

Functional network analysis of gene-phenotype connectivity associated with temozolomide

Jia Shi^{1,*}, Bo Dong^{1,*}, Peng Zhou^{1,*}, Wei Guan¹ and Ya Peng¹

¹Department of Neurosurgery, The Third Affiliated Hospital of Soochow University, Changzhou City, 213003, China

*These authors have contributed equally to this work

Correspondence to: Wei Guan, email: 13961195153@126.com
Ya Peng, email: pengyatg@163.com

Keywords: temozolomide, glioma, TP53

Received: June 26, 2017

Accepted: August 17, 2017

Published: September 12, 2017

Copyright: Shi et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Rationale: Glioma has a poor survival rate in patients even with aggressive treatment. Temozolomide (TMZ) is the standard chemotherapeutic choice for treating glioma, but TMZ treatment consistently leads to high resistance.

Aim: To investigate the underlying mechanisms of TMZ action with new therapeutic regimens in glioma.

Methods and results: The biological effects of TMZ mainly depend on the three following DNA repair systems: methylguanine methyltransferase (MGMT), mismatch repair (MMR) and base excision repair (BER). Based on related genes in these three systems, web-based tools containing data compiled from open-source databases, including DrugBank, STRING, WebGestalt and ClueGO, were queried, and five common genes along with the top fifteen pathways, including the glioma pathway, were identified. A genomic analysis of the six genes identified in the glioma pathway by cBioPortal indicated that TMZ might exert biological effects via interaction with the tumor protein P53 (TP53) signaling axis. Finally, a survival analysis with the six genes in glioma cases (low-grade glioma and glioblastoma multiforme) was conducted using OncoLnc, which might provide directions for the future exploration of prognosis in glioma.

Conclusions: This study indicates that a functional network analysis resembles a "BioGPS", with the ability to draw a web-based scientific map that can productively and cost-effectively associate TMZ with its primary and secondary biological targets.

INTRODUCTION

Glioma is the most common primary brain tumor in adults and has the highest degree of malignancy [1]. Currently, the standard therapy for glioma is maximal surgical resection followed by fractionated radiotherapy along with recurrent adjuvant treatment using alkylating agents such as nitrosoureas, procarbazine, and more recently, temozolomide (TMZ) [2, 3]. TMZ is a monofunctional cytostatic agent that has been synthesized at Aston University since 1984 [4]. TMZ has been reported as a clinical treatment for glioblastoma multiforme (GBM) and metastatic melanoma [5]. Recently, GBM has also been studied in refractory acute leukemia with promising results [6, 7].

The action of TMZ undergo DNA damage and DNA repair system two stages. In general, DNA repair systems are in charge of maintaining the genome integrity of cellular organism as they are able to neutralize DNA damage induced by various chemical and physical agents [8]. Therefore, for a better understanding of the molecular mechanisms of the biological effects of TMZ and target cell resistance to these agents it is necessary to consider underlying pathways of DNA repair. Like many other alkylating agents, the cytotoxic effects of TMZ is believed to manifest largely by the formation of methylation of the O6 position of guanine [9]. Consequently, the primary mechanisms of resistance to TMZ is a function of the activity of the DNA repair enzyme O(6)-alkylguanine-

Table 1: Characterization of Temozolomide using DrugBank

DB_ID	Name	Group	Category	Indication
DB00853	Temozolomide	Approved	Alkylating Agents Antineoplastic Agents Imidazoles Immunosuppressive Agents Noxae Toxic Actions Triazenes	1. For the treatment of adult patients diagnosed with anaplastic astrocytoma who has progressed after nitrosourea and procarbazine therapy. 2. Concomitantly with radiation therapy for treatment of newly diagnosed glioblastoma multiforme. 3. Used as maintenance therapy for glioblastoma multiforme.

Table 2: Identification of direct targets of Temozolomide using DrugBank

DB_ID	Name	Target	Kind	Pharmacological action	Action	Organism
DB00853	Temozolomide	DNA	Nucleotide	Yes	Cross-linking/ Alkylation	Human

DNA-alkyltransferase, also called methylguanine methyltransferase (MGMT) [11]. During the process of DNA repair, MGMT also induces DNA mismatch repair system and base excision repair pathway [10, 11]. Thus, the biological effects of TMZ depend on at least three DNA repair systems [9]. (1) *MGMT* is a DNA repair protein encoded by the *MGMT* gene and is involved in the cellular defense against toxicity and mutagenesis from alkylating agents [12]. (2) The mismatch repair (MMR) system consists of a protein complex that contributes to the repair of biosynthetic errors generated during DNA replication. The MMR system recognizes and repairs not only base mismatches but also insertion and deletion loops [13]. (3) The base excision repair (BER) system, similar to the MMR system, is a protein complex that controls damage from cellular metabolic lesions (such as oxidation or the methylation of DNA bases) and from base modifications induced by physical or chemical agents [14].

Prior to the last decade, TMZ was shown to significantly improve outcomes in patients with GBM when administered concomitantly with radiotherapy and as maintenance strategy thereafter [15]. Nevertheless, despite intense therapy, the average survival rate of GBM patients is only 12-18 months [16, 17]. One critical barrier to the effective treatment of malignant glioma is resistance to TMZ, which has attracted significant scientific interest. As demonstrated previously, the biological effects of TMZ mainly depend on three DNA repair systems. However, the primary and secondary targets of TMZ and the interactions between the target genes of these three DNA repair systems remain unclear.

With the advancements in multicenter genomics studies, including microarrays, proteomics and high-throughput screening assays, exploratory network-based

studies have been used to analyze massive amounts of data and numerous human diseases. In this study, we first employed DrugBank in a broad study of TMZ and drug-target information. Based on the three DNA repair systems, we identified five common targets using the STRING database. Then, the targets and associated genes were further explored through pathway enrichment analysis using the WebGestalt and ClueGO databases; the glioma pathway was screened, and related genes were identified to further explore genomic alterations using the cBio Cancer Genomics Portal (cBioPortal) database. Overall, our study based on these functional network analyses provides valuable information for exploring the underlying mechanisms of TMZ action against glioma and for identifying potential targets to overcome TMZ resistance in glioma.

RESULTS

Characterization of TMZ bioactivity using DrugBank and the visualization of TMZ linkage networks using STRING

DrugBank was first queried using TMZ as an input, which resulted in the output DB00853 categorizing TMZ with alkylating agents, antineoplastic agents, imidazoles, immunosuppressive agents, noxae, toxic actions and triazenes (Table 1). Furthermore, grouping TMZ as a Food and Drug Administration (FDA)-approved drug, the indications of TMZ were divided into 3 detailed classes: (1) treatment of adult patients with anaplastic astrocytoma after nitrosourea and procarbazine therapy; (2) concomitant administration with radiation treatment for patients with newly diagnosed GBM; and (3) maintenance therapy for patients with GBM (Table 1).

Table 3: Targets of Temozolomide by STRING

Searched_Drug(1/1)				
DB_ID	Name	Target	Gene	Entrez_ID
DB00853	Temozolomide	MGMT	MGMT	4255
DB00853	Temozolomide	Mismatch repair	MRC1	4360
			ATP9B	374868
			MSH2	4436
			MSH3	4437
			MSH6	2956
			EXO1	9156
			MLH1	4292
			PMS2	5395
DB00853	Temozolomide	Base excision repair	RAD1	5810
			RAD9A	5883
			HUS1	3364
			NTHL1	4913
			PARP1	142
			PARP2	10038
			PARP3	10039
			PNKP	11284
			POLL	27343
			TDG	6996
			FEN1	2237
			USP47	55031
			APEX1	328
			APEX2	27301

Table 2 illustrates that the direct target of TMZ is DNA (Table 2). As mentioned previously, the biological effects of TMZ primarily depend on three DNA repair systems: MGMT, MMR and BER. The MMR and BER systems consist of protein complexes; the MMR system includes the following 8 genes: MRC1, ATP9B, MSH6, MSH3, EXO1, MLH1, PMS2 and MSH2, while the BER system includes the following 14 genes: RAD1, RAD9A, HUS1, NTHL1, PARP1, PARP2, PARP3, PNKP, APEX1, POLL, TDG, FEN1, USP47 and APEX2 (Table 3). Next, we analyzed MGMT-related genes using STRING, and a total of 50 target protein interactions were identified (Supplementary Table 1). No gene common to all 3 DNA repair systems was detected. Therefore, we expanded our search and further analyzed MMR/BER-related protein-protein interactions (PPIs); we ultimately identified five genes in common (MSH2, MSH6, TOP2B, TP53 and XRCC3) that were related to the three DNA repair systems

(Figure 1). It should be noted that all five of the common genes have been reported to be directly associated with glioma [18-22], suggesting potential roles of these genes in mediating the action of TMZ against glioma.

