Am J Transl Res 2023;15(2):799-816
www.ajtr.org /ISSN:1943-8141/AJTR0144598

Original Article

Identification of potential biomarkers in cancer testis antigens for glioblastoma

Feng Li[1*], Chang Liu[2,3*], Weixia Nong[1,4], Lina Lin[1], Yingying Ge[1,4], Bin Luo[1,4], Shaowen Xiao[2], Qingmei Zhang[1,4], Xiaoxun Xie[1,4]

[1]Department of Histology and Embryology, School of Basic Medicine Science, Guangxi Medical University, Nanning, Guangxi, China; [2]Department of Neurosurgery, The First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi, China; [3]Postdoctoral Research Station, School of Basic Medicine Science, Guangxi Medical University, Nanning, Guangxi, China; [4]Key Laboratory of Basic Research on Regional Diseases (Guangxi Medical University), Education Department of Guangxi Zhuang Autonomous Region, Nanning, Guangxi, China. [*]Equal contributors.

Received June 7, 2022; Accepted December 16, 2022; Epub February 15, 2023; Published February 28, 2023

Abstract: Objective: To screen and validate cancer testis antigens (CTAs) as potential biomarkers and explore their molecular mechanisms in glioblastoma (GBM). Methods: Ribonucleic acid sequencing (RNA-seq) and bioinformatics analyses were utilized to screen the highly expressed CTAs in GBM. Correlation analysis was used to identify potential biomarkers associated with tumor purity and prognosis. Immunohistochemistry was applied for detection of protein expression. Protein-protein interaction (PPI) network construction, functional enrichment analysis, and binding domain prediction were performed to investigate the underlying molecular mechanisms of GBM. Results: A total of 8 highly expressed CTAs were identified in GBM. One of them was PDZ-binding kinase (PBK). PBK messenger RNA (mRNA) was most highly expressed in GBM and associated with tumor purity and prognosis, PBK protein expression was also significantly increased in GBM tissues and correlated with p53 expression. Functional enrichment analysis revealed that the PBK related genes were predominantly enriched in cell cycle pathway with 38 genes enriched. The proteins encoding by these 38 genes were performed by binding domain prediction analysis, which demonstrated 15 proteins interacting with PBK. Most of these proteins were up regulated in GBM. Conclusion: PBK is highly expressed in GBM. It may serve as a potential biomarker for GBM targeting therapy and the cell cycle modulator by interacting with certain key molecules of cell cycle in GBM.

Keywords: Glioblastoma, cancer testis antigen, PDZ-binding kinase

Introduction

Glioblastoma (GBM), which originates from glial cells, is a malignancy of the central nervous system, accounting for approximately 12-15% of all intracranial tumors [1]. The incidence of GBM increases from 0.15/100,000 in children to 15.03/100,000 in patients aged 75-84 [2]. GBM is currently treated with a combination of surgery and adjuvant therapies such as radiotherapy and temozolomide (TMZ) chemotherapy, which may encount undesired effects [3, 4]. It was reported that the radiotherapy and the TMZ chemotherapy followed by surgery provided a mean survival time of only about 14 to 16 months [5]. Therefore, the most challenging mission is to develop novel therapies, one of which is targeted therapy based on specific tumor antigen in GBM.

Cancer testis antigens (CTAs) are a group of tumor-associated antigens. Their expression spectrum consists of multiple malignant tissues, rather than other normal tissues except for testis tissue [6]. With their unique restricted expression pattern, CTAs may serve as an ideal candidate for tumor immunotherapy because they are more specific and less toxic than chemotherapy agents [7, 8]. Moreover, there are data that demonstrated the role of CTAs in immune escape, tumor invasion, and metastasis [9]. To date, more than 40 CTA families with

multiple members and splicing variants have been identified and included in a database called CTdatabase (http://www.cta.lncc.br) [10, 11], which provides detailed information on these CTAs in a variety of tumors, including gliomas.

In the current study, in order to screen the CTAs highly expressed in GBM, we defined one of the CTAs, PDZ-binding kinase (PBK) via multiple database mining combined with our clinic sample analysis. PBK was found to be highly expressed in GBM and correlated with tumor purity, GBM patient prognosis, and p53. In addition, the functional implication of PBK in GBM may involve the cell cycle through interacting with other molecules in the cell cycle signaling pathway.

Materials and methods

Screening of highly expressed CTA (he-CTA) from databases in GBM

Firstly, the list of CTAs was obtained in the CTdatabase (http://www.cta.lncc.br/). Gene Expression Profiling Interactive Analysis (GEPIA, http://gepia.cancer-pku.cn/index.html), a visualization website based on the Cancer Genome Atlas (TCGA, https://tcga-data.nci.nih.gov/tcga/) and Genotype-Tissue Expression (GTEx, http://commonfund.nih.gov/GTEx/) data, was used to screen he-CTAs in GBM. Volcanic map of differentially expressed CTAs was plotted using R software. The he-CTA was defined as log2 fold change (FC) ≥1 and *p*-value <0.05.

Clinic sample collection and ribonucleic acid sequencing (RNA-Seq) analysis

GBM tissue samples stored at -80°C were collected from the First Affiliated Hospital of Guangxi Medical University. The diagnosis of GBM was confirmed by pathological examinations, and normal brain tissues were collected from autopsies of tumor-free patients. Patients' clinicopathological characteristics and follow-up information were obtained from the electronic records in the hospital. All participants provided written informed consent, and the protocol and informed consent form were approved by the Medical Ethics Committee of Guangxi Medical University. The studies were conducted in accordance with the ethical guidelines of Declaration of Helsinki. Samples were divided into two groups, one for RNA-Seq analysis (GBM tissue samples = 5; Normal brain tissue samples = 5), and the other for immunohistochemical (IHC) analysis (GBM tissue samples = 45). Total RNA was extracted using TRIZOL reagent (Invitrogen, Cat No. 15596018) according to the manufacturer's protocol. Transcriptome RNA-Seq was performed using Illumina high-throughput RNA sequencing by Aksomics Biology Technology Co. Ltd (China). The he-CTA was defined as log2 FC ≥1 and *p*-value <0.05.

Correlation analysis between he-CTA and tumor purity in GBM

The correlation between he-CTA expression and tumor purity in GBM was analyzed via the LinkedOmics (http://www.linkedomics.org/login.php) website. The correlation coefficient R >0.3 and *p*-value <0.05 were selected as thresholds.

Survival analysis

To evaluate the prognostic value of he-CTAs in GBM, Kaplan-Meier survival plots were generated for GBM patients using the University of Alabama at Birmingham Cancer data analysis portal (UALCAN, http://ualcan.path.uab.edu/). The expression level of genes was normalized as transcripts per million (TPM). Samples were classified into high expression (TPM above upper quartile) and low expression (TPM equal to or below upper quartile). Genes with *p*-values <0.05 were considered prognostic.

Protein expression analysis

First, our clinical samples our clinical samples underwent immunohistochemistry (IHC) assays as previously reported [12]. Antigen retrieval was performed in EDTA (pH 8.0). After inactivation of endogenous peroxidase incubated with 3% H_2O_2. The sections were then incubated overnight at a temperature of 4°C with rabbit monoclonal antibody anti-PBK (Abcam, Cat No. ab75987), diluted at 1:75 as directed, using the rabbit's pre-immune serum as a negative control.

After biotinylated secondary antibody (ORIGENE, Cat No. PV-600), the slides were incubated with 3,3'-diaminobenzidine (DAB) and counterstained with Hematoxylin. For IHC data

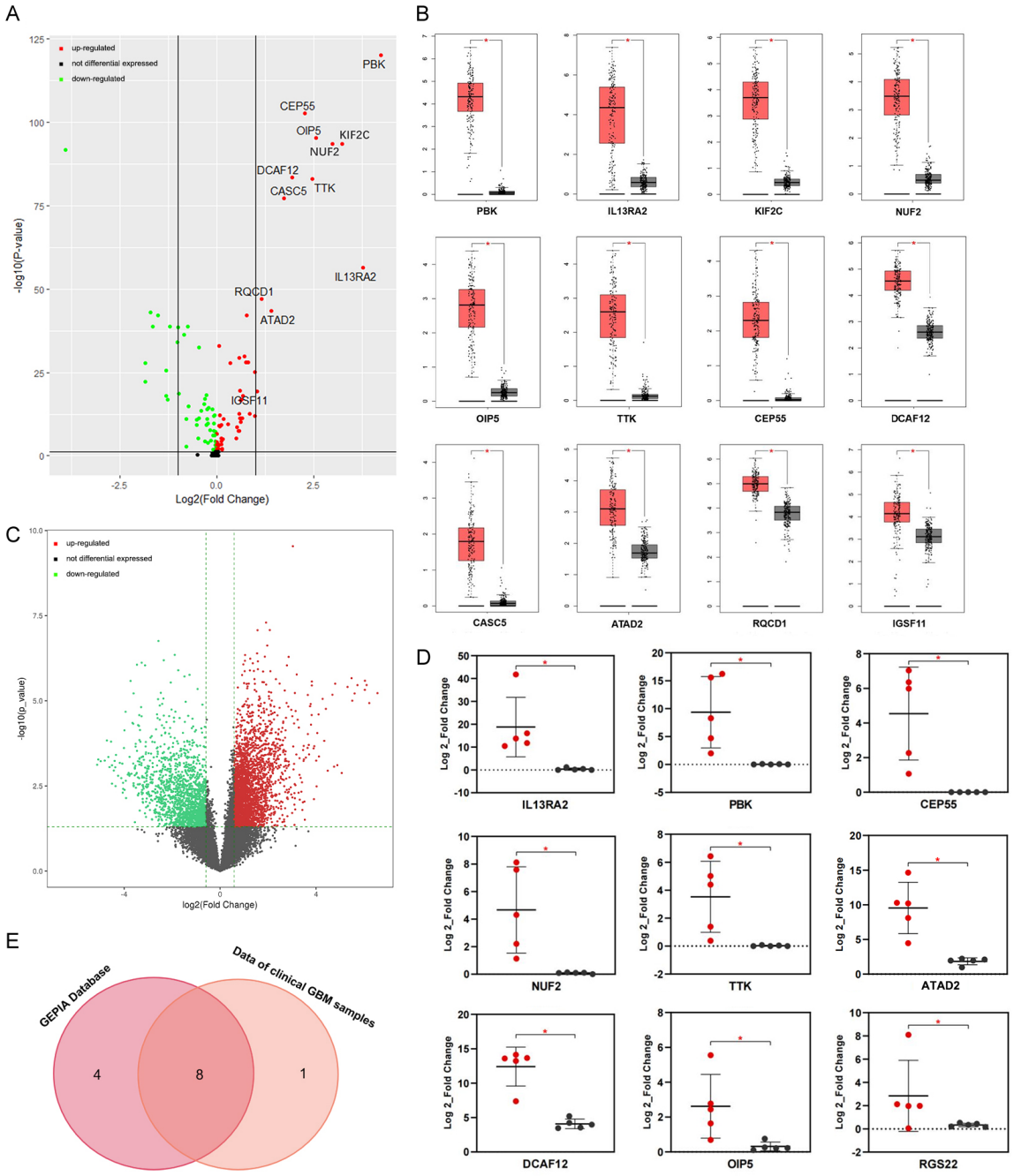


Figure 1. Identification of he-CTAs in GBM based on GEPIA and clinical samples. A. Volcano plot visualized differential expressed CTAs between GBM and normal brain in GEPIA. Information of 162 GBM and 207 normal brain samples came from the TCGA and GTEx databases. B. Expression of 12 he-CTAs mRNA in 163 GBM (red box) and 207 normal brain tissue (black box) in the TCGA and GXTs database. C. Volcano plot visualized differential expressed genes with the comparison of GBM and normal brain tissue samples analyzed by RNA-seq. D. He-CTA mRNA expression of 5 GBM (red box) and 5 normal brain tissue (black box) analyzed by RNA-seq, RGS22: Regulator of G Protein Signaling 22. E. Venn diagrams presented the number of he-CTAs identified with different analyses. *: *p*-value <0.05.

