


# Construction and validation of cell cycle-related prognostic genetic model for glioblastoma

Runpeng Zhou, MD<sup>a</sup>, Kai Zhang, MD<sup>b</sup>, Tingting Dai, MD<sup>b</sup>, Zeshang Guo, MD<sup>c</sup>, Tian Li, MD<sup>d,\*</sup> , Xinyu Hong, PhD<sup>c</sup>

## Abstract

Glioblastoma (GBM) is a common primary malignant brain tumor and the prognosis of these patients remains poor. Therefore, further understanding of cell cycle-related molecular mechanisms of GBM and identification of appropriate prognostic markers and therapeutic targets are key research imperatives. Based on RNA-seq expression datasets from The Cancer Genome Atlas database, prognosis-related biological processes in GBM were screened out. Gene Set Variation Analysis (GSVA), LASSO-COX, univariate and multivariate Cox regression analyses, Kaplan–Meier survival analysis, and Pearson correlation analysis were performed for constructing a predictive prognostic model. A total of 58 cell cycle-related genes were identified by GSVA and analysis of differential expression between GBM and control samples. By univariate Cox and LASSO regression analyses, 8 genes were identified as prognostic biomarkers in GBM. A nomogram with superior performance to predict the survival of GBM patients was established regarding risk score, cancer status, recurrence type, and mRNAsi. This study revealed the prognostic value of cell cycle-related genes in GBM. In addition, we constructed a reliable model for predicting the prognosis of GBM patients. Our findings reinforce the relationship between cell cycle and GBM and may help improve the prognostic assessment of patients with GBM. Our predictive prognostic model, based on independent prognostic factors, enables tailored treatment strategies for GBM patients. It is particularly useful for subgroups with uncertain prognosis or treatment challenges.

**Abbreviations:** AUC = area under the ROC curve, DECCRGs = differentially expressed cell cycle-related genes, GBM = glioblastoma, GEO = Gene Expression Omnibus, GSVA = Gene Set Variation Analysis, LASSO = Least Absolute Shrinkage and Selection Operator, PCA = principal component analysis, ROC = receiver operating characteristic, TCGA = the Cancer Genome Atlas.

**Keywords:** glioblastoma, GSVA, nomogram, prognosis

## 1. Introduction

Gliomas account for up to 30% to 40% of all brain tumors and 80% of all malignant brain tumors. Among them, glioblastoma is the most common and aggressive primary malignant brain tumor in adults, accounting for more than half of all malignant gliomas. Glioblastomas are characterized by high recurrence and mortality rates, and a low cure rate.<sup>[1]</sup> The 5-year survival rate of patients with malignant glioma is < 30% for grade III (anaplastic astrocytoma) and 5% for grade IV (glioblastoma) tumors.<sup>[2–4]</sup> Age, tumor grade, clinical treatment strategies, and the molecular phenotype of the tumor markedly affect the prognoses of patients with gliomas. The treatment efficacy for gliomas is suboptimal, and the prognosis of these patients is highly variable.<sup>[5,6]</sup> The standard treatment for glioblastoma typically includes gross total resection surgery, focal radiotherapy targeting the tumor area, temozolomide chemotherapy, and a certain dose of radiotherapy as the standard

treatment.<sup>[7,8]</sup> Although these treatment modalities can extend patients' survival to some extent, all glioblastomas eventually progress or recur.<sup>[7]</sup> Existing prognostic models often rely on clinical pathological features, but they have certain limitations, such as low predictive accuracy and lack of personalized prediction.<sup>[9]</sup> Therefore, establishing an individualized prognostic scoring system based on important prognostic factors may facilitate the provision of individualized risk assessment and clinical treatment.

Gene Set Variation Analysis (GSVA), a gene set enrichment (GSE) method that estimates variation of pathway activity over a sample population in an unsupervised manner, can be used to assess the results of chipset gene enrichment. It evaluates the differences in the signal pathways by transforming the gene expression matrix among samples into gene sets expression matrix (signaling pathways) with the same biological function between samples. As a bioinformatics analysis method, GSVA can explain the causes of phenotypic differences according to

RZ, KZ, and TD contributed to this article equally.

The authors have no funding and conflicts of interest to disclose.

Consent for publication is not applicable to this study.

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate are not applicable to this study.

<sup>a</sup> Department of Neurosurgery, Pu'er People's Hospital, Pu'er, China,

<sup>b</sup> Department of Neurosurgery, Changzheng Hospital, Naval Medical University, Shanghai, China, <sup>c</sup> Department of Neurosurgery, The First Bethune Hospital of Jilin University, Changchun, China, <sup>d</sup> School of Basic Medicine, Fourth Military Medical University, Xi'an, China.

\* Correspondence: Tian Li, School of Basic Medicine, Fourth Military Medical University, Xi'an 710032, China (e-mail: fmmult@foxmail.com).

Copyright © 2024 the Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Zhou R, Zhang K, Dai T, Guo Z, Li T, Hong X. Construction and validation of cell cycle-related prognostic genetic model for glioblastoma. *Medicine* 2024;103:40(e39205).

Received: 24 November 2023 / Received in final form: 3 July 2024 / Accepted: 16 July 2024

<http://dx.doi.org/10.1097/MD.00000000000039205>

the signal pathways.<sup>[10]</sup> The foundation of the gene-rich integrated gene set (signaling pathway) was mainly derived from KEGG,<sup>[11]</sup> BIOCARTEA, and REACTOME,<sup>[11,12]</sup> 3 general reference databases for signaling pathways. A large amount of biological information data, such as genome, metabolic pathways, and signaling pathways, is contained in above databases.<sup>[13]</sup> The analysis and investigation of signal pathways and biological processes related to GBM are primarily based on KEGG in this study. In this study, GSEA could be utilized to compare differences in gene expression among different patients, aiding us in identifying changes in biological processes or pathways related to the disease. Specifically, GSEA analysis could be employed to determine whether gene sets associated with tumor development were significantly enriched in glioblastoma patients.

This study was based on the TCGA<sup>[14]</sup> database and a combination of bioinformatics methods, GSEA analysis, and modeling analysis were used to identify key genes in relation to the differential function of GSEA in glioblastoma. The GSEA analysis helped us pinpoint specific genes that play a crucial role in the functional differences observed in glioblastoma. Subsequently, we employed LASSO Cox regression analysis to develop a risk score model, which integrated the expression levels of these key genes to predict patient outcomes. By utilizing these advanced computational methods, we were able to construct a comprehensive prognostic model. The accuracy of the model in predicting the prognosis of patients with glioblastoma was verified internally and externally.

## 2. Materials and methods

### 2.1. Data sources

Gene expression profiles and clinical information for the training cohort (<https://portal.gdc.cancer.gov/projects/TCGA-GBM>) of 166 patients with glioblastoma were collected from the Cancer Genome Atlas (TCGA) database (<https://cancergenome.nih.gov>).<sup>[14]</sup> In the TCGA-GBM dataset, there were 104 male patients (62.7% of the total), 56 female patients (33.7% of the total), and 6 cases (3.6% of the total) where the gender was not reported. Concerning age at diagnosis, 9 patients were under 40 years old (5.4% of the total), 49 patients were aged between 40 and 60 (29.5% of the total), 71 patients were over 60 years old (42.8% of the total), and 37 cases had unreported age data (22.3% of the total). Regarding treatment status, 52 patients received treatment (31.3% of the total), 11 patients did not receive treatment (6.6% of the total), and treatment information was not reported for 103 cases (62% of the total).<sup>[15]</sup> GSE13041 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE13041>) and GSE43378 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi>) were downloaded from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>) as the validation dataset. The flowchart of this study was shown in Figure 1

