*CORRESPONDENCE
Tonia Carter,
✉ carter.tonia@marshfieldresearch.org

Gene signatures associated with prognosis and chemotherapy resistance in glioblastoma treated with temozolomide

Tonia Carter[1]*, Robert K. Valenzuela[1],
Srinivasulu Yerukala Sathipati[1] and Rafael Medina-Flores[2]

[1]Center for Precision Medicine Research, Marshfield Clinic Research Institute, Marshfield, WI, United States, [2]Department of Pathology (Neuropathology), Marshfield Clinic, Marshfield, WI, United States

Background: Glioblastoma (GBM) prognosis remains extremely poor despite standard treatment that includes temozolomide (TMZ) chemotherapy. To discover new GBM drug targets and biomarkers, genes signatures associated with survival and TMZ resistance in GBM patients treated with TMZ were identified.

Methods: GBM cases in The Cancer Genome Atlas who received TMZ (n = 221) were stratified into subgroups that differed by median overall survival (mOS) using network-based stratification to cluster patients whose somatic mutations affected genes in similar modules of a gene interaction network. Gene signatures formed from differentially mutated genes in the subgroup with the longest mOS were used to confirm their association with survival and TMZ resistance in independent datasets. Somatic mutations in these genes also were assessed for an association with OS in an independent group of 37 GBM cases.

Results: Among the four subgroups identified, subgroup four (n = 71 subjects) exhibited the longest mOS at 18.3 months (95% confidence interval: 16.2, 34.1; p = 0.0324). Subsets of the 86 genes that were differentially mutated in this subgroup formed 20-gene and 8-gene signatures that predicted OS in two independent datasets (Spearman's rho of 0.64 and 0.58 between actual and predicted OS; $p <$ 0.001). Patients with mutations in five of the 86 genes had longer OS in a small, independent sample of 37 GBM cases, but this association did not reach statistical significance (p = 0.07). Thirty-one of the 86 genes formed signatures that distinguished TMZ-resistant GBM samples from controls in three independent datasets (area under the curve ≥ 0.75). The prognostic and TMZ-resistance signatures had eight genes in common (*ANG*, *BACH1*, *CDKN2C*, *HMGA1*, *IFI16*, *PADI4*, *SDF4*, and *TP53INP1*). The latter three genes have not been associated with GBM previously.

Conclusion: *PADI4*, *SDF4*, and *TP53INP1* are novel therapy and biomarker candidates for GBM. Further investigation of their oncologic functions may provide new insight into GBM treatment resistance mechanisms.

KEYWORDS

chemotherapy resistance, glioblastoma, oncogene, somatic variant, temozolomide

1 Introduction

Glioblastoma (GBM) has a low median overall survival (mOS) of approximately 15 months even with the standard first-line therapy of surgery followed by radiotherapy and temozolomide (TMZ) chemotherapy (Stupp et al., 2005). This treatment is not curative as the tumor eventually reoccurs (Wen et al., 2020). TMZ's cytotoxicity is due to its modification of genomic DNA to generate O^6-methylguanine, which is mispaired with thymine during DNA replication (Newlands et al., 1997). Repetitive futile cycles of the mismatch repair pathway to repair the mismatched base pair produce DNA strand breaks, inducing cell cycle arrest and eventually leading to apoptosis (Karran et al., 1993; Newlands et al., 1997). A DNA repair enzyme encoded by the O^6-methylguanine-DNA methyltransferase (MGMT) gene repairs O^6-methylguanine by removing the TMZ-added methyl group, counteracting the cytotoxic effects of TMZ (Singh et al., 2021). The epigenetic silencing of MGMT by the methylation of CpG sites in its promoter is associated with longer OS in GBM patients treated with TMZ (Hegi et al., 2005), and MGMT promoter methylation status is both a known prognostic biomarker in GBM and a predictive biomarker of response to TMZ treatment (Weller et al., 2009). MGMT promoter methylation is present in more than 90% of GBM patients who have longer-term survival compared with approximately 30% of all GBM patients (Stupp et al., 2014). However, the signaling pathways that control MGMT activity in GBM are incompletely understood, and GBM has a dismal prognosis irrespective of the MGMT promoter methylation status (Wick et al., 2014). Sustained efforts have been made to develop and test new drugs for GBM, but GBM has remained refractory to most new therapies (Khasraw et al., 2022; van Solinge et al., 2022).

The identification of new drug targets has the potential to increase therapeutic options that improve outcomes in GBM. GBM tumor progression is promoted by the genetically heterogeneous nature of the tumor (Brennan et al., 2013), the invasive growth of tumor cells (Venkataramani et al., 2022), the low immunogenicity of the tumor (Pearson et al., 2020), and mechanisms of treatment resistance (Singh et al., 2021). Taken together, these characteristics draw attention to the biological complexity of treatment response in GBM. To begin to identify prospective drug targets for GBM, this hypothesis-generating study aimed to uncover gene signatures associated with survival in GBM patients who received TMZ because the genes that make up those signatures have the potential to provide insight into the numerous biological pathways that influence patient outcomes, facilitate the discovery of targetable components within those pathways, and also serve as prognostic biomarkers for GBM. The first step of the approach was to stratify GBM patients into subgroups that differed by mOS using network-based stratification (NBS) (Hofree et al., 2013), an algorithm that clusters together patients with genetic variants in similar regions of a gene interaction network. Next, the genes that were differentially mutated in the subgroup(s) with a mOS longer than that reported for standard TMZ chemotherapy were used to determine biological pathways associated with survival. Finally, the group of differentially mutated genes was used to identify gene signatures associated with OS and TMZ resistance in independent GBM datasets.

2 Materials and methods

2.1 Study subjects, samples, and datasets

To identify gene networks associated with mOS after TMZ treatment, public data on simple somatic variants and clinical characteristics from The Cancer Genome Atlas (TCGA) were analyzed for the 221 GBM cases (out of a total of 606 TCGA GBM cases) who had received concomitant and/or adjuvant TMZ chemotherapy and had data available on OS and simple somatic variants from primary tumor (the numbers of included and excluded subjects are shown in Supplementary Figure S1). The simple somatic variants included single nucleotide variants and small insertions/deletions and multiple base substitutions ≤ 200 base pairs. Datasets GSE108474 and GSE7696 from the Gene Expression Omnibus data repository were used to confirm the association between the mutated genes and OS. Tumor microarray gene expression data were available for 81 GBM patients who received TMZ in the GSE108474 dataset (Gusev et al., 2018) and 43 GBM patients who received TMZ in the GSE7696 dataset (Murat et al., 2008). To determine whether any of the genes that were correlated with OS were also mutated in an independent group of GBM cases treated with TMZ, somatic variant data were generated from the treatment-naïve, primary tumors of 37 patients (out of a total of 118 patients) who were diagnosed with GBM from 2004 to 2012 at Marshfield Clinic Health Systems (MCHS) in Marshfield, Wisconsin, and had received concomitant and/or adjuvant TMZ chemotherapy (Carter et al., 2018). Of the 118 patients, DNA could be extracted from the archived tumor specimens of 74 patients. Exome sequencing could be performed successfully on 37 of the 74 DNA samples. To determine whether the genes correlated with OS could also discriminate between TMZ-resistant and control GBM cell lines or tumors, gene expression data in the GSE151680, GSE193957, and GSE145128 datasets were analyzed. GSE151680 had RNA-seq data for three TMZ-resistant and three control samples for each of two GBM cell lines (U87 and U251); GSE193957 had microarray data for three TMZ-resistant and three control samples from the U87 cell line (Choo et al., 2023); GSE145128 had microarray data for treatment-naïve tumor and the matched tumor that recurred after TMZ treatment in seven GBM patients (Kebir et al., 2023).

