

RESEARCH

Open Access



Identification of a key glioblastoma candidate gene, *FUBP3*, based on weighted gene co-expression network analysis

Jianmin Li^{1*}, Zhao Zhang^{1†}, Ke Guo^{1†}, Shuhua Wu², Chong Guo¹, Xinfan Zhang¹ and Zi Wang¹

Abstract

Background: Glioblastoma multiforme (GBM) is the most common aggressive malignant brain tumor. However, the molecular mechanism of glioblastoma formation is still poorly understood. To identify candidate genes that may be connected to glioma growth and development, weighted gene co-expression network analysis (WGCNA) was performed to construct a gene co-expression network between gene sets and clinical characteristics. We also explored the function of the key candidate gene.

Methods: Two GBM datasets were selected from GEO Datasets. The R language was used to identify differentially expressed genes. WGCNA was performed to construct a gene co-expression network in the GEO glioblastoma samples. A custom Venn diagram website was used to find the intersecting genes. The GEPIA website was applied for survival analysis to determine the significant gene, *FUBP3*. OS, DSS, and PFI analyses, based on the UCSC Cancer Genomics Browser, were performed to verify the significance of *FUBP3*. Immunohistochemistry was performed to evaluate the expression of *FUBP3* in glioblastoma and adjacent normal tissue. KEGG and GO enrichment analyses were used to reveal possible functions of *FUBP3*. Microenvironment analysis was used to explore the relationship between *FUBP3* and immune infiltration. Immunohistochemistry was performed to verify the results of the microenvironment analysis.

Results: GSE70231 and GSE108474 were selected from GEO Datasets, then 715 and 694 differentially expressed genes (DEGs) from GSE70231 and GSE108474, respectively, were identified. We then performed weighted gene co-expression network analysis (WGCNA) and identified the most downregulated gene modules of GSE70231 and GSE108474, and 659 and 3915 module genes from GSE70231 and GSE108474, respectively, were selected. Five intersection genes (*FUBP3*, *DAD1*, *CLIC1*, *ABR*, and *DNM1*) were calculated by Venn diagram. *FUBP3* was then identified as the only significant gene by survival analysis using the GEPIA website. OS, DSS, and PFI analyses verified the significance of *FUBP3*. Immunohistochemical analysis revealed *FUBP3* expression in GBM and adjacent normal tissue. KEGG and GO analyses uncovered the possible function of *FUBP3* in GBM. Tumor microenvironment analysis showed that *FUBP3* may be connected to immune infiltration, and immunohistochemistry identified a positive correlation between immune cells (CD4 + T cells, CD8 + T cells, and macrophages) and *FUBP3*.

*Correspondence: jml6309@126.com

†Zhao Zhang and Ke Guo contributed as co-first authors.

¹ Department of Neurosurgery, Binzhou Medical University Hospital, Binzhou, Shandong Province, People's Republic of China

Full list of author information is available at the end of the article



Conclusion: *FUBP3* is associated with immune surveillance in GBM, indicating that it has a great impact on GBM development and progression. Therefore, interventions involving *FUBP3* and its regulatory pathway may be a new approach for GBM treatment.

Keywords: GBM, GEO, Immune-related gene, Molecular biomarkers, Survival analysis, WGCNA

Introduction

Glioblastoma multiforme (GBM) is the most aggressive type of brain tumor, arising from the astrocytes, the neuroepithelial tissue group of the brain. GBM, also known as grade IV astrocytoma, is the most dangerous type of astrocytoma [1, 2]. The survival time for GBM patients is no more than 15 months [3]. Although surgery is generally used for GBM therapy [4], the cancer often relapses because of the lack of an effective prevention for GBM [5]. Effective biomarkers play a vital role in tumor treatment. Although there are many biomarkers closely related to GBM, including insulin-like growth factor 1 (IGF-1), vascular endothelial growth factor (VEGF), isocitrate dehydrogenase 1 (IDH1), and epidermal growth factor receptor (EGFR), no effective tumor markers have been discovered [6, 7]. Therefore, it is crucial to uncover appropriate and effective biomarkers to predict the prognosis for glioblastoma patients.

Over the last decade, gene set analysis has become the first choice for gaining insights into underlying complex biology of diseases through gene expression and gene association studies. It also reduces the complexity of statistical analysis and enhances the explanatory power of the obtained results [8]. In recent years, high-quality microarray techniques and high-throughput sequencing have been widely applied in clinical medicine and have created a new generation of molecular diagnostics based on DNA sequencing, RNA sequencing, and epigenetics [9]. Gene profiles can be obtained from public datasets such as Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA), and the limitations between the combinations of different sample sizes and microarray platforms can be overcome by integrated bioinformatics methods. In order to find out the relationship among the selected genes, identification of key genes and gene modules responsible for a particular stress/condition, analysis of gene co-expression networks need to be carried out. Weighted Gene Co-expression Network Analysis (WGCNA) is a latest and popular technique used to decipher co-expression patterns among genes. The WGCNA approach typically deals with the identification of gene modules by using the gene expression levels that are highly correlated across samples [10]. And WGCNA has been successfully applied in many fields to identify candidate biomarkers and therapeutic targets [11].

In this study, we performed WGCNA for two RNA-Seq datasets derived from GEO and reconstructed gene co-expression networks, then we separately identified the most downregulated gene modules of the two datasets. A Venn diagram and GEPIA were used to find the intersecting genes of the two modules and further identify significant genes by survival analysis, and the significance of *FUBP3* was verified by Overall Survival (OS), Disease-Specific Survival (DSS), and Progression-Free Interval (PFI) analyses. We evaluated the expression of *FUBP3* in GBM tissues and normal tissue adjacent to tumors by immunohistochemistry. We then considered the possible mechanism of action of *FUBP3* in GBM by KEGG and GO analyses and tumor microenvironment analysis based on TCGA tumor datasets, and immunohistochemistry identified a positive correlation between immune cells and *FUBP3*.