Analysis of functional attributes associated with TMZ-mediated changes in gene sets using WebGestalt and ClueGO

To evaluate the functional features of TMZ-mediated gene sets, we performed a Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis using WebGestalt. The top 15 KEGG pathways were identified based on genes associated with the three DNA repair systems (Table 4). These pathways included BER (11 genes), platinum drug resistance (11 genes), pathway in cancer (19 genes), bladder cancer (8 genes), non-small cell lung cancer (8 genes), MMR (6 genes),

endometrial cancer (7 genes), colorectal cancer (7 genes), central carbon metabolism in cancer (7 genes), melanoma (7 genes), breast cancer (8 genes), glioma (6 genes), microRNA (miRNA) in cancer (10 genes), prostate cancer (6 genes) and apoptosis (7 genes). All the identified KEGG enrichment pathways represented biological functions that had statistically significant associations with the TMZ gene sets, thus motivating further investigation. A broad grouping of the functional analysis results indicated that the TMZ-linked genes were mainly associated with (1) DNA repair and drug resistance pathways, (2) cancer-related signaling pathways, and (3) signaling pathway cascades with potential mechanistic underpinnings, including the control of cancer cell development and survival via miRNA-mediated carbon metabolism and the regulation of apoptosis.

Next, to validate the pathways identified by WebGestalt, we additionally performed KEGG pathway enrichment using ClueGO. The glioma pathway was also screened for statistical significance in this analysis (Figure 2A and Table 5). A broad search revealed that 6 genes in the glioma pathway were connected to TMZ-linked genes: CDKN2A, EGFR, HRAS, KRAS, PTEN and TP53 (Figure 2B and Table 5). Notably, among these 6 selected genes (identified by both WebGestalt and ClueGO), TP53 was found to participate in numerous signaling pathways, including apoptosis and several cancer-related pathways (Figure 2C), suggesting a critical role for TP53 in diverse biological effects mediated by TMZ.

Mining genomic alterations associated with TMZ-related genes in glioma using cBioPortal

To further validate the relationship between TMZ-related genes and the glioma pathway, cBioPortal was used to explore genomic alterations in genes connected with TMZ treatment of glioma. A total of 8 glioma studies were included in cBioPortal [23-27]. Three studies were provisionally embraced; thus, emphasis was directed to the remaining 5 studies. A query of the 6 selected genes (CDKN2A, EGFR, HRAS, KRAS, PTEN and TP53) identified in the glioma pathway was performed, and these genes were analyzed in the 5 glioma studies. Alterations ranging from 2.1% to 91.5% were detected for submitted gene sets (Figure 3A). A summary of multiple gene alterations detected across each set of tumor samples from the study of Brennan CW showed the most pronounced genomic changes among the 5 glioma studies [26]. In this study, 257 cases (91.5%) had alterations in all 6 genes; the frequency of alterations in each gene is shown in Figure 3B. For CDKN2A (61%), most alterations were classified as deep deletions, with a few cases of truncating mutations. For EGFR (53%), the majority of alterations were amplifications, with a small fraction of missense mutations. For PTEN (31%) and TP53 (22%), the gene changes included deep deletions and truncating, missense

and inframe mutations. For KRAS and HRAS, few alterations were detected with 1.8% and 1.1%, respectively (Figure 3B). An analysis of interactions between genes showed that KRAS and TP53 (Pearson's correlation: 0.36) as well as PTEN and TP53 (Pearson's correlation: 0.32) exhibited co-expression in the GBM samples in the study of Brennan CW (Table 6), while CDKN2A and TP53 (p -value<0.001) as well as EGFR and TP53 (p -value=0.001) exhibited mutual exclusivity (Table 7), suggesting a central axis function for TP53 under TMZ control in the glioma pathway.

We additionally used cBioPortal to perform interactive analysis and construct networks of genes that were altered in cancer. We created a network containing all neighbors of the 6 query genes (Figure 4A). To reduce the complexity of the analysis, we applied genomic alteration frequency within the selected glioma studies as a filter, such that only neighbors with a high alteration frequency are shown (Figure 4B). First, the 6 selected genes were determined to be associated with CD4 when a filter of $\geq 16.7\%$ alteration was applied to neighbors. In comparison, 8 genes, including CD4 and PDGFRA, were evident using a filter of $\geq 13.2\%$ alteration. Ten gene clusters, including CD4, PDGFRA, NF1 and PIK3R1, were demonstrated when the filter was reduced to 11% alteration, while 11 genes, including CD4, PDGFRA, NF1, PIK3R1 and KIT, were revealed with a filter of 10% alteration. The comprehensive and pruned networks revealed the potential interactions as well as the complexity and the variability of differences in the interactions between TMZ-linked genes and the GBM samples in the Brennan CW study. Moreover, drugs that are applied for specific genes, including cancer drugs, FDA-approved drugs and others, are also shown in cBioPortal. Figure 4B lists specific cancer drugs acting on EGFR, TP53 and CDKN2A (Figure 4B); currently, TMZ is not used to target any of these genes. This network analysis provides a molecular basis for future clinical applications of TMZ targeting selected genes.

Low-grade gliomas (LGGs) in the brain are slower growing than their high-grade counterparts. LGGs account for 10-20% of all primary brain tumors, and the median survival of LGG patients ranges from 4.7 to 9.8 years [28]. Related literature revealed that TMZ is also indicated for LGG patients with high-risk features [29]. Figure 3A shows that among the LGG studies in cBioPortal, the study of Johnson BE observed the most pronounced genomic changes (Figure 3A, column 2) [2, 3]. In this study, gene sets were altered in 90.2% of 61 cases. TP53 (90%) exhibited the most prominent changes, which were classified as missense mutations with a few cases of truncating mutations. For CDKN2A, alterations also included missense and truncating mutations. For EGFR, HRAS, KRAS and PTEN, few alterations existed in LGG cases (Figure 5A). Results from co-expression and mutual exclusivity analyses were not statistically significant (data not shown). Next, a

network analysis was performed that included neighbors of the 6 query genes (Figure 5B). Seven genes, including FAT1, were revealed using a filter of $\geq 13.1\%$ alteration; 10 gene clusters (including FAT1, INPPL1, PIK3CA and TNRC6A) were shown when the filter was reduced to 8.1% alteration.

Analysis of survival associated with TMZ-related genes in glioma according to OncoLnc

To explore the association between TMZ-related genes and the survival of patients with glioma, OncoLnc, an integrated data mining system, was used. The 6 selected genes (CDKN2A, EGFR, HRAS, KRAS, PTEN and TP53) identified in the glioma pathway were used to conduct a survival analysis with clinical profiles in glioma. The results indicated that in LGG, a high level of PTEN expression predicted longer survival with statistical significance (log-rank p -value=0.00521) when patients were classified according to the mean value of the mRNA expression level (Figure 6A-6F). In contrast, in GBM, no gene was statistically significantly associated with patient survival (Figure 7A-7F), which might be attributed to 2 reasons: (1) as shown in Figure 7A-7F, the average survival time for GBM patients was only 500 days, which is consistent with relevant studies; and (2) the OncoLnc analysis was based on a relatively

small sample (a total of 152 cases), and the findings could easily be spurious.

DISCUSSION

Since the landmark findings regarding the antineoplastic activity of TMZ reported in 1984 [4], more than 5529 publications on TMZ have appeared on PubMed. A wide range of biological and cellular activities have been identified for TMZ, demonstrating the fascinating nature of this compound regarding a plethora of related diseases. However, tumor resistance to TMZ is still a critical barrier to the effective treatment of glioma. As such, new analytical means or platforms are required to bridge TMZ with its primary or secondary targets and thereby illustrating the underlying mechanisms of TMZ and its clinical outcomes.