2.2 Network-based stratification analysis

NBS, performed using pyNBS (Huang et al., 2018), was used to generate somatic mutation profiles from TCGA GBM somatic variant data and stratify those profiles into subgroups associated with OS. The NBS algorithm uses network propagation to integrate the somatic mutation profiles with a gene interaction network and non-negative matrix factorization to cluster the integrated profiles into a pre-determined number of groups from $k = 2$ up to $k = 12$. For each pre-determined number of clusters, network propagation and non-negative matrix factorization were applied to 1,000 different subsets of the TCGA GBM somatic variant data, with each subset containing 80% of the patients and 80% of the mutated genes sampled at random without replacement, followed by consensus clustering of the aggregated results of the 1,000 subsets into a single

cluster result. The STRING database (version 11.5) (Szklarczyk et al., 2023) of known and predicted protein-protein interactions was the source of the human gene interaction network for NBS. Only interactions having a confidence score > 0.7 (indicating high-confidence interactions; 16,795 nodes and 252,013 edges out of 19,385 nodes and 5,969,249 edges) were input to NBS. The protein nodes represented 16,127 genes. The alpha tuning parameter, which controls the distance that a mutation signal can diffuse through the gene network during propagation, was set to 0.7 for network propagation using the STRING network, as recommended in the original NBS report. Default settings were accepted for other NBS parameters. Kaplan-Meier analysis and the log rank test (Schober and Vetter, 2018) were implemented in pyNBS to generate survival curves for each of the NBS subgroups within k consensus clusters and determine whether NBS-assigned clusters are associated with OS. Cox proportional hazards regression (Cox, 1972) was also performed to assess whether NBS subgroups were associated with OS independent of other clinical factors known to predict survival in GBM, including age at diagnosis (years), type of surgery at initial pathologic diagnosis (biopsy/tumor resection/other), and radiation therapy (yes/no/unknown).

2.3 Identification of differentially mutated genes in NBS subgroups

To identify the gene interaction network regions that contribute the most to distinguishing the somatic mutation profiles of tumors in different NBS subgroups (k = 4 clusters), the non-parametric Significance Analysis of Microarrays (SAM) (Tusher et al., 2001) method was applied to the integrated mutation profile (consisting of 16,127 genes) that was generated after network propagation, as described in the original NBS report. Each subgroup, in turn, was compared with all other subgroups to detect genes with significantly different network-smoothed mutation states in that subgroup. SAM (version 5.0) was run using the SAMR shiny app with data type chosen as array - two class unpaired, median center the arrays as "yes", statistical test as Wilcoxon rank sum test, number of permutations as 1,000, and random seed as the default value. The threshold for statistical significance was determined by the value of a tuning parameter, delta, which is user-selected based on the false discovery rate. A fold-change ≥ 2 and false discovery rate < 0.05 were selected for this analysis.

2.4 Gene Ontology term enrichment

To determine the Gene Ontology (GO) Biological Process categories that were over- or under-represented among a set of genes compared to a reference list of genes, enrichment analysis was performed directly from the GO website (Gene Ontology, 2021), with the use of the PANTHER over-representation test (Mi et al., 2019b) from the PANTHER gene classification resource (Mi et al., 2019a). The GO database version used was DOI: 10.5281/zenodo.6799722, released 1 July 2022, and the list of background genes comprised the 16,127 genes that formed the human gene interaction network for NBS analysis. The p-values were calculated by Fisher's exact test with false discovery rate correction, and a corrected p-value < 0.05 was considered statistically significant.

2.5 Predicting survival and differentiating between temozolomide-resistant and control samples

We utilized the least absolute shrinkage and selection operator (LASSO) (Tibshirani, 1996) technique to estimate the OS of GBM patients. LASSO employs L1 regularization, a process that facilitates the regularization of certain coefficients that contribute to the output evaluation. This regularization technique effectively aids in the feature selection process. For the determination of the tuning parameter λ, we selected the minimum λ after conducting 100 iterations of 10-fold cross-validation (10-CV). In assessing the performance of the model, we employed Spearman's rank correlation coefficient ρ) (Schober et al., 2018) and the mean absolute error (Willmott and Matsuura, 2005) as measurement metrics.

For the prediction of TMZ resistance in patients with GBM, we employed several standard machine learning methods available in Weka (Hall et al., 2009), including Naïve Bayes, simple logistic, sequential minimal optimization, random forest, and J48. The assessment of these methods was conducted based on evaluation metrics such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC).

2.6 Tumor DNA extraction

For MCHS patients, GBM tumor DNA was extracted from formalin-fixed and paraffin-embedded (FFPE) tissue for exome sequencing. A matched germline DNA sample was not available. Sections with a thickness of 10 μM were cut from each tissue block and, after discarding the first three sections, three to four sections, depending on tissue size, were sampled in triplicate and placed in micro-centrifuge tubes for DNA extraction. Tissues were extracted within 12 h of cutting. DNA extractions were performed using the GeneRead™ DNA FFPE kit (Qiagen, Valencia, CA) according to the manufacturer's recommendations. The quality (260/280 ratio) of each DNA sample was assessed using a NanoDrop™ spectrophotometer (ThermoFisher Scientific, Waltham, MA). Initial DNA quantity was determined using either a BR (broad range) or HS (high sensitivity) kit on a Qubit 2 Fluorometer (ThermoFisher Scientific). DNA integrity (fragmentation) and final concentration were determined with Genomic DNA ScreenTape on a TapeStation 2,200 (Agilent Technologies, Santa Clara, CA) prior to sample pooling. DNA from triplicate samples were pooled as necessary to obtain the appropriate concentration and volume for downstream testing.

2.7 Exome sequencing

DNA library preparation and exome sequencing were performed by Admera Health (South Plainfield, NJ). DNA sample quality was assessed by Agilent DNA 6000 Nano Reagent on an Agilent 2,100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) and quantified by Qubit DNA HS assay (ThermoFisher Scientific). Library preparation for exome sequencing was performed with KAPA HyperPrep kits (Roche) and IDT xGen

indexes (Illumina Inc., San Diego, CA) following the manufacturer's instructions. Samples were pooled and sequenced on an Illumina NovaSeq S4 sequencer for 150 base pairs read length in paired-end mode, with an output of 260 million reads per sample.