### 2.2. Identification of significantly differential pathways by GSEA

GSEA was performed to identify differential pathways using the GSEA package in R with default parameters. GSEA enhances the ability to detect subtle changes in pathway activity across sample populations.<sup>[10]</sup> The default parameters in the GSEA package were chosen for their established effectiveness in pathway analysis and their widespread use in similar studies, ensuring consistency and comparability with existing literature.<sup>[16,17]</sup> By identifying differentially expressed pathways, we could gain a deeper understanding of the molecular alterations associated with GBM progression and potentially uncover novel therapeutic targets.<sup>[16]</sup> Consistent clustering of principal component analysis (PCA) and nonnegative matrix

factorization (NMF) was verified to determine the correlation between cell cycle-related genes and GBM. DESeq2 R package was used to distinguish DEGs between 161 glioblastoma samples and 5 adjacent non-tumor samples using the threshold  $\log_2(\text{fold change}) > 1$  and  $P\text{-value} < 0.05$ . The cell cycle-related genes were selected based on GSEA analysis, where we downloaded the pathway gene set KEGG\_CELL\_CYCLE from the Molecular Signatures Database (MSigDB) at <https://www.gsea-msigdb.org/gsea/msigdb/>. Differentially expressed cell cycle-related genes (DECCRGs) were obtained by overlapping DEGs with cell cycle-related genes from GSEA. DECCRGs were used for the subsequent analysis. GO<sup>[18]</sup> and KEGG enrichment analysis of DECCRGs were performed using the clusterProfiler R package. This function analyses would help in understanding the biological functions and relevant pathways of DECCRGs in glioblastoma, providing insights into the potential roles and mechanisms of these genes in the pathogenesis of glioblastoma. In this analysis include the identification of key biological processes, molecular functions, and pathways that are dysregulated in glioblastoma, which can further our understanding of the disease at a molecular level.

### 2.3. Construction of the prognostic risk score

Univariate analysis was conducted with the COX regression model to investigate the prognostic values of cell cycle-related genes ( $P$  value  $< 0.05$  and hazard ratio (HR)  $\neq 1$ ). Then, the lambda value of each candidate survival-associated gene was calculated by LASSO-COX analysis using the glmnet and survival packages in R. In this study, the model construction process involved using 10-fold cross-validation to determine the value of  $\lambda$ , and the  $\lambda$  value with the smallest partial likelihood deviance was selected as the optimal  $\lambda$ . Finally, the risk score was calculated using the following formula:

$$\text{Risk score} = \text{ExpGene1} \times \text{Coef1} + \text{ExpGene2} \times \text{Coef2} + \text{ExpGene3} \times \text{Coef3} \quad (1)$$

where ExpGene is the expression level of the gene and Coef is the corresponding lambda value.<sup>[19]</sup> For the TCGA data, the median value of risk scores in all patients for whom survival information was available was deemed as the cutoff. The same criterion was used to determine the cutoff value in the GEO database. The cutoff value was chosen as it represents a division point in the distribution of risk scores.<sup>[19]</sup> Then, patients with GBM were assigned to the “high risk” or “low risk” group accordingly.

### 2.4. Correlation between prognosis and risk score

Kaplan–Meier survival curves were generated to illustrate the correlation between prognosis and risk score. Univariate and multivariate Cox regression analysis was conducted to identify independent risk factors for survival. In the univariate Cox analysis, variables including pathological subtype, age, gender, recurrence type, cancerstatus, mRNAsi, and risk score were incorporated. Subsequently, factors meeting the criteria of a  $P\text{-value} < 0.05$  and a HR different from 1 in the univariate Cox analysis were included in the multivariate Cox analysis.

### 2.5. Statistical analysis

The Kaplan–Meier method was used for survival analysis and the between-group differences were assessed using the log-rank test. The predictive accuracy of the risk model was determined by receiver operating characteristic (ROC) curve analysis. All data were analyzed using R (version 4.0.0).  $P$  values  $< 0.05$  were considered indicative of statistical significance unless specified otherwise.

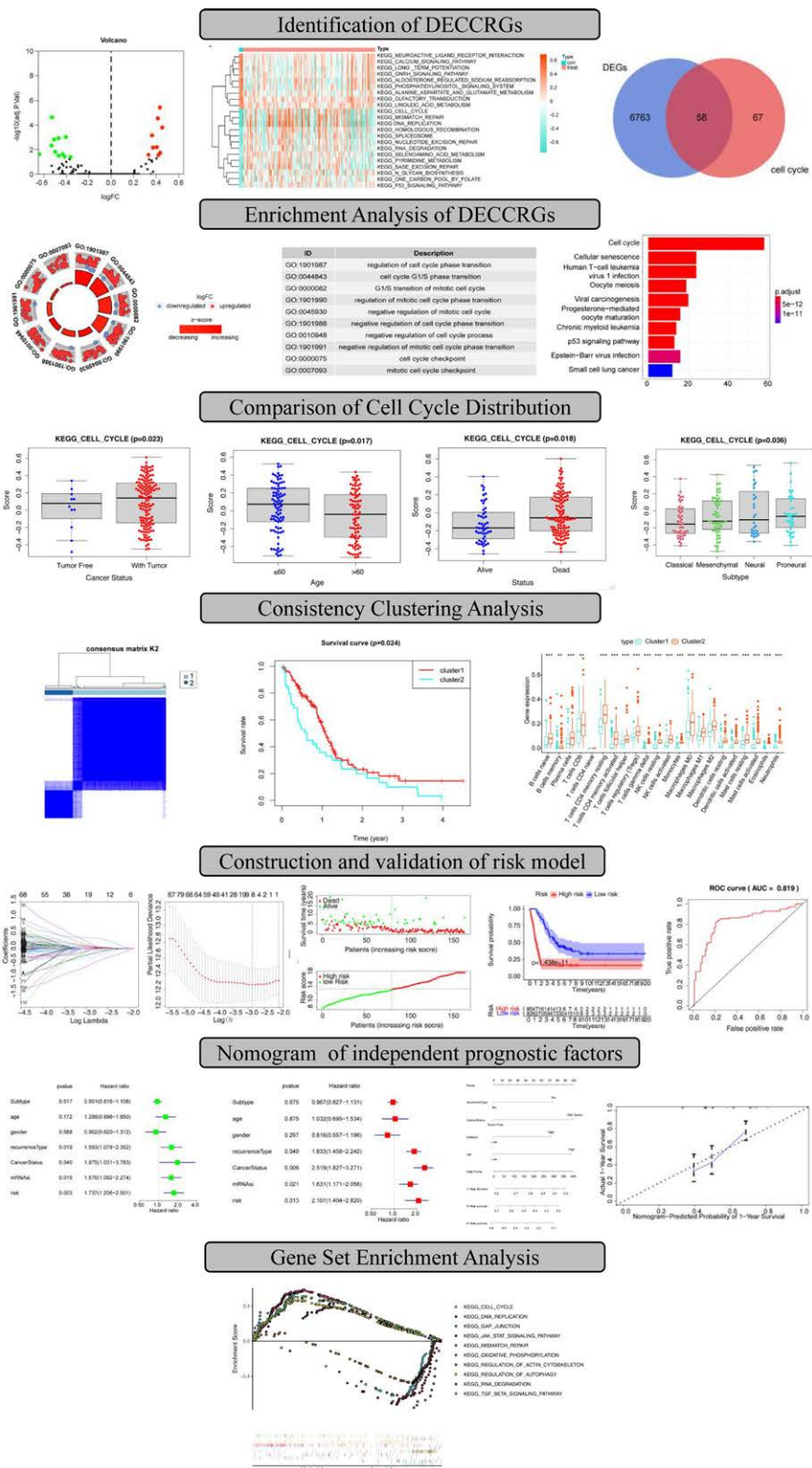


Figure 1. Flowchart.