In this study, we performed a functional network analysis using a set of web-based tools. First, we demonstrated the feasibility of analyzing the connectivity between TMZ and cancer using DrugBank, STRING, WebGestalt and ClueGO. As reported in other studies, TMZ acted mainly through 3 DNA repair systems: MRMT, MMR and BER. Based on these 3 DNA repair systems, 5 common genes (MSH2, MSH6, TOP2B, TP53 and XRCC3) were identified using STRING (Figure 1). Notably, TMZ has been reported to induce apoptosis in melanoma cells, and the inactivation of MGMT results in a

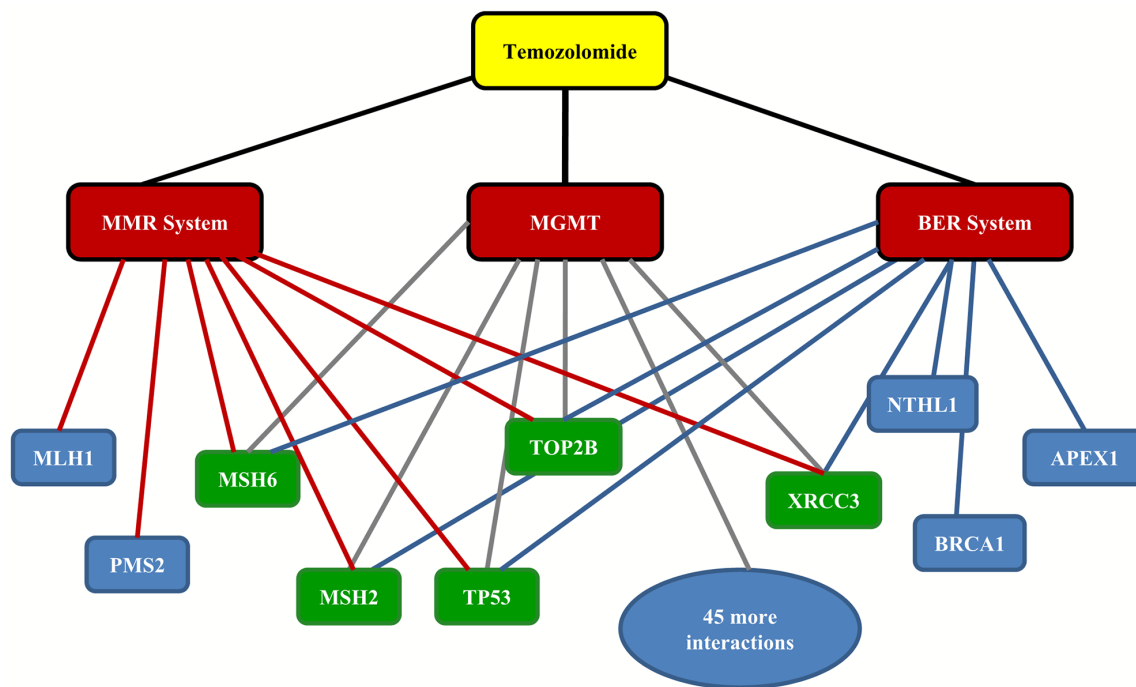


Figure 1: Drug-target protein interaction network of temozolomide. Five protein targets (in green) in common among three DNA repair systems were identified using STRING : MSH2, MSH6, XRCC3, TOP2B and TP53. MGMT: methylguanine methyltransferase. MMR: mismatch repair. BER: base excision repair.

Table 4: List of enriched Temozolomide associated gene sets identified using WebGestalt Analysis

Pathway Name	#Gene	Gene (corresponding gene set)	statistics
Base excision repair	11	PARP2;PARP3;PARP1;FEN1;APEX2;POLL;APEX1;NTHL1;OGG1;TDG;XRCC1	C=33; O=11; E=0.21; R=51.99; <i>p</i> value=0
Platinum drug resistance	11	CDKN2A;GSTP1;MSH6;BIRC5;MLH1;MSH2;MSH3;BRCA1;TOP2A;TOP2B;TP53	C=75; O=11; E=0.48; R=22.87; <i>p</i> value=7.61E-13
Pathway in cancer	19	CDKN2A;CDKN2B;RASSF1;DAPK1;EGFR;GSTP1;MSH6;HIF1A;HRAS;BIRC5;KRAS;MLH1;MSH2;MSH3;PTEN;RARB;BRCA2;TP53;CDH1	C=397; O=19; E=2.55; R=7.46; <i>p</i> value=8.94E-13
Bladder cancer	8	CDKN2A;RASSF1;DAPK1;EGFR;HRAS;KRAS;TP53;CDH1	C=41; O=8; E=0.26; R=30.43; <i>p</i> value=1.23E-10
Non-small cell lung cancer	8	CDKN2A;RASSF1;EGFR;FHIT;HRAS;KRAS;RARB;TP53	C=56; O=8; E=0.36; R=22.28; <i>p</i> value=1.70E-09
Mismatch repair	6	MSH6;MLH1;MSH2;MSH3;PMS2;EXO1	C=23; O=6; E=0.15; R=40.69; <i>p</i> value=4.61E-09
Endometrial cancer	7	EGFR;HRAS;KRAS;MLH1;PTEN;TP53;CDH1	C=52; O=7; E=0.33; R=21.00; <i>p</i> value=2.98E-08
Colorectal cancer	7	MSH6;BIRC5;KRAS;MLH1;MSH2;MSH3;TP53	C=62; O=7; E=0.40; R=17.61; <i>p</i> value=1.05E-07
Central carbon metabolism in cancer	7	EGFR;HIF1A;HRAS;IDH1;KRAS;PTEN;TP53	C=67; O=7; E=0.43; R=16.30; <i>p</i> value=1.80E-07
Melanoma	7	CDKN2A;EGFR;HRAS;KRAS;PTEN;TP53;CDH1	C=71; O=7; E=0.46; R=15.38; <i>p</i> value=2.71E-07
Breast cancer	8	EGFR;ESR1;HRAS;KRAS;PTEN;BRCA1;BRCA2;TP53	C=146; O=8; E=0.94; R=8.55; <i>p</i> value=3.29E-06
Glioma	6	CDKN2A;EGFR;HRAS;KRAS;PTEN;TP53	C=66; O=6; E=0.42; R=14.18; <i>p</i> value=3.38E-06
MicroRNAs in cancer	10	CDKN2A;RASSF1;DNMT1;EGFR;HRAS;KRAS;PTEN;BRCA1;TIMP3;TP53	C=299; O=10; E=1.92; R=5.22; <i>p</i> value=1.46E-05
Prostate cancer	6	EGFR;GSTP1;HRAS;KRAS;PTEN;TP53	C=89; O=6; E=0.57; R=10.51; <i>p</i> value=1.93E-05
Apoptosis	7	PARP2;PARP3;PARP1;HRAS;BIRC5;KRAS;TP53	C=140; O=7; E=0.90; R=7.80; <i>p</i> value=2.63E-05

The row lists the following statistics: C: the number of reference genes in category; O: the number of genes in the gene set and also in the category; E: the expected number in the category; R: ratio of enrichment; *p* value: *p* value from hypergeometric test.

high level of resistance to TMZ and impairs the expression of MSH2/MSH6 through the over expression of P53 [30]. Furthermore, other studies have shown that the expression levels of TOP2A/Bare significantly higher in human GBM and that TOP2B transcription is corrected in PDGF (+) PTEN (-/-) or PDGF (+) PTEN (-/-) P53 (-/-) models by susceptibility to cancer drugs [20]. For XRCC3, a double-strand break repair gene, the thr241Met polymorphism of XRCC3 has been associated with susceptibility to developing astrocytomas and GBM [31, 32]. Therefore, these findings suggested the hypothesis that TMZ might

exert anti-tumor effects through MSH2/MSH6/TOP2B/XRCC3 in glioma patients via a regulated interaction with the TP53 signaling axis.

As supporting evidence, the KEGG enrichment analysis conducted using WebGestalt and ClueGO identified the glioma pathway as significantly altered by TMZ-related genes (Table 4). The association between glioma (GBM and LGG) and the beneficial effects exerted by TMZ in cancer was further observed and evaluated using cBioPortal with 6 genes (CDKN2A, EGFR, HRAS, KRAS, PTEN and TP53) identified in the glioma pathway.