2.8 Exome sequence data processing

The quality of sequencing reads were checked with FastQC (version 0.11.2) before the reads were aligned to a human genome reference sequence (GRCh38. d1. vd1) using the Burrows-Wheeler Aligner (version 0.7.5a) (Li and Durbin, 2009). Duplicate reads generated by polymerase chain reaction (PCR) were marked and removed using MarkDuplicates in Picard (version 1.73), and local realignment of reads followed by base quality recalibration were performed using the Genome Analysis Toolkit (GATK version 4.1.4.1) (Van der Auwera et al., 2013). Variants were called from recalibrated reads using the Mutect2 algorithm in GATK (McKenna et al., 2010). To improve the quality of variant calls, variants derived from normal tissue, known as a Panel of Normals, was employed as a reference by Mutect2 to detect sequencing and alignment artifacts (Cibulskis et al., 2013). The Panel of Normals for a tumor-only variant calling pipeline, consisting of variants from approximately 5,000 TCGA exome sequencing samples of individuals without cancer, was downloaded from the online Genomic Data Commons data portal for input to Mutect2 (Zhang et al., 2021). The CollectSequencingArtifactMetrics and FilterByOrientationBias commands in GATK were used to filter called variants for sequence artifacts that arise from the deamination of cytosines by formaldehyde in FFPE tissue or the oxidation of guanine to 8-oxoguanine (Costello et al., 2013). Next, PureCN (version 1.22.2) was used to classify variants by somatic/germline status based on an assessment of tumor purity, ploidy, loss of heterozygosity, and copy number (Oh et al., 2020). Variants that were flagged as unreliable or were assigned a somatic posterior probability of less than 0.8 by PureCN were removed. Variants not located in the target region or supported by fewer than five reads were also removed. Next, filtered variants were annotated using Variant Effect Predictor (version 105) (McLaren et al., 2016). Only rare variants, defined as having an alternate allele frequency < 1% in the GnomAD database (Karczewski et al., 2020), were considered as somatic variants.

2.9 *MGMT* promoter methylation

DNA methylation assays were performed by EpigenDx (Hopkinton, MA) using their ASY470-FS2 protocol, which involved bisulfite sequencing of eight CpG sites in the differentially methylated region 2 (Malley et al., 2011) of the *MGMT* promoter. DNA was available from 31 of the 37 MCHS patients to perform the assays. For each patient, 300 ng of GBM tumor DNA was bisulfite treated using the EZ DNA Methylation kit (Zymo Research, Inc., Irvine, CA). Bisulfite-treated DNA was purified according to the manufacturer's protocol and eluted to a final volume of 46 µL. PCRs were performed using 1 µL of bisulfite treated DNA and 0.2 µM of each primer. One primer was biotin-labeled and HPLC purified in order to purify the final PCR product using sepharose beads. PCR product was bound to

Streptavidin Sepharose High Performance beads (GE Healthcare Life Sciences), after which the immobilized PCR products were purified, washed, denatured with a 0.2 µM NaOH solution, and rewashed using the Pyrosequencing Vacuum Prep Tool (Pyrosequencing, Qiagen), as per the manufacturer's protocol. Next, 0.5 µM of sequencing primer was annealed to the purified single-stranded PCR products. 10 µL of the PCR products were sequenced by pyrosequencing on the PSQ96 HS System (Pyrosequencing, Qiagen), following the manufacturer's instructions. The methylation status of each CpG site was determined individually as an artificial C/T single nucleotide polymorphism using QCpG software (Pyrosequencing, Qiagen). The methylation level at each CpG site was calculated as the percentage of the methylated alleles divided by the sum of all methylated and unmethylated alleles. The mean methylation level was calculated using methylation levels of all measured CpG sites within the targeted region. Each experiment included non-CpG cytosines as internal controls to detect incomplete bisulfite conversion of the input DNA. In addition, a series of unmethylated and methylated DNA were included as controls in each PCR. PCR bias testing was performed by mixing unmethylated control DNA with *in vitro* methylated DNA at different ratios (0%, 5%, 10%, 25%, 50%, 75%, and 100%), followed by bisulfite modification, PCR, and pyrosequencing analysis. A mean methylation level > 7% was used as the cutoff for scoring samples as unmethylated or methylated (Kristensen et al., 2016).

2.10 Statistical methods

The median and 95% confidence interval for OS and time to progression were determined by Kaplan-Meier analysis, and comparisons of OS or time to progression among subgroups were performed using the log rank test, which generated *p*-values for the comparisons (Schober and Vetter, 2018). Multivariate analyses of OS were performed using Cox proportional hazard regression (Cox, 1972). The correlation between OS and time to progression and between actual and predicted OS was determined by Spearman's rank correlation coefficient (Spearman's ρ) (Schober et al., 2018). The test statistic for Spearman's rank correlation coefficient was calculated as:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

with *r* being the sample correlation coefficient and *n* being the number of subjects with no missing data for the pair of variables (Kendall and Gibbons, 1990). The two-tailed *p*-value is 2 x P (T > t) where T follows a Student's t distribution with *n* – 2 degrees of freedom. The proportion of subjects with a methylated *MGMT* promoter was compared among the subgroups in the TCGA dataset using the Chi-squared test and between the two subgroups in the smaller, independent dataset of 37 GBM cases using Fisher's exact test. Because this was a hypothesis-generating study, *p*-values were not used for hypothesis testing but for identifying genes that warrant further exploration as GBM biomarkers or drug targets. The results and *p*-values are to be considered explorative, and no threshold level of statistical
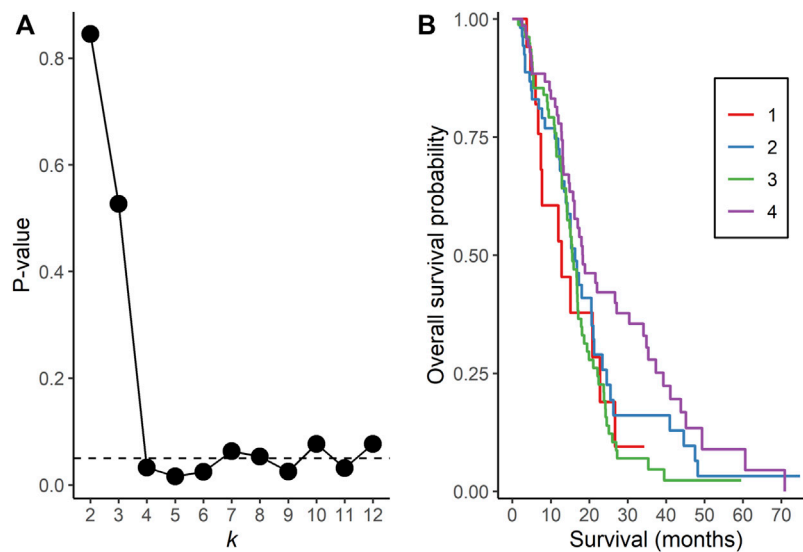
FIGURE 1
Network Based Stratification analysis of TCGA GBM tumors. **(A)** *p*-values from the log rank test comparing overall survival time between assigned subgroups for cluster sizes *k* = 2 to 12. The horizontal, dotted line represents a *p*-value of 0.05. **(B)** Kaplan-Meier curves of overall survival in each subgroup for *k* = 4 clusters.

significance should be fixed. Correction for multiple testing was not performed. Survival analysis, Cox proportional hazards regression, Spearman's rank correlation, the Chi-squared test, and Fisher's exact test were performed using the R software program (R-Core, 2021).

