

A three-gene signature for prognosis in patients with MGMT promoter-methylated glioblastoma

Wen Wang^{1,2,3,7}, Lu Zhang⁵, Zheng Wang^{2,3,7}, Fan Yang^{2,3,7}, Haoyuan Wang^{6,7}, Tingyu Liang^{2,3,7}, Fan Wu^{3,4,7}, Qing Lan¹, Jiangfei Wang^{2,4,7}, Jizong Zhao^{2,1,8}

¹Department of Neurosurgery, The Second Affiliated Hospital of Soochow University, Suzhou, China

²Department of Neurosurgery, Beijing Tiantan Hospital, Capital Medical University, Beijing, China

³Department of Molecular Neuropathology, Beijing Neurosurgical Institute, Capital Medical University, Beijing, China

⁴Brain Tumor Center, Beijing Institute for Brain Disorders, Beijing, China

⁵Department of Ophthalmology, School of Medicine, Shandong University, Jinan, China

⁶Department of Neurosurgery, Zhujiang Hospital, Southern Medical University, Guangzhou, China

⁷Chinese Glioma Cooperative Group (CGCG), Beijing, China

⁸China National Clinical Research Center for Neurological Diseases, Beijing, China

Correspondence to: Jizong Zhao, **email:** zhaojz205@163.com

Keywords: signature, MGMT, prognosis, glioblastoma, RNA-Seq

Received: February 27, 2016

Accepted: August 09, 2016

Published: August 31, 2016

ABSTRACT

Glioblastoma is the most malignant tumor and has high mortality rate. The methylated promoter of MGMT results in chemotherapy sensitivity for these patients. However, there are still other factors that affected the prognosis for the glioblastoma patients with similar MGMT methylation status. We developed a signature with three genes screened from the whole genome mRNA expression profile from Chinese Glioma Genome Atlas (CGGA) and RNAseq data from The Cancer Genome Atlas (TCGA). Patients with MGMT methylation in low risk group had longer survival than those in high risk group (median overall survival 1074 vs. 372 days; $P = 0.0033$). Moreover, the prognostic value of the signature was significant difference in cohorts stratified by MGMT methylation and chemotherapy ($P=0.0473$), while there is no significant difference between low and high risk group or unmethylated MGMT patients without chemotherapy. Multivariate analysis indicated that the risk score was an independent prognosis factor ($P = 0.004$). In conclusion, our results showed that the signature has prognostic value for patients with MGMT promoter-methylated glioblastomas based on bioinformatics analysis.

INTRODUCTION

Gliomas are the most common and lethal primary tumors of central nervous system [1]. Glioblastoma (GBM) is the most malignant tumor and has high mortality rate despite maximal tumor resection with concomitant chemoradiotherapy. It is reported that the median survival for GBM patients remained about 14 months [2].

Many biomarkers, such as mutations of PTEN, IDH1, TP53 and methylation of O(6)-methylguanine DNA methyltransferase (MGMT) promoter have been identified for GBM. PTEN is a tumor suppressor and mutation of PTEN could up-regulated the AKT pathway which play a critical role in cell cycle regulation, apoptosis, and cell

migration in GBM [3]. The mutation of IDH1, which may alter DNA methylation patterns in GBM [4], may serve as an early driving mutation of GBM [5]. The methylated promoter of MGMT causes an impaired ability for cells to recover from damage induced by chemotherapeutic agents and results in chemotherapy sensitivity for GBM patients [3, 6, 7]. However, due to variable overall survival of similar MGMT methylation status, there are still other factors that affect the prognosis for GBM patients with MGMT promoter-methylated.

In our study, we obtained whole genome mRNA expression profiling microarray data from Chinese Glioma Genome Atlas (CGGA) as training set and two validation datasets, CGGA mRNA-sequencing data and The Cancer

Genome Atlas (TCGA) mRNA sequencing data. By applying significance analysis of microarray (SAM) and Cox regression analysis, we screened three target genes and thus generated a signature based on these genes. Patients with MGMT promoter-methylated were divided into low risk and high risk group based on the cutoff (median risk score) and the clinical outcomes and molecular features were quite different between two groups.

RESULTS

A three-gene prognostic signature identified and validated in three datasets

A total of 41 GBM with MGMT promoter methylation, 60 GBM with MGMT promoter unmethylation and 5 non-cancerous brain tissue (NBT) samples were included in the comparison, respectively. Moreover, we compared NBT and methylated or unmethylated patients

using SAM analysis and 688 mRNAs (760 probes) were differentially expressed in these samples (false discovery rate, FDR < 0.05) (Supplementary Table S1). We evaluated the prognostic value of the 688 mRNAs in training set with a univariate Cox regression analysis. Finally, three genes (FPR3, IKBIP and S100A9) were significantly associated with overall survival (OS) (Figure 1).