analysis, the staining intensity was quantified using the following scores: 0, no staining; 1, weak staining; 2, moderate staining; and 3, strong staining. The percentage of positive

Table 1. He-CTAs in GBM based on a combined analysis of the GEPIA database and clinical samples

Gene symbol (No. CTA)*	Chromosome localization	Expression (Fold Change)	*P*-value
IL13RA2 (CT19)	Xq13.1-q28	3.752	1.44E-05
PBK (CT84)	8p21.2	3.0497	2.18E-04
CEP55 (CT111)	10q23.33	2.289	3.59E-04
NUF2 (CT106)	1q23.3	2.176	6.88E-04
TTK (CT96)	6q13-q21	1.893	3.31E-03
ATAD2 (CT137)	8q24.13	1.830	1.86E-04
DCAF12 (CT102)	9p13.3	1.377	6.87E-05
OIP5 (CT86)	15q15.1	1.339	4.36E-03

Official Full Name: IL13RA2: Interleukin 13 Receptor subunit Alpha 2; PBK: PDZ Binding Kinase; CEP55: Centrosomal Protein 55; NUF2: NUF2 component of NDC80 kinetochore complex; TTK: TTK protein kinase; ATAD2: ATPase family AAA domain containing 2; DCAF12: DDB1 and CUL4 associated factor 12; OIP5: Opa Interacting Protein 5. *: CT numbers are based on the list in CTdatabase.

GBM cells was defined as follows: 0, 0-5%; 1, 6-25%; 2, 26-50%; 3, 51-75%; 4, 76-100%. According to the sum scores, 0 point, negative (overall score = '-'); 1-3 point, weak expression (overall score = '+'); 4-5 points, moderate expression (overall score = '++'); 6-7 points, strong expression (overall score = '+++'). Assessment of positive immunoreactivity was achieved by two independent pathologists using a blind method. At least five independent foci of neoplastic infiltration in each tissue specimen were observed using an optical microscope (Leica Corporation, Germany) at a magnification, of × 400.

Then, two online tools, UALCAN and the Human Protein Atlas (HPA, www.proteinatlas.org) were used to verify the protein expression. UALCAN provided a protein expression analysis by using data from the Clinical Proteomic Tumor Analysis Consortium (CPTAC) dataset. HPA, on the other hand, can download images of immunohistochemically stained tissue according to the usage policy.

Protein-protein interaction (PPI) network construction

Setting Pearson's correlation coefficient >0.7 and *p*-value <0.05, we further identified the gene co-expressed with the he-CTA that related to tumor purity and survival of GBM patients by using the LinkedOmics website (http://www.linkedomics.org/). Their interactions were then predicted using String11.0 (http://string.embl.de/) database with 0.9 as the threshold of a minimum required interaction score. Cytoscape (version 3.6.1, http://www.cytoscape.org/) software was applied to visualize the network.

Functional enrichment analysis

Gene Ontology (GO) analysis was undertaken from molecular function (MF), biological process (BP), and cellular component (CC). KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway enrichment analysis of the hub genes involved in the PPI network was performed to reveal enriched biological functions. Database for Annotation, Visualization and Integrated Discovery (DAVID, https://davidd.ncifcrf.gov/) online tools was used to conducted the functional analysis. Then a bubble chart was draw by the Ehbio-information online map (http://www.ehbio.com/ImageGP/index.php/Home/Index/GOenrichmentplot.html) to present the analysis results.

Domain-domain interaction (DDI) construction

In order to explore the potential mechanism of genes in GBM, motif analysis of proteins that may interact with each other, was performed with Sequence Similarity DataBase (SSDB) motif search in KEGG online database (http://www.kegg.jp/). Then, the database of 3D interacting domains (3DID, https://3did.irbbarcelona.org/) to predict the protein domain interactions.

Statistical analysis

Statistical analyses and graphical representation were carried out using Statistical Program for Social Sciences (SPSS) version 22.0 software and GraphPad Prism 6.0 for Windows. The correlation between PBK expression and clinicopathological parameters was analyzed using Chi-square test or Fisher's exact probability test. Data are expressed as the mean ± standard deviation. We considered *p*-value <0.05 as statistically significant.

Result

Eight he-CTAs were identified in GBM

Based on the mRNA expression data of 245 CTAs in GEPIA, a volcano plot was constructed to visualize their expression distributions in

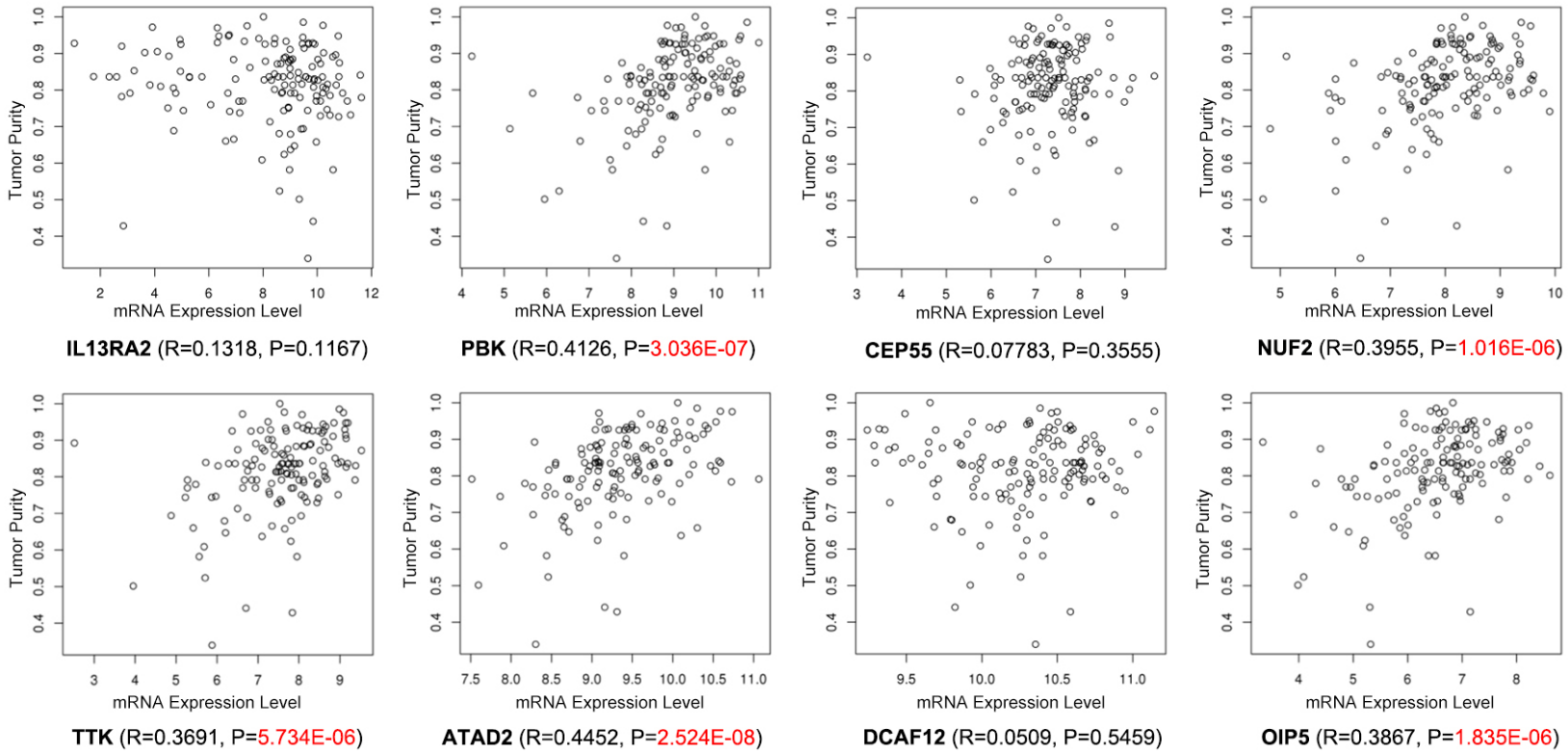


Figure 2. The correlation between he-CTAs mRNA expression and tumor purity in GBM. Data of 151 samples obtained in the LinkedOmics database, Spearman-Statistic, R-value represents correlation coefficient.