In LASSO models (Tibshirani, 1996) used for predicting OS based on gene signatures, mean absolute error was used as a measure of the average magnitude of the error produced by a LASSO model. Mean absolute error (Willmott and Matsuura, 2005) was determined by calculating the magnitude of the difference between predicted OS and observed OS for each subject (absolute error for each subject) and taking the average of the absolute errors for the entire group of subjects. Therefore, the lower the value of mean absolute error for a LASSO model, the smaller the differences between OS predicted by the model and observed OS. To assess the performance of machine learning methods to classify TMZ resistant and control samples using gene signatures, values of accuracy, sensitivity, specificity, and the AUC were determined for each machine learning method (Kourou et al., 2015). Accuracy indicated the proportion of correct predictions made by a machine learning method and was calculated as the number of correct predictions/ total number of predictions. Sensitivity measured how well a machine learning method could correctly identify TMZ resistant samples and was calculated as True Positives/(True Positives + False Negatives). Specificity measured how well a machine learning method could correctly identify control samples and was calculated as True Negatives/(True Negatives + False Positives). The AUC measured the ability of a machine learning method to distinguish between TMZ resistant and control samples. True positives were TMZ resistant samples that were correctly predicted to be TMZ resistant; false positives were control samples that were falsely predicted to be TMZ

resistant; true negatives were control samples that were correctly predicted to be controls; false negatives were TMZ resistant samples that were falsely predicted to be controls.

3 Results

3.1 Glioblastoma stratification by somatic mutation profiles and identification of genes associated with survival

NBS analysis of somatic mutation data for the 221 TCGA GBM tumors indicated that *k* = 4 or more clusters were associated with OS based on the log rank test (Figure 1A), and further analyses to identify the gene interaction subnetworks of relevance to OS in TMZ-treated GBM were conducted using the NBS subgroups within *k* = 4 clusters. mOS, as determined by Kaplan-Meier analysis, was highest in subgroup 4 and lowest in subgroup 1 (*p* = 0.0324; Table 1; Figure 1B). Cox proportional hazards regression using data for all 221 subjects showed that subgroups 1 and 3 (compared with subgroup 4 as a reference) were associated with an increased risk of death, independent of other known clinical prognostic factors (Table 2). When the regression was restricted to the 206 subjects without *IDH1/2* variants, similar results were observed (Table 2). Time to progression after initial treatment, an important indicator of GBM treatment efficacy, and thus, of TMZ resistance (Han et al., 2014), was also compared among the four subgroups. Data were available on time to progression for 215 subjects (17, 51, 78, and 68 in subgroups 1, 2, 3, and 4, respectively), and time to progression was positively correlated with overall survival (Spearman's rank correlation = 0.60, $p < 3.15 \times 10^{-22}$; Supplementary Figure S2). The median and 95% confidence interval for time to progression was 6.0 (3.5, 7.7) months for subgroup 1, 7.5 (6.0, 9.2) months for subgroup

TABLE 1 Overall survival (months) based on Kaplan-Meier analysis of four TCGA GBM subgroups as assigned by NBS.

Subgroup	N	Number of deaths	Median overall survival (95% confidence interval)	P^a
1	17	12	12.8 (7.5, NA)	0.0324
2	54	39	16.3 (14.0, 21.3)	
3	79	59	15.6 (14.1, 17.9)	
4	71	47	18.3 (16.2, 34.1)	

NA, not available.
[a]Comparison among all four groups using the log-rank test.

TABLE 2 Cox proportional hazards regression of TCGA GBM.

Characteristic	All subjects ($N = 221$)		Subjects without *IDH1/2* variants ($N = 206$)	
	Odds ratio (95% confidence interval)	P	Odds ratio (95% confidence interval)	P
NBS subgroup				
4	Reference	1.0000	Reference	1.0000
1	2.05 (1.07, 3.93)	0.0308	2.06 (1.04, 4.09)	0.0379
2	1.23 (0.80, 1.89)	0.3531	1.18 (0.76, 1.83)	0.4609
3	1.60 (1.07, 2.37)	0.0205	1.54 (1.02, 2.31)	0.0380
Age at diagnosis (years)	1.02 (1.01, 1.03)	0.0068	1.02 (1.00, 1.03)	0.0564
Type of surgery at initial diagnosis				
Biopsy	Reference	1.0000	Reference	1.0000
Resection	1.19 (0.78, 1.80)	0.4291	1.21 (0.79, 1.86)	0.3890
Other	12.06 (1.48, 98.12)	0.0200	13.83 (1.68, 114.09)	0.0147
Radiation therapy				
Yes	Reference	1.0000	Reference	1.0000
No	2.70 (0.81, 9.02)	0.1056	2.81 (0.84, 9.23)	0.0948
Unknown	0.35 (0.05, 2.54)	0.3003	0.33 (0.05, 2.37)	0.2686

2, 5.5 (4.9, 6.9) months for subgroup 3, and 7.1 (5.7, 9.3) months for subgroup 4, as determined by Kaplan-Meier analysis. $p = 0.0103$ for the comparison based on the log-rank test, indicating that time to progression was significantly different among the four subgroups, with subgroups 1 and 3 having a shorter median time to progression (suggestive of greater TMZ resistance) than subgroup 4.

Fifty (22.6%) of the 221 subjects had missing data on *MGMT* methylation status; therefore, this variable was not included as a covariate in the regression model. Of the 171 subjects with *MGMT* methylation status available, 7 (46.7%) of 15 in subgroup 1, 18 (43.9%) of 41 in subgroup 2, 24 (37.5%) of 64 in subgroup 3, and 34 (66.7%) of 51 in subgroup 4 had a methylated *MGMT* promoter ($p = 0.0169$; Chi-squared test). Therefore, each subgroup had a mixture of subjects with and without *MGMT* promoter methylation, and the subgroup with the longest mOS had a greater proportion of subjects with a methylated *MGMT* promoter. Differentially mutated genes numbered 6,683, 447, 507, and 86 genes in subgroups 1, 2, 3, and 4, respectively, and GO enrichment analyses of these genes identified statistically significant biological process terms that varied by subgroup (Supplementary Tables S1–4).

3.2 Gene Ontology analysis and validation of association with overall survival using gene signatures

Because subgroup 4 had a mOS > 15 months, which is the mOS reported for TMZ-treated GBM patients, and also was significantly longer than the mOS of subgroups 1 and 3, additional analyses focused on subgroup 4. Statistically significant terms from GO enrichment analysis (Supplementary Table S4) for the 86 differentially mutated genes of this subgroup (Supplementary Table S5) were related to the regulation of intrinsic apoptotic signaling by a *TP53* mediator or in response to DNA damage, the regulation of cyclin-dependent protein serine/threonine kinase activity, DNA damage checkpoint signaling, the negative regulation of transcription, cellular senescence, the regulation of cysteine-type endopeptidase activity in apoptosis, and the positive regulation of release of cytochrome c from mitochondria. To determine whether gene signatures associated with OS in independent groups of GBM patients could be identified from among the 86 genes, microarray gene expression data from the
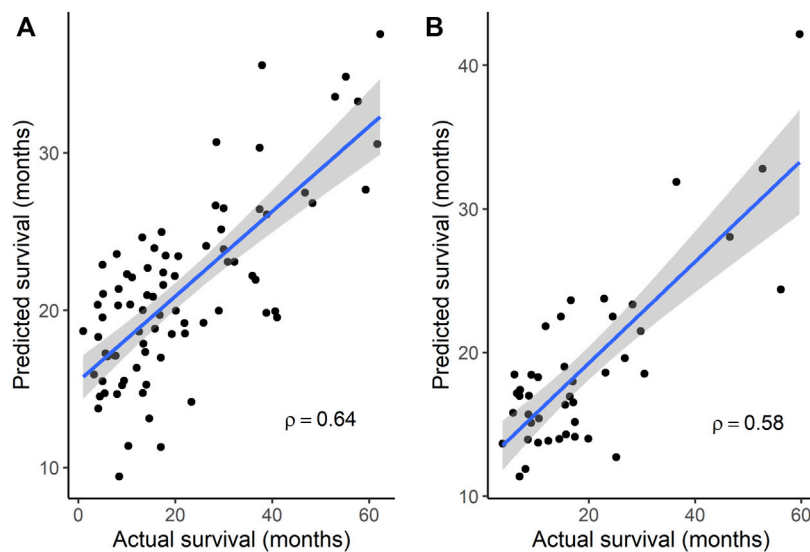
FIGURE 2
Correlation between actual survival and survival predicted from gene signatures. **(A)** A 20-gene signature was used to predict survival in data set GSE108474 (N = 81). $p = 8.83 \times 10^{-11}$ for Spearman's ρ. **(B)** An 8-gene signature was used to predict survival in data set GSE7696 (N = 43). $p = 4.84 \times 10^{-5}$ for Spearman's ρ. In each plot, the line through the points is the linear regression line, and the shading represents the 95% confidence region for the regression line.

