Prognostic Impact of a Lymphocyte Activation-Associated Gene Signature in GBM based on Transcriptome Analysis

Yujia Lan
Harbin Medical University
Erjie Zhao
Harbin Medical University
Xinxin Zhang
Harbin Medical University
Xiaojing Zhu
Harbin Medical University
Linyun Wan
Harbin Medical University
Suru A
Harbin Medical University
Yihan Wang
Harbin Medical University
Yanyan Ping (pingyanyan@hrbmu.edu.cn)
Harbin Medical University

Research Article

Keywords: Glioblastoma multiforme (GBM), Lymphocyte,

Posted Date: November 24th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-107112/v1

Prognostic impact of a lymphocyte activation-associated gene signature in GBM based on transcriptome analysis

Yujia Lan[1,†], Erjie Zhao[1,†], Xinxin Zhang[1,†], Xiaojing Zhu[1], Linyun Wan[1], Suru A[1], Yihan Wang[1,*], and Yanyan Ping[1,*]

[1] College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, Heilongjiang 150086, China.

[†] **Contributed equally to this work.**

[*] **Corresponding authors:**

Yanyan Ping, PhD, College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China. Email: pingyanyan@hrbmu.edu.cn

Yihan Wang, PhD, College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China. Email: wangyihan@hrbmu.edu.cn

Abstract

Background Glioblastoma multiforme (GBM) is the most aggressive primary central nervous system malignant tumor that has poor prognosis. Lymphocyte activation played important roles in cancer and therapy.

Objective To identify an efficient lymphocyte activation-associated gene signature that could predict the progression and prognosis of GBM.

Methods We used univariate Cox proportional hazards regression and stepwise regression algorithm to develop a lymphocyte activation-associated gene signature in the training data set (TCGA, n = 525). Then, the signature was validated in two data sets, including GSE16011 (n = 150) and GSE13041 (n = 191) by the Kaplan Meier method. Univariate and multivariate Cox proportional hazards regression models were used to adjust for clinicopathological factors.

Results In the training data set, we identified a lymphocyte activation-associated gene signature (*TCF3*, *IGFBP2*, *TYRO3* and *NOD2*), which classified the patients into high-risk and low-risk groups with significant differences in overall survival (median survival 15.33 months vs 12.57 months, HR = 1.55, 95% CI = 1.28-1.87, logrank test $P < 0.001$). In the other two data sets, this signature showed similar prognostic values. Further, univariate and multivariate Cox proportional hazards regression models analysis indicated that the signature was an independent prognostic factor for GBM patients.Moreover, we found that there were differences in lymphocyte activity between the high- and low-risk groups of GBM patients among all data sets. Furthermore, by longitudinal analysis, the lymphocyte activation-associated gene signature could significantly predict the survival of patients with some features, such as IDH-wildtype patients and patients with radiotherapy. In addition, the signature could improve prognostic power of age.

Conclusions In summary, our results suggested that the lymphocyte activation-associated gene signature is a promising factor for the survival of patients, which may be helpful for diagnosis, prognosis, and treatment of GBM patients.

Introduction

Glioblastoma multiforme (GBM) as grade IV diffuse glioma, is one of the most aggressive and lethal brain cancers. GBM often occurs in the cerebral hemispheres, which can quickly spread into the other parts of brain [1, 2]. Unfortunately, the recurrence rate for GBM is high. IDH status is a classical marker of GBM， patients with IDH mutation are tend to have a better survival [3-5]. In previous studies, GBM patients can be divided into four distinct molecular subtypes based on gene expression profiling, including proneural, neural, classical and mesenchymal [6, 7]. The standard treatment for GBM is surgery, followed by radiation and chemotherapy [1, 2, 8]. In the last few years, some immunotherapies have been introduced in GBM, especially for pateins with EGFR mutations. However, the median overall survival of GBM patients is only 15 months, which suggested that the drug treatments might be ineffective for most of patients [9].

The lymphocyte activation was a process where the lymphocytes stimulate by specific antigen or nonspecific mitogens, which resulting in synthesis of protein and production of lymphokines [10, 11]. This process would affect proliferation and differentiation of various effector and memory cells. The effector cells were in response to first-time exposure to an antigen, which was called primary immune response. And, memory cells were in response to a second immune response, which is known as immunological memory. Lymphocyte activation was destroyed in cancer and is important for immunotherapy [12, 13]. Some lymphocyte activation-associated markers were found to be associated with favorable overall survival and effective for the treatments of patients, for example, lymphocyte activation gene 3 (LAG-3) in cancers [14-16]. In addition, the lymphocyte activation combing with radiotherapy will be a new treatment regimen for cancer patients [17, 18]. Therefore, it would be valuable to explore a lymphocyte activation-related signature for prognosis of GBM patients, which would be helpful for the management of patients.

In this study, we developed a lymphocyte activation-associated gene signature which can predict the overall survival of GBM patients in training data set and validated its prognostic value in another two independent cohorts. The signature also was an independently prognostic factor after adjusting other clinicopathologic factors. Moreover, this signature not only predicted the overall survival of patients with IDH wild-type glioblastoma, but also patients after radiotherapy. These findings indicated the signature could serve as an effective prognostic biomarker for patients with GBM.

Methods

Datasets description

Three gene expression data sets (TCGA, GSE16011 [19] and GSE13041 [20]) were downloaded from The Cancer Genome Atlas (TCGA) portal (https://portal.gdc.cancer.gov) and Gene Expression Omnibus (GEO), which contained patient outcome and clinicopathological factors. These data were filtered to exclude patients without matched clinical information. In total, 866 GBM patients were included in our study. The TCGA data set was treated as training cohorts, and the other two datasets (GSE16011 and GSE13041) were served as independent validation cohorts. In the training data set, the median survival of GBM patients was 12.39 months (range from 0.1 to 127.5). And there were 144 classical, 155 mesenchymal, 83 Neural, and 99 proneural GBM samples. In addition, other clinicopathological characteristics of TCGA GBM patients were summarized in Table 1. The median survival of GBM patients was 58 years old (range from 26 to 90) and 58 years old (range from 26 to 90) in the validation data sets, respectively (Table 2).

Identifying the lymphocyte activation-associated gene signature

Differential expression analysis was performed based on transcriptome profile of GBM patients by the "limma" R package in training data set. Genes with the cutoff criteria of |log2-fold change| ≥1 and FDR < 0.000001 between tumor and normal tissues were regarded as differentially expressed genes (DEGs). Using univariate Cox proportional hazards regression algorithm, we filtered out the DEGs which was strong correlation with survival of patients with GBM ($P < 0.05$). Among these prognostic-associated DEGs, we focused on the genes involved in lymphocyte activation process. The lymphocyte activation-related genes were collected by GO terms in Gene Oncology. Then, we performed stepwise regression analysis to select genes associated with overall survival in the training data set. There were four genes (*TCF3*, *IGFBP2*, *TYRO3* and *NOD2*) were identified in our study (Supplementary Figure 1). Finally, a risk score model was developed based on the gene expression weighted by regression coefficients of univariable Cox regression: Risk score = (-0.6246210 × expression level of *TCF3*) + (0.2992 × expression level of *IGFBP2*) + (0.2421068 × expression level of *TYRO3*) + (0.2155469 × expression level of *NOD2*). Based on the risk score model, the patients were assigned to the high- and low-risk groups according to the median risk score in all data sets.

Statistical analysis

For survival analysis, overall survival was used as the end points. The Kaplan-Meier method was performed for the visualization purposes and the differences between survival curves were calculated by log-rank test. Univariate and multivariate Cox proportional hazards regression models were applied to estimate the prognostic capability of the lymphocyte activation-associated gene signature. The Cox proportional hazards regression model was used to calculated hazard ratio (HR) and 95% confidence intervals (CI). The *P*-values smaller than 0.05 were considered to be statistically significant. The concordance index (C-index) was used to compare the specific prognostic efficacy among the signature, age, sex, and the combined model using a logistic regression with above three variables. All of the statistical analyses were performed using R software (http://www.r-project.org), version 3.5.1.

Results

Development of a prognostic signature related to lymphocyte activation

The transcriptional profile of GBM patients and corresponding clinical information were obtained from TCGA. There were 525 tumor samples and 10 normal samples in the training set. 2,364 differentially expressed genes were identified between tumor and normal samples by using limma package (FDR < 0.000001 and |log2-fold change| ≥1), including 933 up-regulated and 1,431 down-regulated genes (Supplementary Figure 2). Then we applied univariate Cox proportional hazards regression analysis to these DEGs and found 173 differentially expressed genes with prognostic power. By subjecting these genes to the function of lymphocyte activation, which was obtained from Gene Oncology (GO), we obtained 12 lymphocyte activation-associated genes for subsequent analysis. Finally, we obtained four genes (*TCF3*, *IGFBP2*, *TYRO3* and *NOD2*) that were significantly related to overall survival by using stepwise regression algorithm. We derived lymphocyte activation-associated gene signature to calculate the risk score for every patient based on the expression levels of these four genes weighted by regression coefficients of Cox proportional hazards regression modeling (Supplementary Figure 1).

