Development of a Prognostic Five-Gene Signature with Radiotherapy Guidance Significance for Diffuse Lower-Grade Glioma Patients Based on Large-Scale Sequencing Data

Qiang Zhang
Department of Clinical laboratory
Shun-Bin Luo
Department of Clinical Pharmacy
Fu-Chen Xie
Department of urinary surgery
Xiao-Jun Liu
Pathology department
Ren-ai Xu (wmu_xra@163.com)
Wenzhou Medical University First Affiliated Hospital https://orcid.org/0000-0003-0572-0763

Research

Keywords: lower-grade gliomas, signature, prognostic, survival, PCG

Posted Date: June 5th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-32321/v1

Abstract

Background: Diffuse lower-grade gliomas (LGGs) are infiltrative and heterogeneous neoplasms. Gene signature including multiple protein coding genes (PCGs) is widely used as tumor markers. This study aimed to construct a multi-PCG signature to predict survival for LGG patients.

Methods: LGG data including PCG expression profiles and clinical information were downloaded from The Cancer Genome Atlas (TCGA) and the Chinese Glioma Genome Atlas (CGGA). Survival analysis, receiver operating characteristic (ROC) analysis and random survival forest algorithm (RSFVH) were used to identify the prognostic PCG signature.

Results: From the training (n = 524) and test (n = 431) datasets, a five-PCG signature which can classify LGG patients into low- or high-risk group with significantly different overall survival (Log Rank P < 0.001) was screened out and validated. In terms of prognosis predictive performance, the five-PCG signature is stronger than other clinical variables and IDH mutation status. Moreover, the five-PCG signature could further divide radiotherapy patients into two different risk groups. GO and KEGG analysis found PCGs in the prognostic five-PCG signature were mainly enriched in cell cycle, apoptosis, DNA replication pathways.

Conclusions: The new five-PCG signature is a reliable prognostic marker with radiotherapy guidance significance for LGG patients and has a good prospect in clinical application.

Introduction

Glioma is a fatal tumor that derives from gial cell and grows in the central nervous system, including diffuse gliomas and nondiffuse gliomas[1]. Diffuse gliomas are the most frequently occurring intracranial malignant tumors, encompassing various histologic types (astrocytic or oligodendroglial) and malignancy grades(World Health Organization(WHO) grade II, III and IV) tumors. Astrocytomas and oligodendrogliomas in the low-grade (WHO II)and intermediate-grade(WHO III) are incorporated into diffuse lower-grade glioma (LGG), and perform better than IV grade glioblastoma in both malignancy and prognosis. However, it is difficult to predict the clinical outcome of LGG patients because LGG are a highly heterogeneous group of tumors. Firstly, there is a difference in the speed of tumor progression within LGG. Some are relatively inert, while others quickly progress to high-grade glioma or glioblastoma. Secondly, therapeutic sensitivity varies in LGG patients. Some people have effective treatments, while others have poor treatment results. Finally, LGG patients differ greatly in the prognosis, ranging from 1 to 15 years[2]. Due to the limitations of histologic classification of LGG, finding molecular markers that can accurately predict prognosis and treatment response has become an urgent task[3].

In recent years, significant progress has been made in the study of molecular pathology of gliomas, and a series of molecular markers have been discovered that are helpful for clinical diagnosis, prognostic judgment and treatment guidance, such as IDH1/2 gene mutation, chromosome 1p/19q co-deletion, MGMT promoter methylation, etc[4]. Especially, the revised 2016 WHO classification of CNS tumors made fundamentally changes and classified diffuse gliomas based on IDH mutation and 1p/19q codeletion status. This innovative measure highlights the important role of novel and reliable gene biomarkers in the diagnosis and prognosis of gliomas.

With the development of next-generation sequencing technology, a large amount of high-throughput sequencing data and a variety of bioinformatics methods have prompted researchers to further understand tumorigenesis and find prognostic markers. Xin Hu et al selected a prognostic 35-gene signature from 374 glioma patients carrying the 1p/19q co-deletion[5]. Fan Wu constructed a six-gene signature that could classify *IDH*-mutant GBM patients into high or low risk of poor outcome using 33 samples from the Chinese Glioma Genome Atlas RNA-sequencing data and 21 cases from Chinese Glioma Genome Atlas microarray data[6]. Xiangyang Deng found a four-gene immune prognostic signature for predicting prognosis in LGGs through analyzing 511 LGG samples from the TCGA database and 172 LGG samples from CGGA dataset[7]. Therefore, gene signature has become the research focus of glioma prognostic markers.

In the present study, the protein coding genes (PCG) expression data from a total of 955 LGG patients were collected from the Cancer Genome Atlas (TCGA) database and the Chinese Glioma Genome Atlas (CGGA). We aimed to mine the large queue of gene expression data and clinical information to identify a prognostic PCG signature and explore its significance of treatment guidance.

Materials And Methods

Data collection of diffuse lower-grade glioma (LGG) patients

The clinicopathological information and protein coding genes (PCG) expression data of LGG patients were obtained from the TCGA (http://cancergenome.nih.gov/; https://xenabrowser.net/datapages/) and treated as the training dataset. Another independent dataset used as validation or test dataset was downloaded from CGGA database (http://www.cgga.org.cn/). LGG cases with clinical survival information including survival status and survival time were selected for building the prognostic model. Clinical details of LGG patients in the training and test datasets were shown in Table 1. Genes with missing expression values in > 20% samples were removed and FPKM expression values of 0 were set to 0.01 in the given samples for log 2 transformed in subsequent analysis[8].

Table 1
Clinical characteristics of the LGG patients.