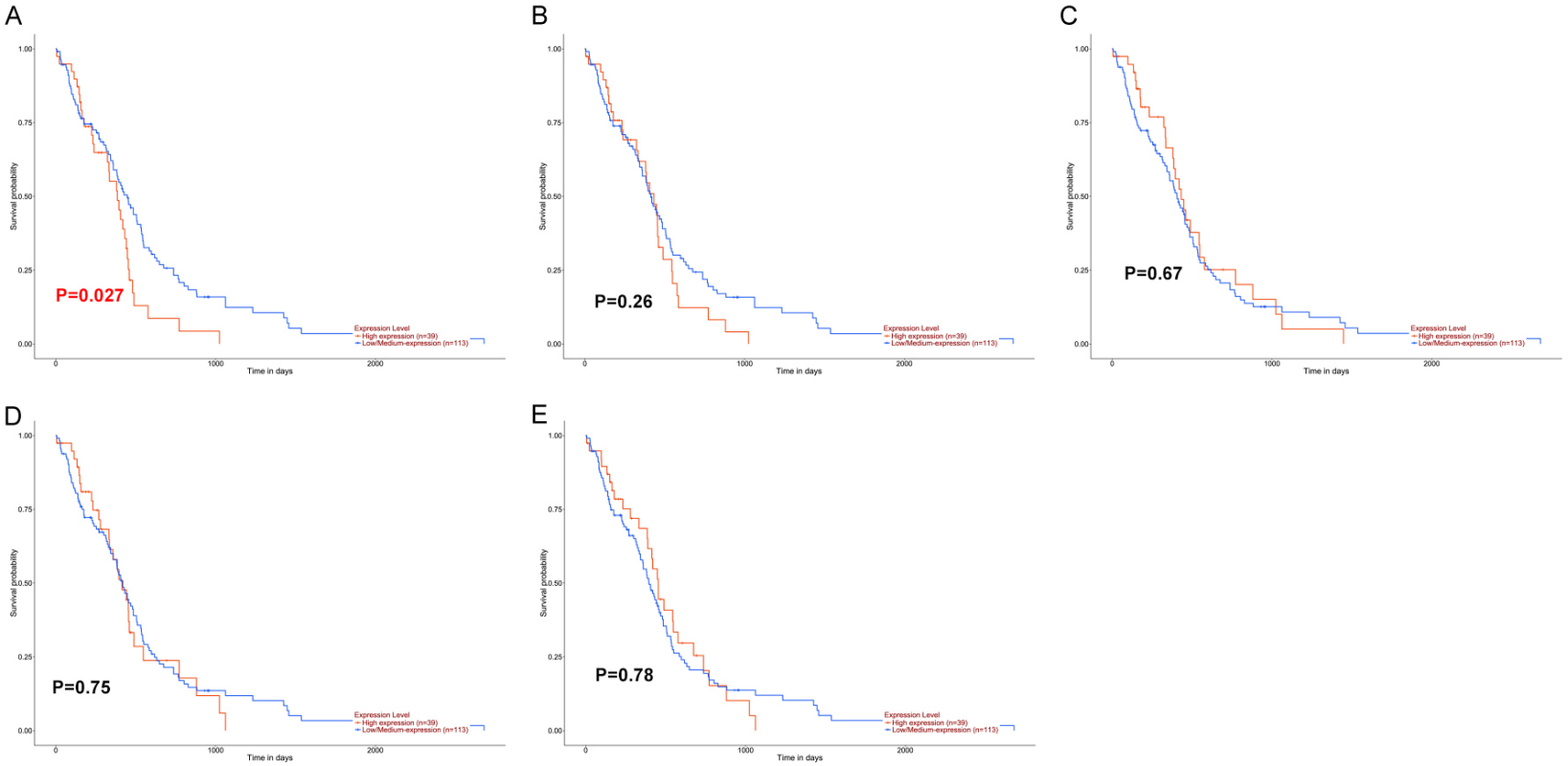


Figure 3. Prognostic value of CTAs expression in patients with GBM. High expression: TPM values above upper quartile; Low/Medium expression: TPM values below upper quartile.

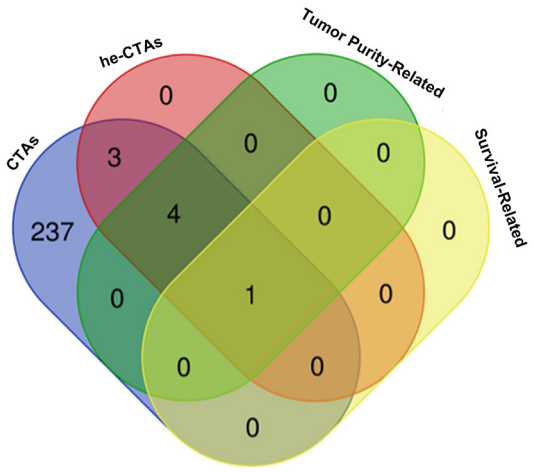


Figure 4. Summary of identified CTA involved in the high expression, tumor purity, and prognosis in GBM. Different colors of the Venn diagram represented different analysis results, Blue: 245 CTAs in the CT Database; yellow: 12 CTAs significantly up-regulated in GBM; red: 7 CTAs significantly positively correlated with tumor purity of GBM; green: 1 CTA related to GBM patient prognosis. The intersection set CTA obtained from the four analyses was PBK.

GBM (**Figure 1A**). A total of twelve he-CTAs [PBK, interleukin 13 receptor subunit alpha 2 (IL13RA2), kinesin family member 2C (KIF2C), NUF2 component of NDC80 kinetochore complex (NUF2), Opa interacting protein 5 (OIP5), TTK protein kinase (TTK), centrosomal protein 55 (CEP55), DDB1 and CUL4 associated factor 12 (DCAF12), cancer susceptibility candidate 5 (CASC5), ATPase family AAA domain containing 2 (ATAD2), RCD1 required for cell differentiation 1 homolog (RQCD1) and immunoglobulin superfamily member 11 (IGSF11)] were identified, among which PBK was the most highly expressed one (**Figure 1B**).

Then, we examined our RNA-seq result. **Figure 1C** demonstrated the expression of genes in which 2771 genes were up-regulated and 1536 genes were down-regulated. Among those up-regulated genes, we screened out a total of nine he-CTAs which were highly expressed in 5 cases of GBM compared to 5 cases of normal brain tissues (**Figure 1D**). Taken the intersection of the results of GEPIA database mining and our RNA-seq (**Figure 1E**; **Table 1**), eight he-CTAs (IL13RA2, PBK, CEP55, NUF2, TTK, ATAD2, DCAF12, and OIP5) were finally selected for further investigation.

Five he-CTAs were correlated with tumor purity in GBM

Tumor purity is a confounding factor that might influence gene expression changes. When selecting he-CTAs as molecular targets, not only their expression level should be taken into account, but also the tumor purity. Therefore, we further analyzed the association between the expression of eight he-CTAs identified above and the tumor purity of GBM. The result showed that the expression of he-CTAs in 5 of 8 (PBK, ATAD2, NUF2, OIP5, and TTK) was positively correlated with the tumor purity, suggesting they may be more specifically expressed in GBM cells (**Figure 2**).

PBK possessed prognostic value in GBM

To understand whether the identified he-CTAs correlated with tumor purity were also related to the prognosis of GBM patients, survival data of 152 GBM patients were applied to perform a Kaplan-Meier survival analysis. The result showed that only 1 of 5 he-CTAs, called PBK, was associated with the overall survival of GBM patients. The patients with PBK high expression had a shorter survival period than those with poorly expressed PBK. In contrast, the expression of NUF2, TTK, ATAD2, and OIP5 did not affect the patients' survival (**Figure 3**). Taken together, PBK is a he-CTA in relation to tumor purity and survival of GBM patients (**Figure 4**).

PBK protein was up-regulated and correlated with p53 in GBM

Considering tumor-associated protein may serve as the candidate for the targeted therapy of tumors, we tested PBK protein by IHC in our clinic samples, which showed that the PBK protein was mainly located in the cytoplasm, membrane, and nucleus (**Figure 5A**). According to the staining intensity and percentage of positive cells, patients were classified into two groups (PBK high or PBK low) based on the IHC score. There were 14 (31.1%) patients with PBK low expression and 31 (68.9%) patients with PBK high expression. Then, we further analyzed the correlation between PBK protein expression and clinicopathological parameters of GBM patients. Chi-square analysis showed that only the p53 protein expression was asso-

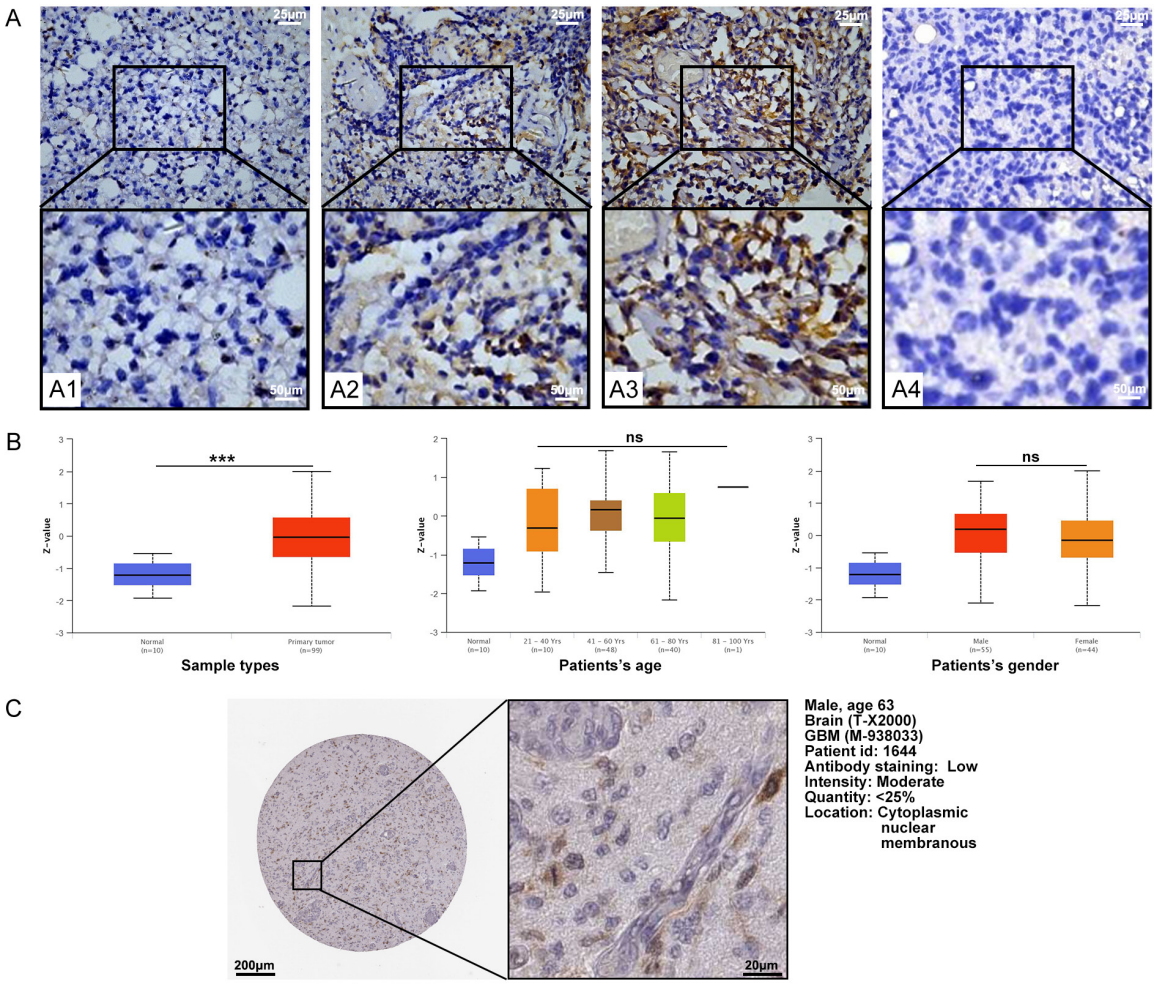


Figure 5. PBK protein expression in GBM. A. Representative IHC staining of PBK protein in GBM tissues. A1, weak expression; A2, moderate expression; A3, strong expression; A4, GBM tissue probed with pre-immune serum as the negative control. B. Expression of PBK in normal brain tissue and GBM based on sample type, patients' age, and patients' gender in UALCAN. Statistical significance is indicated in the graph plots. ***: *p*-value <0.001. C. IHC staining of PBK in GBM based on the HPA database.

ciated with PBK protein expression (**Table 2**). Moreover, to get more information about PBK protein expression in GBM, we downloaded PBK protein expression data pertaining to GBM from the UALCAN and HPA. As shown in **Figure 5B**, the data of UALCAN revealed a significant increase of PBK protein in GBM when compared to normal brain tissues (*p*-value <0.05), and there was no significant correlation between PBK protein expression and patients' age or gender. In the HPA, 12 GBM tissue samples were tested by PBK antibody, and two of them exhibited positive staining (**Figure 5C**).