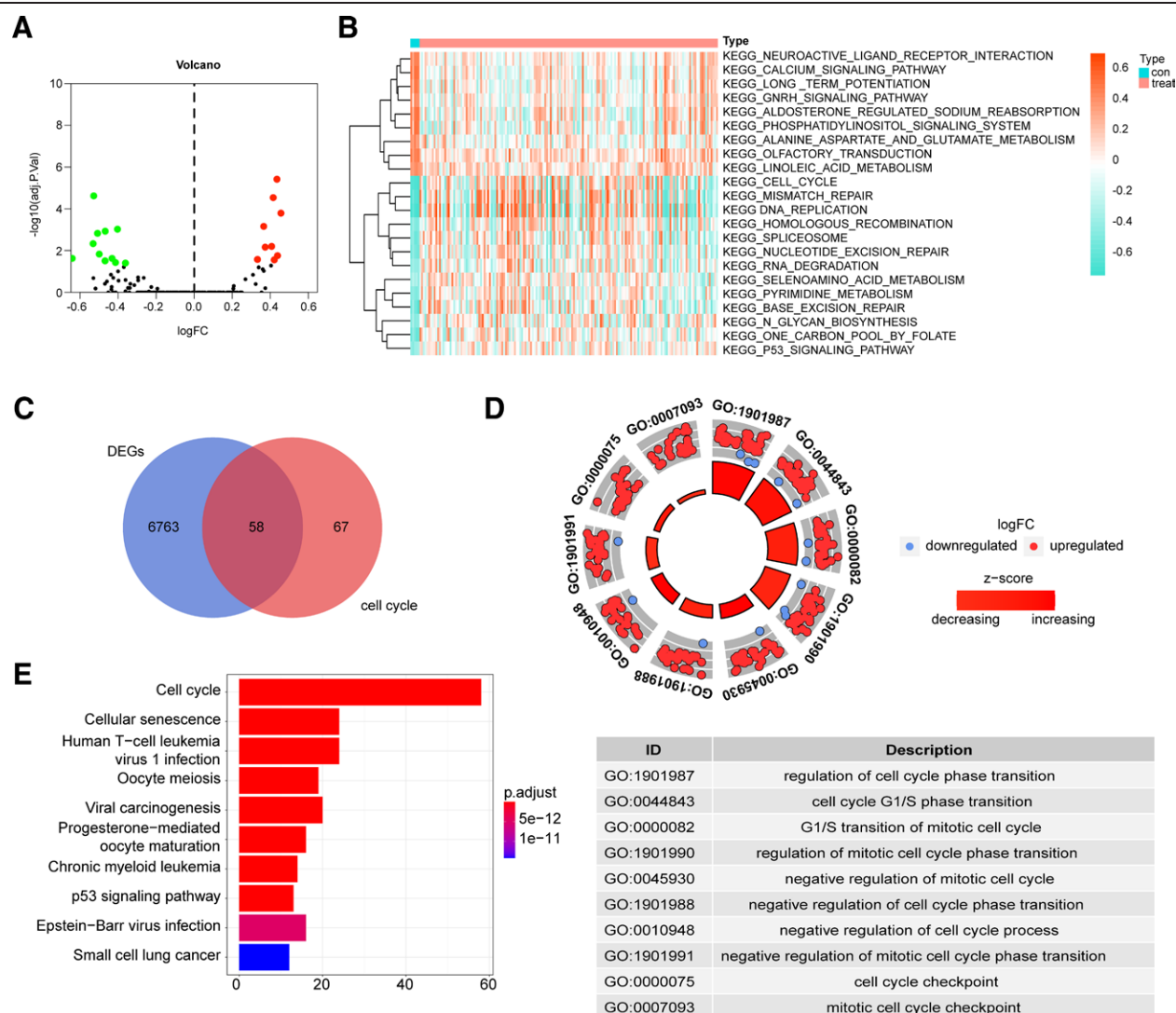
### 3. Results

#### 3.1. Cell cycle pathway was highly correlated with clinicopathological features of GBM

To screen out the most significant differentially expressed pathways in GBM, enrichment scores of pathways were obtained by the GSEA algorithm in the TCGA database. The results identified

that 20 pathways that were significantly different between glioblastoma samples and adjacent non-tumor samples (11 pathways were downregulated and 9 pathways were upregulated in glioblastoma) with  $\log_2FC > 1$  and  $P$  value  $< 0.05$ , such as DNA replication were associated with resistance to GBM<sup>[20]</sup> (Fig. 2A). The cell cycle pathway showed the most significant difference between GBM and control samples (Fig. 2B), indicating a





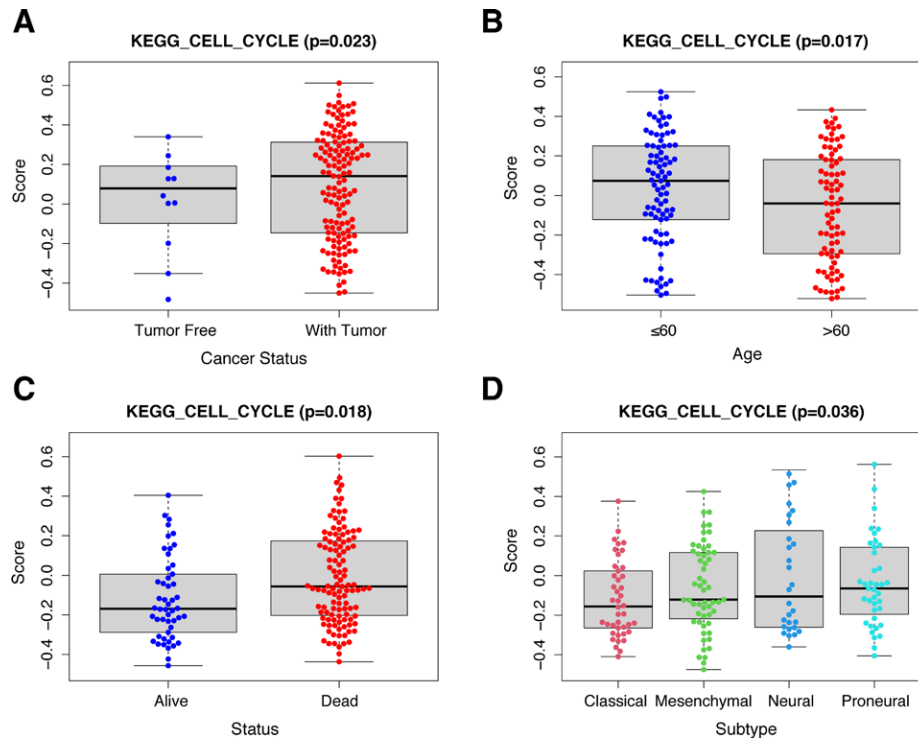
**Figure 2.** GSVA algorithm using the TCGA database. Volcano plot showing the differentially expressed pathways in GBM (A). Green dots represent significantly downregulated pathways and red dots represent significantly upregulated pathways in GBM with high expression. The heatmap (B) showing the expression level. (C) Venn diagram of overlapping 6821 DEGs and 125 cell cycle-related genes. (D, E) GO and KEGG functional enrichment analysis of DECCRGs.

crucial role of the cell cycle pathway in the genesis of GBM. A total of 58 cell cycle-related differential genes (DECCRGs) were obtained by overlapping 6821 DEGs with 125 cell cycle-related genes (Fig. 2C) and were used for the subsequent analyses. Functional enrichment analysis showed that the top GO terms were associated with cell cycle transition, such as regulation of cell cycle phase transition, cell cycle G1/S phase transition, G1/S transition of mitotic cell cycle, regulation of mitotic cell cycle phase transition, and negative regulation of mitotic cell cycle (Fig. 2D). Interestingly, apart from the cell cycle, we found that DECCRGs were also involved in infection, which has a complex interaction with the immune system, such as the human T-cell leukemia virus 1 infection pathway, viral carcinogenesis pathway, and p53 signaling pathway (Fig. 2E). Having been reported in the literature, the crucial role of the p53 signaling pathway in GBM is highlighted, with the inhibition of this pathway by SNRPG being shown to enhance the sensitivity of GBM cells to TMZ, thereby aiding in overcoming chemotherapy resistance.<sup>[21]</sup> These pathways may play a role in the development or progression of GBM, and targeting them could offer potential therapeutic strategies for inhibiting GBM progression.