Characteristic	TCGA(n = 524)	CGGA(n = 444)
Survival status		
Living	**387**	**242**
Dead	**137**	**189**
Age (years)		
≤ 40	261	221
> 40	263	210
Sex		
female	237	193
Male	287	238
Grade		
G2	255	180
G3	269	251
Unknown	1	
IDH mutation status		
Mutant	91	96
Wildtype	34	297
Unknown	399	38
Chemo status		
No		124
Yes		265
Unknown		42
Radio status		
No	173	86
Yes	284	314
Unknown	67	31

The process of developing the prognostic signatures in the training dataset

Using Kaplan–Meier (KM) and receiver operating characteristic (ROC) analysis, we identified the PCGs significantly associated with patients' OS with AUC > 0.6 from the TCGA group. Then we reduced the number of the PCGs by the random survival forest algorithm (RSFVH). Further, prognostic models were constructed as follows:

Risk Score = $\sum^{N}_{i=1}$ (Expression$_i$ * coefficient $_i$)

where N is the number of PCG, Expression is the PCG expression value and coefficient is the PCG expression in Cox regression analysis. The final prognostic PCG signature was screened out with the largest AUC value in all the constructed models[9].

Statistical and bioinformatics analysis

Kaplan–Meier analysis was used to assess the two survival risk groups seperated by the median risk score. Cox regression analysis was performed to explore the independence of the signature. ROC and TimeROC were used to analyse survival prediction performance. Function prediction of prognostic PCGs was analyzed by clusterProfiler[10]. R program (www.r-project.org) with R packages including pROC, TimeROC, randomForestSRC and survival was used to perform the above analyses.

Results

The process of developing the prognostic signatures in the training dataset

All 955 patients diagnosed with LGG were collected from the TCGA (n = 524) and CGGA (n = 431), and a total of 16246 expressed PCGs were identified. From Table 1, we found that the median age of the enrolled patients was 40 years (11–87 years) and that there were more male patients than female patients, indicating that LGG is more likely to occur in adult males. When focusing on the survival status and survival time of these patients, we found more than 1/3 of patients (326 of 955) had died and the median survival time was only h 2.11 years (0.2–14.15 years). In addition, we also obtained IDH mutation status, radiotherapy and chemotherapy information for further analysis (Table 1).

After Kaplan–Meier and ROC analysis, a total of 1702 PCGs were discovered (Red dots in Fig. 1A), which were significantly associated with OS and had a good ability to predict survival (KM $P < 0.05$&AUC > 0.6, Supplementary table S1). Further, we screened out 11 prognostic PCGs by RSFVH analysis based on importance scores (Fig. 1B). Then we brought the prognostic PCGs into the risk prediction model and got 2^{11}-1 = 2047 possible signatures in the training dataset. ROC analyses were performed in all the 2047 signatures to find out the signature with the strongest predictive ability (Supplementary table S2). The final signature including five PCGs (ABCC3, SMC4, EMP3, WEE1, HIST1H2BK) was screened out with the maximum AUC (AUCsignature = 0.739; Fig. 1C, Table 2). The selected risk model is as follows: Risk score = (1.32 × expression value of ABCC3) + (1.94 × expression value of SMC4) + (1.55 × expression value of EMP3) + (1.84 × expression value of WEE1) + (1.57 × expression value of HIST1H2BK).

Table 2
Survival analysis of the PCGs in the prognostic signature in TCGA dataset

Database ID	HR	COX *P*	KM P	AUC
ABCC3	1.32	<<0.001 1	< 0.001	0.64
SMC4	1.94	< 0.001	< 0.001	0.67
EMP3	1.55	< 0.001	< 0.001	0.62
WEE1	1.84	< 0.001	< 0.001	0.69
HIST1H2BK	1.57	< 0.001	< 0.001	0.69

The performance of PCG signature in predicting LGG patient survival

We used the risk model to calculate risk scores for each patient. The median risk score divided patients in the training dataset into either the high-risk ($n = 262$) or low-risk group ($n = 262$). The Kaplan–Meier analysis results showed patients in

the low-risk group lived longer than patients in the high-risk group (median survival time: 12.18 years *vs.* 3.84 years, $P < 0.001$; Fig. 2A). Then we tested the prognostic value of the PCG signature in another large independent LGG dataset (CGGA, $n = 431$). After the median risk score in CGGA separated patients into high- or or low-risk group, Kaplan–Meier analysis found the five-year survival of patients with high risk scores was lower than that of patients with low risk scores (5-year survival: 34.16% vs. 77.05%, log-rank test $P < 0.001$; Fig. 2B). We showed the relationship of PCG expression, risk score and survival information in Fig. 3. With the increase of gene expression value, risk scores and death toll increased in the training (Fig. 3A) and test dataset (Fig. 3B),

The five-PCG signature is an independent predictive factor

In the training and test groups ($n = 524/431$), we found the PCG signature was related with clinical variables such as IDH mutation status and grade by chi-square test (Table 3). Then we further performed univariate and multivariable Cox regression analysis to test the predictive independence of the PCG signature. Multivariable Cox regression results verified that the PCG signature was an independent predictive factor and could independently predict patients' clinical outcome in training or test datasets (High- vs. Low-risk, HR training = 1.70, 95% CI 1.31–2.21, P < 0.001, n = 524; HR test = 3.01, 95% CI 2.12–4.27, P < 0.001, n = 431, Table 4).

Table 3
Association of the PCG signature with clinical characteristics in LGG patients

Variables	TCGA group		*P*	CGGA group		*P*
	Low risk *	High risk *		Low risk *	High risk *	
Age (years)			0.02			0.99
≤ 40	144	117		111	110	
> 40	118	145		105	105	
Sex			0.99			0.88
Female	119	118		98	95	
Male	143	144		118	120	
Grade			< 0.001			< 0.001
G2	169	88		111	69	
G3	92	174		105	146	
Unknown	1	0		0	0	
IDH mutation status			< 0.001			< 0.001
Mutant	69	22		183	114	
Wildtype	9	25		13	83	
Unknown	184	215		20	18	
Radiotherapy			< 0.001			0.31
No	119	54		47	39	
Yes	113	171		157	157	
Unknown	30	37		12	19	
Chemotherapy						0.04
No				74	50	
Yes				124	141	
Unknown				18	24	