With the three genes, we developed a risk score for each patient which was calculated based on a linear combination of the mRNA expression level weighted by the regression coefficient (β) derived from the univariate Cox regression analysis. Patients with MGMT promoter methylated (41 patients) in the training set were divided into low risk and high risk group based on the cutoff (median risk score) and patients in low risk group had longer survival time than high risk group (median OS 1074 vs. 372 days; P = 0.0033; Figure 2A).

We used the same β value obtained from the CGGA training set to calculate the risk score in two

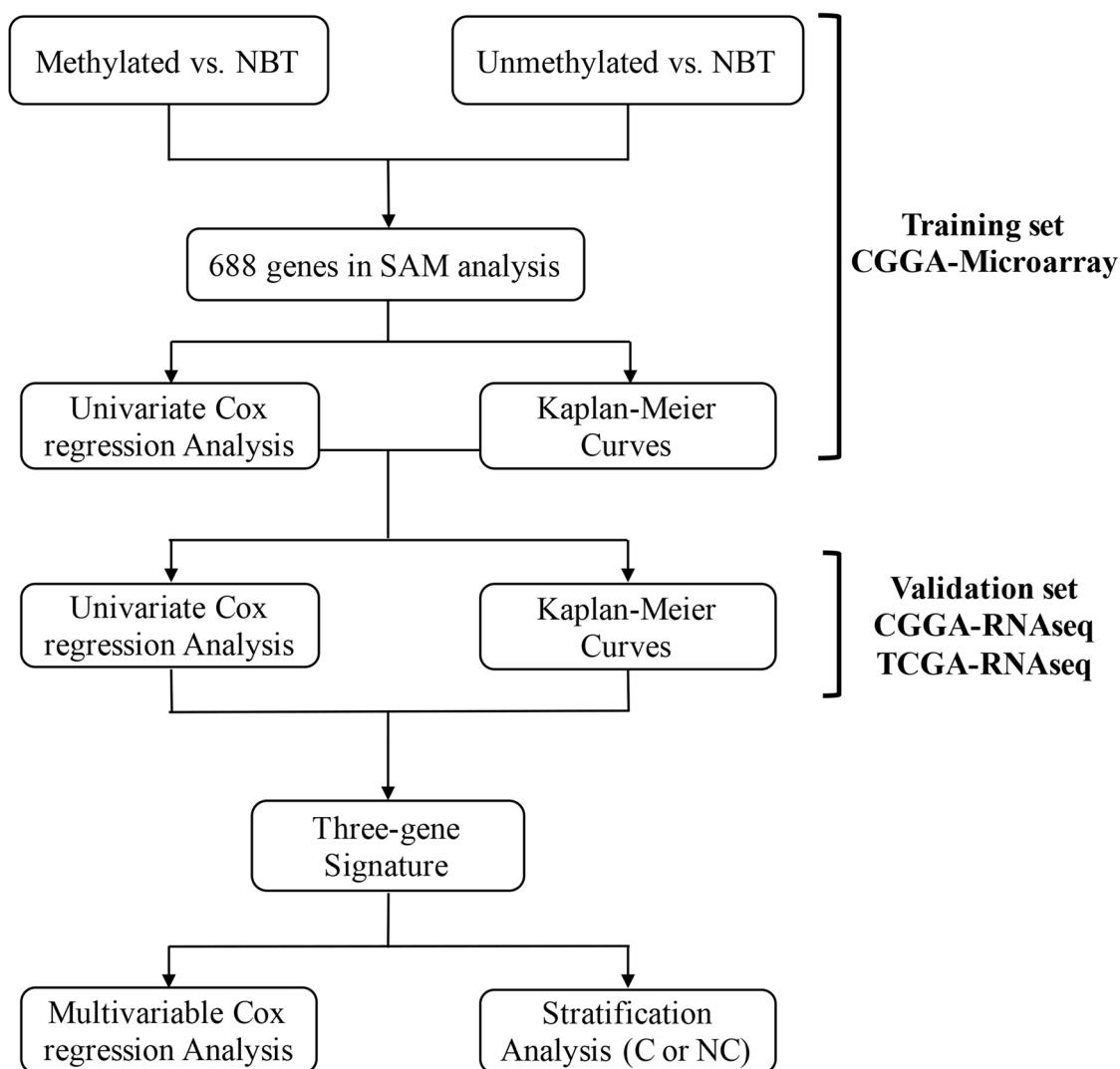


Figure 1: Flow chart indicating the process used to select target genes included in the analysis.

validation datasets. Patients with MGMT methylation were also divided into two groups, low risk group and high risk group, based on the median risk score. The CGGA (59 patients) and TCGA (55 patients) RNAseq datasets showed similar results (median OS 1074 vs. 372; $P = 0.0001$; Figure 2B; median OS 489 vs. 342; $P = 0.0315$; Figure 2A).

Assessment of prognosis value of three-gene signature in related with chemotherapy

We evaluated the prognostic value of the signature for the GBM patients with MGMT promoter methylation. Furthermore, because MGMT promoter methylation status is highly associated with sensitivity to chemotherapy in GBM patients [6, 7], we assessed the predictive value of this signature for chemotherapy.