We calculated the lymphocyte activation-associated gene signature risk score for each patient in the training cohort, and the patients were divided into high-risk (n = 262) and low-risk (n=263) groups according to the median risk score (2.752) as the cutoff. Patients within high-risk group had significantly shorter overall survival time than those within low-risk group (median survival 15.33 months vs 12.57 months, HR = 1.55, 95% CI = 1.28-1.87, logrank test $P < 0.001$, Fig.1a). *IGFBP2* showed relatively higher expression in the high-risk group (Fig.1b). Moreover, the low-risk patients presented a trend of long survival time (Fig.1c).

Prognostic value of the lymphocyte activation-associated gene signature in validation sets

Next, we validated the prognostic value of the lymphocyte activation-associated gene signature in another two independent validation sets, GSE13041 (n = 191) and GSE16011 (n = 150). The risk score for each patient in these two sets was calculated with the same formula used in the training set. In GSE16011, the overall survival time of patients within high-risk group were significantly shorter than those cases with within low-risk group (median survival 7.08 months vs 10.32 months, HR = 1.57, 95% CI = 1.13-2.17, logrank test $P = 0.069$, Fig.2a). The expression of *IGFBP2* and *TYRO3* had relatively higher expression in the high-risk group (Fig.2b). We also found the low-risk patients presented a trend of long survival time in GSE16011 (Fig.2c). Moreover, the signature also significantly predicted the overall survival of patients with GBM in GSE13041 (median survival of high-risk patients 11.97 months vs median survival of low-risk patients 14.47 months, HR = 1.42, 95% CI = 1.05-1.92, logrank test $P = 0.02$, Fig.3a-c).

The signature independently predicted overall survival of GBM patients

To assess whether the survival prediction ability of the lymphocyte activation-associated gene signature was independent of other clinicopathologic factors in GBM, univariate and multivariable Cox regression analysis were performed. The covariables included age, sex, IDH status, subtypes, and the prognostic signature. We found that the signature (HR = 1.46, 95% CI 1.16 to 1.83, $P = 0.001$), age (HR = 1.03, 95% CI 1.02 to 1.04, $P < 0.001$), sex (HR = 1.30, 95% CI 1.04 to 1.63, $P = 0.02$) and proneural types (HR =

1.64, 95% CI 1.18 to 2.27, $P = 0.003$) independently predicted poorer OS of GBM patients in training data set (Table 3). In the validation dataset GSE16011, we also found the lymphocyte activation-associated gene signature had independently prognostic value (HR = 1.36, 95% CI 1.35 to 1.98, $P = 0.029$, Table 4). In addition, this signature independently predicted the survival of patients with GBM, with marginal significance (HR = 1.33, 95% CI 0.98 to 1.81, $P = 0.06$, Supplementary Table 1). These results indicated that the predictive ability of the lymphocyte activation-associated gene signature was independent of clinicopathological factors for OS in GBM.

The differences in lymphocyte activity between the high- and low-risk groups of patients

The GBM patients were divided into low- and high-risk groups by the lymphocyte activation-associated gene signature. Lymphocytes include natural killer cells, T cells and B cells. Therefore, we compare lymphocyte activity in these two groups of patients. In TCGA, five types of lymphocytes show significant differences in activity between high- and low-risk patients by wilcox test ($P < 0.05$), including memory B cells, naive B cells, naive CD4 T cells, follicular helper T cells and regulatory T cells (Fig.4a and Supplementary Figure 3). And the lymphocyte activity of high-risk patients was significantly lower than those of low-risk patients. The similar phenomena were observed in GSE13041 data set (Fig.4b). However, activated CD4 memory T cells, resting CD4 memory T cells, CD8 T cells and gamma delta T cells showed significant differences between two groups. In GSE16011, there were three types of cells showed different activity between low- and high-risk patients ($P < 0.05$, Supplementary Figure 4). These findings suggested that there were differences in lymphocyte activity between the high- and low-risk groups of GBM patients.

Stratification analysis of the lymphocyte activation-associated gene signature

IDH mutation is one of the most critical genomic alterations in GBM. IDH-wildtype GBM survived shorter than those with IDH1-mutation GBM. The lymphocyte activation-associated gene signature significantly predicted overall survival of IDH-wildtype GBM patients by Kaplan–Meier survival analysis (logrank test, $P = 0.0043$), while it did not predict the survival of IDH-mutation patients (Fig. 5a). The IDH-wildtype patients were divided into high-risk and low-risk groups using the same cutoff in training set. The high-risk group of IDH-wildtype patients had significantly shorter OS than those within low-risk group (median survival 12.93 months vs 14.93 months, HR = 1.39, 95% CI = 1.11-1.74, $P = 0.0043$, Fig.5a). The similar phenomenon was also found in GSE16011 data set. The high-risk patients had shorter survival than low-risk patients (median survival 7.14 months vs 10.68 months, HR = 1.95, 95% CI = 1.25-3.02, $P = 0.0025$) among IDH-wildtype patients (Fig.5b).

We next explored whether the lymphocyte activation-associated gene signature was effective for GBM patients within four transcriptome subtypes (proneural, neural, classical, and mesenchymal) established by The Cancer Genome Atlas network using Kaplan–Meier survival analysis. In mesenchymal and proneural subtypes, the signature significantly predicted the overall survival of patients (logrank test P value = 0.013 for mesenchymal subtypes and P value = 0.041 for proneural subtypes, Fig.6). The patients within high-risk groups had shorter survival time than those within low-risk groups (median survival 11.83 months vs 15.3 months for mesenchymal subtypes and 9 months vs 13.17 months for proneural subtypes). In addition, the signature did not have prognostic ability of patients within neural and classical subtypes (Supplementary Figure 5).

Then, we explored whether the radiotherapy was effective for GBM patients. The patients were divided into two groups (high-risk group and low-risk groups). We observed high risk score significantly predicted poor overall survival of GBM patients with radiotherapy in GSE16011 data set (median survival 8.76 months vs 14.64 months, HR = 1.86, 95% CI = 1.27-2.72, P = 0.0011) and GSE13041 (median survival 6.03 months vs 12.53 months, HR = 3.18, 95% CI = 1.13-8.94, P = 0.022, Fig.7).

Comparison of prognostic power between clinical factors and the signature

We performed C-index calculation to evaluate the prognostic impact of clinical factors (age and sex) and the signature in three data sets. Another prognostic model was constructed by combining our signature with clinical factors. There were significantly differences between age and the signature ($P < 0.05$). We also found the same phenomena between age and sex. Moreover, for age, the clinical factors combined with the signature had a higher C-index (0.646 for TCGA, 0.665 for GSE16011 and 0.600 for GSE13041) than age histological grade alone (0.645 for TCGA, 0.659 for GSE16011 and 0.595 for GSE13041, Supplementary Table 2). These results suggesting that the lymphocyte activation-associated gene signature could add complementary value to known clinical factors.

DISCUSSION

In this study, we identified a lymphocyte activation-associated gene signature that predicted the overall survival of patients with GBM. The prognostic value of the signature was evaluated among three independent data sets by Kaplan–Meier and univariate and multivariate Cox analyses. Moreover, the high-risk and low-risk patients separated by the signature showed significant differences in immune cell activity. This signature was an independent prognostic factor after adjusting for some clinicopathological factors. In addition, the lymphocyte activation-associated gene signature also could predict the survival benefits of patients with IDH wild-type status and those after radiotherapy.

We developed a prognostic signature including four genes (*TCF3*, *IGFBP2*, *TYRO3* and *NOD2*). There were four genes in this signature. *IGFBP2* was one of the Insulin-like growth factor binding proteins (IGFBPs), which are a family of proteins binding to Insulin-like growth factors. The expression of *IGFBP2* increased in peripheral blood mononuclear cells and participated in lymphocyte proliferation [21]. In various autoimmune diseases, *IGFBP2* can be as potential biomarkers and therapeutic targets [22]. Moreover, *IGFBP2* predicted prognosis of gastric cancer patients. *NOD2* may play an important role in the pathogenesis of some diseases, such as, oral lichen planus [23], crohn's disease [24], and inflammatory bowel disease [25]. The expression of *NOD2* is higher in activated/memory CD4+ T cells [26], which represented a new diagnostic and treatment target of disease [23]. Another marker, *TCF3* played a role in germinal center B - cell development and promoted cell growth, which contributed to proliferative phenotype in burkitt lymphoma [27-29]. Also, *TCF3* promotes survival in lymphoid cells [29].