GSE108474 and GSE7696 datasets and exome sequence data from the 37 GBM tumors of MCHS patients were analyzed. In the GSE108474 dataset, 20 genes (represented by 24 probes) out of 86 were selected as features that could predict survival in the 81 TMZ-treated GBM patients (Supplementary Table S6). In the GSE7696 dataset, 8 genes (represented by 10 probes) out of 86 were selected as features that could predict survival in the 43 TMZ-treated GBM patients (Supplementary Table S6). The gene signatures from the two datasets had the *HMGA1* gene in common. Spearman's rank correlation between actual and predicted OS was 0.64 ($p = 8.83 \times 10^{-11}$) and 0.58 ($p = 4.84 \times 10^{-5}$) in the GSE108474 and GSE7696 datasets, respectively (Figures 2A, B).

3.3 Somatic mutations and overall survival in an independent sample

The clinical characteristics of the 37 MCHS patients are shown in Supplementary Table S7. All 37 patients started, and 34 of the patients completed, concomitant TMZ and/or at least one cycle of adjuvant TMZ. mOS (95% confidence interval), representing the length of time between date of surgery for GBM and the date of death or end of study follow-up (31 December 2015), was 15.7 (12.9, 19.1) months (Supplementary Table S3), as determined by Kaplan-Meier analysis. Because the tumors were archived clinical specimens, a matched specimen of normal DNA was unavailable, and a tumor-only bioinformatics procedure was used to call somatic variants. The mean sequencing depth was between 30 X and 103 X for 28 tumors and between 14 X and 30 X for the other nine tumors. A total of 999 rare, somatic single nucleotide variants or small insertions/deletions in 905 genes were called (Supplementary Table S8), with an average of 27 (range of 1–64) variants detected per tumor. Three
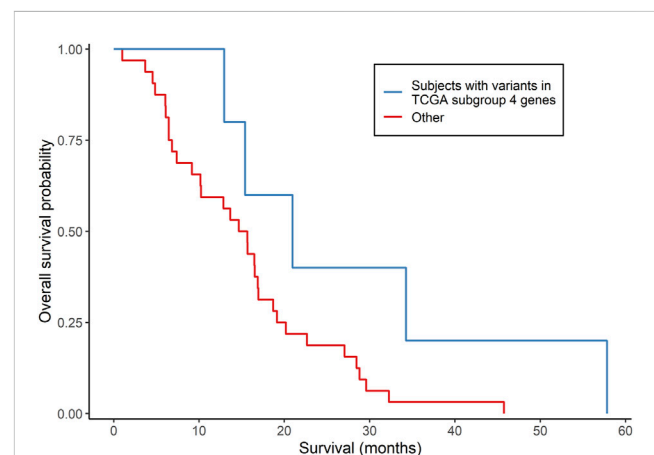

FIGURE 3
Comparison of survival according to subgroups defined by somatically mutated genes in an independent group of GBM patients. Kaplan-Meier analysis compared subjects who had somatic variants in genes that were differentially mutated in TCGA subgroup 4 (N = 5) with all other subjects (N = 32). $p = 0.0697$ from the log-rank test.

hundred and thirty-four (33.4%) of the variants have been reported previously in the COSMIC database of somatic mutations (version 96) (Tate et al., 2019) and, in 21 tumors, 30 variants were detected in 26 genes reported to be mutated recurrently in GBM (Brennan et al., 2013; Frattini et al., 2013). Five patients had five variants in four (*FUBP1*, *L3MBTL1*, *LCOR*, and *USP42*) of the 86 genes that were differentially mutated in TCGA GBM subgroup 4. Three (*FUBP1*, *L3MBTL1*, and *USP42*) of the four genes were among the 20 genes predictive of OS in the GSE108474 dataset. Similar to TCGA GBM subgroup 4, these five patients had a higher mOS than the other 32 patients (21.0 *versus* 15.3 months, respectively), but this

TABLE 3 Discrimination between TMZ-resistant and control GBM samples using gene signatures in three independent datasets.

Dataset	Method	Accuracy (%)	Sensitivity	Specificity	Area under the curve
GSE151680	Naïve Bayes	83.3	0.75	1.00	0.97
	Simple logistic	83.3	0.83	0.83	0.91
	Sequential minimal optimization	100.0	1.00	1.00	1.00
	Random forest	91.7	1.00	0.85	0.94
	J48	75.0	080	0.74	0.75
GSE193957	Naïve Bayes	100.0	1.00	1.00	1.00
	Simple logistic	100.0	1.00	1.00	1.00
	Sequential minimal optimization	100.0	1.00	1.00	1.00
	Random forest	100.0	1.00	1.00	1.00
	J48	83.3	1.00	0.75	0.83
GSE145128	Naïve Bayes	78.6	0.83	0.75	0.79
	Simple logistic	50.0	0.50	0.50	0.51
	Sequential minimal optimization	71.4	0.71	0.71	0.71
	Random forest	78.6	0.75	0.83	0.84
	J48	35.7	0.37	0.33	0.33

difference did not reach statistical significance ($p = 0.0697$; log-rank test; Figure 3). No subjects had *IDH1/2* variants. The *MGMT* promoter was methylated in two (40.0%) of the five patients and 18 (56.3%) of the other 32 patients ($p = 0.4174$; Fisher's exact test). Overall, the results indicated that gene signatures selected from the 86 genes were associated with GBM survival in two of the three independent datasets.