In our study, we compared cell cycle scores across different clinicopathological features using appropriate statistical methods.<sup>[22,23]</sup> Specifically, we applied ANOVA to compare cell cycle

scores among groups stratified by cancer status, age, status, and subtype. post hoc multiple comparison corrections were also performed to control for false positives. By utilizing these statistical methods, we were able to identify significant differences in cell cycle scores across the stratified groups. Furthermore, we assessed the significance levels of these differences to underscore the impact of various clinicopathological features on cell cycle scores. In short, we compared cell cycle scores across different clinicopathological features and identified significant differences in the cell cycle scores among groups stratified by cancer status (Fig. 3A), age (Fig. 3B), status (Fig. 3C), and subtype (Fig. 3D). The higher cell cycle scores in the proneural subtype, as well as in tumor status, age less than or equal to 60 years, and death status, across these 4 clinical subgroups, indicate a significant impact on GBM progression. The higher cell cycle scores may reflect an increased rate of cell proliferation, which could be associated with the invasiveness and malignancy of the tumor.<sup>[24]</sup> Therefore, the results suggest that the abnormal activation of the cell cycle may be an important factor in the progression of GBM, providing valuable clues for further research on the role of cell cycle regulation in the development and treatment of GBM.

Based on 125 cell cycle-related genes expression, when rank = 2, meaning when GBM patients were separated into 2 groups, NMF



**Figure 3.** Cell cycle score were significantly different among groups stratified by cancer status (A), age (B), status (C), subtype (D).

results revealed the improvement of categorization (Fig. 4A). The criteria included: consistent sample numbers in each group, a gradual and smooth rise in the cumulative distribution function curve, and stronger intragroup links compared to weaker inter-group links after clustering.<sup>[23]</sup> The consistency distribution for  $k$  values ranging from 2 to 12 was displayed in the empirical cumulative distribution function plot. Based on the NMF rank survey, 2 subgroups were selected by the standard of rank = 2, named cluster1 and cluster2 (Fig. 4B). Furthermore, to validate the effectiveness of unsupervised clustering, the PCA could clearly demonstrate the distinctions between the 2 clusters. The PCA analysis showed that all samples were PC1 negative (Fig. 4C). Most cluster1 cases were PC1 positive, and the majority of cluster2 cases were PC2 positive. These results indicated significantly different spatial distribution of samples disaggregated by NMF grouping. In terms of clinical outcomes, comparisons between cluster1 and cluster2 showed significant differences in Kaplan–Meier survival analysis (Fig. 4D), suggesting that the DECCRGs were responsible for the prognostic differences. Given the individual immune cell scores, we further analyzed the differences in immune cell composition of various NMF subgroups in the training cohort and identified distinct immune cell composition between cluster1 and cluster2 (Fig. 4E).

### 3.2. Construction and validation of cell cycle-related risk score model in GBM

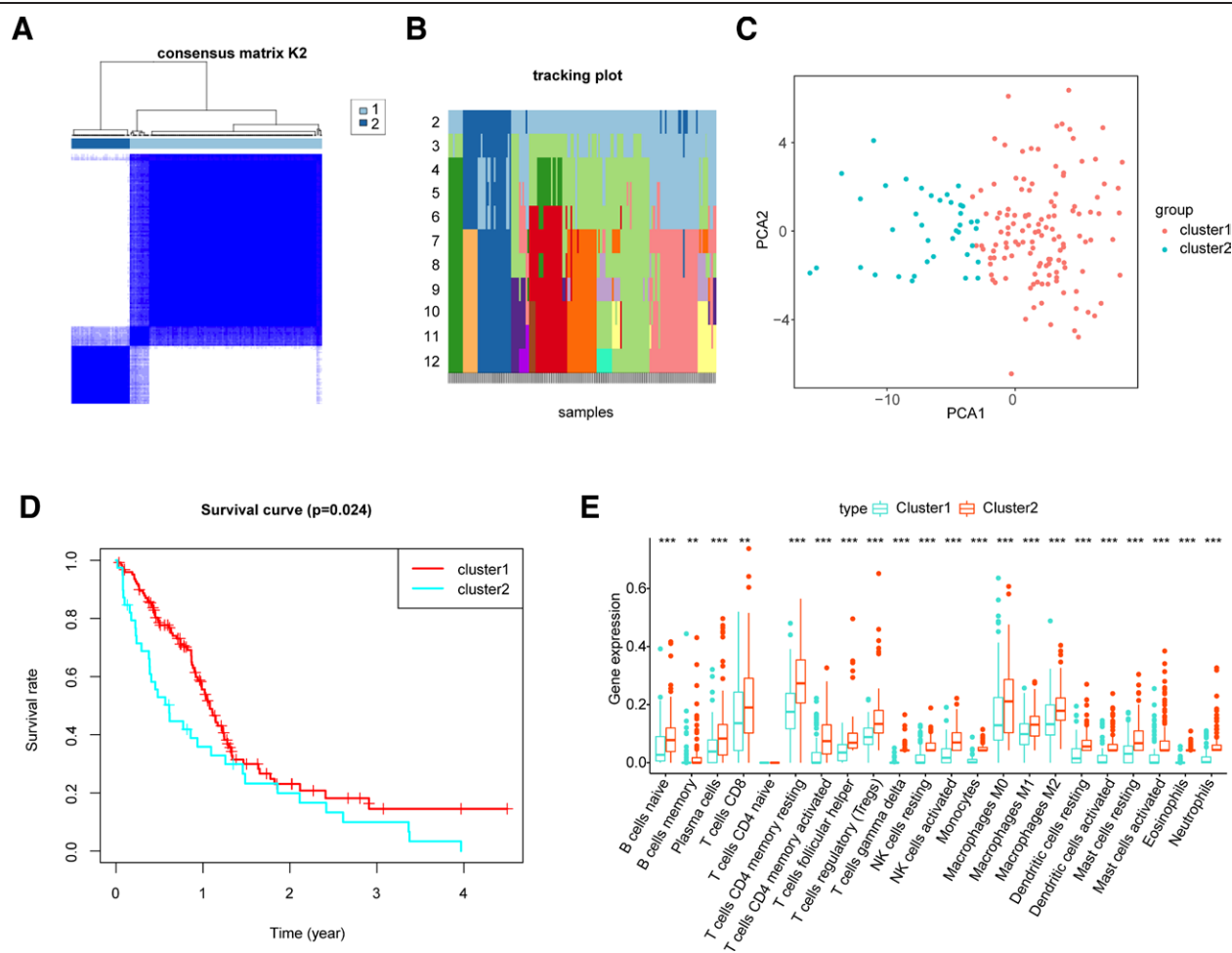
Thereafter, the prognostic value of 58 DECCRGs in GBM was evaluated by univariate Cox and LASSO regression analyses. On univariate Cox regression analysis, *ANAPC4*, *CCNB1*, *CDC25B*, *CHEK1*, *PLK1*, *SMAD3*, *YWHAG*, and *YWHAZ* were identified as being significantly related to the survival of GBM patients. To obtain a more robust prognostic signature in this study, we utilized 10-fold cross-validation to determine the optimal value of  $\lambda$  in the LASSO algorithm, selecting the  $\lambda$  with the smallest partial likelihood deviation as the optimal  $\lambda$ . These 8 DECCRGs, namely *ANAPC4*, *CCNB1*, *CDC25B*, *CHEK1*, *PLK1*, *SMAD3*, *YWHAG*, and *YWHAZ*, were then identified through further stepwise selection (Fig. 5A and B)