PBK-related genes mainly functioned in the cell cycle

To explore the underlying molecular mechanism of PBK in GBM, PBK-related genes were screened out using LinkedOmics. A total of 333 genes were positively correlated with PBK in GBM (**Figure 6A**). Then, we mapped those genes to the STRING11.0 website and constructed a PPI network. The established network consisted of 256 nodes and 3373 edges which were visualized by Cytoscape software (**Figure 6B**).

Subsequently, GO enrichment and KEGG pathway analysis were performed using the 256 genes in the PPI network. The top 10 clusters of significantly-enriched terms of BP, CC, MF, and KEGG pathway were shown as bubble charts (**Figure 6C-E**). Of the BP terms, mitotic sister chromatid separation was the top enriched term. The most enriched term for CC was chromosome separation, and for MF was microtu-

Table 2. Correlation between PBK protein expression and clinical parameters in GBM patients

Parameters	PBK Expression n (%)		χ^2	*P*-value
	High	Low		
Gender			1.07	0.30
Male	24 (77.4)	8 (57.1)		
Female	7 (22.6)	6 (42.9)		
Age (years)			1.66	0.20
<40	6 (19.4)	6 (42.9)		
≥40	25 (80.6)	8 (57.1)		
KPS score			0.091	0.76
<70	14 (45.2)	7 (50.0)		
≥70	17 (54.8)	7 (50.0)		
p53 Protein			5.66	0.04*
-	5 (19.2)	4 (40.0)		
+	21 (80.8)	6 (60.0)		
MGMT Protein			0.36	0.69
-	10 (50.0)	5 (62.5)		
+	10 (50.0)	3 (37.5)		
GFAP Protein			0.65	0.46
-	1 (4.3)	1 (12.5)		
+	22 (95.7)	7 (87.5)		
Ki-67 (%)			0.10	0.75
<10	12 (38.7)	4 (71.4)		
≥10	19 (61.3)	10 (28.6)		
Tumor Size (cm)			0.55	0.46
<5	17 (54.8)	6 (42.9)		
≥5	14 (45.2)	8 (57.1)		

High PBK protein expression (++/+++); low PBK protein expression (+); statistically significant (P<0.05). KPS: Karnofsky Performance Scale; MGMT: O-6-methylguanine-DNA methyltransferase; GFAP: Glial Fibrillary Acidic Protein. *: *p*-value <0.05.

bule movement. The dominant category for KEGG was related to the cell cycle in which there were 38 genes enriched. And the pathway map was constructed based on the cell cycle pathway map in KEGG (**Figures 6F**, S1). Collectively, these results implied that the main function of PBK-related genes may involve in cell division and cell cycle.

Cell cycle molecules might interact with PBK and up-regulate in GBM

As there were 38 genes enriched in the cell cycle pathway presented above, we next focused on it. By searching the dataset of hsa: 55872 in the KEGG online database, Pkinase motifs in the amino acid sequence of PBK were demonstrated in **Figure 7A**. There was a total of 15 identified proteins that might interact with PBK in the cell cycle signal pathway, and their three-dimensional structures and gene information were shown in **Figure 7B-F** and **Table 3**, respectively.

Then, we further explore the expression of the 15 proteins aforementioned with available public databases. The mRNA expression of these genes except for CDC7 was highly expressed in GBM compared with normal brain tissues (Figure S2). Regarding the protein expression, the data from UALCAN shows 9 [BUB1 mitotic checkpoint serine/threonine kinase (BUB1), BUB1 mitotic checkpoint serine/threonine kinase B (BUB1B), cyclin dependent kinase 1 (CDK1), cyclin dependent kinase 2 (CDK2), checkpoint kinase 1 (CHEK1), checkpoint kinase 2 (CHEK2); cyclin A2 (CCNA2), cyclin B1 (CCNB1), and cyclin B2 (CCNB2)] of 15 cell cycle molecules were elevated in GBM with statistical significance in contrast with normal brain tissues (**Figure 8**). Moreover, IHC staining images from the HPA demonstrated 10 [cell division cycle 7 (CDC7), CDK1, CDK2, CHEK2, polo like kinase 1 (PLK1), TTK, CCNA2, CCNB1, CCNB2, and cyclin E2 (CCNE2)] of 15 cell cycle molecules with various extents in GBM tissues (**Figure 9**). The Number of samples and subcellular localization of these 10 proteins and PBK were shown in **Table 4**.

Discussion

For the past decades, therapeutic progress for GBM has been limited, and the prognosis for GBM patients remains unsatisfactory. In recent years, targeted therapies and immunotherapies are emerging to shed some light on tumor treatment. CTAs appear to be important cancer therapeutic targets based on their tumor restricted expression pattern, immunogenicity, and putative role in oncogenesis. Vaccines targeting several CTAs, such as MAGEA1, MAGEA3, and NY-ESO-1, have reached the clinical trial stage and achieved a certain curative effect in multiple extracranial tumors [8, 10, 13, 14]. While there are still many limitations and difficulties for application of CTAs in GBM treatment. This is mainly due to the unclear expression profile of most CTA in GBM.

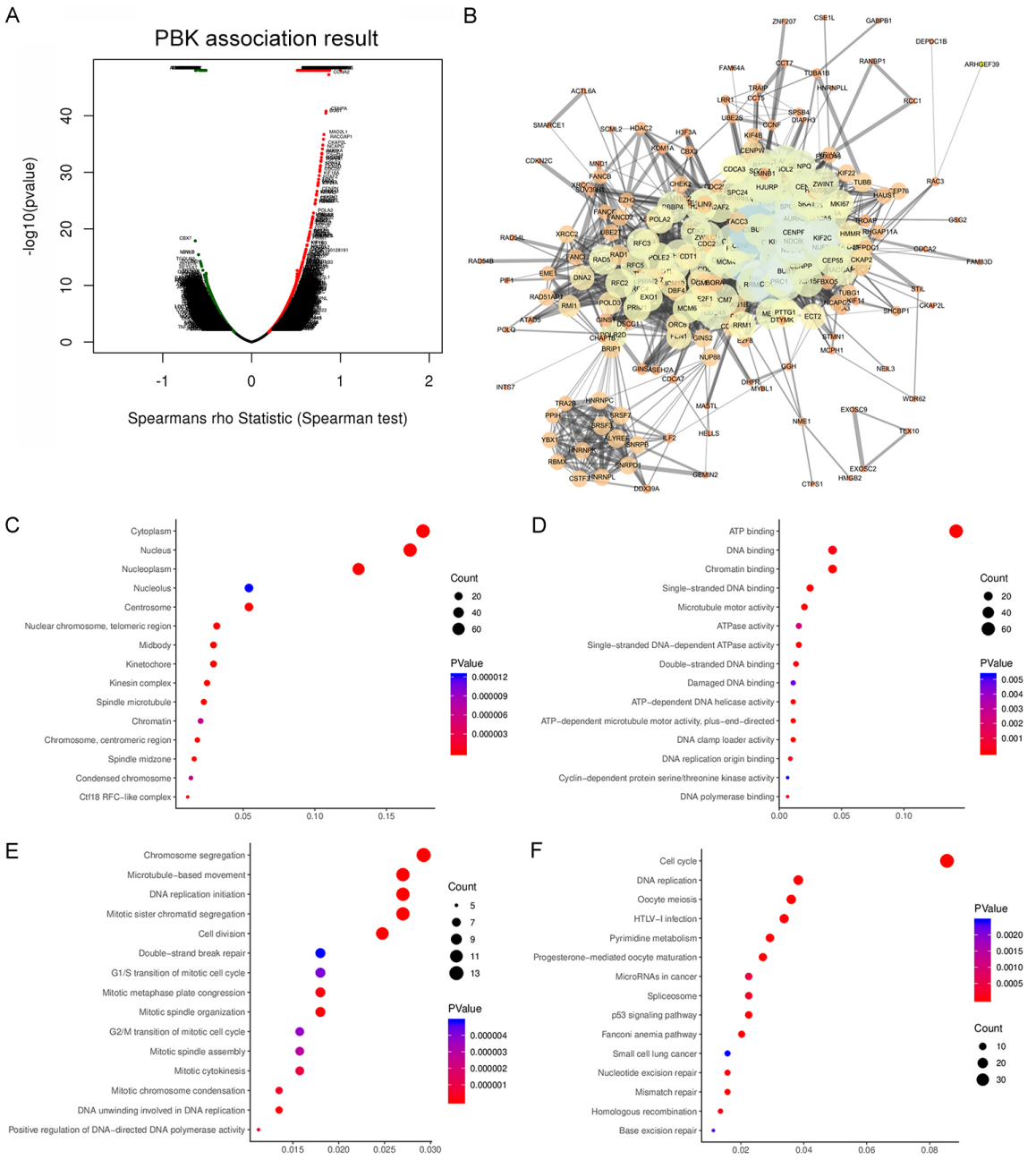


Figure 6. Enrichment analysis of genes positively related to PBK. A. Volcano map shows the genes that positively (green dots) and negatively (red dots) correlated with PBK. The horizontal axis represents spearman's correlation coefficient, and the vertical axis represents the *p*-value (-Log10); Image downloaded from the LinkedOmics website. B. PPI network of genes positively related with PBK. Nodes with a higher degree are displayed in the larger size and bright orange color and nodes with a lower degree are shown in smaller size and dark blue color. The minimum required interaction score is set to 0.9. C-F. Bubble chart of GO and KEGG pathways enrichment analysis. Top 15 significant Cellular components (C), Molecular function (D), Biological process (E), and KEGG pathway (F) terms of key genes involved in the PPI network.