Materials and methods

WGCNA analysis of GEO

RNA sequence data from human glioblastoma samples were obtained from GEO datasets (<http://www.ncbi.nlm.nih.gov/geo/>). A microarray platform data table was used, annotating a series of matrix files with official gene symbols, and a gene expression matrix file was obtained for DEG analysis. The key search words were (glioma) AND “Homo sapiens”[porgn: __txid9606]. Finally, two datasets from the search results were included in our study, GSE70231 and GSE108474. These datasets contain glioblastoma and normal brain samples (GSE70231: GBM, $n=21$; controls, $n=6$, and GSE108474: GBM, $n=221$; controls, $n=28$) and were used to perform differential analysis. A series of matrix files and data sheets for the microarray platform were downloaded from the GEO website. The “limma” R package was used for DEG analysis. The DEG threshold was set as $|\log_2FC| > 1$, $P < 0.05$. The DEG sets for GSE70231 and GSE108474 were then selected (GEO70231_diff and GEO108474_diff).

The “WGCNA” R package was used to construct a co-expression network for all genes in the GBM and normal samples. Samples were used to construct Pearson’s correlation matrices, then the formula $a_{mn} = |c_{mn}|^\beta$ (where a_{mn} is adjacency between gene m and gene n , c_{mn} is Pearson’s correlation, and β is the soft-power threshold) was used to construct a weighted adjacency matrix. This was then transformed into a TOM (topological overlap

measure) matrix to evaluate the network connectivity, and a clustering dendrogram of the TOM matrix was constructed by average connected hierarchical clustering [12]. The threshold was set to 0.9 in GSE70231 and 0.95 in GSE108474. Then the *P*-values of the module eigengenes constructed by each sample and module gene were calculated by the WGCNA algorithm. The Pearson's correlation coefficient was used to identify the connection between clinical traits and different modules, and the module with the highest Pearson's correlation coefficient was selected for subsequent analysis. For our study, the most downregulated gene modules (GSE70231_turquoise and GSE108474_turquoise) in the GBM samples compared with healthy controls were selected.

Key gene screening

A custom Venn diagram website (<http://bioinformatics.psb.ugent.be/webtools/Venn/>) was used to find the intersecting genes between GSE70231_diff, GSE108474_diff, GSE70231_turquoise, and GSE108474_turquoise.

Verifying intersecting genes using GEPIA

The GEPIA website (<http://gepia.cancer-pku.cn/>) is a newly developed interactive web server for analyzing the RNA sequencing expression data of 9,736 tumors and 8,587 normal samples from the TCGA and GTEx projects, using a standard processing pipeline. In our study, GEPIA was used for survival analysis of the intersecting genes and to identify the significant key gene, *FUBP3*. First, we set the intersection gene used for normalizing in the "Gene" field, then we selected the OS survival method, chose GBM in the "Dataset Selection" field, and clicked "add" to build a dataset list in the "Datasets" field.

Revalidating the key gene

The UCSC Cancer Genomics Browser (<https://genome-cancer.ucsc.edu>) offers interactive visualization and exploration of TCGA genomic, phenotypic, and clinical data, as produced by The Cancer Genome Atlas Research Network. The survival and survminer packages of R were used for overall survival(OS) analysis of *FUBP3* in GBM. The limma, survivor, and survminer R packages were used for the other survival analyses (DSS and PFI). The gene expression and clinical datasets were downloaded for 33 common types of cancer from UCSC based on TCGA datasets. Then *FUBP3* was used for three survival analyses, OS, DSS, and PFI, to revalidate the results found by the GEPIA website.

Pathway analysis

The "pathway analysis" has been widely used in many ways including the analysis of Gene Ontology (GO) terms (also referred to as a "gene set"), physical interaction

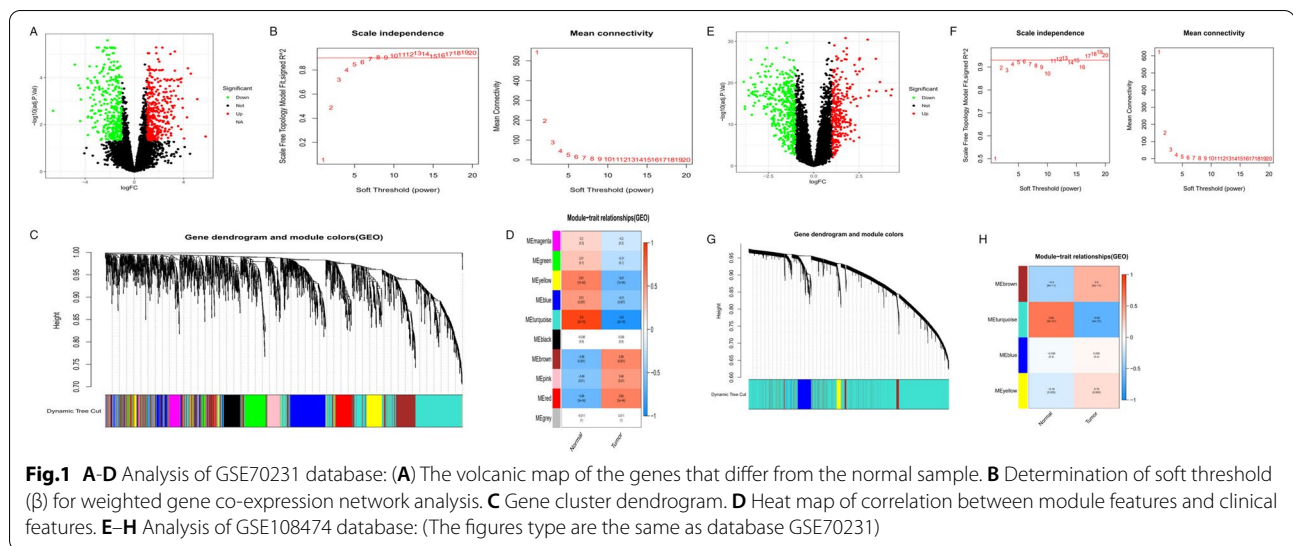
networks (e.g., protein–protein interactions), kinetic simulation of pathways, steady-state pathway analysis (e.g., flux-balance analysis), and in the inference of pathways from expression and sequence data. And in our study, we focused on methods that exploit pathway knowledge in public repositories such as gene ontology (GO) or Kyoto Encyclopedia of Genes and Genomes (KEGG) to identify pathways that may be affected in a condition by correlating information in at least one pathway knowledge base with gene expression patterns for the condition [13]. In recent years, the effective clustering of functional genes, mainly based on GO and KEGG, has been applied widely in DNA- and protein-related research [14]. The KEGG and GO databases can be downloaded from GSEA's official website (<https://www.gsea-msigdb.org/gsea/register.jsp>). In our study, GO and KEGG analyses were used to reveal the possible function of *FUBP3* in GBM. The DEG functional analyses were performed using the clusterProfiler, org.Hs.eg.db, enrich-plot, and ggplot2 R packages, with *P* < 0.05 considered statistically significant.