Twenty-one patients received recommended chemotherapy, while 17 patients didn't in the CGGA microarray dataset. Then we classified the 21 patients with three-gene signature into low and high risk group, and found a significant difference ($P=0.0473$) in OS as shown in Figure 2B. It also showed significant difference

between low risk group with unmethylated group with chemotherapy ($P=0.0197$) (Figure 2B). Moreover, there is no significant difference between low and high risk group or unmethylated MGMT patients without chemotherapy (Figure 2B). We further validated the findings in the TCGA dataset (Figure 2C). The low risk group with chemotherapy treated showed the best outcomes, while the high risk group without chemotherapy treated showed the worst in combined analysis (Figure 2B, 2C). It turned out that the signature could specially recognize a group from the MGMT promoter methylation patients with chemotherapy treated, which showed similar outcomes to the patients with unmethylated MGMT promoter.

Clinical and molecular features of low and high risk GBM patients with MGMT promoter methylation

The expression levels of the three genes showed significant difference between patients with MGMT promoter methylation and non-cancerous brain tissue samples (Figure 3A). Moreover, this difference was also observed between low and high risk group (Figure 3B).

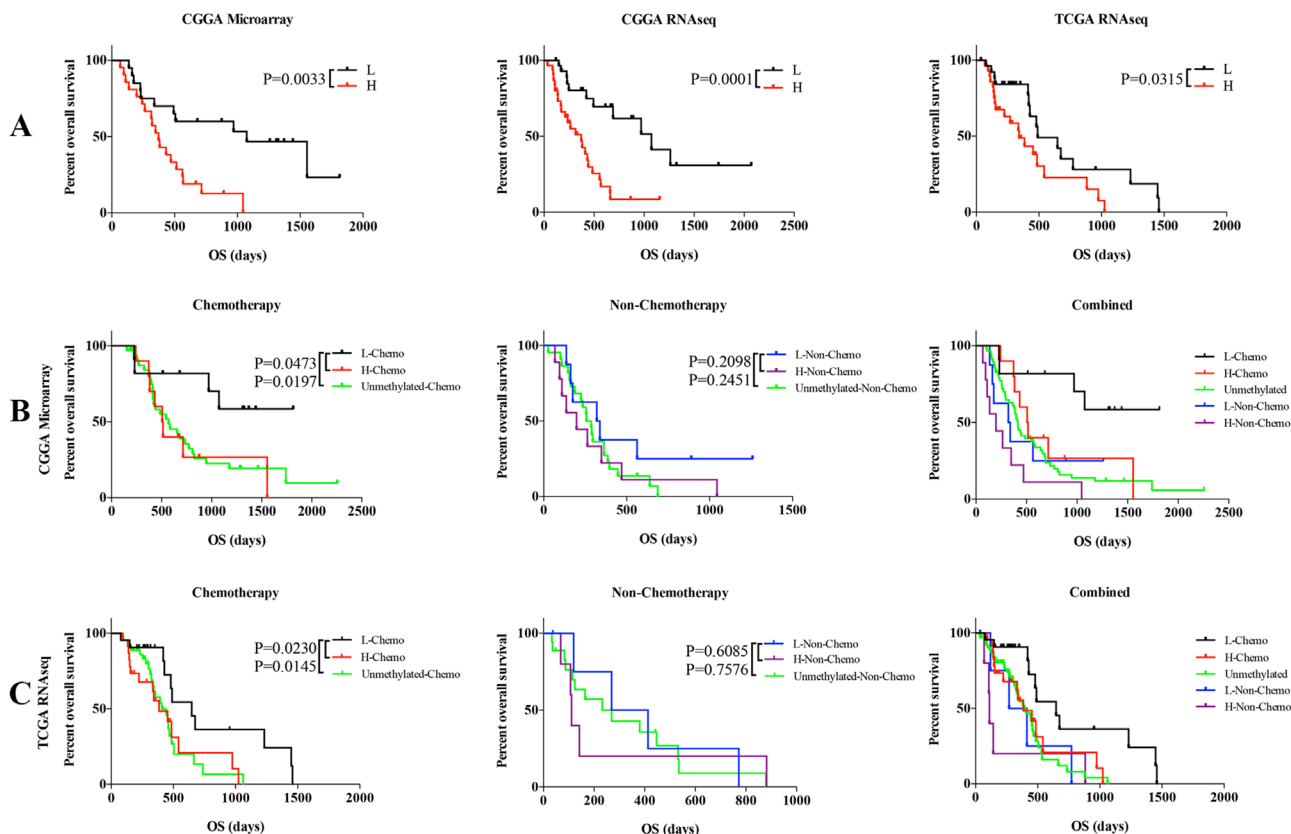


Figure 2: Comparison of prognosis between low and high risk group with MGMT promoter methylation GBM patients and unmethylated GBM patients. A. Survival among GBM patients in different groups stratified by low and high risk group in three datasets. B. Survival among GBM patients in different groups stratified by the signature and chemotherapy in CGGA microarray dataset. C. Survival among GBM patients in different groups stratified by the signature and chemotherapy in TCGA RNA sequencing dataset. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

We assessed the independence of the three-gene signature in the CGGA microarray dataset. It showed that the signature was significant associated with the OS ($P = 0.004$) along with radiotherapy and chemotherapy in the univariate cox regression analysis. On multivariate analysis, it also showed the signature was an independent prognosis factor ($P = 0.004$) (Table 1). In TCGA and CGGA RNAseq dataset, the results indicated that the risk

score was an independent prognostic factor ($P = 0.007$; $P = 0.015$) (Supplementary Table S2).