and defined as prognostic DECCRGs (Fig. 5A and B). The formula for calculating the risk score of each patient can be derived based on the LASSO regression coefficients. Then the risk score of each GBM patient in the training set was calculated as the expression of *ANAPC4* \*  $-0.082$  + expression of *CCNB1* \*  $0.015$  + expression of *CDC25B* \*  $-0.013$  + expression of *CHEK1* \*  $0.008$  + expression of *PLK1* \*  $0.115$  + expression of *SMAD3* \*  $0.048$  + expression of *YWHAG* \*  $0.699$  + expression of *YWHAZ* \*  $0.60$ . In clinical applications, the risk scoring model based on cell cycle-related genes can serve as a valuable tool for predicting the prognosis of GBM patients. Previous studies have combined single-factor Cox regression and LASSO to build a prognostic model related to copper death in GBM, guiding the prognosis and personalized treatment of GBM.<sup>[25]</sup> A comparison between the performance of the scoring model and existing prognostic models for glioblastoma can provide insights into their potential advantages and clinical utility.

In comparison to existing prognostic models for glioblastoma, the risk scoring model may demonstrated superior prognostic accuracy and discrimination. According to the median value of risk scores, the GBM patients in the TCGA training set were divided into high- and low-risk groups (Fig. 5C). Increase in the risk scores was associated with higher rate of death (Fig. 5D). The high-risk group had worse survival compared to the low-risk group (Fig. 5E). The area under the ROC curve (AUC) for 1-, 3-, and 5-year survival were 0.819, 0.852, and 0.837, respectively (Fig. 5F, G, and H), suggesting an excellent performance of the risk score model in predicting the prognosis of GBM patients. Furthermore, on testing the risk score model in GSE13041 and GSE43378 validation sets, the AUC values were 0.696 and 0.677, respectively, (Fig. 5I and J), demonstrating the reliability of the risk score model in predicting the survival of GBM patients.

### 3.3. Development of the cell cycle-related nomogram for GBM

Next, on multivariate analysis, the risk scores, mRNasi, cancerstatus, and recurrencetype were identified as independent



**Figure 4.** GBM patients were separated into 2 groups. (A) Consensus clustering matrix for  $k = 2$ , which was the optimal cluster number in the cohort. (B) Tracking plot for  $k = 2$  to 12 in the samples. (C) Principal component analysis (PCA) plot of the samples. (D) Kaplan-Meier curves for cluster1 and cluster2. (E) Comparison of the gene expressions between different NMF subgroups in the training cohort. The gene expressions of 22 immune cells are displayed in boxplots.

prognostic factors in GBM (Fig. 6A and B). Using these independent prognostic factors (risk scores, mRNA<sub>si</sub>, cancerstatus, and recurrence), nomograms for predicting the 1-, 3-, and 5-year survival of GBM patients were established (Fig. 6C). Calibration curves showed that the predicted probability of overall survival approximated the actual overall survival (Fig. 6D and F), indicating good performance of the nomogram. Thus, this models could help doctors more accurately assess a patient's risk profile, which was expected to improve patient survival and prognosis.<sup>[26]</sup>

### 3.4. Prognostic cell cycle-related genes may affect the progression of GBM by regulating cell proliferation

Next, we performed GSEA analysis to unravel the potential molecular mechanisms. The results showed significant enrichment of the KEGG pathways participating in the cell proliferation, such as cell cycle, DNA replication and JAK/STAT (Fig. 7), in the high-risk group. For instance, inhibition of JAK or STAT phosphorylation decreases antiapoptotic protein levels, leading to apoptosis in GBM cells.<sup>[27]</sup> Studies have found that curcumin inhibits STAT activation and downstream target genes involved in cell proliferation, migration, and invasion, reducing GBM cell migration, invasion, and proliferation.<sup>[28]</sup> These results suggested that prognostic cell cycle-related genes may regulate the development and progression of GBM via cell proliferation. Therefore, further exploration of how these cell cycle-related

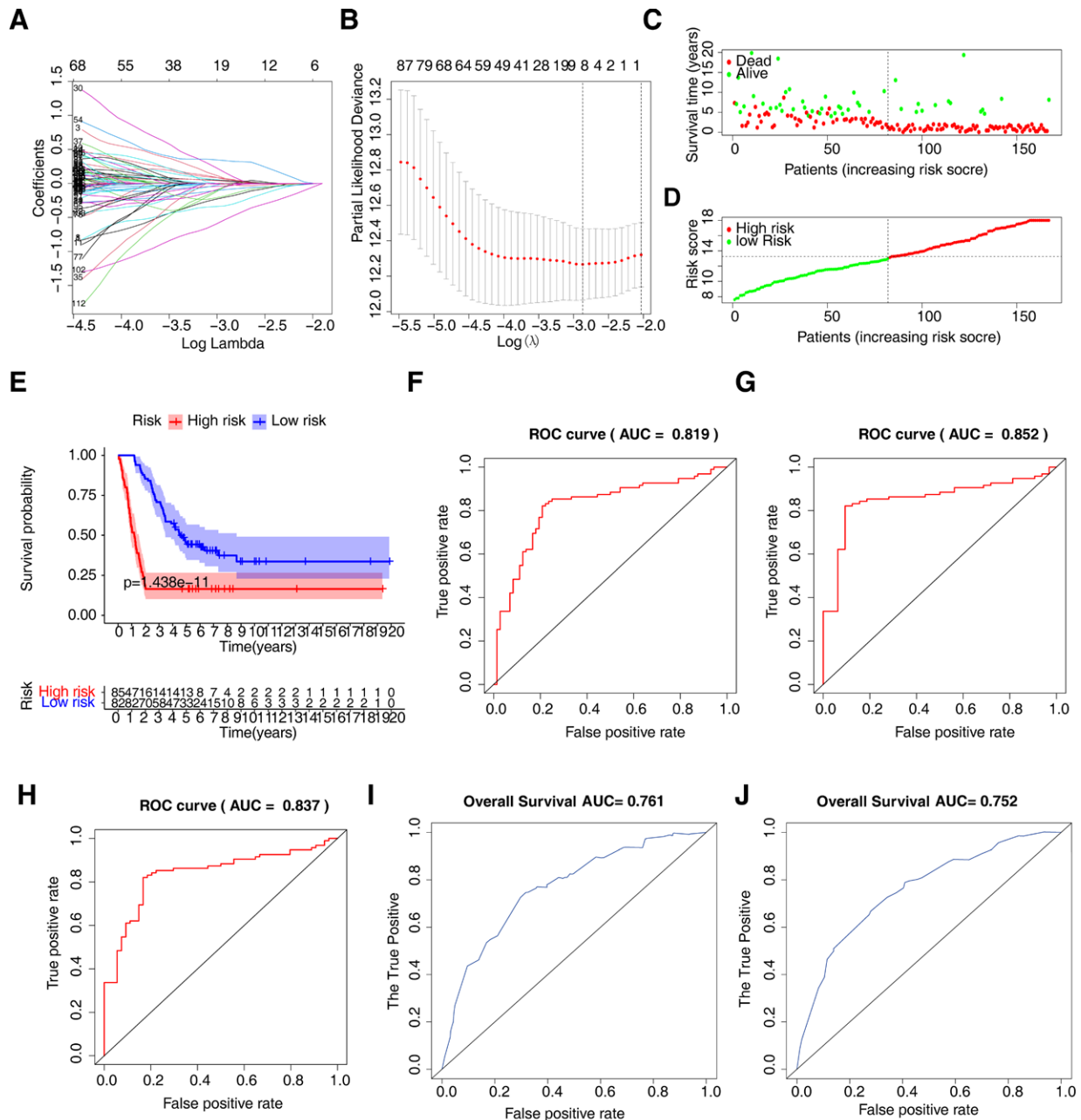
genes affect the growth and progression of glioblastoma would contribute to a deeper understanding of the underlying molecular mechanisms.