Table 4
Cox regression analysis of the PCG signature with LGG survival

		Univariable analysis				Multivariable analysis			
Variables		HR	95% CI of HR		P	HR	95% CI of HR		P
			lower	upper			lower	upper	
TCGA dataset(n = 524)									
Age	> 40 vs.≤40	2.82	1.96	4.04	< 0.001	1.99	0.52	7.60	0.32
Sex	Male vs. Female	1.14	0.81	1.60	0.45	2.00	0.66	6.09	0.22
IDH status	Wildtype vs. Mutant	5.53	2.07	14.82	< 0.001	0.94	0.22	4.07	0.94
LGG Grade	G3 vs. G2	3.31	2.28	4.79	< 0.001	0.79	0.22	2.81	0.72
PCG-signature	High risk vs. low risk	6.86	4.26	11.04	< 0.001	1.70	1.31	2.21	< 0.001
CGGA set (n = 431)									
Age	> 40 vs.≤40	1.19	0.89	1.58	0.24	1.10	0.82	1.48	0.54
Sex	Male vs. Female	1.00	0.75	1.34	0.98	1.14	0.85	1.54	0.38
IDH status	Wildtype vs. Mutant	2.24	1.64	3.07	< 0.001	1.48	1.06	2.07	0.02
LGG Grade	G3 vs. G2	2.62	1.89	3.64	< 0.001	2.58	1.81	3.66	< 0.001
PCG-signature	High risk vs. low risk	3.68	2.69	5.03	< 0.001	3.01	2.12	4.27	< 0.001

Predictive Performance Comparison between the five-PCG signature with other clinical variables

We performed ROC analysis to compare the predictive performance of the five-PCG signature with other clinical variables including IDH mutation status, age and grade. Figure 4A, B showed the PCG signature outperformed above clinical variables both in the training/test sets (AUC signature 0.739/0.678 *vs.* AUCIDH 0.712/0.585; AUCgrade 0.625/0.632; AUCage 0.57/0.527). Further, TimeROC analysis found the AUC values of the signature from 1 to 5 years were greater than that of IDH mutation status, grade or age, indicating that the PCG signature had better survival prediction when integrating the TCGA and CGGA datasets (Fig. 4C).

Radiotherapy stratification analysis

In order to evaluate whether the signature can guide radiotherapy in LGG patients, we performed stratification analyses in the combined dataset of TCGA and CGGA. According to the radio-status information of all the 955 LGG patients, we found 598 received radiotherapy, 259 patients did not, and 98 patients had unknown radiotherapy information. For patients after radiotherapy, the five-PCG signature could further divide patients with into low- and high-risk groups with significantly different survival (5 or 10-years survival: 77.70%/39.84 *vs.* 37.10%/17.69%, log-rank test $P < 0.001$, Fig. 5A). Patients without radiotherapy can also be grouped into different risk groups by the five-PCG signature (5 or 10-years survival: 88.39%/71.53 vs. 53.59%/33.15%, log-rank test $P < 0.001$, Fig. 5B).

Function prediction for the five selected PCGs

To explore the role and function of the five selected PCGs screened in this study, we obtained a total of 741 co-expressing PCGs (Pearson coefficient > 0.5/<-0.5, $P < 0.05$) using pearson test in the TCGA and CGGA datasets respectively, and then performed KEGG and GO analysis. The co-expressing genes of the five PCGs were significantly enriched in 425 Go terms

and 21 KEGG pathways (P < 0.05), such as Cell cycle, DNA replication, p53 signaling pathway, indicating the specific pathway or mechanism in which the prognostic PCGs might play a key role(top 20 showed, Fig. 6).

Discussion

Tumor heterogeneity is a major contributing factor in the adverse clinical outcome of gliomas[11]. Consequently, the latest edition (2016 edition) of the WHO glioma classification incorporates molecular features into the classification criteria, thereby improving the homogeneity of clinical outcomes in patients with the same subtype[12]. However, as one of histological subtype of glioma, LGG has substantial variation in patient survival and lacks effective prognostic markers. In the current study, we analyzed the survival and gene expression information of 955 patients with LGG and found the five-PCG signature could be a good prognostic molecular marker. In addition to predicting prognosis of LGG patients, the five-PCG signature has also been found to have a role in guiding radiotherapy.

Tumor heterogeneity and therapeutic advancements have prompted clinicians to make individualized prognosis and treatment choices for cancer patients, thereby achieving precision medicine. Gene expression has always been at the forefront of the development of personalized medicine, especially in the field of cancer. Gene expression signatures obtained by analyzing gene expression profiling have been shown to predict the tumor behavior and to distinguish patients with specific tumor grades and / or prognosis[13] in various types of cancer, such as esophageal squamous cell carcinoma, hepatocellular carcinoma, bladder carcinoma, breast cancer, and glioblastoma. In the current study, we aimed to analyze the gene expression profile and develop an effective gene signature for accurate prognosis prediction of LGG patients. After a detailed analysis of gene expression profiles of 955 patients with LGG from the TCGA training set and CGGA validation set, a five PCGs based prognostic risk model and the five-PCG signature that could distinguish LGG patients with high risk of poor prognosis from patients with low risk were developed. The five-PCG signature has the following two advantages in prognosis prediction: First, it is an independent factor and does not depend on known prognostic factors such as IDH mutation and tumor grade; second, it has excellent prediction performance for its AUC value was higher than IDH mutation and tumor grade.