3.4 Identification of overlapping genes between signatures of survival and of temozolomide resistance
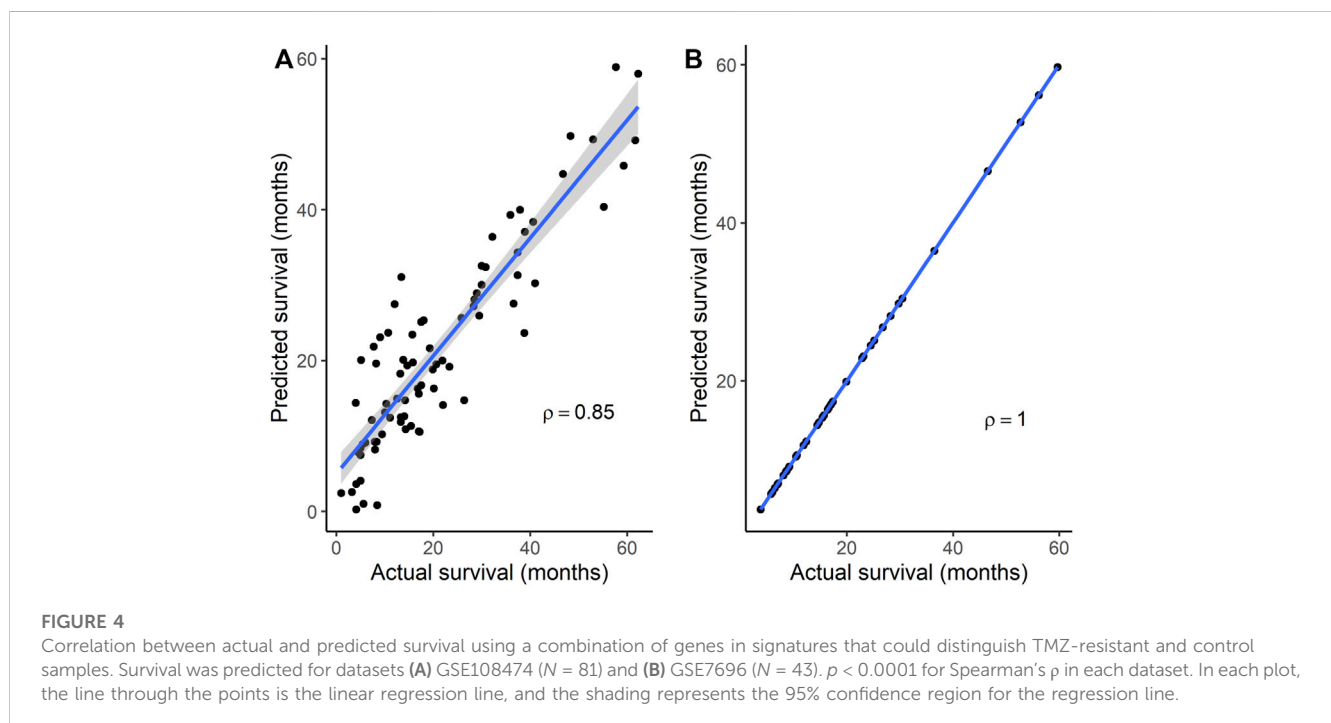
To identify gene signatures associated with TMZ resistance, five prediction algorithms were applied to expression data for the 86 genes from the GSE151680, GSE193957, and GSE145128 datasets. For the GSE151680 and GSE193957 data, the AUC was ≥ 0.75 for all five algorithms, whereas, for GSE145128 data, the AUC was ≥ 0.75 only for the Naïve Bayes and random forest algorithms (Table 3). Feature selection identified signatures consisting of 11 genes in GSE151680 data, 11 genes (represented by 12 probes) in GSE193957 data, and 15 genes (represented by 15 probes) in GSE145128 data that could distinguish between TMZ-resistant and control GBM samples (Supplementary Table S9). Six genes (*DNAJC9*, *HIPK4*, *PCBP4*, *PPP1R13L*, *SCAF8*, and *USP11*) overlapped between pairs of the three gene signatures. To determine whether the three gene signatures could also predict GBM survival, each signature, separately and also combined into one group of 31 unique genes, was used to predict OS in the GSE108474 and GSE7696 datasets. The 66 unique probes that tagged the 31 genes in these two datasets were used as features for prediction. Prediction with

each separate gene signature resulted in a moderate correlation between actual and predicted OS for the two datasets, with slightly higher correlation and lower mean absolute error values for the GSE7696 dataset (Table 4). The correlation increased to 0.85 (Figure 4A) and 1.00 (Figure 4B) for datasets GSE108474 and GSE7696, respectively, when the combined group of 31 genes was used for prediction. The eight genes that overlapped between the set of 31 genes that could distinguish TMZ-resistant from control GBM samples and the set of 27 genes in the two gene signatures predictive of OS in the GSE108474 and GSE7696 datasets were *ANG*, *BACH1*, *CDKN2C*, *HMGA1*, *IFI16*, *PADI4*, *SDF4*, and *TP53INP1*. Thus, these were genes in common between signatures of prognosis and of TMZ resistance.

Of the eight genes, three (*PADI4*, *SDF4*, and *TP53INP1*) have not been investigated previously for associations with GBM. Two *PADI4* variants were detected in subgroups 1 and one *PADI4* variant was detected in subgroup 4 (Supplementary Table S10). A *SDF4* variant was detected in subgroup 1 and a *TP53INP1* variant was detected in subgroup 4. Each of the variants in the three genes was present in a single tumor; therefore, the frequency of each variant in the GBM sample was 0.45% (1/221). The frequency of *PADI4* variants was 1.36% (3/221). Two of the *PADI4* variants and the *TP53INP1* variant have been detected in other cancer types in the COSMIC database (Supplementary Table S10). The *PADI4* p. Ser496Phe variant was detected in one sample of skin carcinoma, the *PADI4* p. Phe602 synonymous variant in two samples of endometrial cancer and one sample of malignant melanoma, and the *TP53INP1* p. Ile66Val variant in one sample of carcinoma of the bile duct. Somatic variants in the three genes have also been reported in several other cancer types in TCGA, with the percentage of cases having a variant ranging from 0.20% to 7.87%, depending on the cancer type (Supplementary Table S11).

TABLE 4 Performance of TMZ-resistance gene signatures in predicting GBM overall survival.

Dataset for survival prediction	Dataset used to select gene signature	Spearman's rank correlation between actual and predicted survival		Mean absolute error
		ρ	P	
GSE108474	GSE151680	0.45	2.82 x 10⁻⁵	10.7
GSE108474	GSE193957	0.52	7.92 x 10⁻⁷	11.1
GSE108474	GSE145128	0.46	1.22 x 10⁻⁵	10.4
GSE7696	GSE151680	0.53	0.00027	7.9
GSE7696	GSE193957	0.74	1.35 x 10⁻⁸	4.6
GSE7696	GSE145128	0.67	8.06 x 10⁻⁷	5.9

FIGURE 4
Correlation between actual and predicted survival using a combination of genes in signatures that could distinguish TMZ-resistant and control samples. Survival was predicted for datasets **(A)** GSE108474 (N = 81) and **(B)** GSE7696 (N = 43). $p < 0.0001$ for Spearman's ρ in each dataset. In each plot, the line through the points is the linear regression line, and the shading represents the 95% confidence region for the regression line.