Table 5: List of enriched Temozolomide associated gene sets identified using ClueGO Analysis

Pathway Name	#Gene	Gene (corresponding gene set)	p value
Base excision repair	11	APEX1, APEX2, FEN1, NTHL1, OGG1, PARP1, PARP2, PARP3, POLL, TDG, XRCC1	45.0E-18
Fanconi anemia pathway	5	BRCA1, BRCA2, ERCC1, MLH1, PMS2	26.0E-6
Apoptosis	7	BIRC5, HRAS, KRAS, PARP1, PARP2, PARP3, TP53	27.0E-6
Pathways in cancer	19	BIRC5, BRCA2, CDH1, CDKN2A, CDKN2B, DAPK1, EGFR, GSTP1, HIF1A, HRAS, KRAS, MLH1, MSH2, MSH3, MSH6, PTEN, RARB, RASSF1, TP53	1.2E-12
Mismatch repair	6	EXO1, MLH1, MSH2, MSH3, MSH6, PMS2	5.1E-9
Colorectal cancer	7	BIRC5, KRAS, MLH1, MSH2, MSH3, MSH6, TP53	95.0E-9
Pancreatic cancer	5	BRCA2, CDKN2A, EGFR, KRAS, TP53	54.0E-6
Endometrial cancer	7	CDH1, EGFR, HRAS, KRAS, MLH1, PTEN, TP53	25.0E-9
Glioma	6	CDKN2A, EGFR, HRAS, KRAS, PTEN, TP53	3.1E-6
Prostate cancer	6	EGFR, GSTP1, HRAS, KRAS, PTEN, TP53	18.0E-6
Thyroid cancer	4	CDH1, HRAS, KRAS, TP53	33.0E-6
Melanoma	7	CDH1, CDKN2A, EGFR, HRAS, KRAS, PTEN, TP53	250.0E-9
Bladder cancer	8	CDH1, CDKN2A, DAPK1, EGFR, HRAS, KRAS, RASSF1, TP53	140.0E-12
Non-small cell lung cancer	8	CDKN2A, EGFR, FHIT, HRAS, KRAS, RARB, RASSF1, TP53	1.4E-9
Breast cancer	8	BRCA1, BRCA2, EGFR, ESR1, HRAS, KRAS, PTEN, TP53	0.0000034

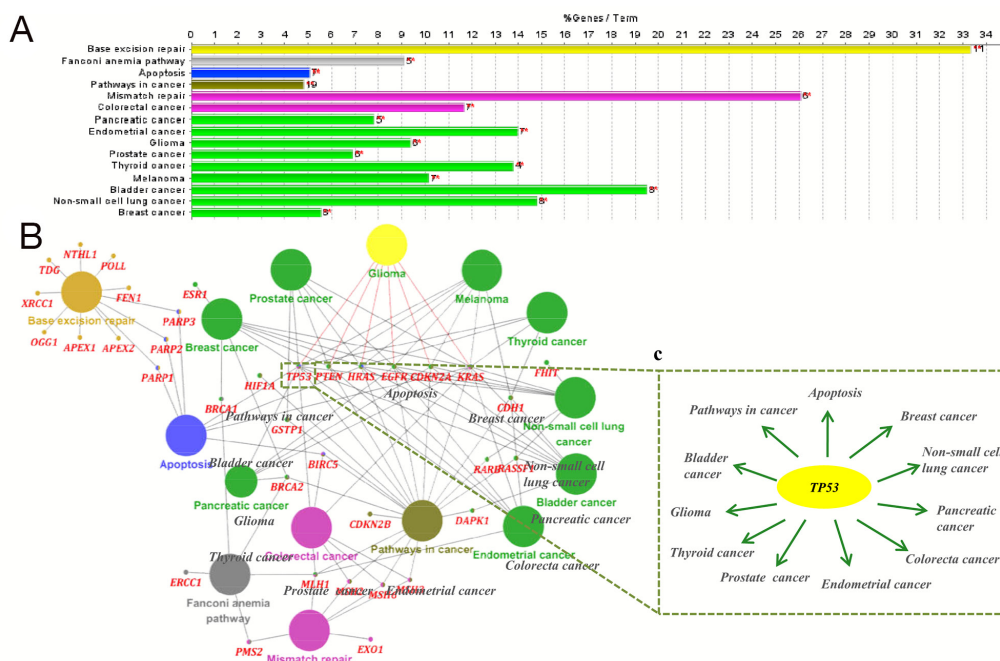


Figure 2: Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of temozolomide-associated gene sets performed using ClueGO. (A) The top 15 statistically enriched KEGG pathways and involved gene numbers (more details can be found in Table 5). **(B)** The biological network of temozolomide-linked genes consists of the top 15 statistically enriched KEGG pathways (large circle with different color), corresponding genes (red) and STRING protein interaction. **(C)** TP53 (in yellow) is shown as an illustration of complex regulatory networks. Among the top 15 KEGG pathways, 11 were associated with TP53.

Table 6: Co-expressions of selected 6 mRNAs (CDKN2A, EGFR, HRAS, KRAS, PTEN and TP53) in Brennan C.W. study

Gene	Correlated gene	Cytoband	Pearson's Correlation	Spearman's Correlation
KRAS	TP53	17p13.1	0.36	0.44
PTEN	TP53	17P13.1	0.32	0.33

Both PTEN and TP53 are tumor suppressor genes [33] that participated in almost all the cancer pathways identified by WebGestalt and ClueGO. In the case of GBM, most of the genetic alterations in CDKN2A, PTEN and TP53 were deletions or mutations (Figure 3B), which resulted in a reduction of their expression in conjunction with the development of carcinogenesis [34, 35]. In contrast, for EGFR, most of the genetic alterations were amplifications (Figure 3B), suggesting an over expression during

the acceleration of glioma [36, 37]. Moreover, the co-expression analysis illustrated synergistic effects between KRAS and TP53 as well as PTEN and TP53 (Figure 3C), and the mutual exclusivity analysis revealed a tendency toward mutual inhibition between CDKN2A and TP53 and between EGFR and TP53 (Table 6). Thus, the results are consistent with the activation of TP53 by TMZ as a major driver of anti-tumor effects in GBM. Regarding LGG, TP53 (90%) exhibited the most prominent alterations