In this study, using bioinformatics analysis and clinical tissue sample detection, we identified a total of 8 CTAs highly expressed in GBM. It is well known that tumor microenvironment plays an important role in determining tumor growth, disease progression, and drug resistance [15], so does GBM microenvironment, which is rather complex containing diverse types of cells.

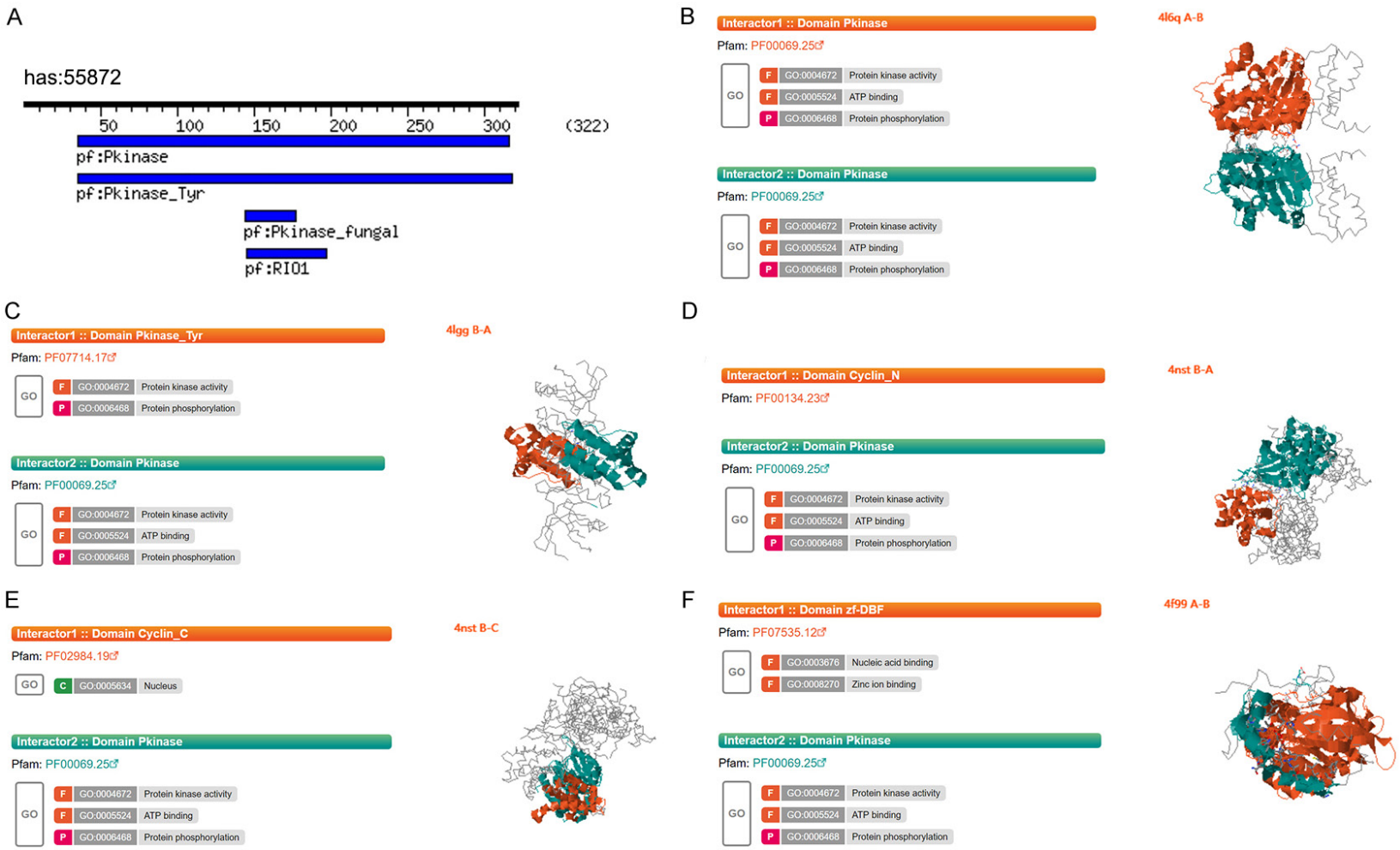


Figure 7. Identification of domain interactions in PBK and cell cycle key molecules. A. The motifs contained in PBK protein. Blue bars represent the location of each motif. B. Structure display of domain Pkinase and domain Pkinase interaction. C. Structure display of domain Pkinase_Try and domain Pkinase interaction. D. Structure display of domain Cyclin_N and domain Pkinase interaction. E. Structure display of domain Cyclin_C and domain Pkinase interaction. F. Structure display of domain zf_DBF and domain Pkinase interaction.

Table 3. Cell cycle proteins interacted with PBK in GBM

Protein symbol	NCBI-Protein ID	Interacted motif	Amino acid Location
BUB1	NP_004327	Pkinase	790-1004
		Pkinase_Tyr	790-926
BUB1B	NP_001202	Pkinase	823-920
		Pkinase_Tyr	825-917
CDC7	NP_001308044	Pkinase	59-207
		Pkinase_Tyr	59-205
		Pkinase	365-465
		Pkinase_Tyr	521-569
CDK1	NP_001307847	Pkinase	4-287
		Pkinase_Tyr	7-200
CDK2	NP_001789	Pkinase	4-286
		Pkinase_Tyr	6-199
CHEK1	NP_001107593	Pkinase	11-264
		Pkinase_Tyr	11-260
CHEK2	NP_009125	Pkinase	22-486
		Pkinase_Tyr	22-478
PKMYT1	NP_004194	Pkinase	110-358
		Pkinase_Tyr	113-353
PLK1	NP_005021	Pkinase	53-305
		Pkinase_Tyr	57-301
TTK	NP_003309	Pkinase	525-791
		Pkinase_Tyr	527-786
CCNA2	NP_001228	Cyclin_N	181-306
		Cyclin_C	309-426
CCNB1	NP_114172	Cyclin_N	173-297
		Cyclin_C	300-417
CCNB2	NP_004692	Cyclin_N	137-261
		Cyclin_C	264-381
CCNE2	NP_477097	Cyclin_N	112-239
		Cyclin_C	241-362
DBF4	NP_006707	zf-DBF	291-334

Official Full Name: BUB1: BUB1 mitotic checkpoint serine/threonine kinase; BUB1B: BUB1 mitotic checkpoint serine/threonine kinase B; CDC7: Cell Division Cycle 7; CDK1: Cyclin Dependent Kinase 1; CDK2: Cyclin Dependent Kinase 2; CHEK1: Checkpoint Kinase 1; CHEK2: Checkpoint Kinase 2; PKMYT1: Protein Kinase Membrane Associated Tyrosine/Threonine 1; PLK1: Polo Like Kinase 1; TTK: TTK protein kinase; CCNA2: Cyclin A2; CCNB1: Cyclin B1; CCNB2: Cyclin B2; CCNE2: Cyclin E2; DBF4: DBF4 zinc finger.

Tumor purity refers to the proportion of tumor cells in tissue and is closely related to clinical or genomic characteristics and biological significance. Ignoring tumor purity significantly affects transcriptome analyzes and molecular means [16]. Therefore, we analyzed the correlation between he-CTA mRNA expression and tumor purity in GBM, and correlation coefficients greater than 0.3 were selected for further survival analysis. Thus, we proposed that PBK may serve as a prognostic biomarker and promising molecular target for GBM.

PBK was initially amplified from the complementary deoxyribonucleic acid (cDNA) library of HeLa cells and linked with the PDZ2 domain of tumor suppressor protein the human homologue of the Drosophila Discs-large (hDlg) [17]. Abe et al. discovered a novel protein kinase in lymphokine-activated killer T (T-LAK) cell, which was found to be the same molecule as PBK by sequence analyses [18]. The PBK gene is located on chromosome 8p21.2, contains 8 exons with 1811 base long encoding a 322 amino acid serine-threonine kinase, which promotes tumor cell growth and participates in the DNA damage repair response mainly through activating mitogen-activated protein kinase (MAPK), extracellular signal-regulated kinase (ERK), c-Jun N-terminal protein kinase 1 (JNK1) and other signaling pathways [19].

Several previous reports have demonstrated that PBK expresses in various malignancies including ovarian, prostate, bladder cancer and lung cancer [20-22], but PBK expression is less reported in GBM. Joel et al. firstly reported that PBK protein is highly upregulated in glioma initiating cells [23]. Then, Hayashi et al. demonstrated that PBK protein is highly expressed in 32 cases of GBM [24]. Quan et al. found that PBK protein was also highly expressed in 29 cases of grade IV glioma tissues [25]. Given that PBK is highly expressed in GBM, our IHC results showed that PBK expression in GBM was significantly correlated with p53. However, others failed to observe a significant association between p53 and PBK protein expression in gliomas [25]. The discrepancy may be due to the different samples (e.g race, gene background etc), and also the number of sample test. There was a study in colorectal cancer cell, showing that

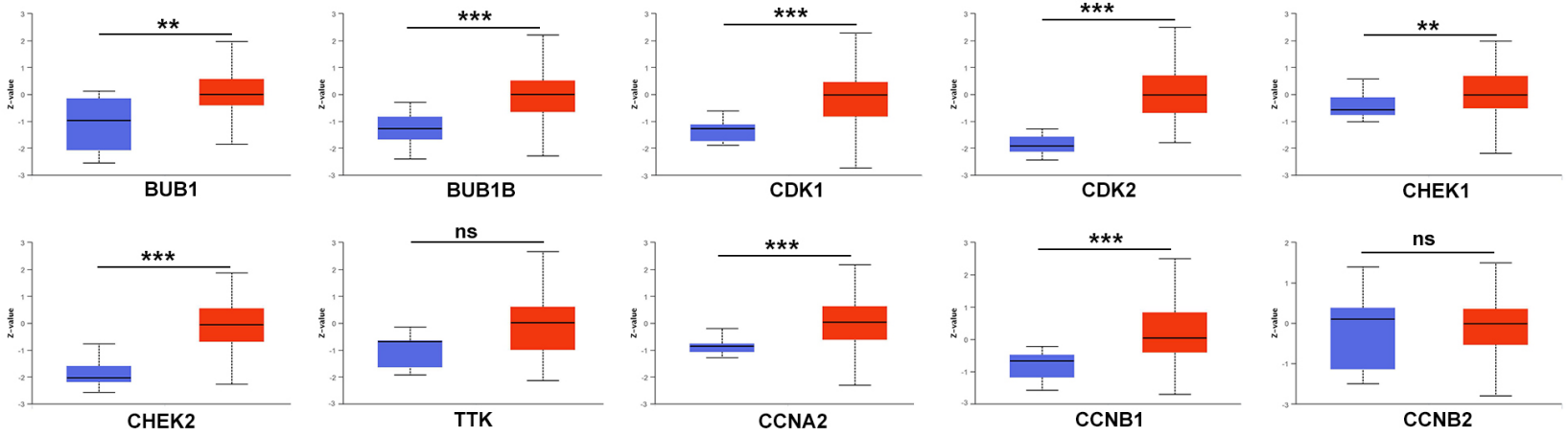


Figure 8. Protein expression of cell cycle molecules interacted with PBK in GBM based on the data of UALCAN. Blue box: 10 Normal brain samples; Red box: 99 GBM samples. ***: *p*-value <0.001, **: *p*-value <0.01, ns: no significance.