## 4. Discussion

The purpose of this study was to identify key genes that affect the prognosis of glioblastoma patients, construct a prognostic model combined with GSVA to predict the survival, and explore the potential regulatory mechanisms in glioma.

Neoplasms remain the main killer worldwide.<sup>[29–34]</sup> In this study, we identified 58 DECCRGs in GBM compared to control tissues. These genes were principally involved in the biological process and signal pathway of cell cycle and DNA replication. The key cellular and molecular mechanisms involved in tumorigenesis are disorder of cell cycle regulation, uncontrolled growth, and reduction of apoptosis. Cell cycle has been shown to play a decisive role in the occurrence and development of GBM.<sup>[35,36]</sup> In both breast cancer (BRCA) and GBM, key roles were played in cell cycle regulation, with lncRNAs TUG1, LINC01355, CASC9, and FOXC2-AS1 impacting cell proliferation and apoptosis in breast cancer, while cell cycle arrest led to reduced proliferation and subsequent cell death in GBM, revealing specific differences in cell cycle regulation between BRCA and GBM.<sup>[37,38]</sup>

Thereafter, we assessed the prognostic value of 58 DECCRGs in GBM using univariate Cox and LASSO regression analyses.

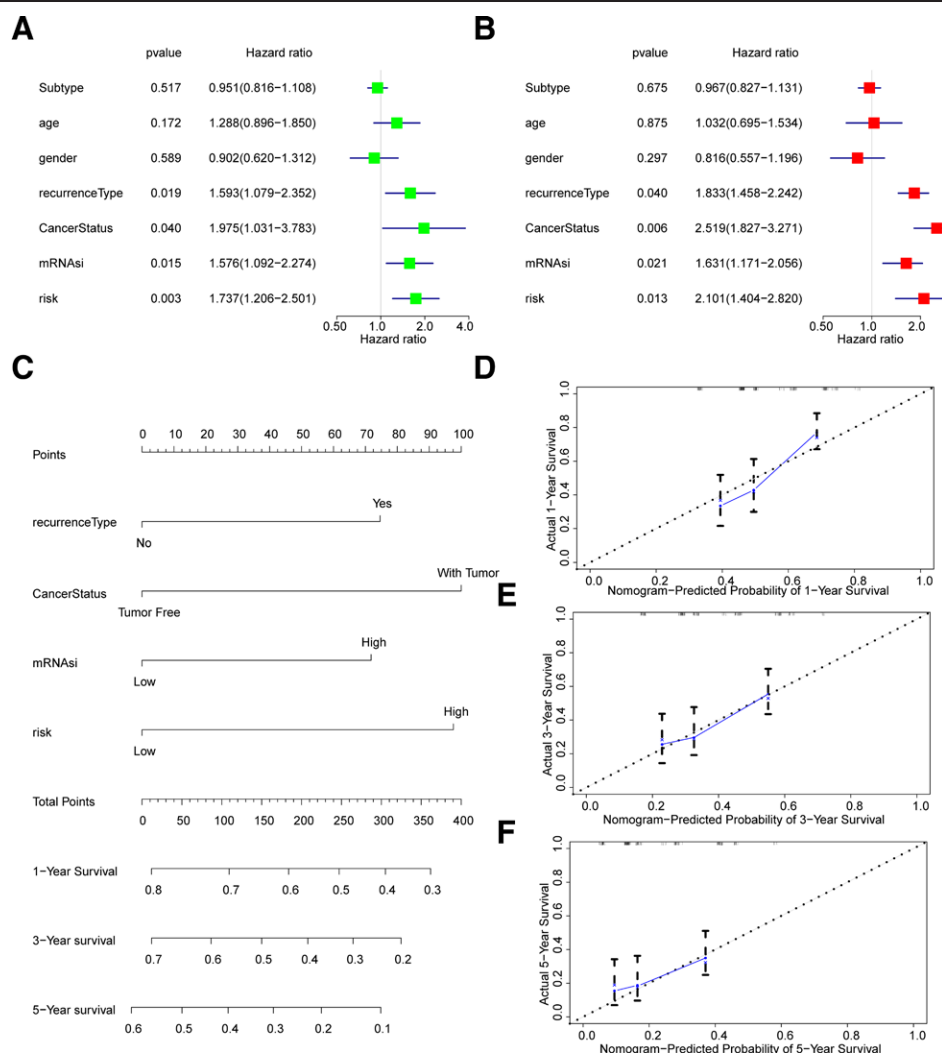


**Figure 5.** Eight DECCRGs were input into LASSO algorithm. (A) LASSO coefficient profiles of the 58 cell cycle-related differentially expressed genes. (B) Selection of optimal  $\lambda$  value in the LASSO model for glioma. (C) Distribution of risk score in the cohort. (D) Scatter plot of patients with different survival statuses. (E) Kaplan-Meier (K-M) survival analyses showing the survival difference between the low- and high-risk groups. (F, G, and H) ROC curves showing the predictive efficiency of the risk model in the 1-, 3-, and 5-year sets. (I and J) ROC curves showing the prediction efficiency of the risk model in the validation set.

On univariate Cox regression analysis, *ANAPC4*, *CCNB1*, *CDC25B*, *CHEK1*, *PLK1*, *SMAD3*, *YWHAG*, and *YWHAZ* were identified as being significantly related to the survival of GBM patients. Based on the results, a risk score model and nomogram for predicting survival of GBM patients were established. These 8 genes are well-established cell cycle factors, which are involved in the regulation of M phase of mitotic cell cycle, cell cycle checkpoint, and cell arrest, etc. It has been found by multiple studies that *CCNB1* regulates the occurrence of cancer, such as hepatocellular carcinoma, pancreatic cancer, and lung adenocarcinoma, by mediating the cell cycle.<sup>[39-41]</sup> In a previous study, the expression of *CCNB1* showed a negative correlation with the therapeutic efficacy of temozolomide (TMZ) in low-grade gliomas.<sup>[42]</sup> Some studies have reported that *CCNB1*, as a biomarker of GBM, overexpression accelerates mitosis and

promotes the proliferation and invasion of GBM tumor cells.<sup>[43]</sup> Therefore, it was speculated that this gene may promote the development of GBM through the regulation of the cell cycle process. In recent years, it has been found that *CHEK1* may play a role in colorectal cancer by affecting cell cycle regulation, and it is closely related to the development and progression of CRC.<sup>[44]</sup> *CHEK1* is an important therapeutic target in gliomas. Inhibition of *CHEK1* may increase the radiosensitivity of gliomas.<sup>[45,46]</sup> Research reports indicate that *CHEK1* plays a crucial role in promoting proliferation and inducing radioresistance in GBM, while its inhibition can enhance apoptosis, suggesting a potential therapeutic target in GBM.<sup>[47]</sup> In addition, studies have reported that the cell cycle-related gene *PLK1* is associated with poor prognosis of cancer, and in addition, the gene may promote cell proliferation and migration of different types of cancer.<sup>[48,49]</sup>





**Figure 6.** Prognostic value of risk score in GBM. (A and B) Univariate and multivariate Cox regression of risk score. (C, D, E, and F) Nomogram to predict the 1-, 2-, and 3-year overall survival rates of patients with GBM. The dashed diagonal line represents the ideal nomogram, and the blue lines represent the 1-, 2-, and 3-year observed nomograms.