We observed that GBM patients in the high risk group had shorter OS than low risk group (Figure 3C). The related clinical information such as gender, age, TCGA subtype, IDH1 mutation radiotherapy and chemotherapy were obtained from CGGA microarray database. Patients in high risk group tended to display older age (>45 years),

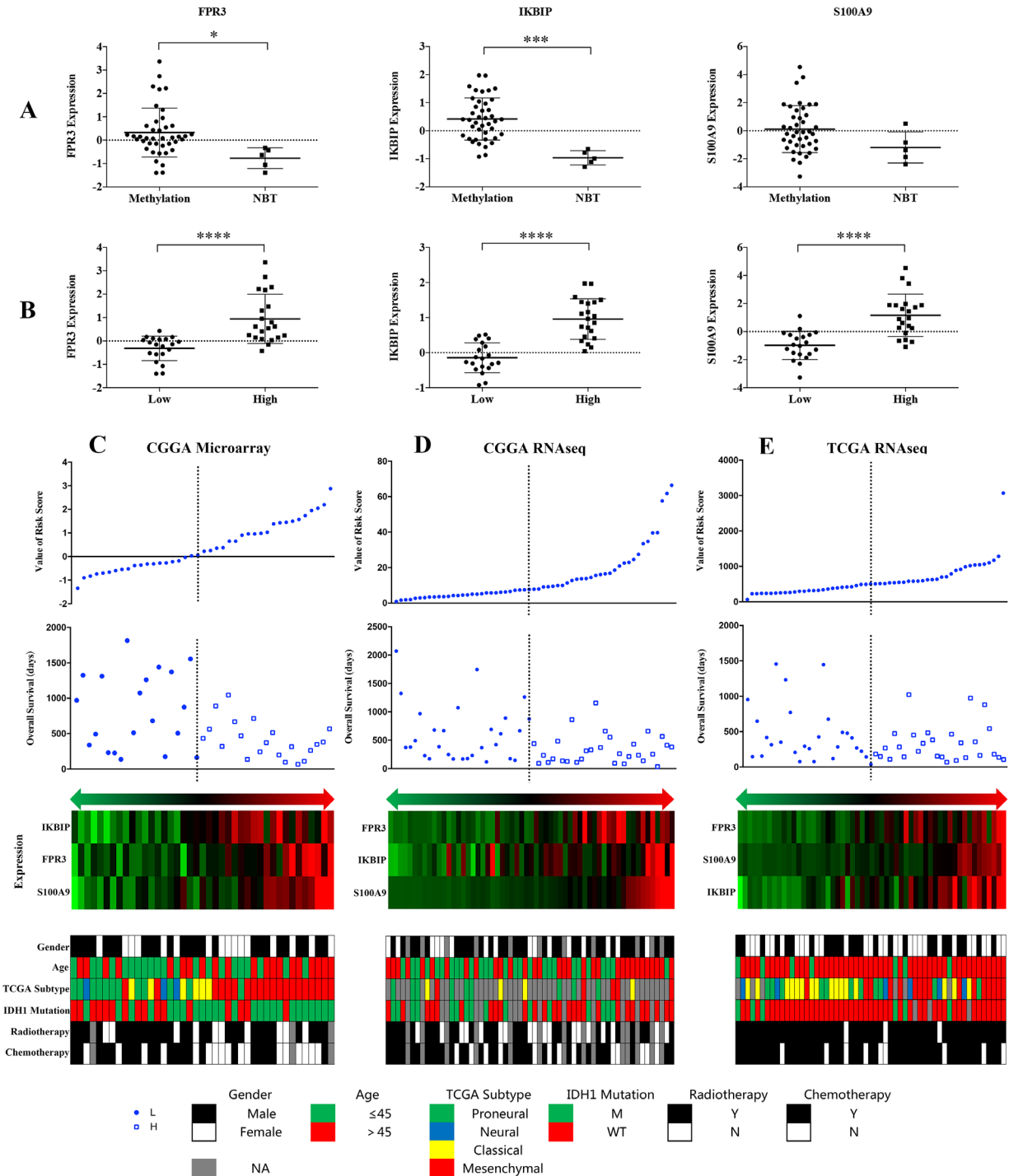


Figure 3: Distribution of risk score, OS, gene expression and clinical or molecular pathological features in CGGA microarray, RNA sequencing and TCGA RNA sequencing datasets.