Notably, the five-PCG signature was found to be a predictive marker for radiotherapy in LGG patients. More specifically, the marker can identify who can benefit from radiotherapy or who is suitable for radiotherapy. As a result, LGG patients have more scientific guidance on whether to accept radiotherapy, and clinicians can also have more standardized guidelines for radiotherapy to facilitate their implementation. This finding shows that the five-PCG signature not only makes the prognosis assessment of patients more precise, but also can play the role of individualized treatment.In addition, we noted that the five PCGs in the signature had positive risk factors, meaning they were all prognostic risk factors. By searching the existing literature, we found that the important role of these genes in prognosis prediction had been reported in a variety of tumors. ATP binding cassette subfamily C member 3(ABCC3), also named multidrug resistance-associated protein 3(MRP3), is an organic anion transporter and contributes to drug resistance of cancer cells[14]. Consistent with the results in this article, the poor prognosis predictive role of ABCC3 has been reported not only in acute myeloid leukemia[15], gastric cancer[16], pancreatic cancer[17] and lung cancer[18], but also in gliomas[19]. In addition to being found as a prognostic marker for gliomas in this article, structural maintenance of chromosomes 4(SMC4) has also been found to be a survival marker for colorectal cancer[20], breast cancer[21] and prostate cancer[22]. Epithelial membrane protein 3(EMP3) is considered to be a tumor suppressor, but this article found that this gene is a prognostic risk gene for LGG. Similar to our results, Haiwei Wang et al also found that EMP3 was associated with the worse prognosis of LGG patients[23] and Xiao-Xia Guo et al discovered EMP3 was also a risk gene in the process of developing a prognostic 4-gene panel for glioblastoma patients[24]. WEE1 G2 checkpoint kinase(WEE1) is reported to be an oncogenic nuclear kinase and a regulator of the G2 checkpoint. Expression of WEE1 has been found to be associated with poor prognosis in a variety of tumor types including gliomas[25]. Two other gene signatures constructed to predict the prognosis of LGG are also consistent with the results of this article, and found that WEE1 is a prognostic risk factor[26, 27]. H2B clustered histone 12(H2BC12 or HIST1H2BK) is a replication-dependent histone and belongs to the histone H2B family. The prognostic role of HIST1H2BK was identified in ovarian cancer[28],

breast cancer[29] and pancreatic ductal adenocarcinoma[30].Although we had some findings on the function of the five prognostic genes by KEGG and GO analysis, further functional exploration of these genes is needed.

Conclusion

Our study developed a prognostic five-PCG signature for LGG patients that can predict individual clinical outcome with high accuracy. Surprisingly, the five-gene signature can also predict radiotherapy response which makes the biomarker have broad clinical value.

Abbreviations

TCGA: The Cancer Genome Atlas; CGGA: the Chinese Glioma Genome Atlas; ROC: receiver operating characteristic; Kaplan–Meier: KM; AUC: area under the ROC curve; GO: Gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; CI: confidence interval; OS: overall survival;

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

None declared.

Funding

Not applicable.

Authors' Contributions

The authors contributed in the following way: Qiang Zhang: data collection, data analysis, interpretation, and drafting; Xiao-Jun Liu,Ren-ai Xu: study design, study supervision, and final approval of the manuscript; Shun-Bin Luo, Fu-Chen Xie: technical support and critical revision of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

None.

Data availability statement

Publicly available datasets can be found here,

TCGA:https://xenabrowser.net/datapages/?cohort=TCGA%20Lower%20Grade%20Glioma%20(LGG)&removeHub=https%3A%2F%2Fxena.treehouse.gi.ucsc.edu%3A443.

CGGA:http://www.cgga.org.cn/download?file=download/20200506/CGGA.mRNAseq_693.RSEM-genes.20200506.txt.zip&type=mRNAseq_693&time=20200506

References

1. Wesseling P, Capper D: **WHO 2016 Classification of gliomas.** *Neuropathol Appl Neurobiol* 2018, **44:**139-150.
2. Cancer Genome Atlas Research N, Brat DJ, Verhaak RG, Aldape KD, Yung WK, Salama SR, Cooper LA, Rheinbay E, Miller CR, Vitucci M, et al: **Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas.** *N Engl J Med* 2015, **372:**2481-2498.
3. Ramaswamy V, Taylor MD: **Fall of the Optical Wall: Freedom from the Tyranny of the Microscope Improves Glioma Risk Stratification.** *Cancer Cell* 2016, **29:**137-138.
4. Chen R, Smith-Cohn M, Cohen AL, Colman H: **Glioma Subclassifications and Their Clinical Significance.** *Neurotherapeutics* 2017, **14:**284-297.
5. Hu X, Martinez-Ledesma E, Zheng S, Kim H, Barthel F, Jiang T, Hess KR, Verhaak RGW: **Multigene signature for predicting prognosis of patients with 1p19q co-deletion diffuse glioma.** *Neuro Oncol* 2017, **19:**786-795.
6. Wu F, Chai RC, Wang Z, Liu YQ, Zhao Z, Li GZ, Jiang HY: **Molecular classification of IDH-mutant glioblastomas based on gene expression profiles.** *Carcinogenesis* 2019, **40:**853-860.
7. Deng X, Lin D, Chen B, Zhang X, Xu X, Yang Z, Shen X, Yang L, Lu X, Sheng H, et al: **Development and Validation of an IDH1-Associated Immune Prognostic Signature for Diffuse Lower-Grade Glioma.** *Front Oncol* 2019, **9:**1310.
8. Xu J, Li Y, Lu J, Pan T, Ding N, Wang Z, Shao T, Zhang J, Wang L, Li X: **The mRNA related ceRNA-ceRNA landscape and significance across 20 major cancer types.** *Nucleic Acids Res* 2015, **43:**8169-8182.
9. Guo JC, Li CQ, Wang QY, Zhao JM, Ding JY, Li EM, Xu LY: **Protein-coding genes combined with long non-coding RNAs predict prognosis in esophageal squamous cell carcinoma patients as a novel clinical multi-dimensional signature.** *Mol Biosyst* 2016, **12:**3467-3477.
10. Yu G, Wang LG, Han Y, He QY: **clusterProfiler: an R package for comparing biological themes among gene clusters.** *OMICS* 2012, **16:**284-287.
11. Cheng J, Meng J, Zhu L, Peng Y: **Exosomal noncoding RNAs in Glioma: biological functions and potential clinical applications.** *Mol Cancer* 2020, **19:**66.
12. Molinaro AM, Taylor JW, Wiencke JK, Wrensch MR: **Genetic and molecular epidemiology of adult diffuse glioma.** *Nat Rev Neurol* 2019, **15:**405-417.
13. Burska AN, Roget K, Blits M, Soto Gomez L, van de Loo F, Hazelwood LD, Verweij CL, Rowe A, Goulielmos GN, van Baarsen LG, Ponchel F: **Gene expression analysis in RA: towards personalized medicine.** *Pharmacogenomics J* 2014, **14:**93-106.
14. Kool M, van der Linden M, de Haas M, Scheffer GL, de Vree JM, Smith AJ, Jansen G, Peters GJ, Ponne N, Scheper RJ, et al: **MRP3, an organic anion transporter able to transport anti-cancer drugs.** *Proc Natl Acad Sci U S A* 1999, **96:**6914-6919.
15. Benderra Z, Faussat AM, Sayada L, Perrot JY, Tang R, Chaoui D, Morjani H, Marzac C, Marie JP, Legrand O: **MRP3, BCRP, and P-glycoprotein activities are prognostic factors in adult acute myeloid leukemia.** *Clin Cancer Res* 2005, **11:**7764-7772.
16. Mao X, He Z, Zhou F, Huang Y, Zhu G: **Prognostic significance and molecular mechanisms of adenosine triphosphate-binding cassette subfamily C members in gastric cancer.** *Medicine (Baltimore)* 2019, **98:**e18347.