In conclusion, our results identified a lymphocyte activation-associated gene signature with prognostic power and offered new insights for the treatments of GBM. However, more data are needed to test the prognostic value of the signature before it is applied clinically. And whether the signature is useful for the prediction of the prognostic benefits of therapy for GBM patients requires researches with enough patients with therapy information. In summary, the results may refine the future treatment of GBM and further help achieve the goal of precision medicine for GBM patients.

Author contributions

YYP, YJL and YHW conceived and designed the project. YJL, EJZ and XXZ acquired the data. YJL and EJZ performed the statistical analysis and analyzed and interpreted all the data. YJL, XJZ, LYW and SRA prepared the figures and tables. YJL and EJZ wrote the paper. YYP and YHW reviewed and revised the manuscript. All authors approved the final manuscript.

Funding

This work was supported in part by the Heilongjiang Postdoctoral Science Foundation (LBH-Z19082), the Special funds for the construction of higher education in Heilongjiang Province [Grant No. UNPYSCT-2018068].

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Alifieris, C. and D.T. Trafalis, *Glioblastoma multiforme: Pathogenesis and treatment.* Pharmacol Ther, 2015. **152**: p. 63-82.
2. Batash, R., et al., *Glioblastoma Multiforme, Diagnosis and Treatment; Recent Literature Review.* Curr Med Chem, 2017. **24**(27): p. 3002-3009.
3. Aldape, K., et al., *Glioblastoma: pathology, molecular mechanisms and markers.* Acta Neuropathol, 2015. **129**(6): p. 829-48.
4. Reuss, D.E., et al., *IDH mutant diffuse and anaplastic astrocytomas have similar age at presentation and little difference in survival: a grading problem for WHO.* Acta Neuropathol, 2015. **129**(6): p. 867-73.
5. Burgenske, D.M., et al., *Molecular profiling of long-term IDH-wildtype glioblastoma survivors.* Neuro Oncol, 2019. **21**(11): p. 1458-1469.
6. Lee, J.H., et al., *Human glioblastoma arises from subventricular zone cells with low-level driver mutations.* Nature, 2018. **560**(7717): p. 243-247.
7. Liang, Y., et al., *Gene expression profiling reveals molecularly and clinically distinct subtypes of glioblastoma multiforme.* Proc Natl Acad Sci U S A, 2005. **102**(16): p. 5814-9.
8. Wick, W., et al., *Treatment of glioblastoma in adults.* Ther Adv Neurol Disord, 2018. **11**: p. 1756286418790452.
9. Zanders, E.D., F. Svensson, and D.S. Bailey, *Therapy for glioblastoma: is it working?* Drug Discov Today, 2019. **24**(5): p. 1193-1201.
10. Kay, J.E., *Mechanisms of T lymphocyte activation.* Immunol Lett, 1991. **29**(1-2): p. 51-4.
11. Hodgkin, P.D., et al., *The importance of efficacy and partial agonism in evaluating models of B lymphocyte activation.* Int Rev Immunol, 1997. **15**(1-2): p. 101-27.
12. Burugu, S., A.R. Dancsok, and T.O. Nielsen, *Emerging targets in cancer immunotherapy.* Semin Cancer Biol, 2018. **52**(Pt 2): p. 39-52.
13. Ware, C.F., *Targeting lymphocyte activation through the lymphotoxin and LIGHT pathways.* Immunol Rev, 2008. **223**: p. 186-201.
14. Wang, Y., et al., *Lymphocyte-Activation Gene-3 Expression and Prognostic Value in Neoadjuvant-Treated Triple-Negative Breast Cancer.* J Breast Cancer, 2018. **21**(2): p. 124-133.
15. Saleh, R.R., et al., *Prognostic Value of Lymphocyte-Activation Gene 3 (LAG3) in Cancer: A Meta-Analysis.* Front Oncol, 2019. **9**: p. 1040.
16. He, Y., et al., *Lymphocyte-activation gene-3, an important immune checkpoint in cancer.* Cancer

Sci, 2016. **107**(9): p. 1193-7.

17. Akiyoshi, T., et al., *Immunogenomic profiles associated with response to neoadjuvant chemoradiotherapy in patients with rectal cancer.* Br J Surg, 2019. **106**(10): p. 1381-1392.
18. Sato, H., P.A. Jeggo, and A. Shibata, *Regulation of programmed death-ligand 1 expression in response to DNA damage in cancer cells: Implications for precision medicine.* Cancer Sci, 2019. **110**(11): p. 3415-3423.
19. Gravendeel, L.A., et al., *Intrinsic gene expression profiles of gliomas are a better predictor of survival than histology.* Cancer Res, 2009. **69**(23): p. 9065-72.
20. Lee, Y., et al., *Gene expression analysis of glioblastomas identifies the major molecular basis for the prognostic benefit of younger age.* BMC Med Genomics, 2008. **1**: p. 52.
21. Hettmer, S., et al., *Effects of insulin-like growth factors and insulin-like growth factor binding protein-2 on the in vitro proliferation of peripheral blood mononuclear cells.* Hum Immunol, 2005. **66**(2): p. 95-103.
22. Ding, H. and T. Wu, *Insulin-Like Growth Factor Binding Proteins in Autoimmune Diseases.* Front Endocrinol (Lausanne), 2018. **9**: p. 499.
23. Ahn, M.Y., et al., *Expression of nucleotide-binding oligomerization domain 1 and 2 in oral lichen planus.* J Dent Sci, 2020. **15**(1): p. 1-8.
24. Niess, J.H., et al., *NOD2 polymorphism predicts response to treatment in Crohn's disease--first steps to a personalized therapy.* Dig Dis Sci, 2012. **57**(4): p. 879-86.
25. Franchi, L., et al., *Intracellular NOD-like receptors in innate immunity, infection and disease.* Cell Microbiol, 2008. **10**(1): p. 1-8.
26. Zanello, G., et al., *Nod2 activates NF-kB in CD4+ T cells but its expression is dispensable for T cell-induced colitis.* PLoS One, 2013. **8**(12): p. e82623.
27. Richter, J., et al., *Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing.* Nat Genet, 2012. **44**(12): p. 1316-20.
28. Dave, S.S., et al., *Molecular diagnosis of Burkitt's lymphoma.* N Engl J Med, 2006. **354**(23): p. 2431-42.
29. Sakata-Yanagimoto, M., et al., *Disease-specific mutations in mature lymphoid neoplasms: recent advances.* Cancer Sci, 2014. **105**(6): p. 623-9.

Figure Legends

Figure 1 The survival analysis of the lymphocyte activation-associated gene signature in training set (TCGA, n=525). (a) Kaplan–Meier curves of overall survival for the four gene signature. The patients were divided into high-risk (red) and low-risk group (grey). (b) The expression profiles of the four genes in the signature. (c) The distribution of the GBM patients' overall survival status.

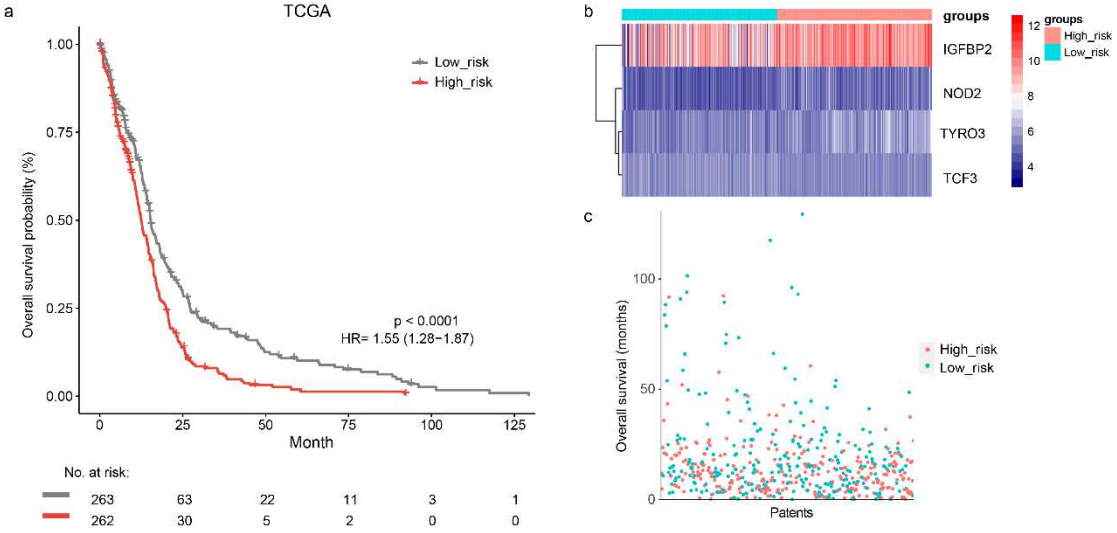

Figure 2 The survival analysis of the lymphocyte activation-associated gene signature in GSE16011 (n = 150). (a) Kaplan–Meier curves of overall survival for the four gene signature. (b) The expression profiles of the four genes in the signature. (c) The distribution of the GBM patients' overall survival status.