4 Discussion

Genes that were differentially mutated in the subgroup with the longest mOS, compared with the other subgroups, were enriched in pathways related to intrinsic apoptosis signaling, cell cycle signaling, the DNA damage response, and cellular senescence, all mechanisms that affect GBM tumor growth and TMZ resistance. Intrinsic apoptosis signaling is initiated by the tumor suppressor p53 (*TP53*) in response to an intracellular stress signal such as DNA damage, with subsequent activation of a caspase cascade leading to apoptotic cell death (Li et al., 1997; Tomita et al., 2006). *TP53* also regulates the expression of genes, including cyclins and cyclin-dependent kinases, necessary for progression through the cell cycle (Engeland, 2018). Somatic variants in *TP53* are frequently detected in GBM (Brennan et al., 2013), and TMZ induces cell cycle arrest and senescence in GBM in a *TP53*-dependent manner (Hirose

et al., 2001). The PI3K/AKT/mTOR pathway, which also regulates the cell cycle and promotes cell proliferation and survival (Porta et al., 2014), has a mutation in at least one pathway gene in > 80% of GBM patients, causing constitutive activation and aberrant function of the pathway (Cancer Genome Atlas Research, 2008). Furthermore, *AKT* activation protects against the cytotoxic effects of TMZ in a GBM cell line (Hirose et al., 2005). The activation of DNA damage response upstream regulators, the serine/threonine kinases *ATR* and *ATM*, is associated with a reduced load of TMZ-induced DNA double strand breaks and increased TMZ resistance in GBM (Eich et al., 2013), as are mutations in *MSH6*, a mismatch repair gene (Yip et al., 2009). Cellular senescence, characterized by stable cell cycle arrest triggered by damaging stress signals, can exert context-dependent anti- or pro-tumorigenic effects on GBM (Beltzig et al., 2022; Salam et al., 2023). The GBM tumor growth arrest associated with cellular senescence induced by TMZ can be

considered anti-tumorigenic (Hirose et al., 2001). In a report describing a pro-tumorigenic effect, tumor microvascular endothelial cells that had been irradiated after removal from GBM tumors were shown to be senescent and TMZ-resistant while also being able to support the proliferation of GBM stem cells (Borovski et al., 2013).

Therapeutic agents that target these pathways are the subject of recently completed and ongoing GBM clinical trials (Wen et al., 2020). Drugs that target the interaction of p53 protein with the E3 ubiquitin ligase, *MDM2*, seek to relieve the inhibition of p53 by *MDM2* and rescue p53 transcriptional activation of pro-apoptotic factors such as *BAX* and *BBC3* (Puca et al., 2009). GBM showed no response to everolimus (Chinnaiyan et al., 2018), which targets mTOR, but other drugs that target mTORC1/2 and cyclin-dependent kinases in the cell cycle machinery are under evaluation (Wen et al., 2020). Inhibitors of the DNA damage response are also being evaluated because DNA repair is partly responsible for TMZ resistance in GBM (Wen et al., 2020). Anti-senescence agents are of interest in GBM because senescent cells are therapy-resistant and can be the precursors of a relapsed tumor (Beltzig et al., 2022). A suggested, two-step strategy involves the use of a chemotherapy drug to induce tumor cell senescence followed by the use of a senolytic to ablate the senescent cells before the tumor progresses and reoccurs (Chojak et al., 2023). TMZ is a potent inducer of senescence (Hirose et al., 2001), and pre-clinical studies of glioma cells have identified other inducers to be nutlin-3a (Villalonga-Planells et al., 2011) and resveratrol (Filippi-Chiela et al., 2013), among others, and senolytics to include BH3 mimetics that target pro-survival *BCL2* family members, which are regulators of apoptosis (Schwarzenbach et al., 2021). The failure of most experimental drugs to extend GBM survival emphasizes the therapeutic challenges in GBM, among which are the genetic heterogeneity within and between tumors, redundant signaling pathways, and gaps in knowledge regarding the contributions of genetic mutations to tumor growth and treatment resistance (Wen et al., 2020).

The identification of gene signatures predictive of OS in independent datasets in this study confirmed the association of GBM survival with the 86 differentially mutated genes in subgroup 4. Subsets of the 86 genes also formed gene signatures that could distinguish TMZ-resistant from control GBM samples, and the pool of 31 genes in these signatures were also predictive of OS, suggesting that some of the genes potentially can function as biomarkers of both prognosis and response to TMZ treatment. To minimize reliance on findings specific to individual datasets, the overlap between signatures of prognosis and TMZ resistance was determined, resulting in eight overlapping genes being identified.

Five (*ANG, BACH1, CDKN2C, HMGA1, IFI16*) of the eight genes have shown associations with GBM in other reports. The expression of angiogenin (*ANG*), a potent stimulator of angiogenesis, was inversely correlated with survival in GBM patients, and *Ang* deficiency prolonged the survival of mice with platelet derived growth factor-induced GBM (Yang et al., 2022). The GBM tumor in these mice showed lower proliferation, less invasiveness, reduced angiogenesis, and higher apoptosis compared with control mice. *ANG* knockdown also decreased cell proliferation and increased apoptosis in GBM cell lines (Yang et al., 2022). The inhibition of *ANG* with neomycin, a

small molecule that prevents binding of *ANG* to its receptor, *PLXNB2*, decreased GBM growth in mice with platelet derived growth factor-induced GBM and a xenograft mouse model, demonstrating the potential of *ANG* inhibitors in GBM therapy (Yang et al., 2022). High expression of *BACH1*, a transcription factor, has been detected in GBM (Yuan et al., 2022). *BACH1* overexpression enhanced the expression of *MGMT* leading to TMZ resistance in a GBM cell line with a hypomethylated *MGMT* promoter and a xenograft mouse model, whereas *BACH1* depletion sensitized TMZ-resistant cells to TMZ (Nie et al., 2016). These effects were blocked by *TP53*, which inhibits *MGMT* by preventing the binding of *SP1* to the *MGMT* promoter, but the antagonistic effects of *TP53* could be overcome by *BACH1* overexpression (Nie et al., 2016). The low expression of *BACH1* in combination with wild-type *TP53* was associated with longer survival in patients who received TMZ, suggesting that the targeting of this pathway holds potential for regulating TMZ resistance (Nie et al., 2016). Co-deletion of *CDKN2C*, a cyclin-dependent kinase inhibitor and tumor suppressor, and *CDKN2A* in GBM cells predicted sensitivity to palbociclib (PD0332991), a selective *CDK4/6* inhibitor that is being evaluated in an ongoing GBM clinical trial (NCT03158389) (Wiedemeyer et al., 2010; Cen et al., 2012). *HMGA1*, a gene that alters chromatin architecture to regulate transcription, is highly expressed in some GBM stem cell lines and its silencing reduces the self-renewal and sphere-forming efficiency of these cells and sensitizes them to TMZ (Colamaio et al., 2016). Moreover, decreased *HMGA1* expression in a xenograft mouse model, resulting from knockdown of an upstream long non-coding RNA regulator, correlated with delayed GBM tumor growth and increased survival (Mineo et al., 2016). The binding of *IFI16*, a sensor of DNA in the innate immune response, to *ARPC1B*, a subunit of the actin-related protein-2/3 complex, resulted in activation of the NF-kappa-B pathway that contributed to promotion of a mesenchymal phenotype transformation and radiotherapy resistance in GBM stem cells and a xenograft mouse model (Gao et al., 2022). This study further showed that GBM cell lines with high expression of *ARPC1B* and *IFI16* were more sensitive to ceralasertib (AZD6738), an inhibitor of *ATR*. *IFI16* was also upregulated in GBM cells treated with both Y15 (a small molecule inhibitor of activated focal adhesion kinase) and TMZ (Huang et al., 2014), a combination that decreases viability and tumor growth in GBM cells (Golubovskaya et al., 2013).