17. Adamska A, Ferro R, Lattanzio R, Capone E, Domenichini A, Damiani V, Chiorino G, Akkaya BG, Linton KJ, De Laurenzi V, et al: **ABCC3 is a novel target for the treatment of pancreatic cancer.** *Adv Biol Regul* 2019, **73:**100634.
18. Zhao Y, Lu H, Yan A, Yang Y, Meng Q, Sun L, Pang H, Li C, Dong X, Cai L: **ABCC3 as a marker for multidrug resistance in non-small cell lung cancer.** *Sci Rep* 2013, **3:**3120.
19. Wang F, Zheng Z, Guan J, Qi D, Zhou S, Shen X, Wang F, Wenkert D, Kirmani B, Solouki T, et al: **Identification of a panel of genes as a prognostic biomarker for glioblastoma.** *EBioMedicine* 2018, **37:**68-77.
20. Feng XD, Song Q, Li CW, Chen J, Tang HM, Peng ZH, Wang XC: **Structural maintenance of chromosomes 4 is a predictor of survival and a novel therapeutic target in colorectal cancer.** *Asian Pac J Cancer Prev* 2014, **15:**9459-9465.
21. Ma RM, Yang F, Huang DP, Zheng M, Wang YL: **The Prognostic Value of the Expression of SMC4 mRNA in Breast Cancer.** *Dis Markers* 2019, **2019:**2183057.
22. Zhao SG, Evans JR, Kothari V, Sun G, Larm A, Mondine V, Schaeffer EM, Ross AE, Klein EA, Den RB, et al: **The Landscape of Prognostic Outlier Genes in High-Risk Prostate Cancer.** *Clin Cancer Res* 2016, **22:**1777-1786.
23. Wang H, Wang X, Xu L, Zhang J, Cao H: **Prognostic significance of age related genes in patients with lower grade glioma.** *J Cancer* 2020, **11:**3986-3999.
24. Guo XX, Su J, He XF: **A 4-gene panel predicting the survival of patients with glioblastoma.** *J Cell Biochem* 2019, **120:**16037-16043.
25. Mir SE, De Witt Hamer PC, Krawczyk PM, Balaj L, Claes A, Niers JM, Van Tilborg AA, Zwinderman AH, Geerts D, Kaspers GJ, et al: **In silico analysis of kinase expression identifies WEE1 as a gatekeeper against mitotic catastrophe in glioblastoma.** *Cancer Cell* 2010, **18:**244-257.
26. Xiao K, Liu Q, Peng G, Su J, Qin CY, Wang XY: **Identification and validation of a three-gene signature as a candidate prognostic biomarker for lower grade glioma.** *PeerJ* 2020, **8:**e8312.
27. Pang FM, Yan H, Mo JL, Li D, Chen Y, Zhang L, Liu ZQ, Zhou HH, Wu J, Li X: **Integrative analyses identify a DNA damage repair gene signature for prognosis prediction in lower grade gliomas.** *Future Oncol* 2020, **16:**367-382.
28. Li N, Zhan X: **Signaling pathway network alterations in human ovarian cancers identified with quantitative mitochondrial proteomics.** *EPMA J* 2019, **10:**153-172.
29. Han J, Lim W, You D, Jeong Y, Kim S, Lee JE, Shin TH, Lee G, Park S: **Chemoresistance in the Human Triple-Negative Breast Cancer Cell Line MDA-MB-231 Induced by Doxorubicin Gradient Is Associated with Epigenetic Alterations in Histone Deacetylase.** *J Oncol* 2019, **2019:**1345026.
30. Yu S, Li Y, Liao Z, Wang Z, Wang Z, Li Y, Qian L, Zhao J, Zong H, Kang B, et al: **Plasma extracellular vesicle long RNA profiling identifies a diagnostic signature for the detection of pancreatic ductal adenocarcinoma.** *Gut* 2020, **69:**540-550.