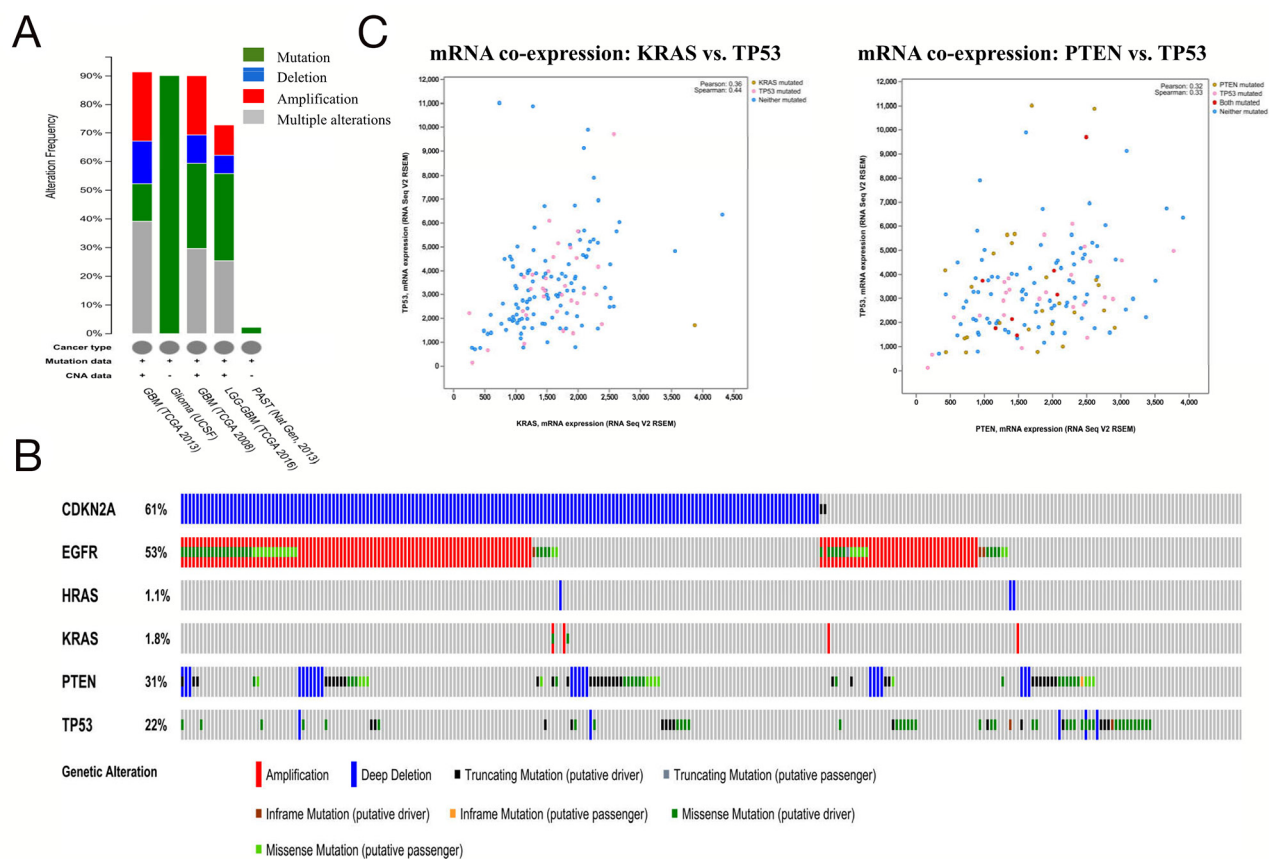


Figure 3: Mining genetic alterations connected with the temozolomide-associated genes CDKN2A, EGFR, HRAS, KRAS, PTEN and TP53 in glioma studies embedded in cBioPortal. (A) Overview of changes in the CDKN2A, EGFR, HRAS, KRAS, PTEN and TP53 genes in the genomics datasets available in 5 different glioma studies. (B) OncoPrint: a visual summary of genomic alterations across a set of glioblastoma multiforme (GBM) samples (data taken from the Brennan CW study, Cell, 2013) based on a query of 6 genes (CDKN2A, EGFR, HRAS, KRAS, PTEN and TP53). Different genomic alterations involving mutations and CNAs (copy number alterations, gene amplifications and homozygous deletions) are summarized, color-coded and displayed as % change in specific affected genes in individual glioma samples. Each row represents a gene, and each column represents a sample. Red bars represent amplifications, blue bars represent homozygous deletions, and green bars represent nonsynonymous mutations. (C) mRNA co-expression of KRAS A. with TP 53 and PTEN with TP53

Table 7: Mutual exclusivity of selected 6 mRNAs (CDKN2A, EGFR, HRAS, KRAS, PTEN and TP53) in Brennan CW study

Gene A	Gene B	p Value	Log Odds Ratio	Association
CDKN2A	TP53	<0.001	-1.353	Tendency towards mutual exclusivity
EGFR	TP53	0.001	-0.939	Tendency towards mutual exclusivity
CDKN2A	EGFR	0.006	0.647	Tendency towards co-occurrence

(Figure 5A). Finally, a survival analysis of the 6 genes was conducted in glioma cases (LGG or GBM) using OncoLnc, and the results indicated that high levels

of PTEN predicted a statistically significantly longer survival in LGG. This finding is in accordance with the diagnostic significance of PTEN mutation as a molecular

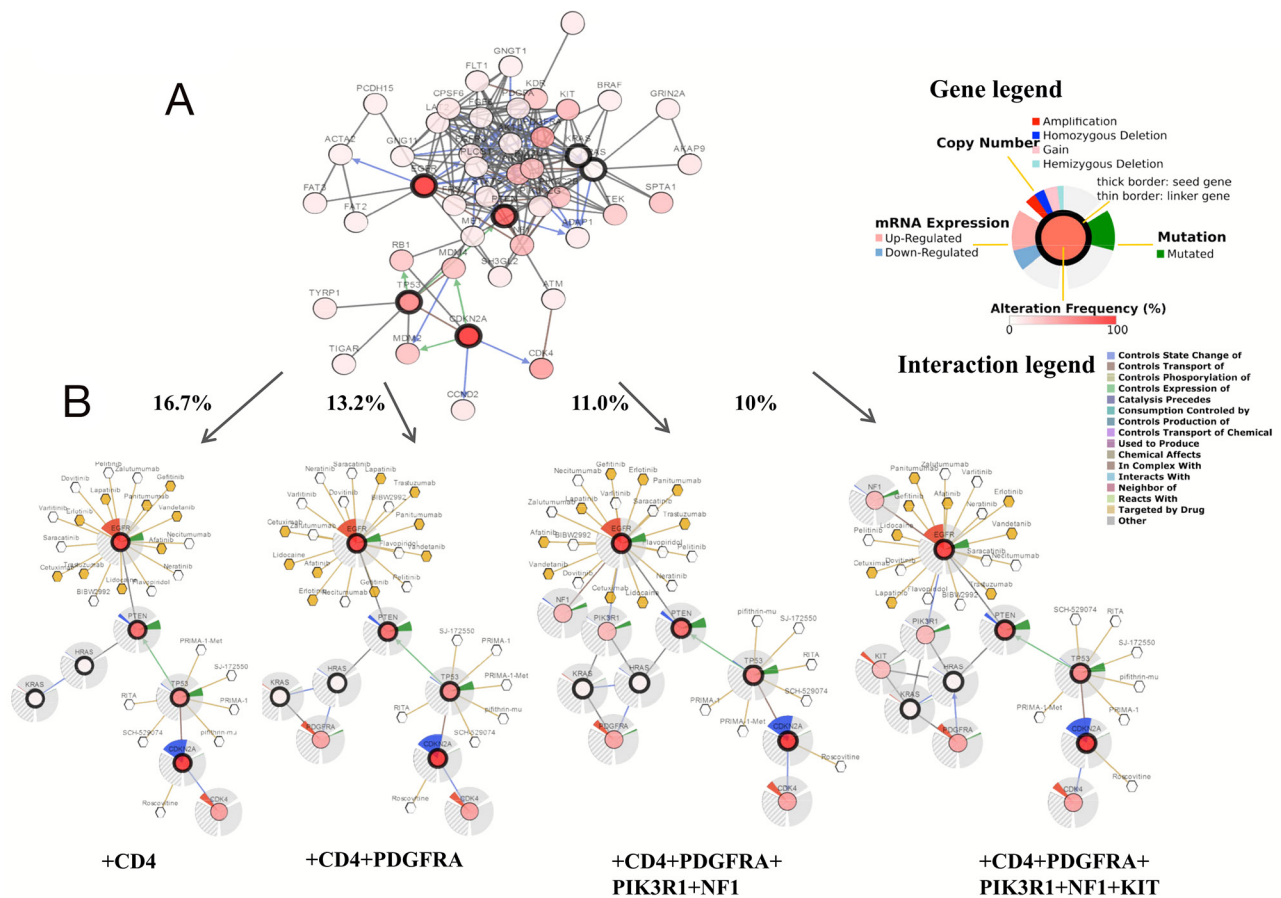


Figure 4: A visual display of gene networks connected to CDKN2A/EGFR/HRAS/KRAS/PTEN/TP53 in glioblastoma multiforme (GBM) (based on the study of Brennan CW, Cell, 2013). (A) Cross-cancer alteration summary for CDKN2A/EGFR/HRAS/KRAS/PTEN/TP53 mined from the cBioPortal for Cancer Genomics. Six selected genes and temozolomide-linked genes were applied as seed genes (indicated with thick black border) to automatically harvest all other genes identified as altered in GBM (data taken from the Brennan CW study, Cell, 2013). (B) Neighboring genes connected to the 6 query genes, filtered by alterations (%). Multidimensional genomic information and drug administration for a specific gene are exhibited for the seed genes (CDKN2A, EGFR, HRAS, KRAS, PTEN and TP53). Darker red indicates increased frequency of alteration (defined as a mutation, homozygous deletion and copy number amplification). The filters used involved the highest genomic alteration frequency within the selected GBM study in addition to the query genes.

marker for poor prognosis in LGG [38, 39]. However, the expression levels of other genes showed no statistically significant associations with survival in either LGG or GBM. Therefore, more large, multicenter clinical trials are urgently required to investigate the association between the expression of TMZ-related genes and provide a molecular biomarker for prognosis in patients with glioma.