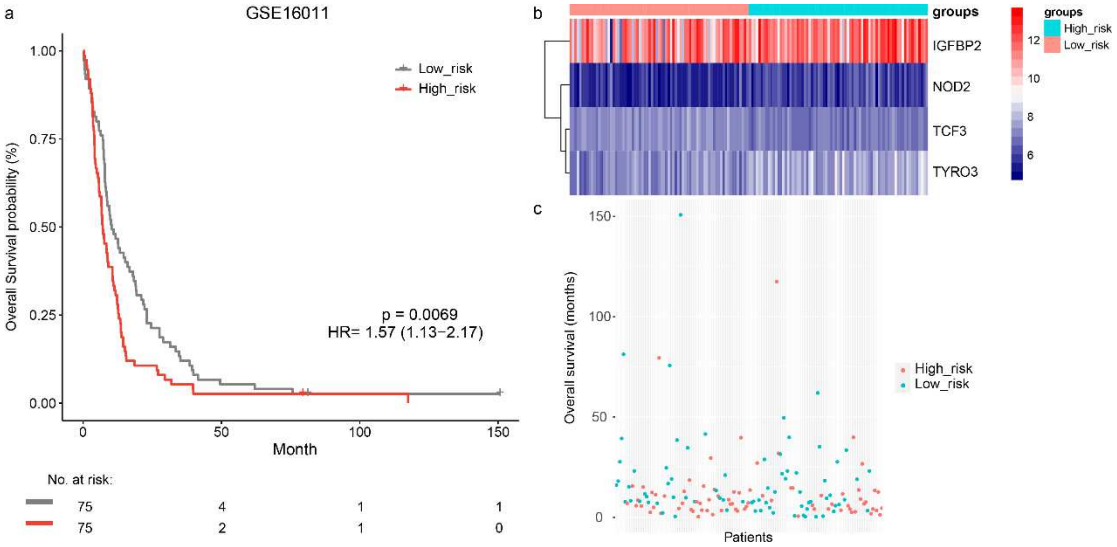

Figure 3 The survival analysis of the lymphocyte activation-associated gene signature in GSE13041 (n = 150). (a) Kaplan–Meier curves of overall survival for the four gene signature. (b) The expression profiles of the four genes in the signature. (c) The distribution of the GBM patients' overall survival status.

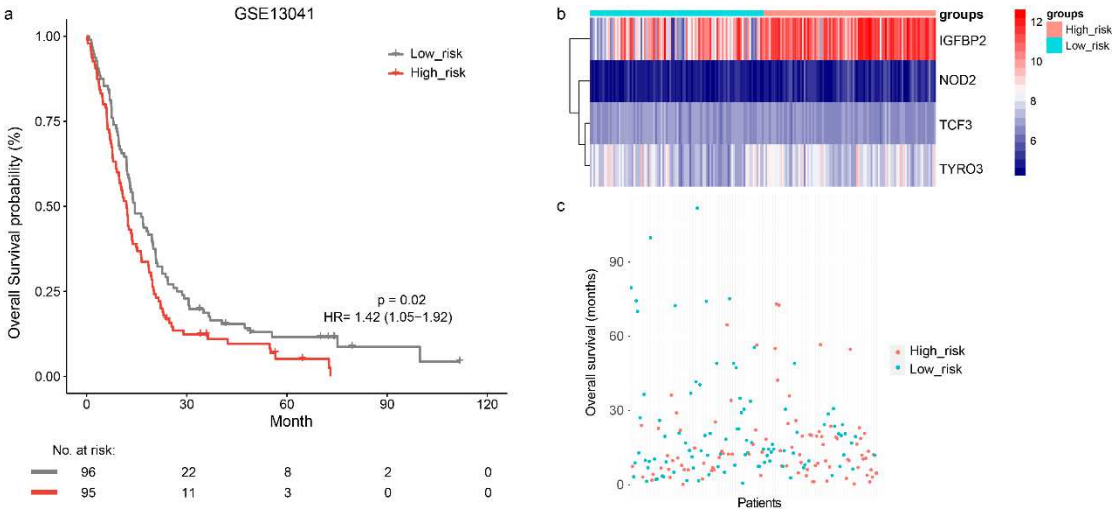

Figure 4 The cell activity of lymphocytes (memory B cells, native B cells, native CD4 T cells, follicular helper T cells and regulatory T cells) in TCGA (a) and GSE13041 (b) data set.

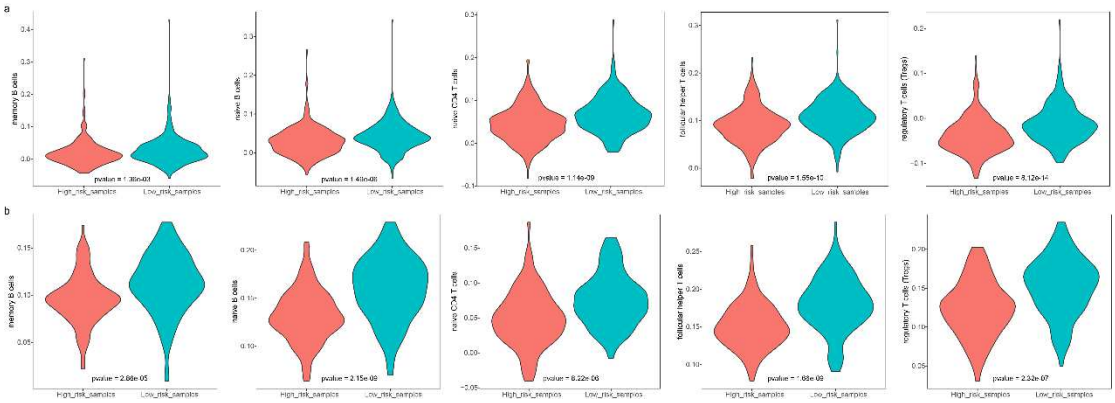

Figure 5 Survival analysis for IDH-wildtype and IDH-mutation GBM patients. Kaplan-Meier estimates overall survival among these patients in TCGA (a) and GSE16011 (b).