Figures

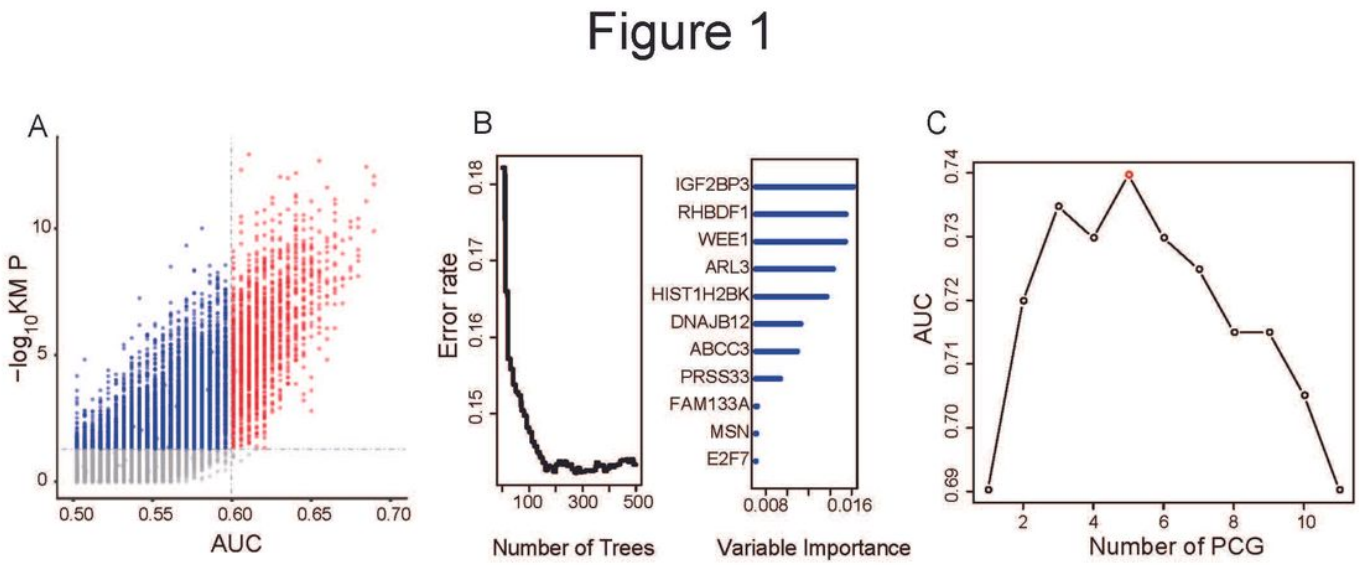

Figure 1

Development of the prognostic signature in the training dataset. (A) The survival associated PCGs in Kaplan–Meier analysis were displayed as red dots in scatter diagram. (B) Random forest supervised classificantion algorithm reduced the prognosis-associated PCGs to 11 PCGs.(C) The prognostic five-PCGsignature was selected because its AUC was the largest (AUC=0.739) among the 211-1=2047 signatures.

Figure 2

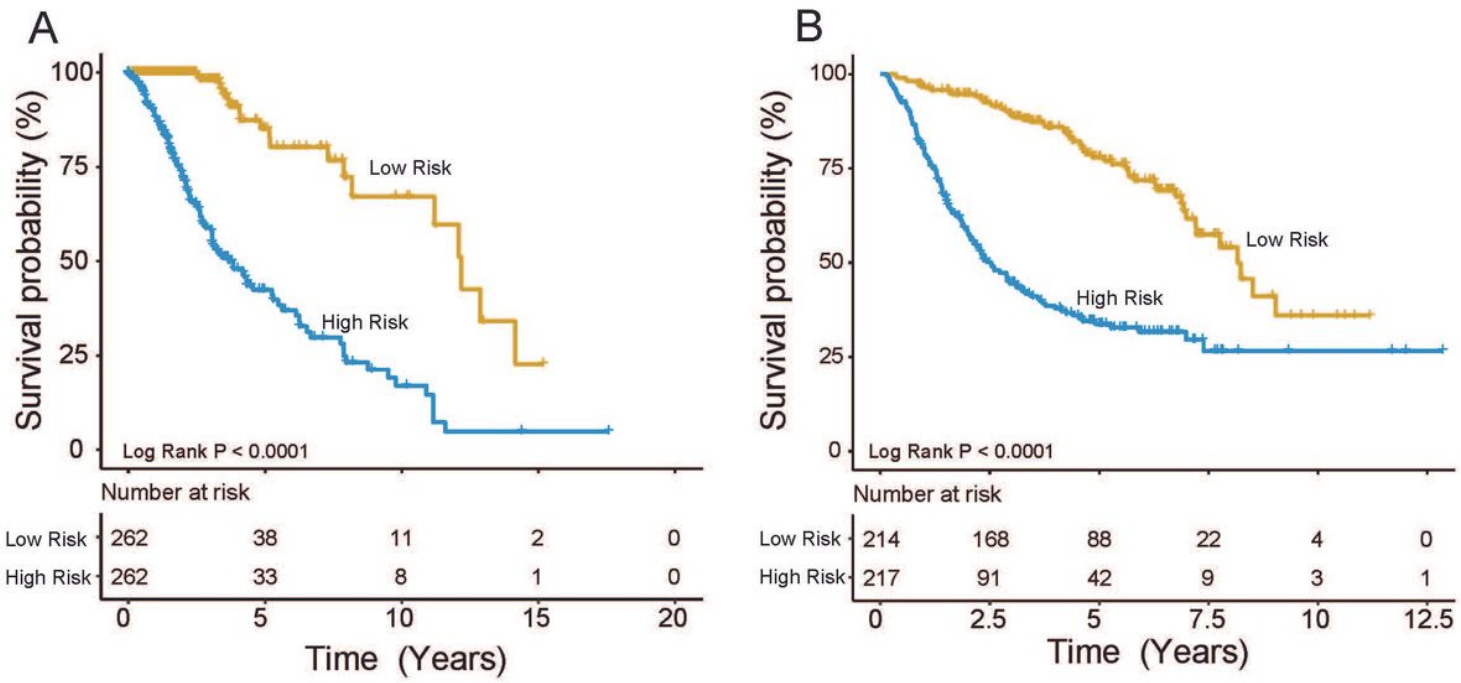

Figure 2

Kaplan–Meier plots indicated LGG patients could be classified into high- and low-risk groups according to the five-gene signature in the training (A) and test (B) datasets.