In summary, the query of publicly available computational databases may significantly advance research by (1) unraveling the critical role of TMZ in glioma and revealing the mechanisms of glioma. Furthermore, by providing a deeper understanding of TMZ, the current analysis can assist in TMZ bio-curation, and new biological experimental designs will significantly accelerate glioma biology research. (2) This approach should facilitate early disease diagnosis and improve the accuracy of disease prognosis. The candidate genes identified by STRING, WebGestalt and the ClueGO database may facilitate the interpretation of genomic results and thus provide

information useful for guiding research. However, several challenges remain for us to investigate and solve: (1) In addition to the glioma pathway, there are other cancer pathways identified by WebGestalt and ClueGO, such as bladder cancer and non-small-cell lung cancer. Therefore, whether the connectivity map shown in this paper to exist between TMZ and glioma can be extended to other solid tumors remains to be investigated. (2) The roles of drug targets detected by STRING must be further explored in the glioma pathway and in other signaling pathways under TMZ control. (3) Genomic alterations in the LGG samples differed from those in the GBM cases; these alterations may play a role in transition from low-grade glioma to high-grade glioma. Therefore, these genomic differences between LGG and GBM can be used to direct future research with reasonable experimental feasibility based on a functional network analysis. Overall, this paper provides a simple yet flexible procedure to test and validate reasonable hypotheses regarding genetic alterations in glioma by

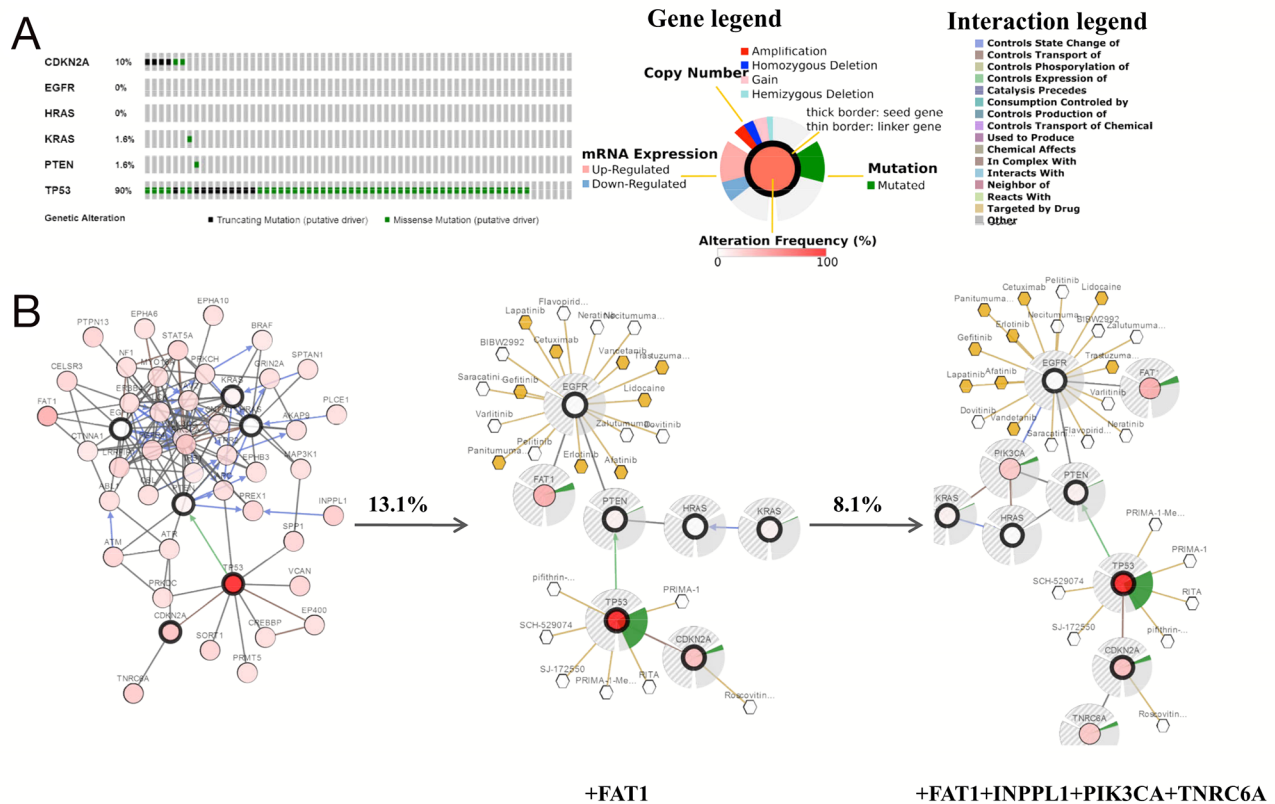


Figure 5: Genetic alterations and a visual display of the gene network connected to CDKN2A/EGFR/HRAS/KRAS/PTEN/TP53 in brain low-grade glioma (LGG) (based on the study of Johnson BE, Science 2014). (A) Oncopoint: a visual summary of genomic alterations across a set of LGG samples (data taken from Johnson BE, Science, 2014) based on a query of 6 genes (CDKN2A, EGFR, HRAS, KRAS, PTEN and TP53). (B) Neighboring genes connected to the 6 query genes, filtered by alterations (%).

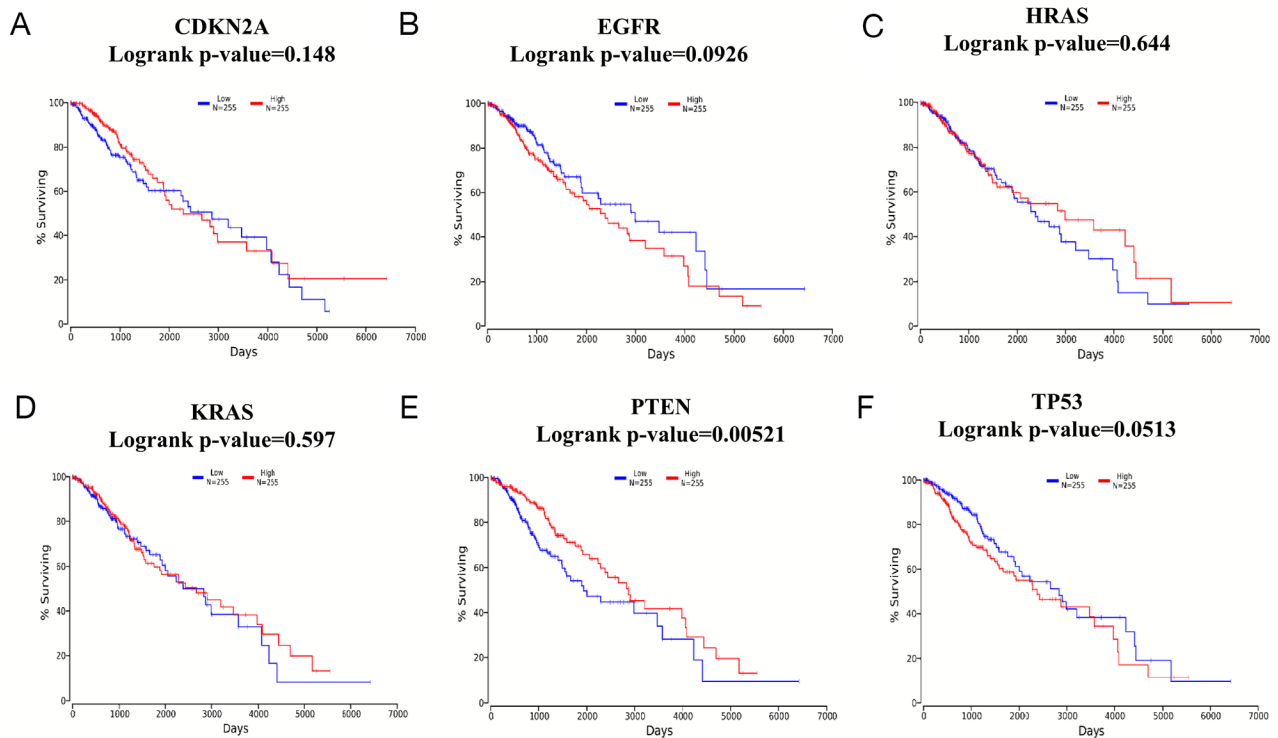


Figure 6: Survival analysis according to CDKN2A/EGFR/HRAS/KRAS/PTEN/TP53 mRNA expression in brain low-grade glioma (LGG) (A-F). A total of 510 LGG samples were included in the OncoLnc database and classified according to the meanvalue of the mRNA expression levels. Blue lines indicate lower levels of mRNA expression, while red lines indicate higher levels of mRNA expression.