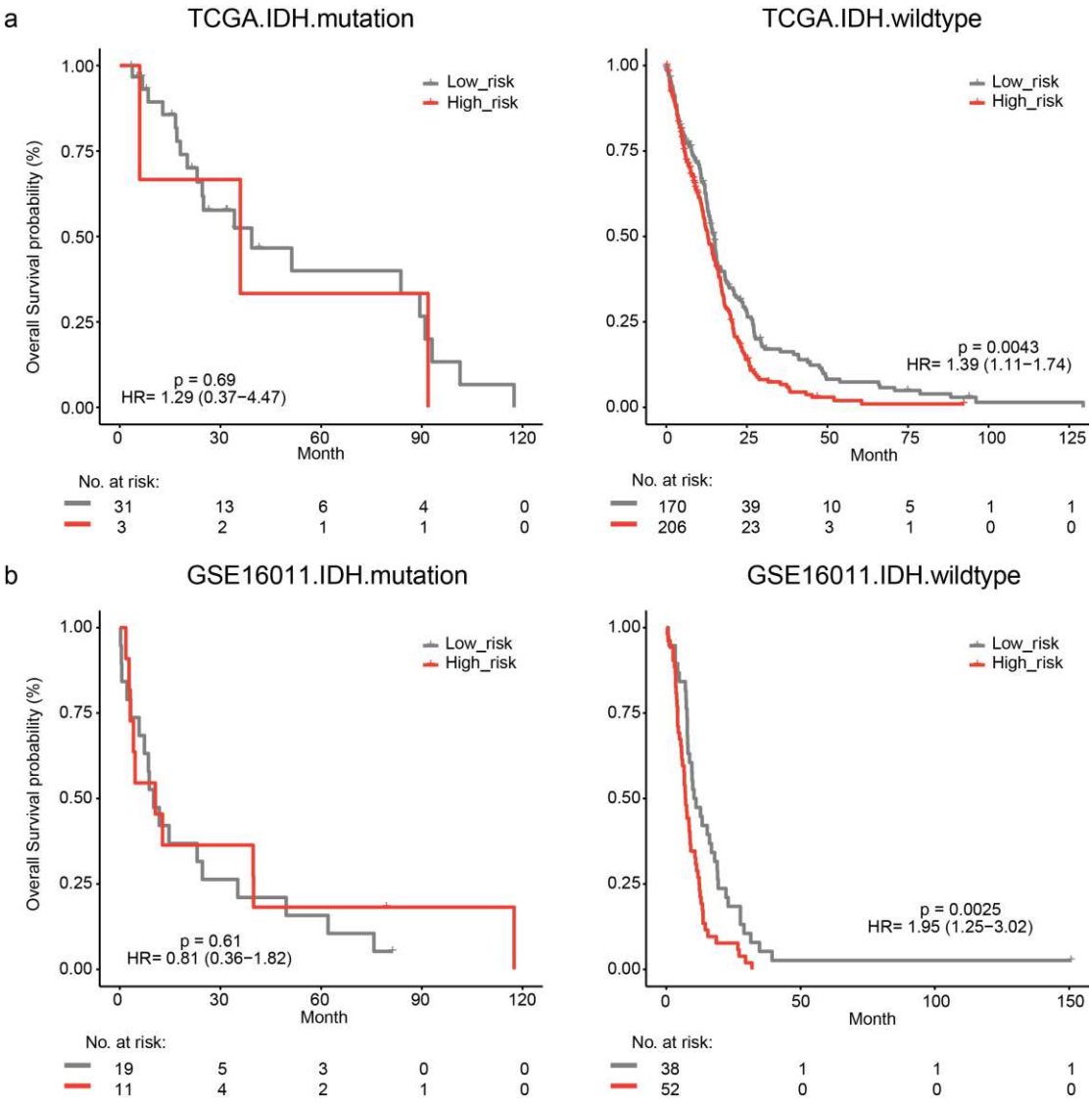

Figure 6 Survival analysis for GBM patients according to TCGA mesenchymal and proneural subtypes.

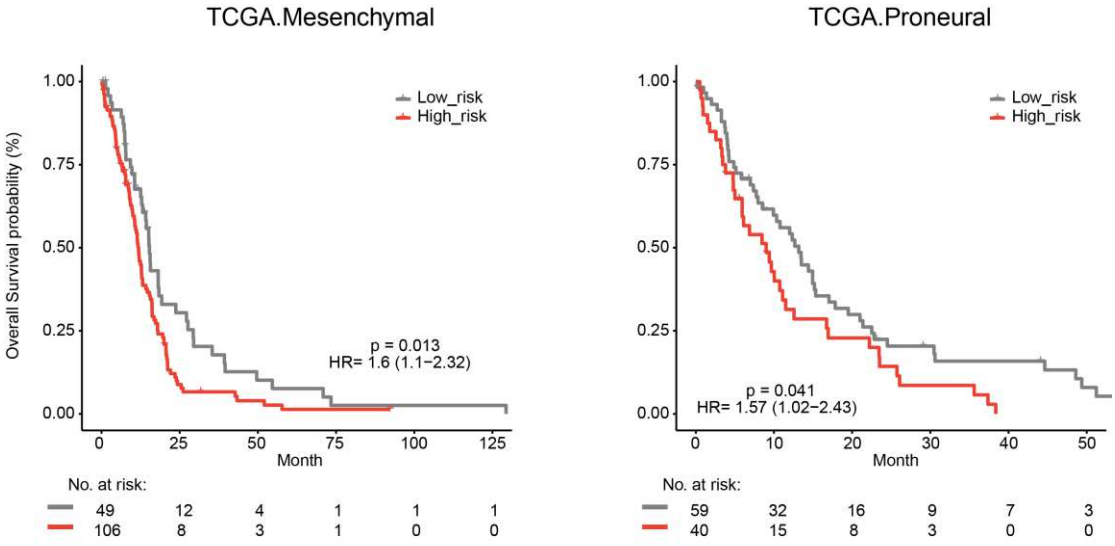

Figure 7 The signature predicts cancer patient outcome after radiotherapy.

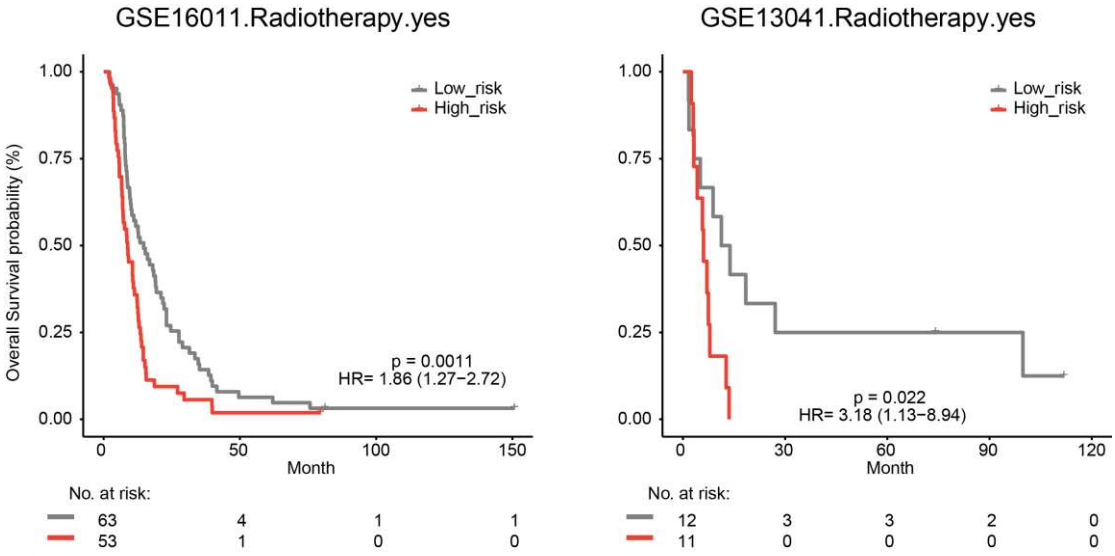

Table Legends

Table 1 Clinical and pathological characteristics of GBM patients in TCGA.

Clinical features	Category	GBM, n = 525
Gender	Female	205
	Male	320
Age	Median (range)	59 (10-89)
Classical Subtype	Classical	144
	Mesenchymal	155
	Neural	83
	Proneural	99
Survival status	Alive	75
	Deceased	449
Follow-up from samples (months)	Median (range)	12.39(0.1 – 127.5)

Table 2 Clinical and pathological characteristics of GBM patients in GSE13041 and GSE16011.

Clinical features	Category	GSE13041 (n=191)	GSE16011 (n=150)
Gender	Female	74	47
	Male	117	103
Age	Median (range)	54 (18-86)	55.44 (14.38-80.65)
Survival status	Alive	15	3
	Deceased	176	147
Follow-up from samples (months)	Median (range)	12.97(0.23 – 111.77)	8.7 (0.24 – 150.72)

Table 3 Multivariate analysis for the lymphocyte activation-associated gene signature of overall survival in TCGA.

Variables		Univariate			Multivariate		
		HR	95% CI	*P* value	HR	95% CI	*P* value
Age		1.032	1.025-1.04	<0.001*	1.028	1.018-1.037	<0.001*
Sex	Male vs Female	1.241	0.993-1.551	0.057	1.3	1.035-1.634	0.024*
IDH status	Mutation vs Wild type	0.353	0.229-0.544	<0.001*	0.831	0.396-1.742	0.623
Signature	High_risk vs Low_risk	1.588	1.275-1.979	<0.001*	1.455	1.159-1.828	0.001*
Subtypes	G-CIMP vs Classical	0.369	0.2271-0.599	<0.001*	0.701	0.319-1.544	0.378
	Mesenchymal vs Classical	1.169	0.876-1.56	0.289	1.117	0.835-1.494	0.455
	Neural vs Classical	1.091	0.788-1.509	0.6	1.177	0.848-1.633	0.331
	Proneural vs Classical	1.314	0.954-1.808	0.094	1.638	1.184-2.268	0.003*

Significant *P* values are labeled with * ($P < 0.05$).

Table 4 Multivariate analysis for lymphocyte activation-associated gene signature of overall survival in GSE16011.