Microenvironment analysis

TIMER webserver (<https://cistrome.shinyapps.io/timer/>) is a comprehensive resource for systematic analysis of immune infiltrates across diverse cancer types. The abundances of six immune infiltrates (B cells, CD4 + T cells, CD8 + T cells, neutrophils, macrophages, and dendritic cells) were estimated by TIMER algorithm, and the gene module was selected to explore the correlation between *FUBP3* expression and the abundance of immune infiltrates.

Immunohistochemical analysis

The human anti-*FUBP3* monoclonal antibody came from Affinity Biosciences (Victoria, Australia), and the secondary antibodies, DAB color detection, enzyme-labeled goat anti-mouse/rabbit IgG polymer, mouse mAb assisted/induced T cell (CD4) antibody, rabbit mAb inhibitory/cytotoxic T cell (CD8) antibody, and mouse macrophage cell (CD68) antibody, were produced by Beijing Zhongshan Jinqiao Biotechnology. All the samples were selected from the Department of Pathology of the Affiliated Hospital of Binzhou Medical University and were double-blind rediagnosed by two senior pathologists. We selected 41 samples of human brain tissue from glioblastoma patients, from wax blocks that were serially sectioned, commonly dewaxed, rehydrated, and EDTA repaired. After intervention, we added H₂O₂ to the slices, after water jet and PBS washes. After serum blocking, human anti-*FUBP3* monoclonal antibody (1:150), mouse mAb assisted/induced T cell (CD4) antibody, rabbit mAb inhibitory/cytotoxic T cell (CD8) antibody, and macrophage cell (CD68) antibody were added to both



the glioblastoma tissue and normal tissue adjacent to the cancer, and incubated overnight at 4 °C. After water jet and PBS rinses the next day, the secondary antibody was added, then the slices were incubated for 30 min at 37 °C, and color developing agent was added to the glioblastoma tissue and the tissue adjacent to the cancer. Then, the tissue was stained retained with hematoxylin and differentiated by hydrochloric acid. Finally, they were washed with alcohol and left to dry. Cells expressing FUBP3 in the nucleus could be identified as the nucleus stained yellow or brownish-yellow, and CD68, CD4, and CD8 were expressed in the cytoplasm of microglia and T cells and appeared as brownish-yellow or light yellow colored granules. Ten independent fields of view images (400 ×) for each section were randomly chosen for image acquisition, and the five highest densities of FUBP3-expressed cells and immune cells were selected. The average value of the number of FUBP3-expressed cells and immune cells (CD4+T cells, CD8+T cells, and CD68+macrophages) in each of the five fields were calculated. Samples with higher counts than average were classified as strong immune invasion, and samples with lower counts than average were classified as weak immune invasion. For the staining intensity, a score of 0 was given for cells without staining, 1 for light yellow, 2 for yellow–brown, and 3 for brown. For the number of positive cells, a score of 0 was given for ≤ 5%, 1 for 6%–25%, 2 for 26%–50%, 3 for 51%–75%, and 4 for >75%. Then the two scores were multiplied, and a final score ≥ 4 was considered positive.

Statistical analysis

SPSS 25.0 statistical software and R software (v 3.6.3) were used for statistical analysis. A statistical significance was identified between multiple groups by one-way

ANOVA analysis of variance with the Tukey HSD test based on data in this study. The correlation analysis of protein expression was performed by spearman correlation analysis, and $P < 0.05$ was considered statistically significant.

Results

Pre-processing of GSM RNA sequencing data

Glioblastoma RNA-Seq data from GSE70231 and GSE108474 were downloaded from GEO datasets. Then we obtained a gene expression matrix file for DEG analysis (GBM patients compared with controls). The “limma” R package was used for the DEG analysis. The DEG threshold was set as $|\log_2FC| > 1$, $P < 0.05$. The GSE70231_diff dataset included 715 DEGs, consisting of 322 downregulated genes and 393 upregulated genes, while GSE108474_diff included 694 DEGs, consisting of 417 downregulated genes and 277 upregulated genes. These are shown in two volcano diagrams (Fig. 1: A, E).

Weighted gene co-expression network analysis and module preservation

WGCNA was used to construct two gene co-expression networks for the glioblastoma and control samples of GSE70231 and GSE108474. The gene dendrogram constructed by dynamic tree was then identified, where each color represents one gene module. Then the module–trait relationships were determined. Finally, the most downregulated gene modules (GSE70231_turquoise and GSE108474_turquoise) from GSE70231 and GSE108474 were selected (Fig. 1: B-D, F-H).