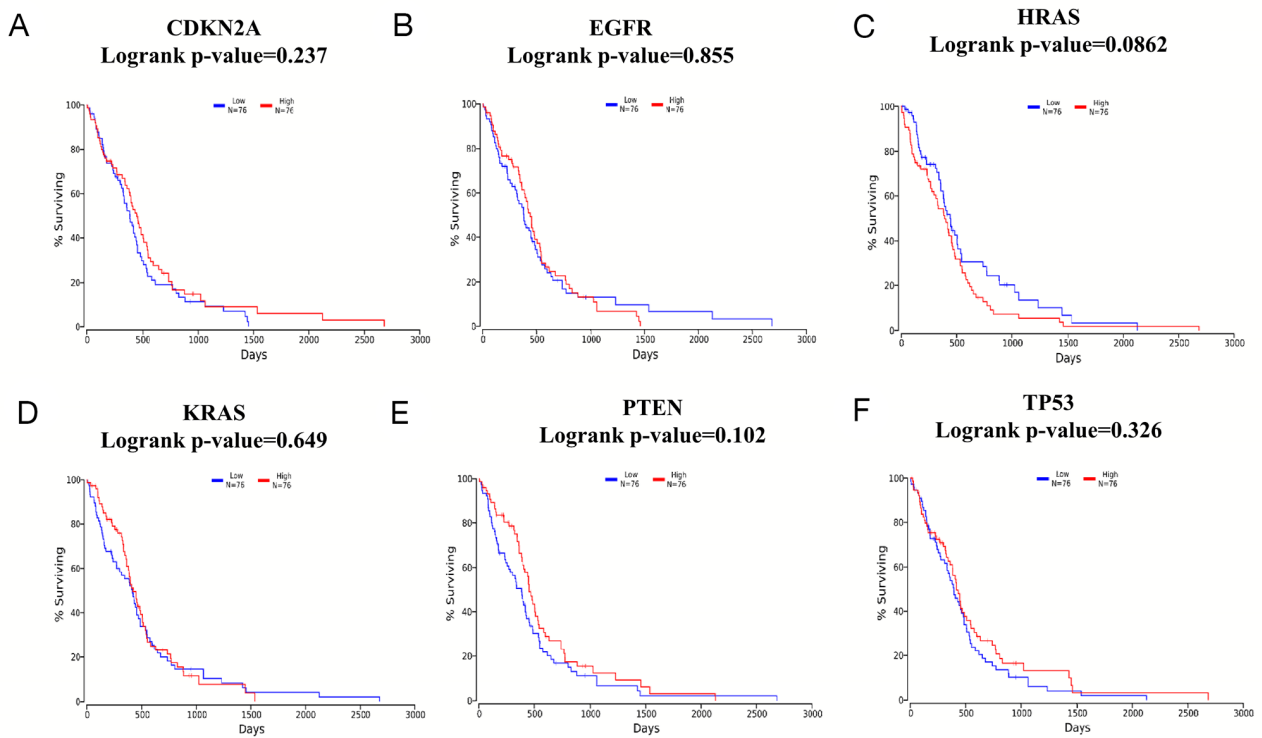


Figure 7: Survival analysis according to CDKN2A/EGFR/HRAS/KRAS/PTEN/TP53 mRNA expression in glioblastoma multiforme (GBM) (A-F). A total of 152 GBM samples were included in the OncoLnc database and classified according to the meanvalue of the mRNA expression level. Blue lines indicate lower mRNA expression levels, while red lines indicate higher mRNA expression levels.

applying available biological information, such as in BioGPS, to assist researchers in translating basic studies to clinical applications.

MATERIALS AND METHODS

Drug-target search

DrugBank is a web-based bioinformatic database containing comprehensive biochemical and pharmacological information about drugs, their pharmacological mechanisms and their targets [40, 41]. The tool contains over 4100 drug entries consisting of over 800 FDA-approved small molecules and biotech drugs as well as over 3200 experimental drugs. In this study, DrugBank was used to probe the category and indications for TMZ and the interaction between TMZ and its targets to provide insights regarding the TMZ target network.

Pathway enrichment analysis and network generation

STRING v10.0 is an online database tool to analyze PPIs for differentially expressed genes and provide interaction information predicted by comparative genomics [42]. In this study, human PPIs for target genes involved in three DNA repair systems were constructed using the STRING database. WebGestalt is a comprehensive web-based integrated data mining system that provides the maximum flexibility for functional enrichment analyses [43]. Biochemical pathways and functions related to the TMZ gene set were specifically queried with a KEGG pathway enrichment analysis in WebGestalt, and the top 15 pathways with an adjusted p -value < 0.01 were selected. ClueGO is a cytoscape plug in app that visualizes non-redundant biological terms for large clusters of gene sets in a functionally grouped network [44, 45]. In our study, the enrichment analysis of gene-GO terms and bio-pathways was statistically validated with the cytoscape plugin ClueGO + Cluepedia app.

Cancer genomics data linked to TMZ and glioma survival analysis

The cBioPortal is an open-access resource for interactively exploring multidimensional cancer genomics datasets [46, 47]. OncoLnc is a tool for studying survival correlations and downloading clinical data combined with expression profiles for mRNAs, miRNAs and long non-coding RNAs (lncRNAs) [48]. In this study, the cBioPortal database was used to analyze connectivity of TMZ-related genes across all glioma studies; these genes were then classified as altered or not altered in glioma samples. The altered genes were then further studied using a Kaplan-Meier analysis to evaluate glioma survival according to gene expression using OncoLnc.

ACKNOWLEDGMENTS

Not applicable.

CONFLICTS OF INTEREST

All authors declare that they have no conflicts of interest.

FUNDING

Not applicable.

REFERENCES

1. von Neubeck C, Seidlitz A, Kitzler HH, Beuthien-Baumann B, Krause M. Glioblastoma multiforme: emerging treatments and stratification markers beyond new drugs. *Br J Radiol.* 2015;88:20150354.
2. Weller M, Cloughesy T, Perry JR, Wick W. Standards of care for treatment of recurrent glioblastoma-are we there yet? *Neuro Oncol.* 2013;15:4-27.
3. Rosso L, Brock CS, Gallo JM, Saleem A, Price PM, Turkheimer FE, Aboagye EO. A new model for prediction of drug distribution in tumor and normal tissues: pharmacokinetics of temozolomide in glioma patients. *Cancer Res.* 2009;69:120-7.
4. Stevens MF, Hickman JA, Stone R, Gibson NW, Baig GU, Lunt E, Newton CG. Antitumor imidazotetrazines. 1. Synthesis and chemistry of 8-carbamoyl-3-(2-chloroethyl)imidazo[5,1-d]-1,2,3,5-tetrazin-4(3 H)-one, a novel broad-spectrum antitumor agent. *J Med Chem.* 1984;27:196-201.
5. McQuade JL, Posada LP, Lecagoonporn S, Cain S, Bassett RL Jr, Patel SP, Hwu WJ, Hwu P, Davies MA, Bedikian AY, Amaria RN. A phase I study of TPI 287 in combination with temozolomide for patients with metastatic melanoma. *Melanoma Res.* 2016;26:604-608.
6. Brandwein JM, Kassis J, Leber B, Hogge D, Howson-Jan K, Minden MD, Galarneau A, Pouliot JF. Phase II study of targeted therapy with temozolomide in acute myeloid leukaemia and high-risk myelodysplastic syndrome patients pre-screened for low O(6)-methylguanine DNA methyltransferase expression. *Br J Haematol.* 2014;167:664-70.
7. Seiter K, Katragadda S, Ponce D, Rasul M, Ahmed N. Temozolomide and cisplatin in relapsed/refractory acute leukemia. *J Hematol Oncol.* 2009;2:21.
8. Khanna KK, Lavin MF, Jackson SP, Mulhern TD. ATM, a central controller of cellular responses to DNA damage. *Cell Death Differ.* 2001;8:1052-65.
9. Dullea A, Marignol L. MGMT testing allows for personalised therapy in the temozolomide era. *Tumour Biol.* 2016;37:87-96.