Variables		Univariate			Multivariate		
		HR	95% CI	*P* value	HR	95% CI	*P* value
Age		1.039	1.024-1.054	<0.001*	1.031	1.024-1.039	<0.001*
Sex	Male vs Female	0.809	0.544-1.204	0.296	0.928	0.62-1.391	0.719
IDH status	Mutation vs wild type	0.579	0.367-0.915	0.019*	0.809	0.493-1.326	0.4
Signature	High_risk vs Low_risk	1.554	1.284-1.88	<0.001*	1.359	1.346-1.979	0.029*

Significant *P* values are labeled with * ($P < 0.05$).

Figures

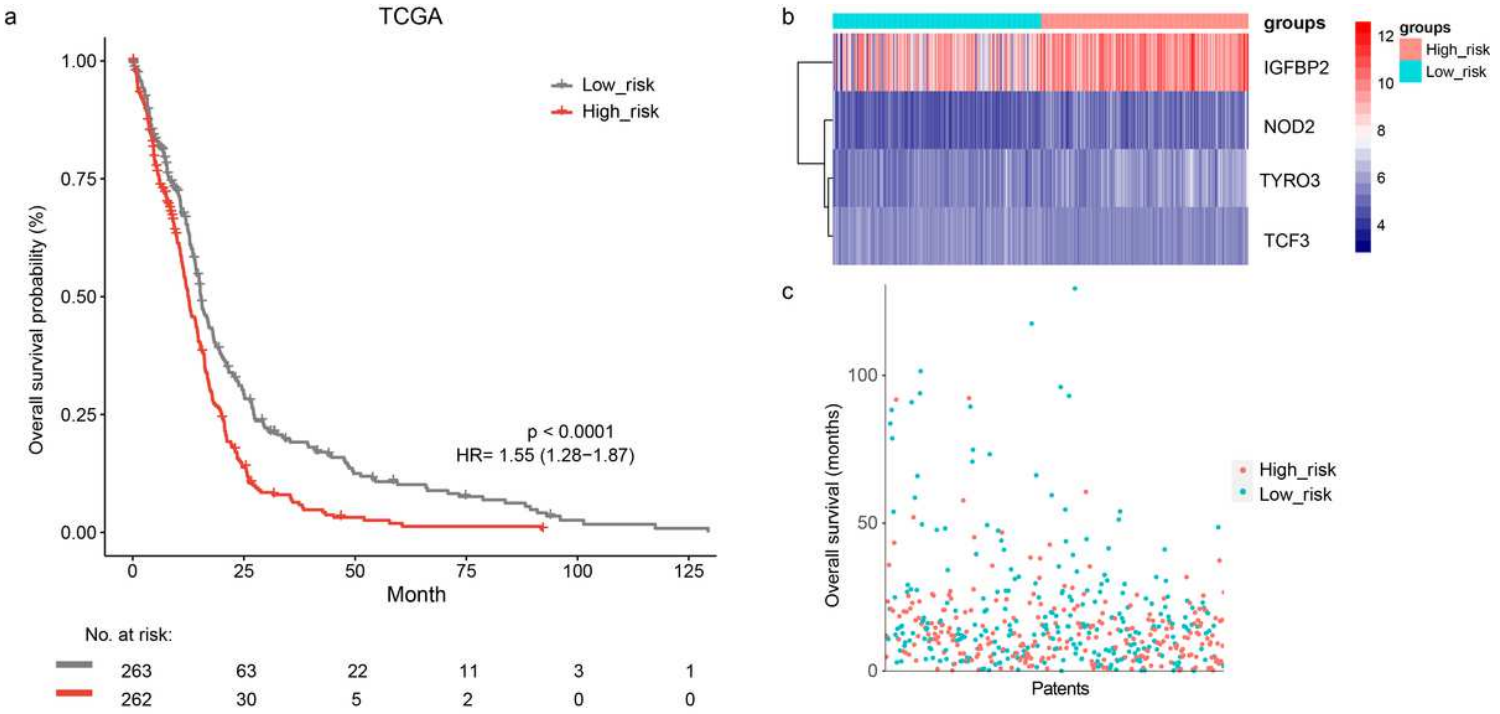

Figure 1

The survival analysis of the lymphocyte activation-associated gene signature in training set (TCGA, n=525). (a) Kaplan–Meier curves of overall survival for the four gene signature. The patients were divided into high-risk (red) and low-risk group (grey). (b) The expression profiles of the four genes in the signature. (c) The distribution of the GBM patients' overall survival status.

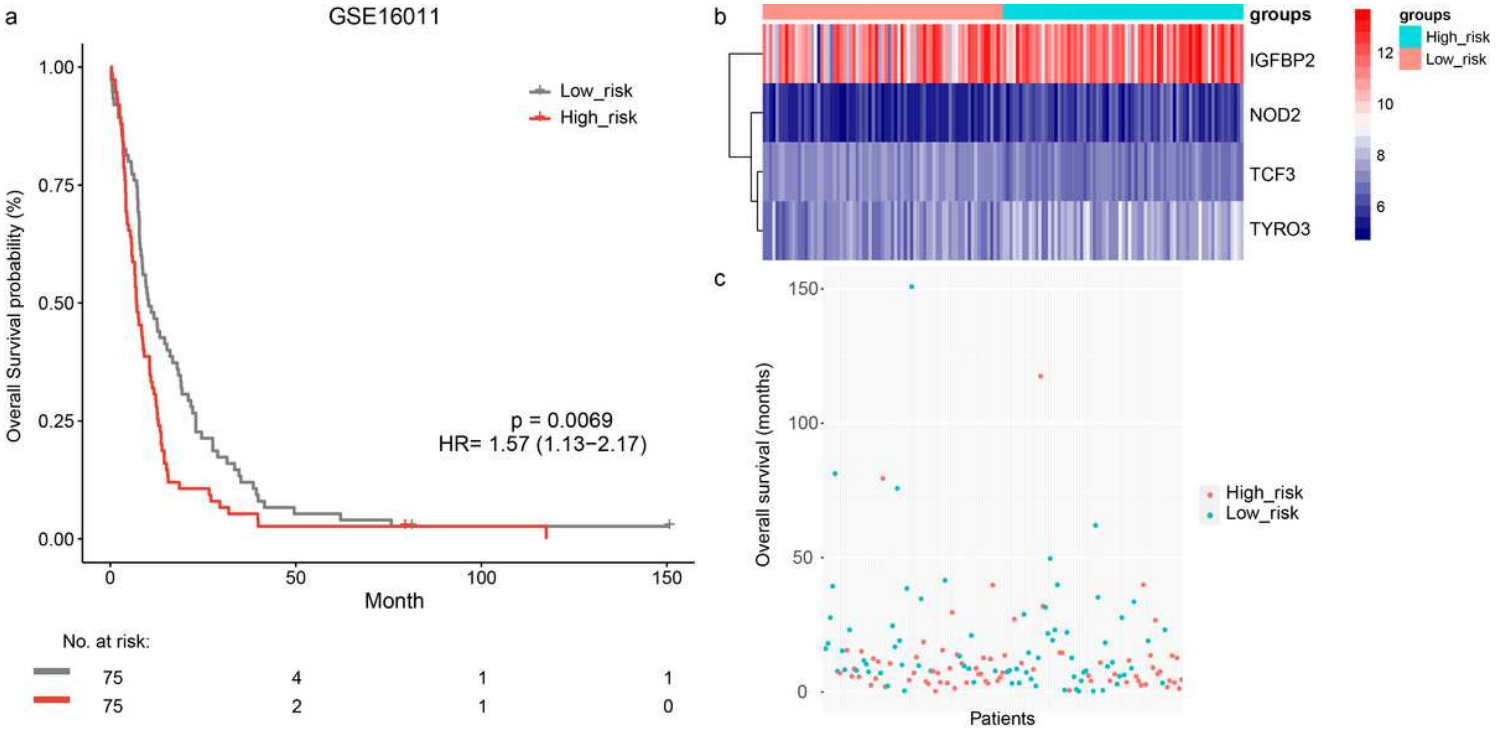

Figure 2

The survival analysis of the lymphocyte activation-associated gene signature in GSE16011 (n = 150). (a) Kaplan–Meier curves of overall survival for the four gene signature. (b) The expression profiles of the four genes in the signature. (c) The distribution of the GBM patients' overall survival status.