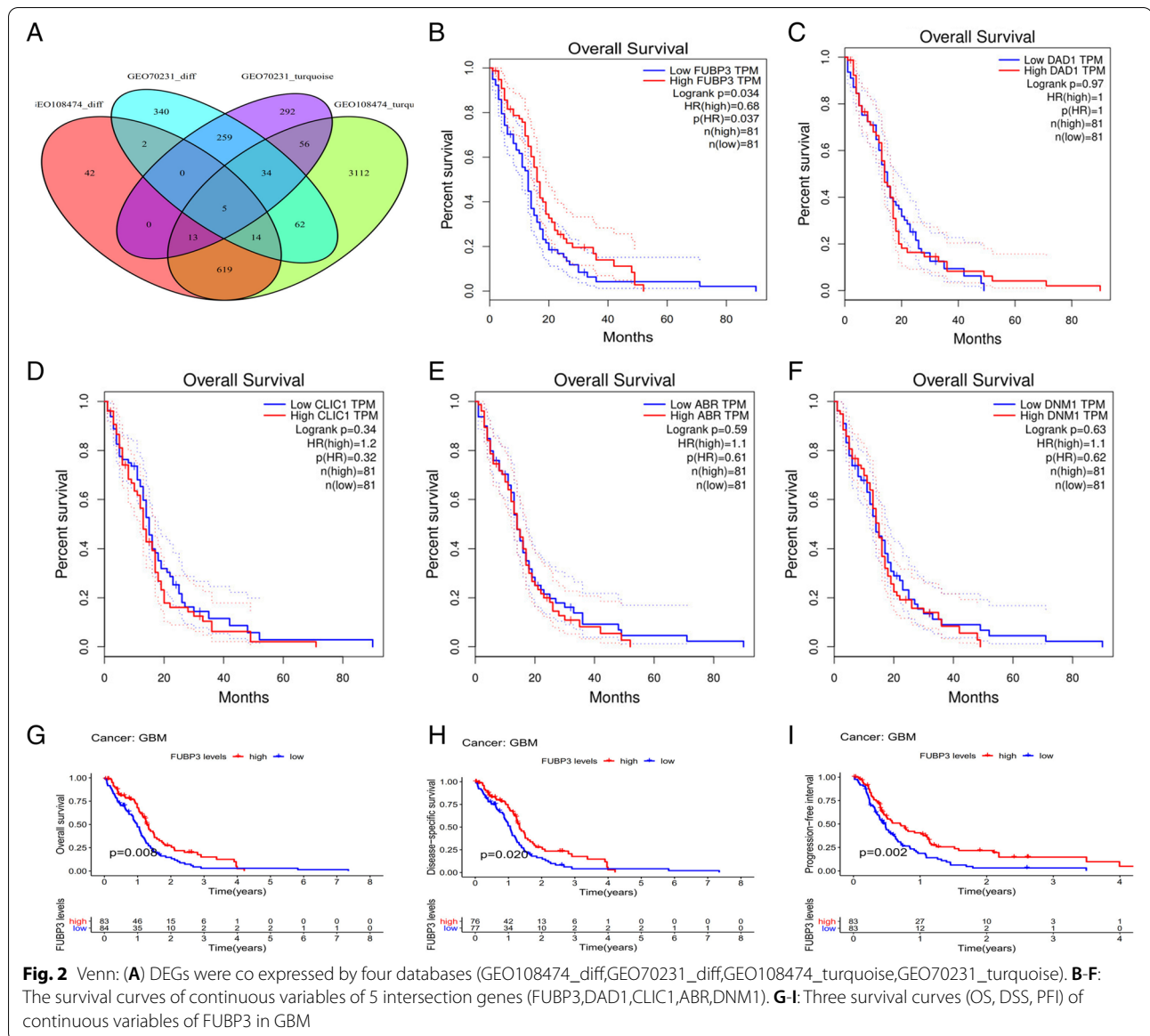


Fig. 2 Venn: (A) DEGs were co expressed by four databases (GEO108474_diff,GEO70231_diff,GEO108474_turquoise,GEO70231_turquoise). **B-F**: The survival curves of continuous variables of 5 intersection genes (FUBP3,DAD1,CLIC1,ABR,DNM1). **G-I**: Three survival curves (OS, DSS, PFI) of continuous variables of FUBP3 in GBM

Screening co-expressed genes

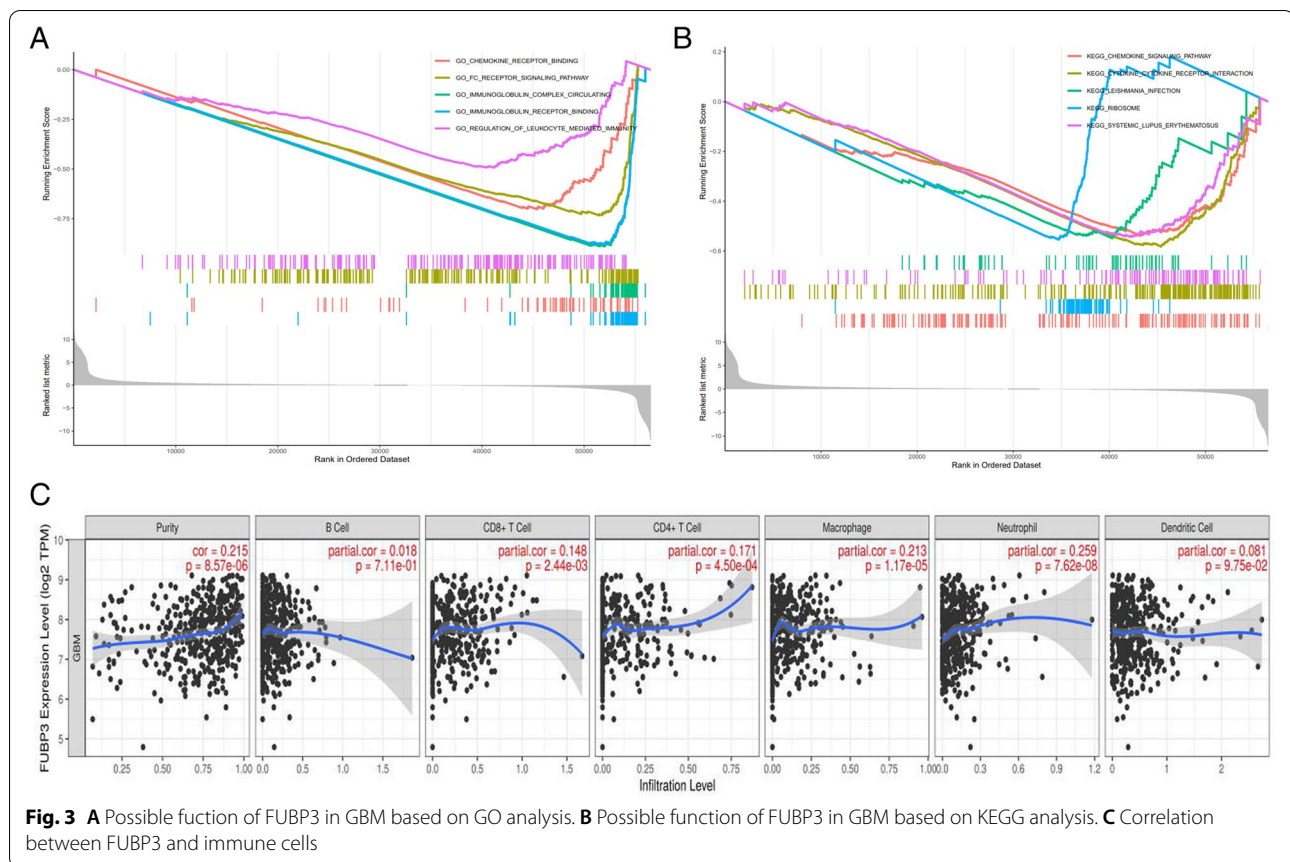
A custom Venn diagram website was used to calculate the intersection of the lists of elements in our study. We entered files (in plain text format) from four elements: GSE70231_diff, GSE108474_diff, GSE70231_turquoise, and GSE108474_turquoise, where GSE70231_diff contain 715 DEGs from GSE70231, GSE108474_diff contain 694 DEGs from GSE108474, GSE70231_turquoise contain the most downregulated gene module from GSE70231 and GSE108474_turquoise contain the most downregulated gene module from GSE108471 respectively based on WGCNA. Finally, five intersecting downregulated genes (*FUBP3*, *DAD1*, *CLIC1*, *ABR*, and *DNM1*) were found in common between the four datasets (Fig. 2: A).