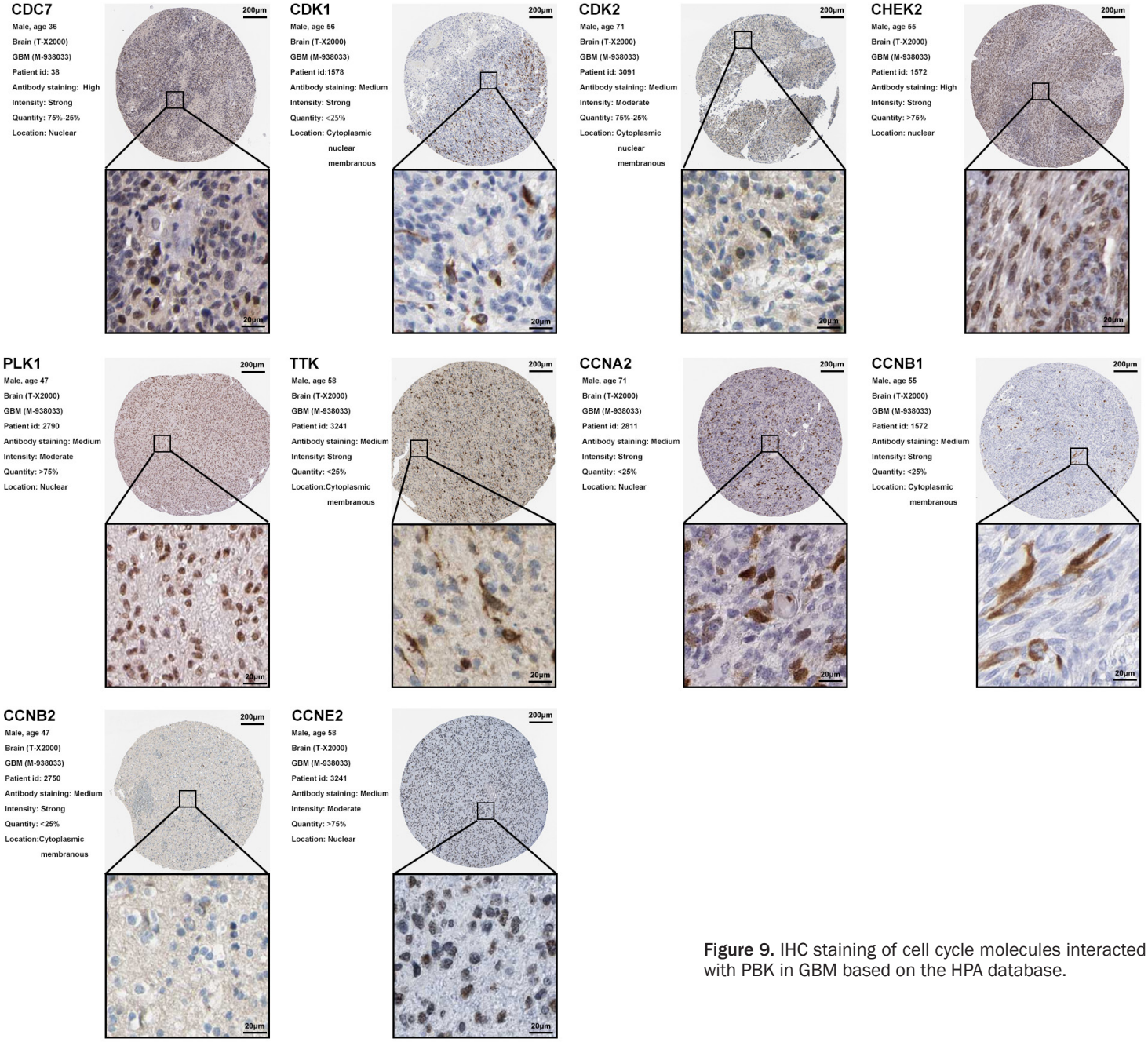


Figure 9. IHC staining of cell cycle molecules interacted with PBK in GBM based on the HPA database.

Table 4. Protein expression of PBK and its related cell cycle genes in GBM from HPA

Protein Symbol (Antibody No.)	ND*	Staining status (number)			Protein location		
		Low	Medium	High	Cytoplasmic	membranous	nuclear
PBK (HPA050656)	10	2	-	-	√	√	√
CDC7 (CAB002669)	-	3	4	5	-	-	√
CDK1 (CAB003799)	1	1	10	-	√	√	√
CDK2 (CAB013115)	5	5	1	-	√	√	√
CHEK2 (HPA001878)	-	3	6	2	-	-	√
PLK1 (HPA053229)	-	1	10	-	-	-	√
TTK (CAB013229)	1	2	8	-	√	√	-
CCNA2 (CAB000114)	5	2	5	-	-	-	√
CCNB1 (CAB003804)	4	2	5	-	√	√	-
CCNB2 (HPA008873)	5	4	2	-	√	√	-
CCNE2 (CAB019374)	-	3	7	2	-	-	√

*; ND, Not Detected.

PBK bound p53 DNA-binding domain to inhibit its transactivation activity, thus influence cell cycle [26]. Whether PBK has such similar regulatory mechanism on p53 in GBM, it merits further investigation.

Our functional enrichment analysis revealed that PBK is mainly involved in biological processes such as chromosome segregation, mitotic sister chromatid separation, and DNA replication initiation in GBM, and is mainly enriched in the cell cycle. Furthermore, PBK phosphorylates p38 MAPK and histone H3, and the phosphotransferase activity of PBK is regulated in a cell cycle-dependent manner [27, 28]. Joel et al. found that knockdown of PBK in vitro inhibited the growth of glioma-initiating cells and the formation of tumor spheres with significant apoptotic cells. All of this suggests that PBK plays an important role in regulating GBM cell cycle.

We predicted possible interactions between PBK and 15 key cell cycle molecules by motif analysis. Therefore, we hypothesized that PBK may influence cell cycle progression through these interactions, ultimately leading to GBM's malignant biological behavior. The association of CCNB1, CCNB2, CDK1 and CHEK1 with PBK has been clearly reported. Previous studies have shown that PBK mRNA expression is associated with cell cycle activation. PBK is upregulated during mitosis and PBK binding and phosphorylation by CDK1/cyclin B1 are required for its mitotic activity. Moreover, PBK makes a kinase-substrate complex with CDK1/Cyclin B1 and protein regulator of cytokinesis 1 on microtubules during mitosis and promotes cytokinesis [18, 29]. Furthermore, PBK is essential for the activation of CCNB2 promoter and promotes tumorigenesis and radio-resistance in GBM [30]. It was also shown that PBK promoted checkpoint signaling in response to replication stress and DNA damage by directly interacting with CHEK1 and cell division cycle 25C [31]. In addition to the above-reported genes, BUB1, BUB1B, CDC7, CCNA2, PLK1, and TTK were identified as genetic biomarkers for several tumors along with PBK [32-34]. In contrast, CCNE2, CDK2, CHEK2, DBF4 zinc finger (DBF4) and protein kinase membrane associated tyrosine/threonine 1 (PKMYT1) have not been previously reported to be associated with PBK. The present study provides new insights into the mechanisms by which PBK supports GBM development, but further validation and further validation is still needed.

Notably, we found for the first time that the cancer testis antigen TTK also interacts with PBK. TTK, also known as CT96 is a dual specificity protein kinase essential for centrosome duplication and mitotic progression [35]. Some TTK inhibitors, such as Mps1-IN-3 and NMS-P715, have been developed to improve survival and prognosis in GBM patients [36, 37]. Our findings indicated that TTK was highly expressed in GBM and correlated with tumor purity. Although there is no correlation between TTK expression and prognosis in GBM patients, it may theoretically serve as potential target for GBM therapy. In addition, we found 5 of 8

CTAs distinctly elevated in GBM associating with the tumor purity. These CTAs may have value in development of multiple targets therapy for GBM considering the heterogeneous expression of CTAs in tumor cells. However, all of them need to be further investigated.

In conclusion, the present study screened CTAs in GBM and found that PBK was significantly highly expressed in GBM tissues correlating with tumor purity and patient prognosis. PBK may interact with key molecules of cell cycle signaling pathway, potentially leading to malignant biological behavior in GBM. These results suggest that PBK could be a valuable prognostic marker and promising therapeutic target for GBM.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 81960453, No. 81860445), Natural Science Foundation of Guangxi Province (2022-GXNSFAA035639), Key Laboratory of Early Prevention and Treatment for Regional High Frequency Tumor (Guangxi Medical University) and Ministry of Education (No. GK2018-09, No. GKE 2019-08, No. GKE-ZZ202006). This study was approved by the Medical Ethics Committee of the First Affiliated Hospital of Guangxi Medical University (approval #20200040).

Disclosure of conflict of interest

None.

Address correspondence to: Qingmei Zhang and Xiaoxun Xie, Department of Histology and Embryology, School of Basic Medicine Science, Guangxi Medical University, Nanning, Guangxi, China. Fax: +86-0771-5358577; E-mail: 940478-522@qq.com (QMZ); xxiaoxun@163.com (XXX)

References

[1] Soomro SH, Ting LR, Qing YY and Ren M. Molecular biology of glioblastoma: classification and mutational locations. J Pak Med Assoc 2017; 67: 1410-1414.

[2] Wirsching HG, Galanis E and Weller M. Glioblastoma. Handb Clin Neurol 2016; 134: 381-397.

[3] Jiapaer S, Furuta T, Tanaka S, Kitabayashi T and Nakada M. Potential strategies overcoming the temozolomide resistance for glioblastoma. Neurol Med Chir (Tokyo) 2018; 58: 405-421.