Table 1: Clinicopathologic factors associated with OS in the Cox regression analysis for patients from the CGGA microarray dataset

Variable	Univariate Cox		Multivariate Cox	
	p-value	HR	p-value	HR
Age	0.870	1.003		
Gender	0.740	0.883		
KPS	0.075	0.968		
IDH1	0.125	0.526		
Radiotherapy	0.001	0.242	0.007	0.264
Chemotherapy	0.002	0.292	0.217	0.565
Risk Score	0.004	1.643	0.004	2.195

Gender, male 1, female 2; IDH1 mutation status, mutated 1, wild-type 0; Radiotherapy, treated 1, untreated 0; Chemotherapy, treated 1, untreated 0.

classical and mesenchymal TCGA subtype and non-chemoradiotherapy. Moreover, we further validated in addition two datasets (Figure 3D, 3E).

Functional annotation of the different prognosis

We performed SAM ($FDR < 0.05$) between low and high risk group in the CGGA microarray dataset and we screened top 500 positively (599 probes) and 500 negatively (588 probes) correlated expression genes with the risk score, respectively. The expression patterns of genes were showed in Figure 4A using a hierarchical clustering analysis. Moreover, we performed the gene enrichment analysis using DAVID (The Database for Annotation, Visualization and Integrated Discovery) [8]. It showed that the biological processes, such as immune response, inflammatory response, regulation of cell death, regulation of apoptosis, biological adhesion and cell adhesion, et al (Figure 4B) were significantly enriched in the high risk group. In contrast, the biological processes, such as cell surface receptor linked signal transduction, neurological system process, intracellular signaling cascade and cell-cell signaling, et al (Figure 4B) were significantly enriched in low risk group. Gene enrichment analysis was further validated by Gene set enrichment analysis (GSEA) [9] and it revealed that the subgroup with high risk score had increased expression of inflammatory response, apoptosis, P53 pathway, hypoxia, epithelial mesenchymal transition (EMT) and TNFA signaling via NFKB (Figure 4C). Nuclear factor kB (NFKB) is a nuclear transcription factor that regulates expression of a large number of genes that are critical for the regulation of apoptosis, tumorigenesis, inflammation, and various autoimmune diseases [10–13].

Overall, the results of GSEA analysis were consistent with GO analysis' and functional annotation of the signature mainly enriched in immune response, apoptosis, cell adhesion pathways. It might partially

explain the poor OS of patients with MGMT promoter methylation in high risk group.

DISCUSSION

It is reported that the status of MGMT promoter methylation is closely associated with chemotherapy sensitivity [6, 7]. However, among the patients with methylation of MGMT, the overall survival is variance significantly. Therefore, we performed the analysis on the whole genome mRNA expression profiling to screen the determinant genes which can predict the overall survival of patients with equivalent MGMT methylated status. As far as we know, it is the first study for the type of patients.

The three genes have different expression between GBMs with non-cancerous tissues, and they all have prognosis values. The risk score based on the three genes divided patients into low and high risk groups and patients in low risk group had longer OS than high risk group. Further, low risk group also showed better OS than high risk groups and unmethylated group in patients with chemotherapy, but no significant difference between low and high or unmethylated MGMT patients without chemotherapy. The signature can specifically predict the prognosis of MGMT promoter methylation patients with chemotherapy treated. It also showed the signature was an independent prognostic factor along with age, gender, IDH1, radiotherapy and chemotherapy.

All of the 3 genes were remarkably associated with prognosis in GBM with MGMT methylation status. FPR3 is a member of the human formyl peptide receptors, which plays an important functional role in the regulation of immune responses and host defense mechanisms [14, 15]. FPR3 is also noted as FPRL2. It is reported that FPR3 can promote calcium mobilization and chemotaxis [16]. GO analysis showed that up-regulation genes were enriched