Verifying intersecting genes using GEPIA

The GEPIA website performs overall survival (OS), disease-free survival (DFS, also called relapse-free survival or RFS) analyses based on gene expression. This analysis selected *FUBP3* as the significant gene for OS ($P < 0.05$) (Fig. 2: B-F).

Revalidating *FUBP3* based on the UCSC Cancer Genomics Browser

The results showed that when the overall survival was restricted to a particular level, the higher the *FUBP3* levels, the longer the GBM patients survived ($P = 0.008$). When the disease-specific survival was restricted to the same level, the higher the *FUBP3*



levels, the longer the GBM patients survived ($P=0.02$). When the progression-free interval was restricted to the same level, the higher the *FUBP3* levels, the longer the GBM patients survived ($P=0.002$) (Fig. 2: G-I).

Exploring the possible function of *FUBP3* in GBM based on KEGG and GO analyses

GO analysis showed that *FUBP3* may be mainly involved in chemokine receptor binding, the FC receptor signaling pathway, immunoglobulin complex circulating, immunoglobulin receptor binding, and regulation of leukocyte mediated immunity. KEGG analysis showed that *FUBP3* may be significantly associated with the chemokine signaling pathway, cytokine receptor interaction, leishmania infection, the ribosome, and systemic lupus erythematosus (Fig. 3: A-B).

Tumor microenvironment analysis based on TIMER

The results showed that as the expression of *FUBP3* increased, the number of CD8+ T cells ($cor=0.148$, $P<0.05$), CD4+ T cells ($cor=0.171$, $P<0.05$), macrophages ($cor=0.213$, $P<0.05$), and neutrophils

($cor=0.259$, $P<0.05$) all increased. The relationships between *FUBP3* and B cells and dendritic cells were insignificant (both $P<0.05$) (Fig. 3: C).

Immunohistochemical verification of *FUBP3* and immune-related cell expression

FUBP3 expression was identified by immunohistochemistry. *FUBP3* was mainly expressed in the nucleus of the cell. In our study, *FUBP3* was expressed in the normal tissue adjacent to the cancer and in the GBM samples to varying degrees. Through one-way analysis of variance comparison, compared with controls, the number of glioblastoma cells expressing *FUBP3* was greater, but the color of the cell nuclei was lighter, meaning that the expression of *FUBP3* in the cell nucleus of GBM tissues decreased ($P<0.05$) (Fig. 4: A-B) (Table 1). The 41 GBM samples were divided into a high *FUBP3* expression group and a low *FUBP3* expression group based on the immunochemical assessment of the intensity of invasion. The infiltration intensities of CD4+ T cells, CD8+ T cells, and CD68+ macrophages in the *FUBP3* high expression group were significantly stronger than those in the *FUBP3* low expression group ($P<0.05$) (Fig. 4: C-H) (Table 2).