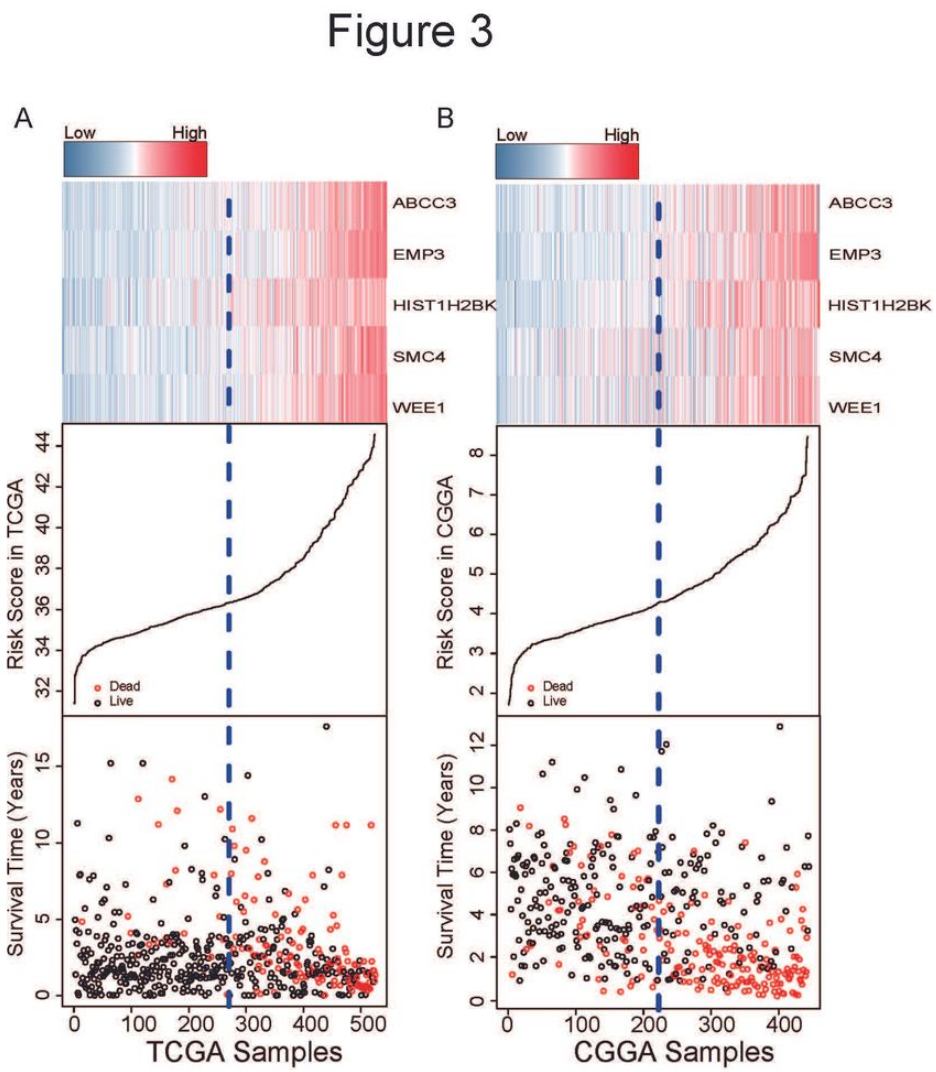

Figure 3

Risk score distribution, survival status and PCGs expression patterns for LGG patients in the training, test dataset.

Figure 4

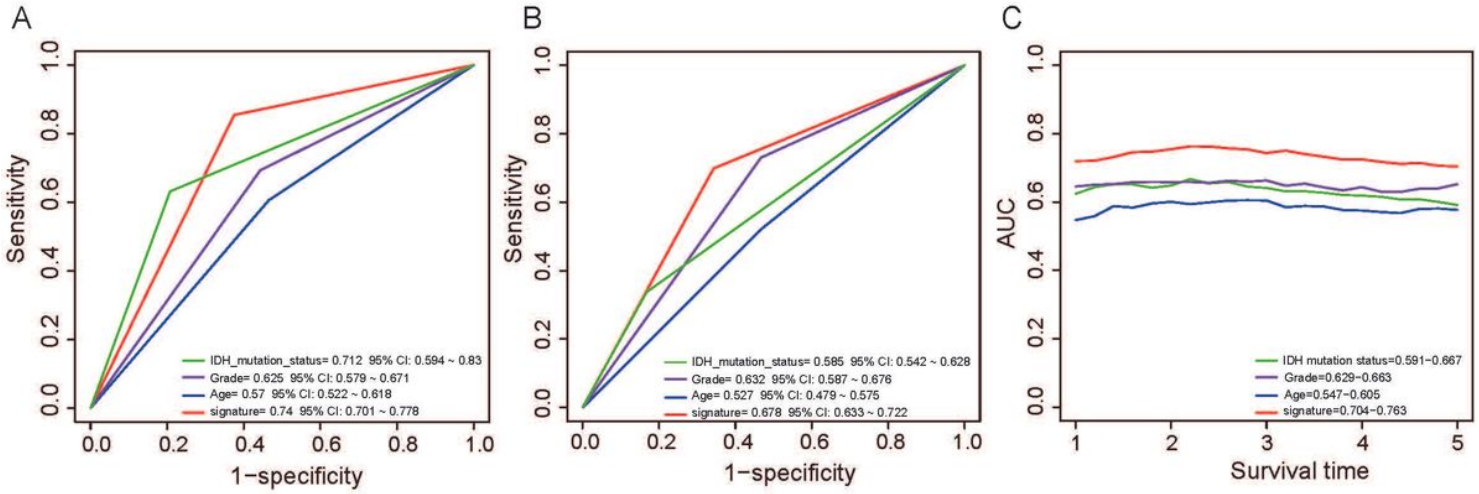

Figure 4

Comparison of the survival predictive power of the signature with grade. age and IDH mutation by ROC in the training (A) and test (B) sets. TimeROC analysis of survival predictive power for the signature,grade. age and IDH mutation(C).