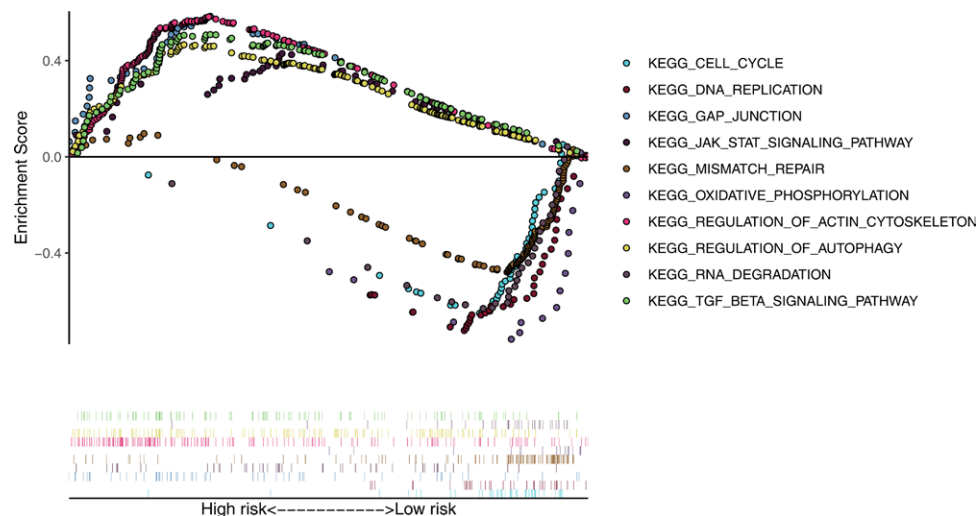
Down-regulation of *PLK1* expression was found to enhance the inhibitory effect of TMZ on glioma stem cells.<sup>[50]</sup> Additionally, dual inhibition of *PLK1* and *STAT3* co-regulates the *MYC* pathway in GBM, leading to reduced *MYC* expression, enhanced cell invasion and apoptosis, and tumor suppression, suggesting a potential therapeutic strategy for targeting GBM.<sup>[51]</sup> These genes, particularly *CCNB1*, *CHEK1*, and *PLK1*, have been implicated in the development and progression of cancer, including gliomas and colorectal cancer. Their potential as therapeutic targets and their impact on treatment outcomes suggest their importance in the prognosis of GBM patients. To summarize, we believe that there is adequate rationale for selection of these 8 cell cycle markers.

Our research results indicated a negative correlation between risk score and prognosis, suggesting that patients with higher risk scores might have poorer prognosis. Furthermore, the K-M curve based on independent prognostic factors could effectively predict patients' survival at different time points, providing new theoretical basis for personalized treatment or better improvement of patients' survival. In addition, the nomogram constructed based on independent prognostic factors can also predict the survival rate of GBM patients at different times. However, due to limitations in the sample resources of the dataset, validation in datasets from more diverse sources might be necessary to ensure the applicability of this model.

Finally, we analyzed the potential mechanism by which these 8 prognostic cell cycle markers regulate GBM. Subsequent GSEA analysis demonstrated that the high-risk group was significantly enriched in the biological processes and signal pathways involved in cell proliferation, metastasis, and invasion. This suggests that patients of high-risk were more likely to exhibit increased cellular growth, metastatic potential, and invasive characteristics, which may contribute to a poorer prognosis or more aggressive disease behavior. Some studies have reported that by promoting cell cycle arrest, it might be possible to slow down or inhibit the proliferation of GBM cells, thereby potentially impacting the aggressiveness and growth of high-risk GBM tumors.<sup>[52]</sup> Therefore, we speculate that these 8 genes regulate metastasis and deterioration through the above biological processes and signal pathways, and ultimately affect the prognosis of patients with GBM.

Some limitations of this study should be acknowledged. The retrospective nature of the study may have induced an element of selection bias. Second, several other factors affect the post-operative outcomes, such as preoperative examination, preoperative comorbidities, operative time, and occurrence of serious postoperative complications, none of which were included in this study. Further studies involving larger datasets are required to establish a more accurate prognostic model for GBM patients. Furthermore, further research with larger datasets





**Figure 7.** KEGG pathway enrichment analysis using GSEA. Cell cycle and DNA replication were significantly enriched in the high-risk group.

and potentially incorporating experimental studies could help in the development of a more robust and accurate prognostic model for GBM patients, which could aid in the improvement of treatment strategies and patient outcomes. Finally, exploration of the therapeutic potential of identified genes, or investigation of the applicability of research models in personalized medicine will provide a roadmap for the advancement of this field in the future.

In summary, this study identified 8 cell cycle-related genes related to prognosis in GBM, explored the potential molecular mechanism, and established a prognostic model for GBM patients.

## Acknowledgments

The manuscript has been edited and proofread by Medjaden Inc., and was supported by the Medjaden Academy & Research Foundation for Young Scientists (Grant No. MJA202306050).

## Author contributions

**Conceptualization:** Zeshang Guo, Xinyu Hong.

**Data curation:** Runpeng Zhou, Kai Zhang.

**Formal analysis:** Runpeng Zhou, Tingting Dai.

**Investigation:** Runpeng Zhou, Kai Zhang, Tian Li.

**Methodology:** Runpeng Zhou.

**Project administration:** Tian Li.

**Resources:** Runpeng Zhou.

**Software:** Runpeng Zhou, Tingting Dai.

**Supervision:** Zeshang Guo.

**Validation:** Runpeng Zhou.

**Visualization:** Zeshang Guo.

**Writing – original draft:** Runpeng Zhou.

**Writing – review & editing:** Kai Zhang, Tingting Dai, Zeshang Guo, Tian Li.

## References

- [1] de Robles P, Fiest KM, Frolkis AD, et al. The worldwide incidence and prevalence of primary brain tumors: a systematic review and meta-analysis. *Neuro Oncol.* 2015;17:776–83.
- [2] Ostrom QT, Cioffi G, Gittleman H, et al. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2012–2016. *Neuro Oncol.* 2019;21(Suppl 5):v1–v100.
- [3] Lapointe S, Perry A, Butowski NA. Primary brain tumours in adults. *Lancet.* 2018;392:432–46.
- [4] Louis DN, Perry A, Reifenberger G, et al. The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta Neuropathol.* 2016;131:803–20.
- [5] Cheng W, Zhang C, Ren X, et al. Treatment strategy and IDH status improve nomogram validity in newly diagnosed GBM patients. *Neuro Oncol.* 2017;19:736–8.
- [6] Wesseling P, Capper D. WHO 2016 classification of gliomas. *Neuropathol Appl Neurobiol.* 2018;44:139–50.
- [7] Yang K, Wu Z, Zhang H, et al. Glioma targeted therapy: insight into future of molecular approaches. *Mol Cancer.* 2022;21:39.
- [8] Tan AC, Ashley DM, López GY, Malinzak M, Friedman HS, Khasraw M. Management of glioblastoma: state of the art and future directions. *CA Cancer J Clin.* 2020;70:299–312.
- [9] Wang W, Lu Z, Wang M, et al. The cuproptosis-related signature associated with the tumor environment and prognosis of patients with glioma. *Front Immunol.* 2022;13:998236.
- [10] Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf.* 2013;14:7.
- [11] Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28:27–30.
- [12] Croft D, O’Kelly G, Wu G, et al. Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res.* 2011;39:D691–697.
- [13] Singh S, Vennila JJ, Snijesh VP, George G, Sunny C. Implying analytic measures for unravelling rheumatoid arthritis significant proteins through drug-target interaction. *Interdiscip Sci.* 2016;8:122–31.
- [14] Bell G, Hey T, Szalay A. Computer science. Beyond the data deluge. *Science.* 2009;323:1297–8.
- [15] Syafruddin SE, Nazarie W, Moidu NA, Soon BH, Mohtar MA. Integration of RNA-Seq and proteomics data identifies glioblastoma multiforme surfaceome signature. *BMC Cancer.* 2021;21:850.
- [16] Wan Z, Zuo X, Wang S, et al. Identification of angiogenesis-related genes signature for predicting survival and its regulatory network in glioblastoma. *Cancer Med.* 2023;12:17445–67.
- [17] Feng Y, Li G, Shi Z, et al. A novel methylation signature predicts radiotherapy sensitivity in glioma. *Sci Rep.* 2020;10:20406.
- [18] Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000;25:25–9.
- [19] Fan W, Wang D, Li G, et al. A novel chemokine-based signature for prediction of prognosis and therapeutic response in glioma. *CNS Neurosci Ther.* 2022;28:2090–103.
- [20] Zhang J, Lian H, Chen K, et al. RECQ1 promotes stress resistance and DNA replication progression through PARP1 signaling pathway in glioblastoma. *Front Cell Dev Biol.* 2021;9:714868.
- [21] Lan Y, Lou J, Hu J, Yu Z, Lyu W, Zhang B. Downregulation of SNRPG induces cell cycle arrest and sensitizes human glioblastoma cells to temozolomide by targeting Myc through a p53-dependent signaling pathway. *Cancer Biol Med.* 2020;17:112–31.
- [22] Du Y, Lin Y, Wang B, et al. Cuproptosis patterns and tumor immune infiltration characterization in colorectal cancer. *Front Genet.* 2022;13:976007.