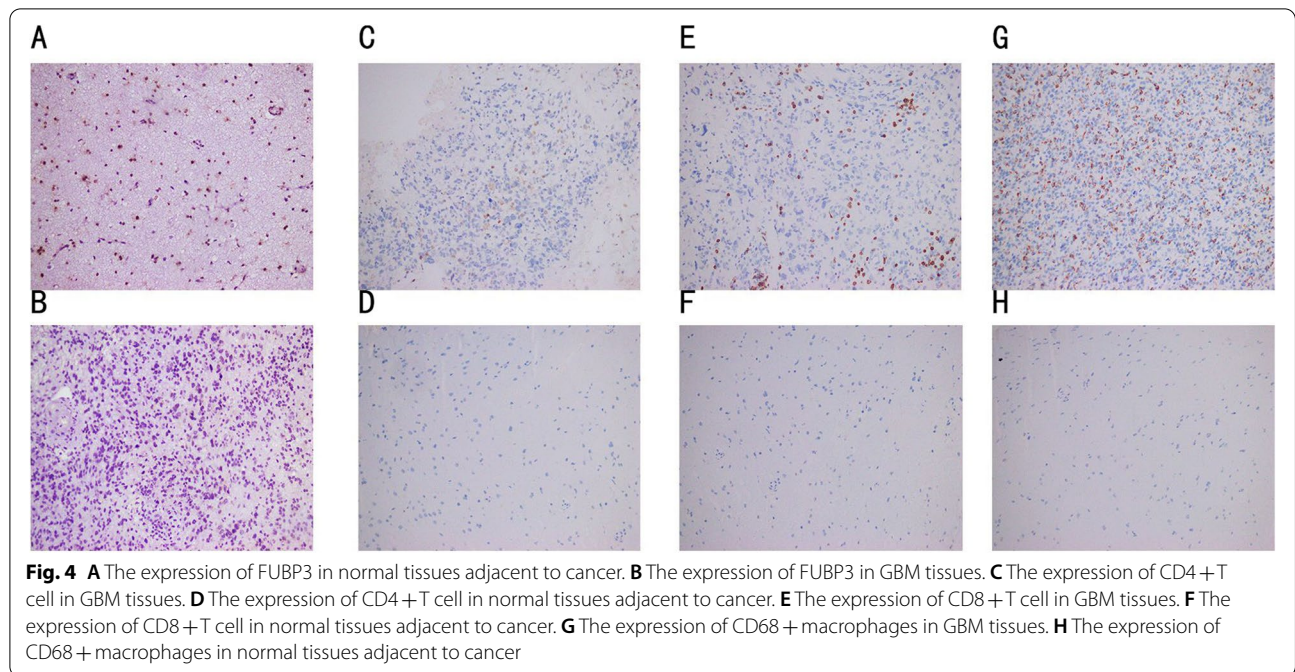


Table 1 Expression of FUBP3 in GBM and normal tissue adjacent to cancer

	n	FUBP3		P
		+	-	
Normal	41	41	0	<0.001*
Tumor	41	15	26	

Notes: *P < 0.05

Abbreviation: FUBP3: expression of FUBP3 in GBM and normal tissue adjacent to cancer

Discussion

Glioblastoma (GBM) is the most common primary malignant brain tumor, accounting for 16% of all primary brain and central nervous system neoplasms [15]. Currently,

the standard therapy for GBM includes surgical resection followed by temozolomide (TMZ) [16], but it is difficult to remove the tumor completely, so the surrounding brain tissue remains with tumor infiltration, leading to tumor progression or relapse [17]. Despite great effort, little progress has been made towards prolonged survival of GBM. Therefore, it is important to identify significant prognostic biomarkers, such as mutations in isocitrate dehydrogenase (IDH) and O6-methylguanine-methyltransferase (MGMT) promoter methylation, to improve the prognosis for GBM patients.

In this study, bioinformatics methods and clinical experiments were used to analyze the DEGs between GBM and normal samples to explore the key candidate genes. After pre-processing of the RNA sequencing

Table 2 Correlation between the expression of FUBP3 and immune cells of GBM tissues

GBM		n	FUBP3		r	P
			+	-		
CD4 +	+	16	13	3	0.742	<0.001*
	-	25	2	23		
CD8 +	+	20	14	6	0.677	<0.001*
	-	21	1	20		
CD68 +	+	17	13	4	0.697	<0.001*
	-	24	2	22		

Notes: *P < 0.05

Abbreviation: Correlation between the expression of FUBP3 and immune cells(CD4 +, CD8 +, CD68 +) of GBM tissues

data, 715 DEGs from GSE70231 were included, consisting of 322 downregulated and 393 upregulated genes, and 694 DEGs from GSE108474 were included, consisting of 417 downregulated and 277 upregulated genes. Then, WGCNA was performed to determine the modules of the most downregulated genes (GSE70231_turquoise and GSE108474_turquoise) from GSE70231 and GSE108474, to reveal the key candidate genes. A custom Venn diagram website was used to find the intersecting genes (*FUBP3*, *DAD1*, *CLIC1*, *ABR*, and *DNM1*). *FUBP3* was selected as the significant gene for OS (overall survival) by the GEPIA website. Then immunohistochemistry was used to identify the expression of *FUBP3* in GBM and normal samples. The results showed that the expression of *FUBP3* in the cell nuclei of GBM tissue decreased compared with that of the normal tissue adjacent to the tumor. The survival analyses (OS, DFS, and PFI) showed that when the survival rate was restricted at the same level, the higher the *FUBP3* levels, the longer the GBM patients survived ($P < 0.05$). Then the possible function of *FUBP3* based on GO and KEGG analyses were evaluated. GO analysis showed that *FUBP3* may be mainly involved in chemokine receptor binding, the FC receptor signaling pathway, immunoglobulin complex circulating, immunoglobulin receptor binding, and regulation of leukocyte mediated immunity. KEGG analysis showed that *FUBP3* may be significantly associated with the chemokine signaling pathway, cytokine receptor interaction, leishmania infection, the ribosome, and systemic lupus erythematosus. Tumor microenvironment analysis showed that *FUBP3* expression was positively correlated with the expression of CD8+ T cells, CD4+ T cells, and macrophages ($P < 0.05$). We further verified the expression of *FUBP3* and immune cells by immunohistochemistry.

FBP3 (encoded by *FUBP3*) is a member of a mammalian three gene family of single-strand nucleic acid binding proteins, also including FBP2 (encoded by *KHSRP*) and FBP (encoded by *FUBP1*). The structure of the protein consists of four regular K homologous motifs that can recognize single-stranded RNA or DNA sequences [18]. The MYC transcription factor plays an important role in cell differentiation, growth, and senescence. Single-stranded DNA of the far upstream element (*FUSE*), 1.7 kb upstream of the major P2 promoter of the human *MYC* gene, is bound by FBP (FUSE binding protein) [19–21], and the specific area p89/XPB/ERCC3 30 to 50 helicase/translocase of transcription factor II H (TFIIH) is stimulated [22–24]. As transcription increases, FIR (FUBP interference inhibitor) is recruited through DNA–protein and protein–protein interactions, and then FIR reduces TFIIH activity to a basic level to suppress transcription, maintaining the balance of the levels of MYC [25–28]. *FUBP3* is closely related to many cancers. For instance,