[4] Koukourakis MI, Mitrakas AG and Giatromanolaki A. Therapeutic interactions of autophagy with radiation and temozolomide in glioblastoma: evidence and issues to resolve. Br J Cancer 2016; 114: 485-496.

[5] Zhang H, Wang R, Yu Y, Liu J, Luo T and Fan F. Glioblastoma treatment modalities besides surgery. J Cancer 2019; 10: 4793-4806.

[6] Jäger E, Chen YT, Drijfhout JW, Karbach J, Ringhoffer M, Jäger D, Arand M, Wada H, Noguchi Y, Stockert E, Old LJ and Knuth A. Simultaneous humoral and cellular immune response against cancer-testis antigen NY-ESO-1: definition of human histocompatibility leukocyte antigen (HLA)-A2-binding peptide epitopes. J Exp Med 1998; 187: 265-270.

[7] Fratta E, Coral S, Covre A, Parisi G, Colizzi F, Danielli R, Nicolay HJ, Sigalotti L and Maio M. The biology of cancer testis antigens: putative function, regulation and therapeutic potential. Mol Oncol 2011; 5: 164-182.

[8] Thomas R, Al-Khadairi G, Roelands J, Hendrickx W, Dermime S, Bedognetti D and Decock J. NY-ESO-1 based immunotherapy of cancer: current perspectives. Front Immunol 2018; 9: 947.

[9] Salmaninejad A, Zamani MR, Pourvahedi M, Golchehre Z, Hosseini Bereshneh A and Rezaei N. Cancer/testis antigens: expression, regulation, tumor invasion, and use in immunotherapy of cancers. Immunol Invest 2016; 45: 619-640.

[10] Grizzi F, Mirandola L, Qehajaj D, Cobos E, Figueroa JA and Chiriva-Internati M. Cancer-testis antigens and immunotherapy in the light of cancer complexity. Int Rev Immunol 2015; 34: 143-153.

[11] Almeida LG, Sakabe NJ, deOliveira AR, Silva MC, Mundstein AS, Cohen T, Chen YT, Chua R, Gurung S, Gnjatic S, Jungbluth AA, Caballero OL, Bairoch A, Kiesler E, White SL, Simpson AJ, Old LJ, Camargo AA and Vasconcelos AT. CTdatabase: a knowledge-base of high-throughput and curated data on cancer-testis antigens. Nucleic Acids Res 2009; 37: D816-D819.

[12] Yan J, Wen J, Wei ZD, Li XS, Li P and Xiao SW. Prognostic and clinicopathological value of melanoma-associated antigen D4 in patients with glioma. Oncol Lett 2018; 15: 4151-4160.

[13] Chianese-Bullock KA, Pressley J, Garbee C, Hibbitts S, Murphy C, Yamshchikov G, Petroni GR, Bissonette EA, Neese PY, Grosh WW, Merrill P, Fink R, Woodson EM, Wiernasz CJ, Patterson JW and Slingluff CL Jr. MAGE-A1-, MAGE-A10-, and gp100-derived peptides are immunogenic when combined with granulocyte-macrophage colony-stimulating factor

and montanide ISA-51 adjuvant and administered as part of a multipeptide vaccine for melanoma. J Immunol 2005; 174: 3080-3086.
[14] Morgan RA, Chinnasamy N, Abate-Daga D, Gros A, Robbins PF, Zheng Z, Dudley ME, Feldman SA, Yang JC, Sherry RM, Phan GQ, Hughes MS, Kammula US, Miller AD, Hessman CJ, Stewart AA, Restifo NP, Quezado MM, Alimchandani M, Rosenberg AZ, Nath A, Wang T, Bielekova B, Wuest SC, Akula N, McMahon FJ, Wilde S, Mosetter B, Schendel DJ, Laurencot CM and Rosenberg SA. Cancer regression and neurological toxicity following anti-MAGE-A3 TCR gene therapy. J Immunother 2013; 36: 133-151.
[15] Perrin SL, Samuel MS, Koszyca B, Brown MP, Ebert LM, Oksdath M and Gomez GA. Glioblastoma heterogeneity and cancer cell plasticity. Biochem Soc Trans 2019; 47: 625-638.
[16] Aran D, Sirota M and Butte AJ. Systematic pan-cancer analysis of tumour purity. Nat Commun 2015; 6: 8971.
[17] Gaudet S, Branton D and Lue RA. Characterization of PDZ-binding kinase, a mitotic kinase. Proc Natl Acad Sci U S A 2000; 97: 5167-5172.
[18] Abe Y, Matsumoto S, Kito K and Ueda N. Cloning and expression of a novel MAPKK-like protein kinase, lymphokine-activated killer T-cell-originated protein kinase, specifically expressed in the testis and activated lymphoid cells. J Biol Chem 2000; 275: 21525-21531.
[19] Han Z, Li L, Huang Y, Zhao H and Luo Y. PBK/TOPK: a therapeutic target worthy of attention. Cells 2021; 10: 371.
[20] Chen JH, Liang YX, He HC, Chen JY, Lu JM, Chen G, Lin ZY, Fu X, Ling XH, Han ZD, Jiang FN and Zhong WD. Overexpression of PDZ-binding kinase confers malignant phenotype in prostate cancer via the regulation of E2F1. Int J Biol Macromol 2015; 81: 615-623.
[21] Singh PK, Srivastava AK, Dalela D, Rath SK, Goel MM and Bhatt ML. Expression of PDZ-binding kinase/T-LAK cell-originated protein kinase (PBK/TOPK) in human urinary bladder transitional cell carcinoma. Immunobiology 2014; 219: 469-474.
[22] Ikeda Y, Park JH, Miyamoto T, Takamatsu N, Kato T, Iwasa A, Okabe S, Imai Y, Fujiwara K, Nakamura Y and Hasegawa K. T-LAK cell-originated protein kinase (TOPK) as a prognostic factor and a potential therapeutic target in ovarian cancer. Clin Cancer Res 2016; 22: 6110-6117.
[23] Joel M, Mughal AA, Grieg Z, Murrell W, Palmero S, Mikkelsen B, Fjerdingstad HB, Sandberg CJ, Behnan J, Glover JC, Langmoen IA and Stangeland B. Targeting PBK/TOPK decreases growth and survival of glioma initiating cells in vitro and attenuates tumor growth in vivo. Mol Cancer 2015; 14: 121.
[24] Hayashi T, Hayakawa Y, Koh M, Tomita T, Nagai S, Kashiwazaki D, Sugimori M, Origasa H and Kuroda S. Impact of a novel biomarker, T-LAK cell-originating protein kinase (TOPK) expression on outcome in malignant glioma. Neuropathology 2018; 38: 144-153.
[25] Quan C, Xiao J, Duan Q, Yuan P, Xue P, Lu H, Yan M, Guo D, Xu S, Zhang X, Lin X, Wang Y, Dogan S, Zhang J, Zhu F, Ke C and Liu L. T-lymphokine-activated killer cell-originated protein kinase (TOPK) as a prognostic factor and a potential therapeutic target in glioma. Oncotarget 2017; 9: 7782-7795.
[26] Hu F, Gartenhaus RB, Eichberg D, Liu Z, Fang HB and Rapoport AP. PBK/TOPK interacts with the DBD domain of tumor suppressor p53 and modulates expression of transcriptional targets including p21. Oncogene 2010; 29: 5464-5474.
[27] Li S, Zhu F, Zykova T, Kim MO, Cho YY, Bode AM, Peng C, Ma W, Carper A, Langfald A and Dong Z. T-LAK cell-originated protein kinase (TOPK) phosphorylation of MKP1 protein prevents solar ultraviolet light-induced inflammation through inhibition of the p38 protein signaling pathway. J Biol Chem 2011; 286: 29601-29609.
[28] Park JH, Lin ML, Nishidate T, Nakamura Y and Katagiri T. PDZ-binding kinase/T-LAK cell-originated protein kinase, a putative cancer/testis antigen with an oncogenic activity in breast cancer. Cancer Res 2006; 66: 9186-9195.
[29] Abe Y, Takeuchi T, Kagawa-Miki L, Ueda N, Shigemoto K, Yasukawa M and Kito K. A mitotic kinase TOPK enhances Cdk1/cyclin B1-dependent phosphorylation of PRC1 and promotes cytokinesis. J Mol Biol 2007; 370: 231-245.
[30] Mao P, Bao G, Wang YC, Du CW, Yu X, Guo XY, Li RC and Wang MD. PDZ-binding kinase-dependent transcriptional regulation of CCNB2 promotes tumorigenesis and radio-resistance in glioblastoma. Transl Oncol 2020; 13: 287-294.
[31] Herbert KJ, Puliyadi R, Prevo R, Rodriguez-Berriguete G, Ryan A, Ramadan K and Higgins GS. Targeting TOPK sensitises tumour cells to radiation-induced damage by enhancing replication stress. Cell Death Differ 2020; 28: 1333-1346.
[32] Liu K, Chen Y, Feng P, Wang Y, Sun M, Song T, Tan J, Li C, Liu S, Kong Q and Zhang J. Identification of pathologic and prognostic genes in prostate cancer based on database mining. Front Genet 2022; 13: 854531.
[33] de Boussac H, Bruyer A, Jourdan M, Maes A, Robert N, Gourzones C, Vincent L, Seckinger A, Cartron G, Hose D, De Bruyne E, Kassambara

A, Pasero P and Moreaux J. Kinome expression profiling to target new therapeutic avenues in multiple myeloma. Haematologica 2020; 105: 784-795.

[34] Yakkundi P, Gonsalves E, Galou-Lameyer M, Selby MJ and Chan WK. Aryl hydrocarbon receptor acts as a tumor suppressor in a syngeneic MC38 colon carcinoma tumor model. Hypoxia (Auckl) 2019; 7: 1-16.

[35] Fisk HA, Mattison CP and Winey M. A field guide to the Mps1 family of protein kinases. Cell Cycle 2004; 3: 439-442.

[36] Tannous BA, Kerami M, Van der Stoop PM, Kwiatkowski N, Wang J, Zhou W, Kessler AF, Lewandrowski G, Hiddingh L, Sol N, Lagerweij T, Wedekind L, Niers JM, Barazas M, Nilsson RJ, Geerts D, De Witt Hamer PC, Hagemann C, Vandertop WP, Van Tellingen O, Noske DP, Gray NS and Wurdinger T. Effects of the selective MPS1 inhibitor MPS1-IN-3 on glioblastoma sensitivity to antimitotic drugs. J Natl Cancer Inst 2013; 105: 1322-1331.

[37] Maachani UB, Kramp T, Hanson R, Zhao S, Celiku O, Shankavaram U, Colombo R, Caplen NJ, Camphausen K and Tandle A. Targeting MPS1 enhances radiosensitization of human glioblastoma by modulating DNA repair proteins. Mol Cancer Res 2015; 13: 852-862.