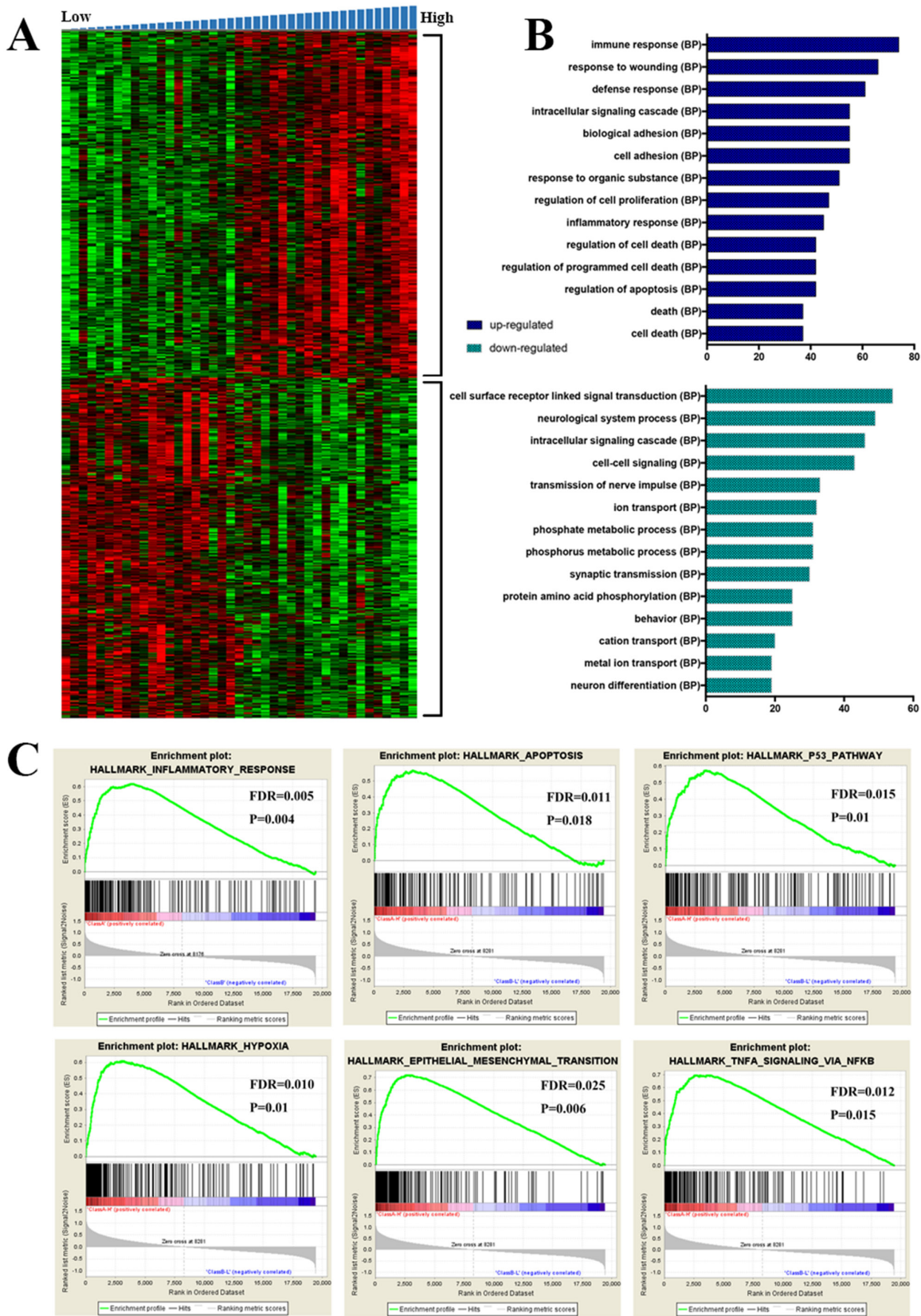


Figure 4: Functional annotation of each risk groups. A. Hierarchical clustering analysis of mRNA expression profiles based on the top 1000 genes. **B.** Biological processes revealed the significant association of the genes with different expression in each group. Column length: gene counts. **C.** The top six enriched pathways in high risk group analyzed by gene set enrichment analysis.

in immune response and inflammatory response pathways. Moreover, it reported that immune response may play important role in GBM recently [17]. IKBIP (I kappa B kinase interacting protein) is a novel p53 target gene with proapoptotic function which is consistency with P53 pathway of GSEA results. It locates on chromosome 12 in close proximity to APAF1 (apoptotic protease-activating factor-1) and the two genes are transcribed in opposite directions [18]. Therefore, we considered it had similar functions to APAF1. APAF1 can encode a cytoplasmic protein that initiates apoptosis and it is a key regulator of mitochondrial apoptotic pathway. It is reported that decreased expression of APAF1 could be interpreted as an event contributing to melanoma chemo-resistance [19]. Moreover, APAF1 positively regulated the 5-FU-induced mitochondria-mediated apoptosis in colorectal cancer cells [20]. GO and GSEA analysis also confirmed that the three-gene signature may have a potential of regulation of apoptosis, cell death. S100A9 is a member of the S100 protein family. It has shown to regulate inflammatory processes [21] and be detected in various human cancers, such as breast cancer, prostate cancer and hepatocellular carcinoma [22–25]. It is reported that S100A9 promotes cell growth and invasion in hepatocellular carcinoma through stimulating MAPK signaling cascades [25]. GO analysis showed that related genes were enriched in biological adhesion and cell adhesion pathways, while EMT pathway in GSEA analysis. Moreover, S100A9 can induce inflammatory cytokines and is associated with overall survival in ER- PgR- breast cancers [22].

There are limitations in our manuscript, only 5 NBT samples were included into SAM analysis and only 41 patients (21 with chemotherapy; 17 non-chemotherapy) collected into our analysis which could cause bias of data analysis. The three-gene signature were selected based on the bioinformatics analysis and it is a totally observational study which may just provide clues for further study of GBM patients with MGMT promoter methylation. We will continue to collect samples and validated these findings experimentally in our future work. However, we further validated screened target genes in CGGA and TCGA RNAseq datasets.