the *FUBP3*–c-Myc axis is activated to promote colorectal cancer progression [29], and *FUBP3* is more frequently expressed in prostate and bladder cancer than in renal cancer. A positive relationship between the expression of *FUBP3* and c-Myc is detectable [30]. Some studies have shown that increased expression of *FUBP1* is a predictor of poor survival in human glioma [31]. Different members of the FUBP family have different functions in different cancers, and *FUBP3* plays different roles in different tumors. Through intersection analysis of the general prognostic factors for *FUBP3*, we found that there was a positive correlation between *FUBP3* and the amount of infiltration of CD4+ T cells, CD8+ T cells, and macrophages in the GBM microenvironment. Therefore, we carried out further verification and found that the expression level of *FUBP3* was lower in GBM samples than in normal tissue adjacent to the tumor, and the higher the expression level of *FUBP3*, the longer the GBM patients survived. The expression of *FUBP3* was positively correlated with the expression of CD8+ T cells, CD4+ T cells, and macrophages in tumor tissue, based on immunohistochemistry. Cancer immunosurveillance has been recognized as a component of the general process of cancer immunoediting, which could be responsible for eliminating tumors [32]. However, the central nervous system (CNS) is thought to be exempt from the effects of the immune system. The brain has physical barriers for protection, and cells in the nervous system respond to inflammation and injury in unique ways [33]. CD8+ T cells constitute an important branch of adaptive immunity, contributing to the clearance of intracellular pathogens and providing long-term protection [34]. CD4+ T cell subsets have the ability to “dedifferentiate” given appropriate environmental signals, allowing individuals to respond to environmental stimuli in a context-dependent manner. A balance of CD4+ T cell subsets is critical to mount responses against pathogen challenges to prevent inappropriate activation, maintain tolerance, and participate in antitumor immune responses [35]. Tumor-associated macrophages (TAMs) represent the most abundant immune cells within the tumor microenvironment and have been associated with adverse outcomes in patients with different types of cancer [36]. There are two main macrophage phenotypes, M1 and M2. Classically activated M1 macrophages promote the antitumor immune response by modulating antigen presentation and secreting proinflammatory cytokines, while activated M2 macrophages play an immunosuppressive role [37]. The TAM phenotypes can be determined by antibodies against TAM-associated biomarkers, such as CD68 (macrophage marker), iNOS (M1 marker), and CD163 (M2 marker). Therefore, due to the blood-brain barrier the central nervous system (CNS) has been recognized as an

immune-free zone [38]. We speculated that FUBP3 was highly expressed in normal brain tissue to maintain the balance of immune cells. Cancer cells with mutated genes can produce mutated proteins that would not normally be produced, and these proteins can be recognized as foreign by the immune system, allowing macrophages to discover cancer cells [39]. In the FUBP3(+) GBM group, the expression could enhance the antigen presentation of CD68+ macrophages through specific molecular pathways (chemokine receptor binding and the FC receptor signaling pathway), stimulating the immune effects of CD4+ and CD8+ T cells to enhance the killing effect on tumor cells. Conversely, the immune monitoring effect is weakened and tumor prognosis is poor for the FUBP3-lacking GBM group.

Collectively, our results suggest that *FUBP3* in GBM is a potential predictor for the malignancy of the tumor. However, this result needs to be studied further in a larger group.

Conclusion

In summary, the purpose of this study was to identify a key gene that may be relevant to the prediction and prognosis of GBM patients. A WGCNA approach with GBM RNA-Seq data was performed to find intersecting genes, and survival analysis was used to determine the significant key gene, *FUBP3*. We speculated that FUBP3 could accelerate the death of glioblastoma cells and increase the survival rate of patients by activating immune cells (CD4+ T cells, CD8+ T cells, and CD68+ macrophages). We found that *FUBP3* is a potential biomarker for the prediction, prognosis, and treatment of GBM. However, the function and specific pathway requires further study.

Abbreviations

GBM: Glioblastoma; WGCNA: Weighted gene co-expression network analysis; TCGA: The Cancer Genome Atlas; GEO: Gene Expression Omnibus; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; OS: Overall survival; DFF: Disease-specific survival; PF: Progression-free interval; TIMER: Tumor Immune Estimation Resource.

Acknowledgements

We appreciate Gene Expression Omnibus and The Cancer Genome Atlas for the open data.

Conflict of interest

All authors declare no conflict of interest in this study.

Authors' contributions

All authors contributed to data analysis, drafting or revising the article, have agreed on the journal to which the article will be submitted, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

Funding

The study is financially supported by National Natural Science Foundation of China (NSFC) (No. 81772637).

Availability of data and materials

RNA sequence data from human glioblastoma samples (GSE70231 and GSE108474) were obtained from GEO datasets (<http://www.ncbi.nlm.nih.gov/geo/>).

Venn diagram website (<http://bioinformatics.psb.ugent.be/webtools/Venn/>).

GEPIA website (<http://gepia.cancer-pku.cn/>).

UCSC Cancer Genomics Browser (<https://genome-cancer.ucsc.edu/>).

GSEA's official website (<https://www.gsea-msigdb.org/gsea/register.jsp>).

TIMER webserver (<https://cistrome.shinyapps.io/timer/>).

Declarations

Ethics approval and consent to participate

The experiments were approved by the Ethics Committee of the Binzhou Medical University Hospital. Written informed consent was obtained from all patients. The experiments were carried out following the Declaration of Helsinki. Tumor surgery follows the principle of maximum safe resection. No animal experiment is applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Neurosurgery, Binzhou Medical University Hospital, Binzhou, Shandong Province, People's Republic of China. ²Department of Pathology, Binzhou Medical University Hospital, Binzhou, Shandong Province, China.

Received: 15 December 2021 Accepted: 16 March 2022

Published online: 12 April 2022

References

- Ohgaki H, Kleihues P. Epidemiology and etiology of gliomas. *Acta Neuropathol.* 2005;109(1):93–108.
- Maher EA, Brennan C, Wen PY, et al. Marked genomic differences characterize primary and secondary glioblastoma subtypes and identify two distinct molecular and clinical secondary glioblastoma entities. *Cancer Res.* 2006;66(23):11502–13.
- Young RM, Jamshidi A, Davis G, Sherman JH. Current trends in the surgical management and treatment of adult glioblastoma. *Ann Transl Med.* 2015;3(9):121.
- Nam JY, de Groot JF. Treatment of glioblastoma. *J Oncol Pract.* 2017;13(10):629–38.
- Gallego O. Nonsurgical treatment of recurrent glioblastoma. *Curr Oncol.* 2015;22(4):e273–81.
- Verhaak RG, Hoadley KA, Purdom E, et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell.* 2010;17(1):98–110.
- Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways [published correction appears in *Nature*. 2013 Feb 28;494(7438):506]. *Nature.* 2008;455(7216):1061–8.
- Das S, McClain CJ, Rai SN. Fifteen years of gene set analysis for high-throughput genomic data: a review of statistical approaches and future challenges. *Entropy (Basel).* 2020;22(4):427 (Published 2020 Apr 10).
- Head SR, Komori HK, LaMere SA, et al. Library construction for next-generation sequencing: overviews and challenges. *Biotechniques.* 2014;56(2):61-passim (Published 2014 Feb 1).
- Das S, Meher PK, Rai A, et al. Statistical approaches for gene selection, hub gene identification and module interaction in gene co-expression network analysis: an application to aluminum stress in soybean (*Glycine max* L.). *PLoS One.* 2017;12(1):e0169605 (Published 2017 Jan 5).
- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* 2008;9:559 (Published 2008 Dec 29).

