo Y, et al. Insights into the epigenetic effects of nanomaterials on cells. *Biomater Sci* 2020;8:763-75.
  14. Wang L, Xia Y, Chen Y, et al. Brain Banks Spur New Frontiers in Neuropsychiatric Research and Strategies for Analysis and Validation. *Genomics. Proteomics Bioinformatics* 2019;17:402-14.
  15. Cadieux B, Ching TT, VandenBerg SR, et al. Genome-wide hypomethylation in human glioblastomas associated with specific copy number alteration, methylenetetrahydrofolate reductase allele status, and increased proliferation. *Cancer Res* 2006;66:8469-76.
  16. Huang RC, Garratt ES, Pan H, et al. Genome-wide methylation analysis identifies differentially methylated CpG loci associated with severe obesity in childhood. *Epigenetics* 2015;10:995-1005.

(English Language Editor: B. Draper)

**Cite this article as:** Du SZ, Chen C, Qin L, Tang XL. Bioinformatics analysis of immune infiltration in glioblastoma multiforme based on data using a methylation chip in the GEO database. *Transl Cancer Res* 2021;10(3):1484-1491. doi: 10.21037/tcr-21-74