- [23] Wang J, Qin D, Tao Z, et al. Identification of cuproptosis-related subtypes, construction of a prognosis model, and tumor microenvironment landscape in gastric cancer. *Front Immunol.* 2022;13:1056932.
- [24] Wang Z, Su G, Dai Z, et al. Circadian clock genes promote glioma progression by affecting tumour immune infiltration and tumour cell proliferation. *Cell Prolif.* 2021;54:e12988.
- [25] Yan T, Yang H, Meng Y, et al. Targeting copper death genotyping associated gene RARRES2 suppresses glioblastoma progression and macrophages infiltration. *Cancer Cell Int.* 2023;23:105.
- [26] Dai L, Zhou P, Lyu L, Jiang S. Systematic analysis based on the cuproptosis-related genes identifies ferredoxin 1 as an immune regulator and therapeutic target for glioblastoma. *BMC Cancer.* 2023;23:1249.
- [27] Weissenberger J, Priester M, Bernreuther C, et al. Dietary curcumin attenuates glioma growth in a syngeneic mouse model by inhibition of the JAK1,2/STAT3 signaling pathway. *Clin Cancer Res.* 2010;16:5781–95.
- [28] Wong SC, Kamarudin MNA, Naidu R. Anticancer mechanism of curcumin on human glioblastoma. *Nutrients.* 2021;13:950.
- [29] Guo W, Qiao T, Li T. The role of stem cells in small-cell lung cancer: evidence from chemoresistance to immunotherapy. *Semin Cancer Biol.* 2022;87:160–9.
- [30] Hu W, Yang Y, Fan C, et al. Clinical and pathological significance of N-Myc downstream-regulated gene 2 (NDRG2) in diverse human cancers. *Apoptosis.* 2016;21:675–82.
- [31] Li T, Qiao T. Unraveling tumor microenvironment of small-cell lung cancer: implications for immunotherapy. *Semin Cancer Biol.* 2022;86(Pt 2):117–25.
- [32] Yang Z, Jiang S, Lu C, et al. SOX11: friend or foe in tumor prevention and carcinogenesis? *Ther Adv Med Oncol.* 2019;11:1758835919853449.
- [33] Sun M, Liu X, Xia L, et al. A nine-lncRNA signature predicts distant relapse-free survival of HER2-negative breast cancer patients receiving taxane and anthracycline-based neoadjuvant chemotherapy. *Biochem Pharmacol.* 2021;189:114285.
- [34] Wu X, Yan F, Wang L, et al. MicroRNA: another pharmacological avenue for colorectal cancer? *Front Cell Dev Biol.* 2020;8:812.
- [35] Meyer M, Reimand J, Lan X, et al. Single cell-derived clonal analysis of human glioblastoma links functional and genomic heterogeneity. *Proc Natl Acad Sci U S A.* 2015;112:851–6.
- [36] Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature.* 2008;455:1061–8.
- [37] Pezuk JA, Brассесco MS, Morales AG, et al. Polo-like kinase 1 inhibition causes decreased proliferation by cell cycle arrest, leading to cell death in glioblastoma. *Cancer Gene Ther.* 2013;20:499–506.
- [38] Zangouei AS, Zangoue M, Taghehchian N, et al. Cell cycle related long non-coding RNAs as the critical regulators of breast cancer progression and metastasis. *Biol Res.* 2023;56:1.
- [39] Bao B, Yu X, Zheng W. MiR-139-5p targeting CCNB1 modulates proliferation, migration, invasion and cell cycle in lung adenocarcinoma. *Mol Biotechnol.* 2022;64:852–60.
- [40] Rong MH, Li JD, Zhong LY, et al. CCNB1 promotes the development of hepatocellular carcinoma by mediating DNA replication in the cell cycle. *Exp Biol Med (Maywood).* 2022;247:395–408.
- [41] Zhang H, Zhang X, Li X, et al. Effect of CCNB1 silencing on cell cycle, senescence, and apoptosis through the p53 signaling pathway in pancreatic cancer. *J Cell Physiol.* 2018;234:619–31.
- [42] Wang Q, He Z, Chen Y. Comprehensive analysis reveals a 4-gene signature in predicting response to temozolomide in low-grade glioma patients. *Cancer Control.* 2019;26:1073274819855118.
- [43] Cui K, Chen JH, Zou YF, et al. Hub biomarkers for the diagnosis and treatment of glioblastoma based on microarray technology. *Technol Cancer Res Treat.* 2021;20:1533033821990368.
- [44] Sun J, Li M, Lin T, et al. Cell cycle arrest is an important mechanism of action of compound Kushen injection in the prevention of colorectal cancer. *Sci Rep.* 2022;12:4384.
- [45] Qin T, Mullan B, Ravindran R, et al. ATRX loss in glioma results in dysregulation of cell-cycle phase transition and ATM inhibitor radiosensitization. *Cell Rep.* 2022;38:110216.
- [46] Lee JK, Chang N, Yoon Y, et al. USP1 targeting impedes GBM growth by inhibiting stem cell maintenance and radioresistance. *Neuro Oncol.* 2016;18:37–47.
- [47] Bai X, Wang J, Huo L, et al. Serine/threonine kinase CHEK1-dependent transcriptional regulation of RAD54L promotes proliferation and radio resistance in glioblastoma. *Transl Oncol.* 2018;11:140–6.
- [48] Kahl I, Mense J, Finke C, et al. The cell cycle-related genes RHAMM, AURKA, TPX2, PLK1, and PLK4 are associated with the poor prognosis of breast cancer patients. *J Cell Biochem.* 2022;123:581–600.
- [49] Ning G, Lu C, Chen Y, Jiang M, Si P, Zhang R. Transcription factor ZEB1 regulates PLK1-mediated SKA3 phosphorylation to promote lung cancer cell proliferation, migration and cell cycle. *Anticancer Drugs.* 2023;34:866–76.
- [50] Liu N, Hu G, Wang H, Li Z, Guo Z. PLK1 inhibitor facilitates the suppressing effect of temozolomide on human brain glioma stem cells. *J Cell Mol Med.* 2018;22:5300–10.
- [51] Wang H, Tao Z, Feng M, et al. Dual PLK1 and STAT3 inhibition promotes glioblastoma cells apoptosis through MYC. *Biochem Biophys Res Commun.* 2020;533:368–75.
- [52] Wang Y, Sakaguchi M, Sabit H, et al. COL1A2 inhibition suppresses glioblastoma cell proliferation and invasion. *J Neurosurg.* 2023;138:639–48.