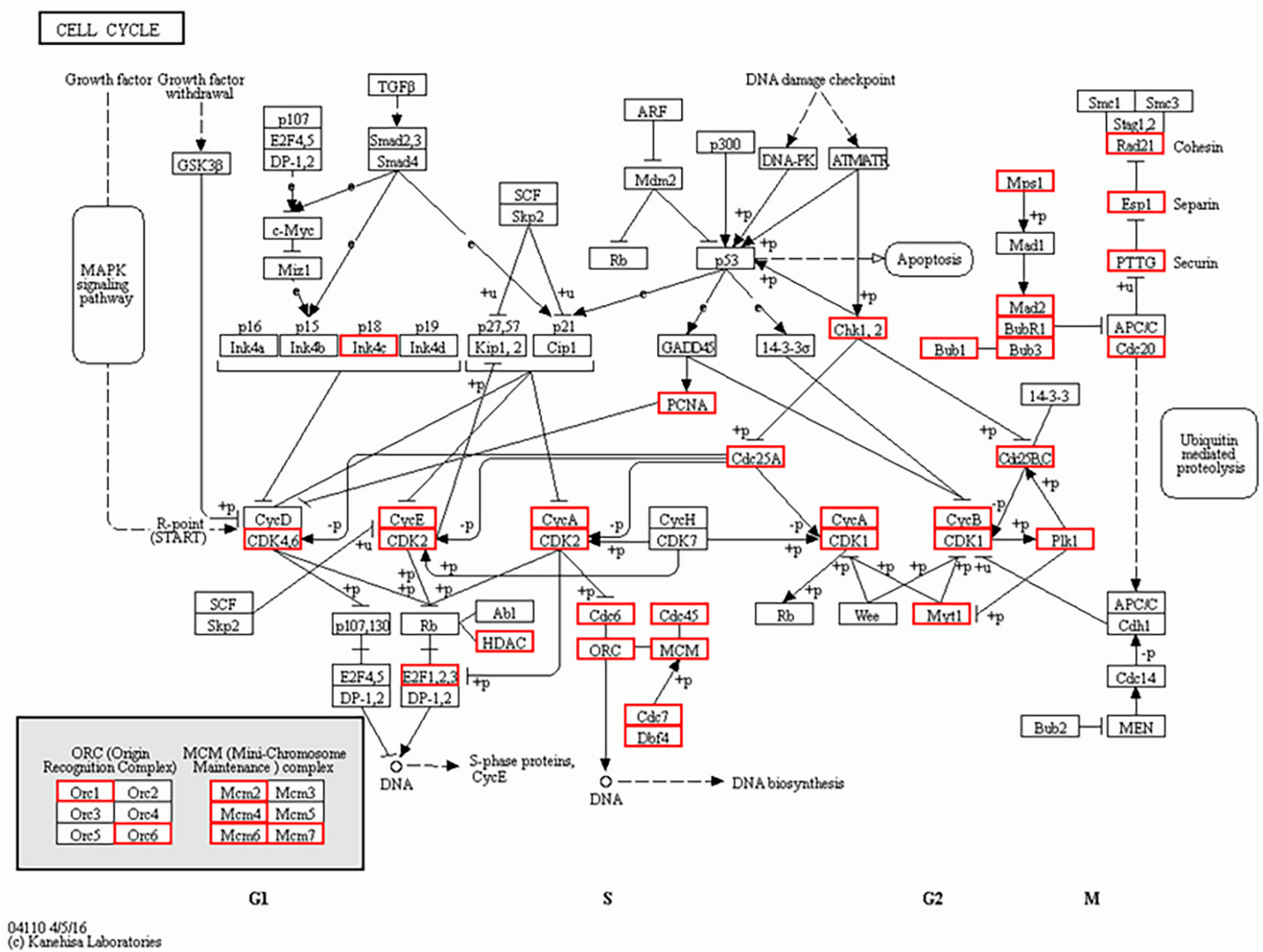
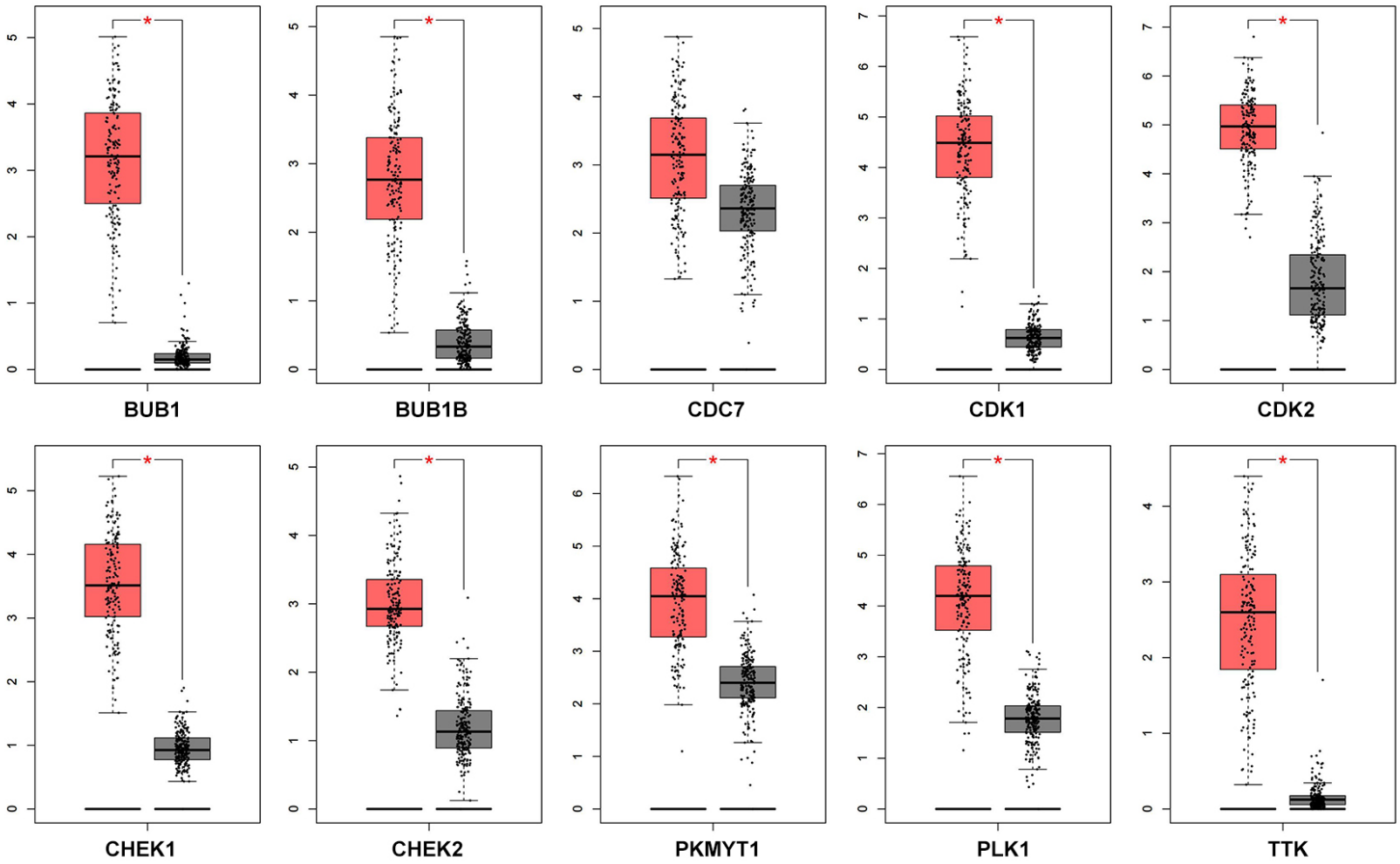


Figure S1. KEGG pathway map for cell cycle (KEGG map 04110) highlighting PBK related genes in PPI network. The genes with red box are genes positively correlated with PBK and also included in the PPI network.


BUB1
BUB1B
CDC7
CDK1
CDK2
CHEK1
CHEK2
PKMYT1
PLK1
TTK

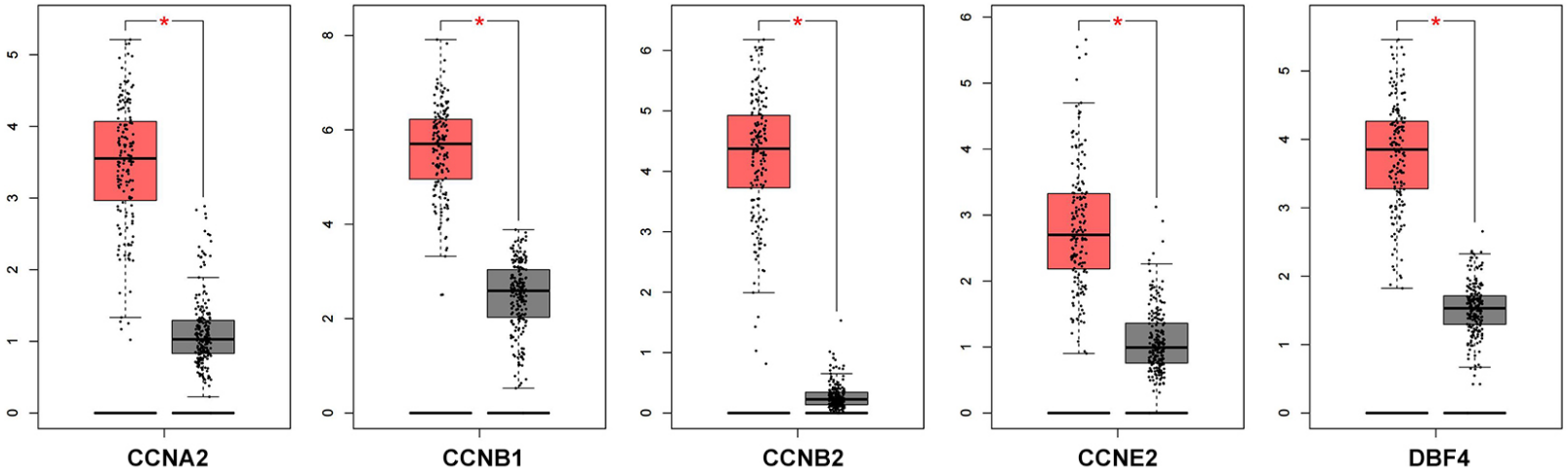


Figure S2. Cell cycle-related molecules mRNA expression of 163 GBM and 207 normal brain tissue in the TCGA and GXTs database. Red box: GBM, black box: Normal brain, notes: Expression of genes.