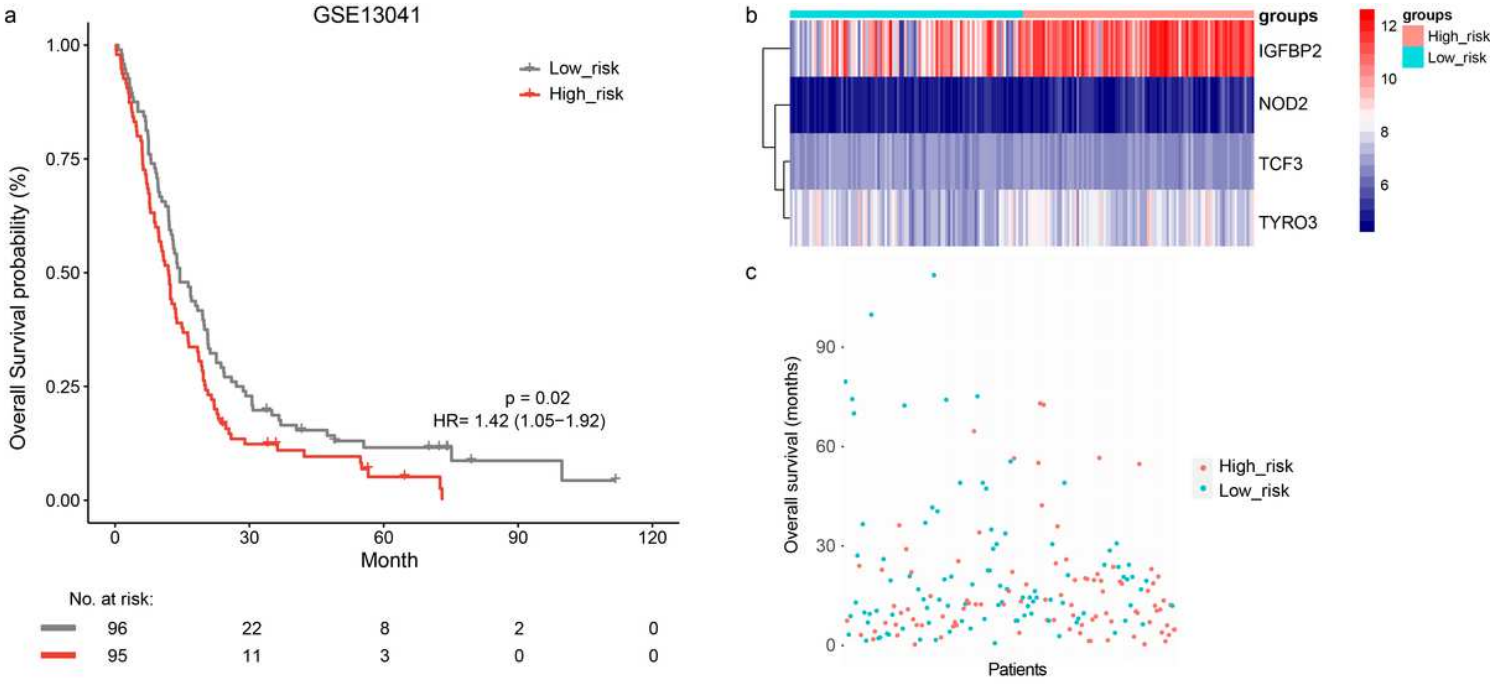

Figure 3

The survival analysis of the lymphocyte activation-associated gene signature in GSE13041 (n = 150). (a) Kaplan–Meier curves of overall survival for the four gene signature. (b) The expression profiles of the four genes in the signature. (c) The distribution of the GBM patients' overall survival status.

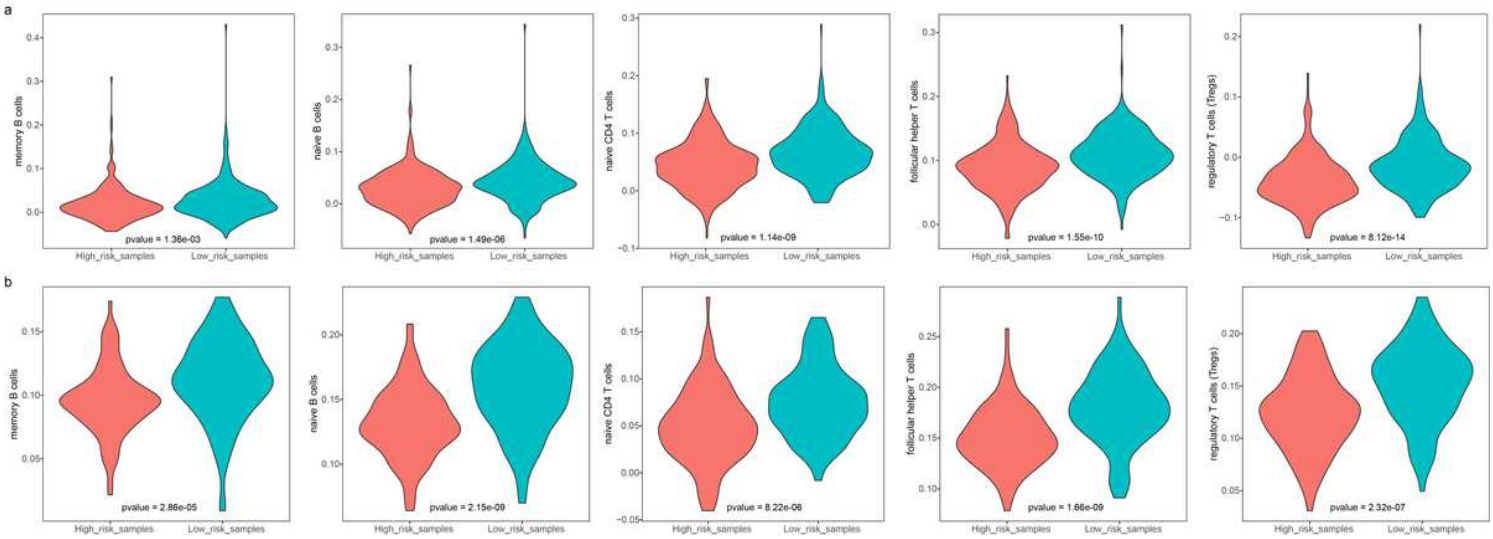

Figure 4

The cell activity of lymphocytes (memory B cells, native B cells, native CD4 T cells, follicular helper T cells and regulatory T cells) in TCGA (a) and GSE13041 (b) data set.