12. Chen S, Yang D, Lei C, et al. Identification of crucial genes in abdominal aortic aneurysm by WGCNA. *PeerJ*. 2019;7:e7873 (Published 2019 Oct 8).
13. Khatri P, Sirota M, Butte AJ. Ten years of pathway analysis: current approaches and outstanding challenges. *PLoS Comput Biol*. 2012;8(2):e1002375.
14. Chen L, Zhang YH, Wang S, Zhang Y, Huang T, Cai YD. Prediction and analysis of essential genes using the enrichments of gene ontology and KEGG pathways. *PLoS One*. 2017;12(9):e0184129 (Published 2017 Sep 5).
15. Thakkar JP, Dolecek TA, Horbinski C, et al. Epidemiologic and molecular prognostic review of glioblastoma. *Cancer Epidemiol Biomarkers Prev*. 2014;23(10):1985–96.
16. National Comprehensive Cancer Network. (2015). Clinical Practice Guidelines in Oncology: Central nervous system cancers [v.1.2015]. Retrieved from https://www.nccn.org/professionals/physician_gls/pdf/cns.pdf.
17. Wilson TA, Karajannis MA, Harter DH. Glioblastoma multiforme: State of the art and future therapeutics. *Surg Neurol Int*. 2014;5:64 (Published 2014 May 8).
18. Zhou W, Chung YJ, Parrilla Castellar ER, et al. Far upstream element binding protein plays a crucial role in embryonic development, hematopoiesis, and stabilizing Myc expression levels. *Am J Pathol*. 2016;186(3):701–15.
19. Avigan MI, Strober B, Levens D. A far upstream element stimulates c-myc expression in undifferentiated leukemia cells. *J Biol Chem*. 1990;265(30):18538–45.
20. Duncan R, Bazar L, Michelotti G, et al. A sequence-specific, single-strand binding protein activates the far upstream element of c-myc and defines a new DNA-binding motif. *Genes Dev*. 1994;8(4):465–80.
21. Michelotti GA, Michelotti EF, Pullner A, Duncan RC, Eick D, Levens D. Multiple single-stranded cis elements are associated with activated chromatin of the human c-myc gene in vivo. *Mol Cell Biol*. 1996;16(6):2656–69.
22. Liu J, Akoulitchev S, Weber A, et al. Defective interplay of activators and repressors with TFIIH in xeroderma pigmentosum. *Cell*. 2001;104(3):353–63.
23. Liu J, Kouzine F, Nie Z, et al. The FUSE/FBP/FIR/TFIIH system is a molecular machine programming a pulse of c-myc expression. *EMBO J*. 2006;25(10):2119–30.
24. Weber A, Liu J, Collins I, Levens D. TFIIH operates through an expanded proximal promoter to fine-tune c-myc expression. *Mol Cell Biol*. 2005;25(1):147–61.
25. Braddock DT, Louis JM, Baber JL, Levens D, Clore GM. Structure and dynamics of KH domains from FBP bound to single-stranded DNA. *Nature*. 2002;415:1051e1056.
26. Cukier CD, Hollingworth D, Martin SR, Kelly G, Díaz-Moreno I, Ramos A. Molecular basis of FIR-mediated c-myc transcriptional control. *Nat Struct Mol Biol*. 2010;17(9):1058–64.
27. Crichlow GV, Zhou H, Hsiao HH, et al. Dimerization of FIR upon FUSE DNA binding suggests a mechanism of c-myc inhibition. *EMBO J*. 2008;27(1):277–89.
28. Hsiao HH, Nath A, Lin CY, Folta-Stogniew EJ, Rhoades E, Braddock DT. Quantitative characterization of the interactions among c-myc transcriptional regulators FUSE, FBP, and FIR. *Biochemistry*. 2010;49(22):4620–34.
29. Gao Q, Zhou R, Meng Y, et al. Long noncoding RNA CMPK2 promotes colorectal cancer progression by activating the FUBP3-c-Myc axis. *Oncogene*. 2020;39(19):3926–38.
30. Weber A, Kristiansen I, Johannsen M, et al. The FUSE binding proteins FBP1 and FBP3 are potential c-myc regulators in renal, but not in prostate and bladder cancer. *BMC Cancer*. 2008;8:369.
31. Ding Z, Liu X, Liu Y, et al. Expression of far upstream element (FUSE) binding protein 1 in human glioma is correlated with c-Myc and cell proliferation. *Mol Carcinog*. 2015;54(5):405–15.
32. Dunn GP, Bruce AT, Ikeda H, Old LJ, Schreiber RD. Cancer immunoevasion: from immunosurveillance to tumor escape. *Nat Immunol*. 2002;3(11):991–8.
33. Gemma C. Neuroimmunomodulation and aging. *Aging Dis*. 2010;1(3):169–72.
34. Mittrücker HW, Visekruna A, Huber M. Heterogeneity in the differentiation and function of CD8⁺ T cells. *Arch Immunol Ther Exp (Warsz)*. 2014;62(6):449–58.
35. Caza T, Landas S. Functional and Phenotypic Plasticity of CD4(+) T Cell Subsets. *Biomed Res Int*. 2015;2015:521957.
36. Bindea G, Mlecnik B, Tosolini M, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity*. 2013;39(4):782–95.
37. Ostuni R, Kratochvill F, Murray PJ, Natoli G. Macrophages and cancer: from mechanisms to therapeutic implications. *Trends Immunol*. 2015;36(4):229–39.
38. Koelzer VH, Canonica K, Dawson H, et al. Phenotyping of tumor-associated macrophages in colorectal cancer: impact on single cell invasion (tumor budding) and clinicopathological outcome. *Oncoimmunology*. 2015;5(4):e1106677.
39. Chen Z, Hambardzumyan D. Immune Microenvironment in Glioblastoma Subtypes. *Front Immunol*. 2018;9:1004 (Published 2018 May 8).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