Figure 5

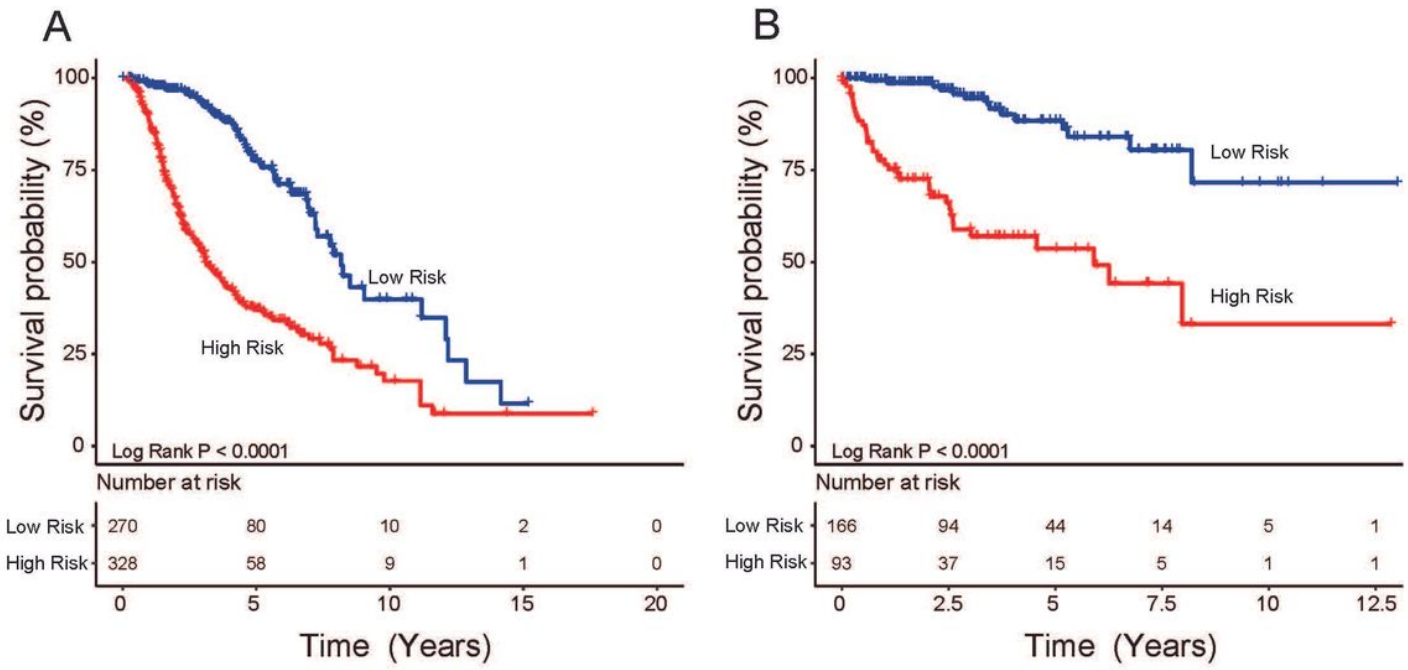

Figure 5

Radiotherapy stratification analysis. The five-PCG signature could further divide patients with radiotherapy(A) or patients without radiotherapy(B) into two groups with significantly different survival.

Figure 6

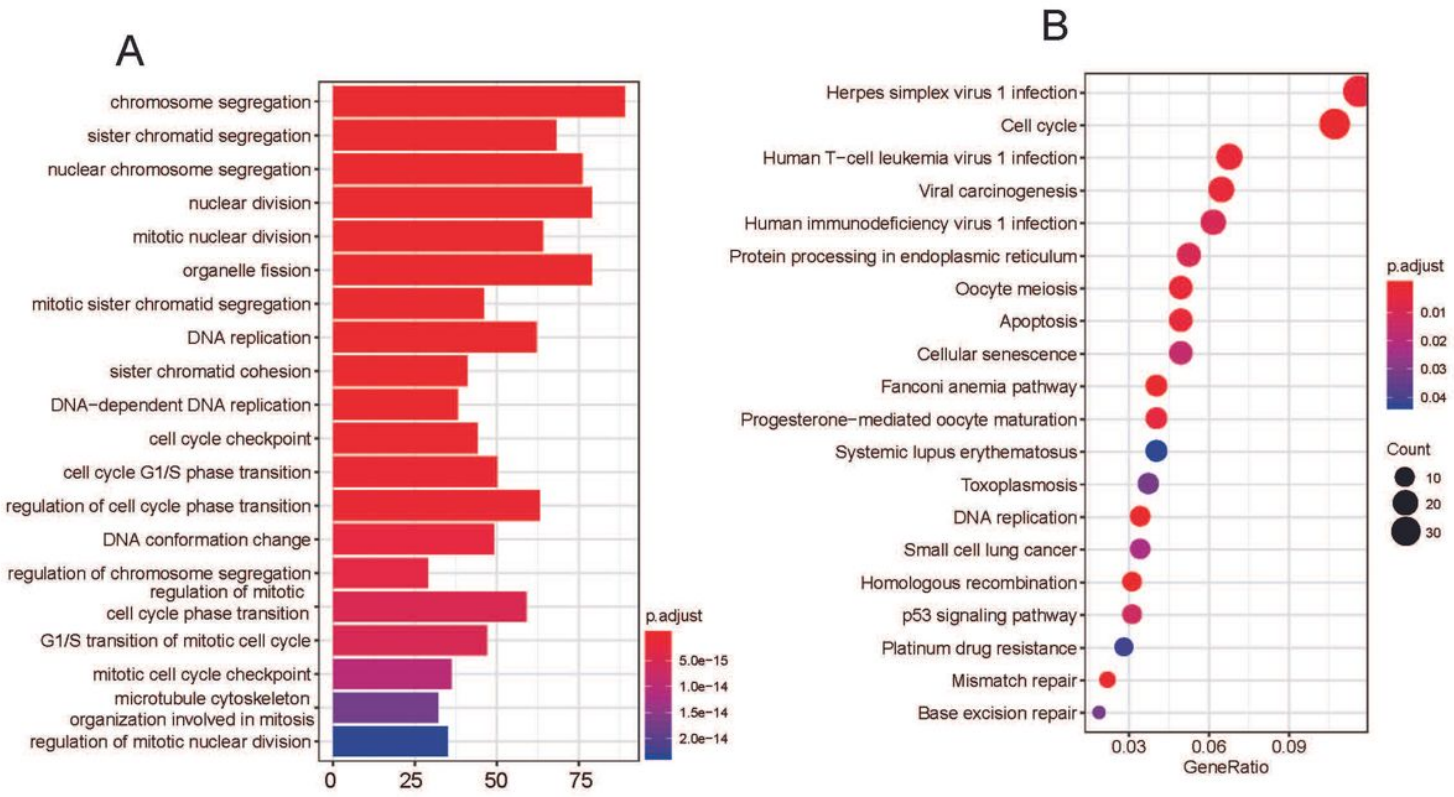

Figure 6

GO and KEGG functional enrichment analysis of the five PCGs in the signature.

Supplementary Files

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

- SupplementaryTableS1.xls
- SupplementaryTableS2.xls