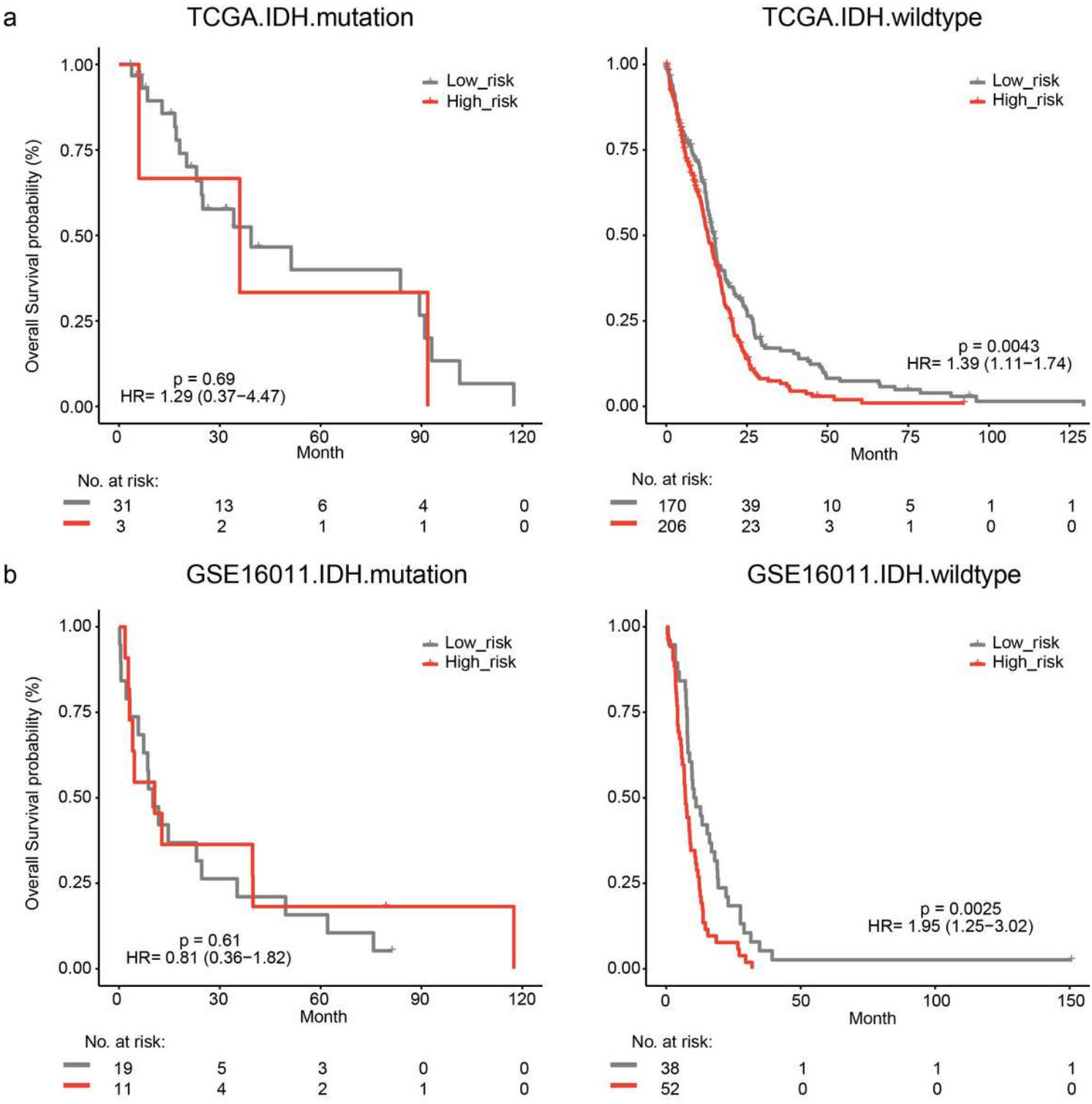

Figure 5

Survival analysis for IDH-wildtype and IDH-mutation GBM patients. Kaplan-Meier estimates overall survival among these patients in TCGA (a) and GSE16011 (b).

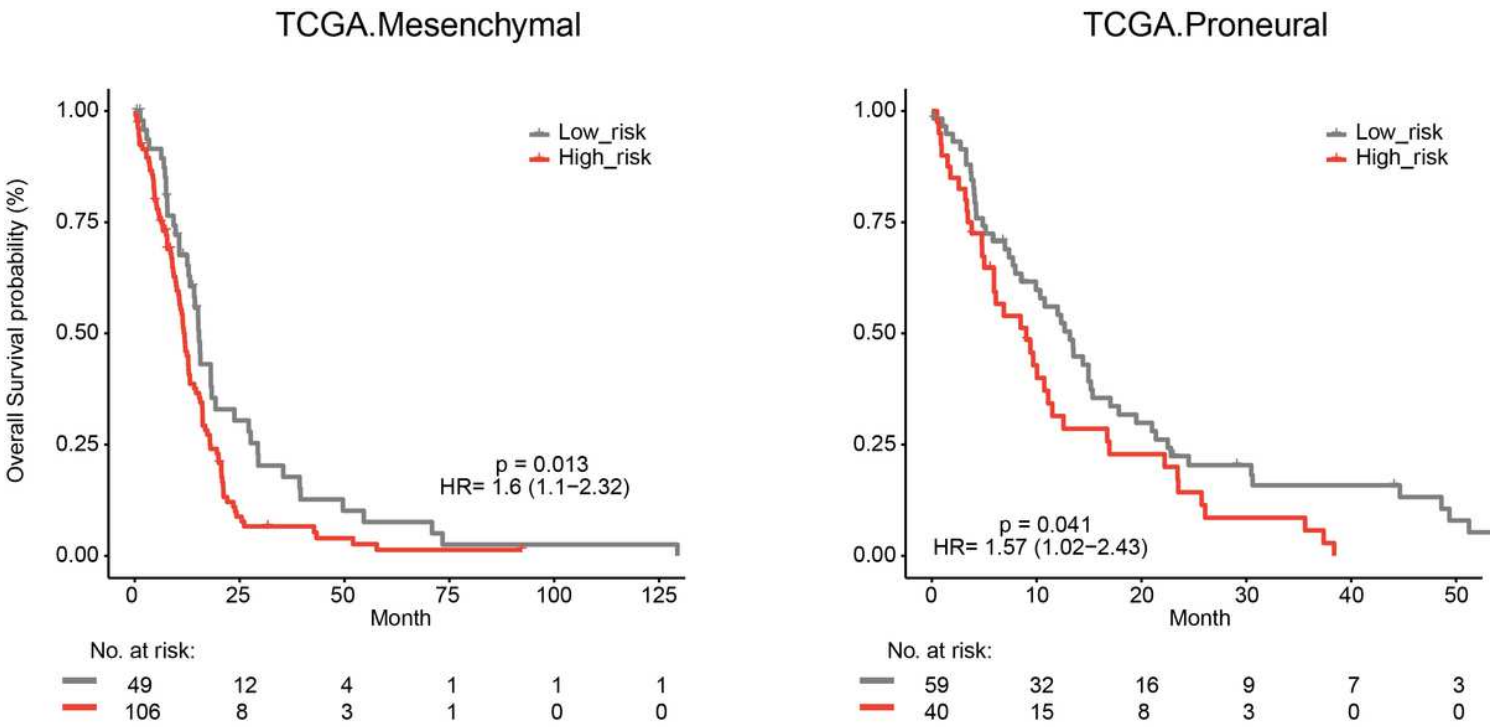

Figure 6

Survival analysis for GBM patients according to TCGA mesenchymal and proneural subtypes.

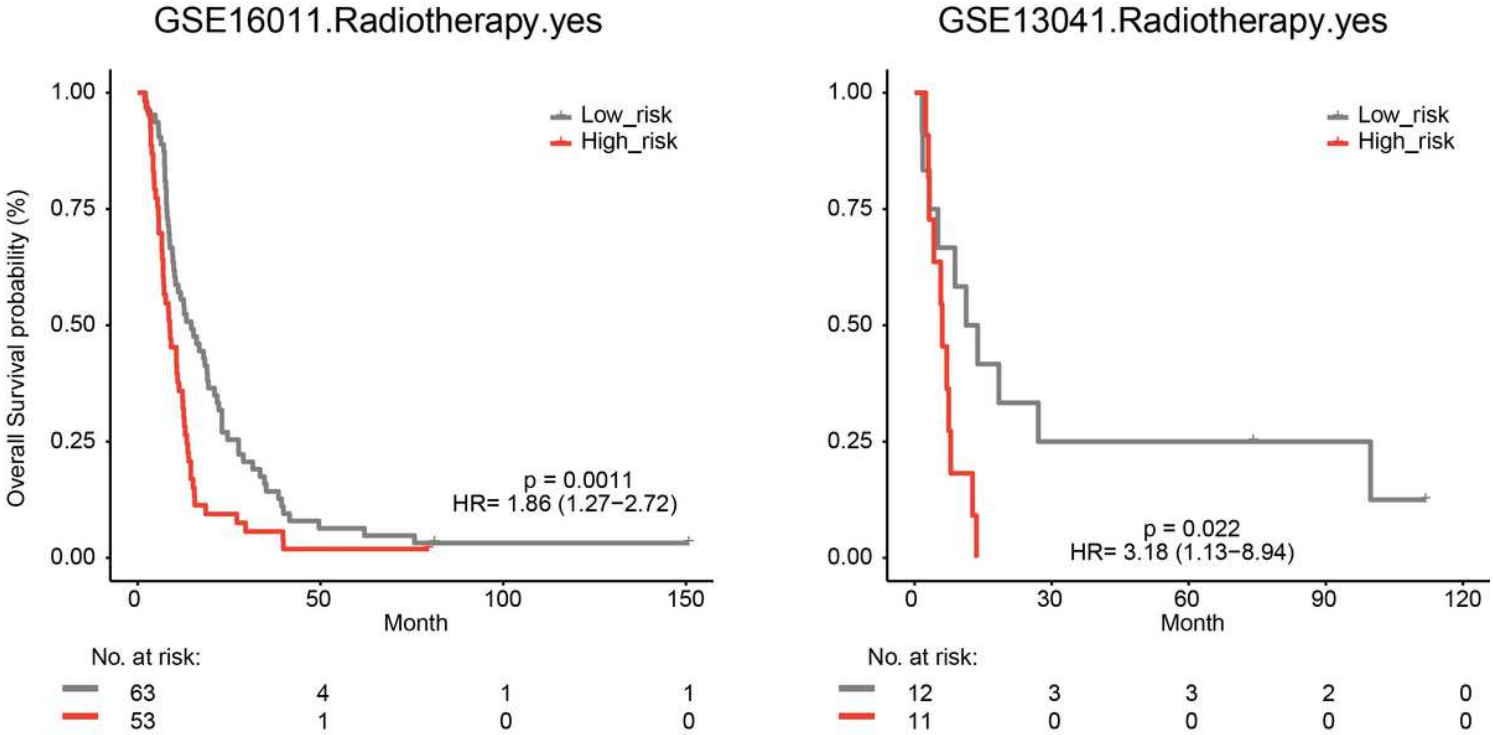

Figure 7

The signature predicts cancer patient outcome after radiotherapy.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformation.pdf