In conclusion, our results showed that the three-gene signature has prognosis value for patients with MGMT promoter-methylated glioblastomas. Due to the poor prognosis in high risk group, clinicians should pay more attention to new treatment therapies. Further study need to validate these findings experimentally in future work.

MATERIALS AND METHODS

Patients and datasets

145 patients (106 microarray; 59 RNA sequencing) from the Chinese Glioma Genome Atlas (CGGA; <http://www.cgga.org.cn/>) were included in our analysis. All

patients' clinical information was download from this website. The RNAseq data were normalized by log2 transformed before analysis. 119 patients from The Cancer Genome Atlas (TCGA; <http://cancergenome.nih.gov/>) were included as validation dataset.

Signature development

We excluded patients without survival data or < 30 days due to they may die of other reasons. The risk score was developed based on a linear combination of the mRNA expression level (expr) weighted by the regression coefficient (β) derived from the univariate Cox regression analysis as previously reported [26–28]. The risk score for each patient was calculated as follows:

$$\text{Risk score} = \beta_{\text{gene1}} \times \text{expr}_{\text{gene1}} + \beta_{\text{gene2}} \times \text{expr}_{\text{gene2}} + \dots + \beta_{\text{genen}} \times \text{expr}_{\text{genen}}$$

The high risk score group presented shorter OS than low risk score group. The significance analysis of microarray (SAM) and Cox analysis was calculated using R software (version 3.2.3) with the samr and survival packages. Moreover, the univariate and multivariate cox regression analysis was performed by SPSS software (version 22; SPSS Inc., Chicago, IL, USA). The Kaplan-Meier analysis was used to evaluate the OS distributions by using GraphPad Prism 6 (GraphPad Software Inc., La Jolla, CA, USA). A two-sided P value of < 0.05 was regarded as statistically significant.

ACKNOWLEDGMENTS AND FUNDING

This work was supported by grants from Ministry of Science and Technology of China grant (2012CB825505 and 2011BAI08B08), National Key Technology Research and Development Program of the Ministry of Science and Technology of China (2013BAI09B03, 2014BAI04B02), National High Technology Research and Development Program (No.2012AA02A508), National Natural Science Foundation of China (No. 91229121), Beijing Science and Technology Plan (No. Z131100006113018).

CONFLICTS OF INTEREST

The authors have declared that no competing interests exist.

REFERENCES

1. Nikiforova MN, Hamilton RL. Molecular diagnostics of gliomas. *Archives of pathology & laboratory medicine*. 2011; 135:558-568.
2. Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJ, Belanger K, Brandes AA, Marosi C, Bogdahn U, Curschmann J, Janzer RC, Ludwin SK, et al.

- Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med.* 2005; 352:987-996.
3. Karsy M, Neil JA, Guan J, Mahan MA, Colman H, Jensen RL. A practical review of prognostic correlations of molecular biomarkers in glioblastoma. *Neurosurg Focus.* 2015; 38:E4.
 4. Noushmehr H, Weisenberger DJ, Diefes K, Phillips HS, Pujara K, Berman BP, Pan F, Pelloski CE, Sulman EP, Bhat KP, Verhaak RG, Hoadley KA, Hayes DN, et al. Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell.* 2010; 17:510-522.
 5. Watanabe T, Vital A, Nobusawa S, Kleihues P, Ohgaki H. Selective acquisition of IDH1 R132C mutations in astrocytomas associated with Li-Fraumeni syndrome. *Acta Neuropathol.* 2009; 117:653-656.
 6. Esteller M, Garcia-Foncillas J, Andion E, Goodman SN, Hidalgo OF, Vanaclocha V, Baylin SB, Herman JG. Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *N Engl J Med.* 2000; 343:1350-1354.
 7. Hegi ME, Diserens AC, Gorlia T, Hamou MF, de Tribolet N, Weller M, Kros JM, Hainfellner JA, Mason W, Mariani L, Bromberg JE, Hau P, Mirimanoff RO, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med.* 2005; 352:997-1003.
 8. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009; 4:44-57.
 9. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005; 102:15545-15550.
 10. Sheikh MS, Huang Y. Death receptor activation complexes: it takes two to activate TNF receptor 1. *Cell Cycle.* 2003; 2:550-552.
 11. Hayden MS, West AP, Ghosh S. NF-kappaB and the immune response. *Oncogene.* 2006; 25:6758-6780.
 12. Gilmore TD. Introduction to NF-kappaB: players, pathways, perspectives. *Oncogene.* 2006; 25:6680-6684.
 13. Liu F, Bardhan K, Yang D, Thangaraju M, Ganapathy V, Waller JL, Liles GB, Lee JR, Liu K. NF-kappaB directly regulates Fas transcription to modulate Fas-mediated apoptosis and tumor suppression. *The Journal of biological chemistry.* 2012; 287:25530-25540.
 14. Lee HY, Lee SY, Shin EH, Kim SD, Kim JM, Lee MS, Ryu SH, Bae YS. F2L, a peptide derived from heme-binding protein, inhibits formyl peptide receptor-mediated signaling. *Biochemical and biophysical research communications.* 2007; 359:985-990.
 15. Devosse T, Guillabert A, D'Haene N, Berton A, De Nadai P, Noel S, Brait M, Franssen JD, Sozzani S, Salmon I, Parmentier M. Formyl peptide receptor-like 2 is expressed and functional in plasmacytoid dendritic cells, tissue-specific macrophage subpopulations, and eosinophils. *Journal of immunology.* 2009; 182:4974-4984.
 16. Migeotte I, Riboldi E, Franssen JD, Gregoire F, Loison C, Wittamer V, Detheux M, Robberecht P, Costagliola S, Vassart G, Sozzani S, Parmentier M, Communi D. Identification and characterization of an endogenous chemotactic ligand specific for FPRL2. *The Journal of experimental medicine.* 2005; 201:83-93.
 17. Cheng W, Ren X, Zhang C, Cai J, Liu Y, Han S, Wu A. Bioinformatic profiling identifies an immune-related risk signature for glioblastoma. *Neurology.* 2016; 86:2226-2234.
 18. Hofer-Warbinek R, Schmid JA, Mayer H, Winsauer G, Orel L, Mueller B, Wiesner C, Binder BR, de Martin R. A highly conserved proapoptotic gene, IKIP, located next to the APAF1 gene locus, is regulated by p53. *Cell death and differentiation.* 2004; 11:1317-1325.
 19. Campioni M, Santini D, Tonini G, Murace R, Dragonetti E, Spugnini EP, Baldi A. Role of Apaf-1, a key regulator of apoptosis, in melanoma progression and chemoresistance. *Experimental dermatology.* 2005; 14:811-818.
 20. Shang J, Yang F, Wang Y, Wang Y, Xue G, Mei Q, Wang F, Sun S. MicroRNA-23a antisense enhances 5-fluorouracil chemosensitivity through APAF-1/caspase-9 apoptotic pathway in colorectal cancer cells. *Journal of cellular biochemistry.* 2014; 115:772-784.
 21. Benedyk M, Sopalla C, Nacken W, Bode G, Melkonyan H, Banfi B, Kerkhoff C. HaCaT keratinocytes overexpressing the S100 proteins S100A8 and S100A9 show increased NADPH oxidase and NF-kappaB activities. *The Journal of investigative dermatology.* 2007; 127:2001-2011.
 22. Bergenfelz C, Gaber A, Allaoui R, Mehmeti M, Jirstrom K, Leanderson T, Leandersson K. S100A9 expressed in ER(-) PgR(-) breast cancers induces inflammatory cytokines and is associated with an impaired overall survival. *British journal of cancer.* 2015; 113:1234-1243.
 23. Hermani A, De Servi B, Medunjanin S, Tessier PA, Mayer D. S100A8 and S100A9 activate MAP kinase and NF-kappaB signaling pathways and trigger translocation of RAGE in human prostate cancer cells. *Experimental cell research.* 2006; 312:184-197.
 24. Choi DK, Li ZJ, Chang IK, Yeo MK, Kim JM, Sohn KC, Im M, Seo YJ, Lee JH, Kim CD, Lee Y. Clinicopathological roles of S100A8 and S100A9 in cutaneous squamous cell carcinoma *in vivo* and *in vitro*. *Archives of dermatological research.* 2014; 306:489-496.
 25. Wu R, Duan L, Cui F, Cao J, Xiang Y, Tang Y, Zhou L. S100A9 promotes human hepatocellular carcinoma cell

- growth and invasion through RAGE-mediated ERK1/2 and p38 MAPK pathways. *Experimental cell research*. 2015; 334:228-238.
26. Cheng W, Ren X, Cai J, Zhang C, Li M, Wang K, Liu Y, Han S, Wu A. A five-miRNA signature with prognostic and predictive value for MGMT promoter-methylated glioblastoma patients. *Oncotarget*. 2015; 6:29285-29295. doi: 10.18632/oncotarget.4978.
27. Bao ZS, Li MY, Wang JY, Zhang CB, Wang HJ, Yan W, Liu YW, Zhang W, Chen L, Jiang T. Prognostic value of a nine-gene signature in glioma patients based on mRNA expression profiling. *CNS neuroscience & therapeutics*. 2014; 20:112-118.
28. Zhang CB, Zhu P, Yang P, Cai JQ, Wang ZL, Li QB, Bao ZS, Zhang W, Jiang T. Identification of high risk anaplastic gliomas by a diagnostic and prognostic signature derived from mRNA expression profiling. *Oncotarget*. 2015; 6:36643-36651. doi: 10.18632/oncotarget.5421.