10. Roos WP, Batista LF, Naumann SC, Wick W, Weller M, Menck CF, Kaina B. Apoptosis in malignant glioma cells triggered by the temozolomide-induced DNA lesion O6-methylguanine. *Oncogene*. 2007;26:186-97.
11. Kaina B, Ochs K, Grosch S, Fritz G, Lips J, Tomicic M, Dunkern T, Christmann M. BER, MGMT, and MMR in defense against alkylation-induced genotoxicity and apoptosis. *Prog Nucleic Acid Res Mol Biol*. 2001;68:41-54.
12. Marchesi F, Turriziani M, Tortorelli G, Avvisati G, Torino F, De Vecchis L. Triazene compounds: mechanism of action and related DNA repair systems. *Pharmacol Res*. 2007;56:275-87.
13. Jiricny J. The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol*. 2006;7:335-46.
14. Leguisamo NM, Gloria HC, Kalil AN, Martins TV, Azambuja DB, Meira LB, Saffi J. Base excision repair imbalance in colorectal cancer has prognostic value and modulates response to chemotherapy. *Oncotarget*. 2017; 8:54199-54214. <https://doi.org/10.18632/oncotarget.14909>.
15. Nanegrungsunk D, Onchan W, Chattipakorn N, Chattipakorn SC. Current evidence of temozolomide and bevacizumab in treatment of gliomas. *Neurol Res*. 2015;37:167-83.
16. Stupp R, Hegi ME, Mason WP, van den Bent MJ, Taphoorn MJ, Janzer RC, Ludwin SK, Allgeier A, Fisher B, Belanger K, Hau P, Brandes AA, Gijtenbeek J, et al. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol*. 2009;10:459-66.
17. Stupp R, Hegi ME, Gilbert MR, Chakravarti A. Chemoradiotherapy in malignant glioma: standard of care and future directions. *J Clin Oncol*. 2007;25:4127-36.
18. McFaline-Figueroa JL, Braun CJ, Stanciu M, Nagel ZD, Mazzucato P, Sangaraju D, Cerniauskas E, Barford K, Vargas A, Chen Y, Tretyakova N, Lees JA, Hemann MT, et al. Minor changes in expression of the mismatch repair protein MSH2 exert a major impact on glioblastoma response to temozolomide. *Cancer Res*. 2015;75:3127-38.
19. Nguyen SA, Stechishin OD, Luchman HA, Lun XQ, Senger DL, Robbins SM, Cairncross JG, Weiss S. Novel MSH6 mutations in treatment-naive glioblastoma and anaplastic oligodendroglioma contribute to temozolomide resistance independently of MGMT promoter methylation. *Clin Cancer Res*. 2014;20:4894-903.
20. Sonabend AM, Carminucci AS, Amendolara B, Bansal M, Leung R, Lei L, Realubit R, Li H, Karan C, Yun J, Showers C, Rothcock R, O J, Califano A, et al. Convection-enhanced delivery of etoposide is effective against murine proneural glioblastoma. *Neuro Oncol*. 2014;16:1210-9.
21. Wang LE, Bondy ML, Shen H, El-Zein R, Aldape K, Cao Y, Pudavalli V, Levin VA, Yung WK, Wei Q. Polymorphisms of DNA repair genes and risk of glioma. *Cancer Res*. 2004;64:5560-3.
22. Verreault M, Schmitt C, Goldwirt L, Pelton K, Haidar S, Levasseur C, Guehenec J, Knoff D, Labussiere M, Marie Y, Ligon AH, Mokhtari K, Hoang-Xuan K, et al. Preclinical efficacy of the MDM2 inhibitor RG7112 in MDM2-amplified and TP53 wild-type glioblastomas. *Clin Cancer Res*. 2016;22:1185-96.
23. Johnson BE, Mazor T, Hong C, Barnes M, Aihara K, McLean CY, Fouse SD, Yamamoto S, Ueda H, Tatsuno K, Asthana S, Jalbert LE, Nelson SJ, et al. Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. *Science*. 2014;343:189-93.
24. Ceccarelli M, Barthel FP, Malta TM, Sabedot TS, Salama SR, Murray BA, Morozova O, Newton Y, Radenbaugh A, Pagnotta SM, Anjum S, Wang J, Manyam G, et al. Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. *Cell*. 2016;164:550-63.
25. Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature*. 2008;455:1061-8.
26. Brennan CW, Verhaak RG, McKenna A, Campos B, Noushmehr H, Salama SR, Zheng S, Chakravarty D, Sanborn JZ, Berman SH, Beroukhim R, Bernard B, Wu CJ, et al. The somatic genomic landscape of glioblastoma. *Cell*. 2013;155:462-77.
27. Jones DT, Hutter B, Jager N, Korshunov A, Kool M, Warnatz HJ, Zichner T, Lambert SR, Ryzhova M, Quang DA, Fontebasso AM, Stutz AM, Hutter S, et al. Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. *Nat Genet*. 2013;45:927-32.
28. Tandon A, Schiff D. Therapeutic decision making in patients with newly diagnosed low grade glioma. *Curr Treat Options Oncol*. 2014;15:529-38.
29. Kumthekar P, Raizer J, Singh S. Low-grade glioma. *Cancer Treat Res*. 2015;163:75-87.
30. Naumann SC, Roos WP, Jost E, Belohlavek C, Lennerz V, Schmidt CW, Christmann M, Kaina B. Temozolomide- and fotemustine-induced apoptosis in human malignant melanoma cells: response related to MGMT, MMR, DSBs, and p53. *Br J Cancer*. 2009;100:322-33.
31. Custodio AC, Almeida LO, Pinto GR, Santos MJ, Almeida JR, Clara CA, Rey JA, Casartelli C. Variation in DNA repair gene XRCC3 affects susceptibility to astrocytomas and glioblastomas. *Genet Mol Res*. 2012;11:332-9.
32. Kiuru A, Lindholm C, Heinavaara S, Ilus T, Jokinen P, Haapasalo H, Salminen T, Christensen HC, Feychting M, Johansen C, Lonn S, Malmer B, Schoemaker MJ, et al. XRCC1 and XRCC3 variants and risk of glioma and meningioma. *J Neurooncol*. 2008;88:135-42.
33. Zheng H, Ying H, Yan H, Kimmelman AC, Hiller DJ, Chen AJ, Perry SR, Tonon G, Chu GC, Ding Z, Stommel JM, Dunn KL, Wiedemeyer R, et al. p53 and Pten control neural and glioma stem/progenitor cell renewal and differentiation. *Nature*. 2008;455:1129-33.

34. Ghasimi S, Wibom C, Dahlin AM, Brannstrom T, Golovleva I, Andersson U, Melin B. Genetic risk variants in the CDKN2A/B, RTEL1 and EGFR genes are associated with somatic biomarkers in glioma. *J Neurooncol.* 2016;127:483-92.
35. Chow LM, Endersby R, Zhu X, Rankin S, Qu C, Zhang J, Broniscer A, Ellison DW, Baker SJ. Cooperativity within and among Pten, p53, and Rb pathways induces high-grade astrocytoma in adult brain. *Cancer cell.* 2011;19:305-16.
36. Guillamo JS, de Bouard S, Valable S, Marteau L, Leuraud P, Marie Y, Poupon MF, Parienti JJ, Raymond E, Peschanski M. Molecular mechanisms underlying effects of epidermal growth factor receptor inhibition on invasion, proliferation, and angiogenesis in experimental glioma. *Clin Cancer Res.* 2009;15:3697-704.
37. Bienkowski M, Piaskowski S, Stoczynska-Fidelus E, Szybka M, Banaszczyk M, Witusik-Perkowska M, Jesien-Lewandowicz E, Jaskolski DJ, Radomiak-Zaluska A, Jesionek-Kupnicka D, Sikorska B, Papierz W, Rieske P, et al. Screening for EGFR amplifications with a novel method and their significance for the outcome of glioblastoma patients. *PLoS One.* 2013;8:e65444.
38. McBride SM, Perez DA, Polley MY, Vandenberg SR, Smith JS, Zheng S, Lamborn KR, Wiencke JK, Chang SM, Prados MD, Berger MS, Stokoe D, Haas-Kogan DA. Activation of PI3K/mTOR pathway occurs in most adult low-grade gliomas and predicts patient survival. *J Neurooncol.* 2010;97:33-40.
39. Sabha N, Knobbe CB, Maganti M, Al Omar S, Bernstein M, Cairns R, Cako B, von Deimling A, Capper D, Mak TW, Kiehl TR, Carvalho P, Garrett E, et al. Analysis of IDH mutation, 1p/19q deletion, and PTEN loss delineates prognosis in clinical low-grade diffuse gliomas. *Neuro Oncol.* 2014;16:914-23.
40. Law V, Knox C, Djoumbou Y, Jewison T, Guo AC, Liu Y, Maciejewski A, Arndt D, Wilson M, Neveu V, Tang A, Gabriel G, Ly C, et al. DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res.* 2014;42:D1091-7.
41. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, Chang Z, Woolsey J. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.* 2006;34:D668-72.
42. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 2015;43:D447-52.
43. Wang J, Vasaikar S, Shi Z, Greer M, Zhang B. WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Res.* 2017.
44. Bindea G, Galon J, Mlecnik B. CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. *Bioinformatics.* 2013;29:661-3.
45. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pages F, Trajanoski Z, Galon J. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics.* 2009;25:1091-3.
46. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C, Schultz N. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal.* 2013;6:pl1.
47. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2012;2:401-4.
48. Anaya J. OncoLnc: linking TCGA survival data to mRNAs, miRNAs, and lncRNAs. *PeerJ.* 2016;2:e67.