The functions of three (*PADI4, SDF4, TP53INP1*) of the eight genes in GBM remain to be explored. *PADI4*, an enzyme that converts arginine residues to citrulline resulting in recognition of the citrullinated proteins by the immune system, is highly expressed in GBM (Sase et al., 2017; Neira et al., 2022). However, reports regarding the role of *PADI4* in cancer are conflicting. Some studies reported high *PADI4* expression in other cancers (Liu et al., 2019) and showed that *PADI4* overexpression is associated with pro-tumorigenic properties such as increased tumor cell proliferation, migration, clone forming ability, and metastasis, and reduced apoptosis (Chang et al., 2022). Other studies observed low *PADI4* expression in cancer (Indeglia et al., 2023) and found that *PADI4* can suppress tumor cell growth (Tanikawa et al., 2009). *PADI4* is also suggested to function as a tumor suppressor that interacts with p53 to regulate the transcription of p53 target genes (Indeglia et al., 2023). The potential of *PADI4* to serve

as a drug target in cancer was demonstrated by the finding that a small molecule inhibitor of *PADI4* activated *SESN2* and other p53 target genes leading to mTORC1 signaling pathway inhibition, perturbation of autophagy, and suppression of tumor cell growth (Wang et al., 2012). In addition, the use of citrullinated proteins as epitopes in cancer immunotherapy is an area of active research (Brentville et al., 2020). The role of *SDF4*, a calcium-binding protein that regulates calcium-dependent cellular activities, in GBM has not been described, but in other cancers, *SDF4* is highly expressed and promotes tumor growth, migration, and distant metastasis (Luo et al., 2016). *SDF4* expression was elevated in a colorectal cancer cell line (Ji et al., 2009) and in pancreatic cancer cells compared with non-neoplastic pancreatic ductal cells (Gronborg et al., 2006). A long non-coding RNA, *LINC00173*, that is upregulated in nasopharyngeal cancer cells, could promote the growth, migration and metastasis of nasopharyngeal cancer cells by interacting with *RAB1B* to facilitate *SDF4* secretion in a *RAB1B*-dependent manner (He et al., 2023). The reversal of *LINC00173*-mediated cancer cell progression by *SDF4* knockdown suggested the potential of the *LINC00173-RAB1B-SDF4* pathway to be a drug target for nasopharyngeal cancer. The overexpression of *TP53INP1*, a tumor suppressor that promotes the transcriptional activity of *TP53* on its target genes, induces G1 cell cycle arrest and enhances *TP53*-mediated apoptosis in cancer cells (Tomasini et al., 2003). *TP53INP1* also interacts directly with components of the autophagy pathway to induce autophagy-dependent cell death (Seillier et al., 2012). In pediatric ependymoma, a malignant glial cell tumor, high expression of the microRNA, miR-124-3p, and low expression of one of its target genes, *TP53INP1*, correlated with shorter progression-free survival, suggesting the potential to target miR-124-3p and *TP53INP1* in new therapeutic approaches for this cancer (Margolin-Miller et al., 2017). The involvement of *TP53INP1* in GBM tumor growth and therapy resistance needs further study.

A study limitation was the small sample sizes of the GEO independent datasets, which lessened the statistical power for detecting correlations between actual and predicted OS and distinguishing TMZ-resistant from control samples. The small size of the MCHS sample also diminished power to detect a significant difference in mOS between study groups. No matched germline DNA sample was available for this dataset, prompting the use of a PureCN tumor-only bioinformatics workflow that can accurately classify variants as somatic *versus* germline. However, the low depth of exome sequencing coverage of some of the MCHS samples likely hampered the detection of somatic variants in these samples. Another limitation was uncertainty about whether time to progression in TCGA GBM data was based on true disease progression that was documented by imaging according to RANO criteria (Wen et al., 2010).

5 Conclusion

In conclusion, targeting the molecular events at the basis of GBM holds potential for the development of new therapeutic strategies in GBM. Somatically mutated genes associated with mOS after TMZ treatment function in the known oncogenic pathways of apoptosis, cell cycle control, the DNA damage response, and cellular senescence. Subsets of these genes form signatures of GBM prognosis or TMZ resistance, and the overlap of these subsets has identified genes that may

influence both TMZ resistance and patient survival in GBM. *PADI4*, *SDF4*, and *TP53INP1* have been little studied in relation to GBM and are new GBM gene candidates. Further investigation is needed to evaluate whether these genes are useful for developing interventions to improve the anticancer activity of TMZ and can serve as drug targets or predictive biomarkers of response to GBM treatment. Additional studies should also explore the molecular mechanisms through which somatic mutations in these genes have an impact on the transition of initially chemotherapy-sensitive tumors to TMZ-resistant tumors.

Data availability statement

The datasets for this article are not publicly available due to concerns regarding participant/patient anonymity. Requests to access the datasets should be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by the Institutional Review Board of Marshfield Clinic Research Institute. The studies were conducted in accordance with the local legislation and institutional requirements. The ethics committee/institutional review board waived the requirement of written informed consent for participation from the participants or the participants' legal guardians/next of kin because study participants were deceased when the study was initiated.

Author contributions

TC: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing–original draft, Writing–review and editing. RV: Formal Analysis, Investigation, Methodology, Software, Validation, Writing–review and editing. SS: Formal Analysis, Investigation, Software, Validation, Visualization, Writing–review and editing. RM-F: Conceptualization, Funding acquisition, Investigation, Resources, Writing–review and editing.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1320789/full#supplementary-material

References

Beltzig, L., Schwarzenbach, C., Leukel, P., Frauenknecht, K. B. M., Sommer, C., Tancredi, A., et al. (2022). Senescence is the main trait induced by temozolomide in glioblastoma cells. *Cancers (Basel)* 14 (9), 2233. doi:10.3390/cancers14092233

Borovski, T., Beke, P., van Tellingen, O., Rodermond, H. M., Verhoeff, J. J., Lascano, V., et al. (2013). Therapy-resistant tumor microvascular endothelial cells contribute to treatment failure in glioblastoma multiforme. *Oncogene* 32 (12), 1539–1548. doi:10.1038/onc.2012.172

Brennan, C. W., Verhaak, R. G., McKenna, A., Campos, B., Noushmehr, H., Salama, S. R., et al. (2013). The somatic genomic landscape of glioblastoma. *Cell* 155 (2), 462–477. doi:10.1016/j.cell.2013.09.034

Brentville, V. A., Vankemmelbeke, M., Metheringham, R. L., and Durrant, L. G. (2020). Post-translational modifications such as citrullination are excellent targets for cancer therapy. *Semin. Immunol.* 47, 101393. doi:10.1016/j.smim.2020.101393

Cancer Genome Atlas Research (2008). Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 455 (7216), 1061–1068. doi:10.1038/nature07385

Carter, T. C., Medina-Flores, R., and Lawler, B. E. (2018). Glioblastoma treatment with temozolomide and bevacizumab and overall survival in a rural tertiary Healthcare practice. *Biomed. Res. Int.* 2018, 6204676. doi:10.1155/2018/6204676

Cen, L., Carlson, B. L., Schroeder, M. A., Ostrem, J. L., Kitange, G. J., Mladek, A. C., et al. (2012). p16-Cdk4-Rb axis controls sensitivity to a cyclin-dependent kinase inhibitor PD0332991 in glioblastoma xenograft cells. *Neuro Oncol.* 14 (7), 870–881. doi:10.1093/neuonc/nos114

Chang, X. T., Wu, H., Li, H. L., Li, H. L., and Zheng, Y. B. (2022). PADI4 promotes epithelial-mesenchymal transition(EMT) in gastric cancer via the upregulation of interleukin 8. *BMC Gastroenterol.* 22 (1), 25. doi:10.1186/s12876-022-02097-0

Chinnaiyan, P., Won, M., Wen, P. Y., Rojiani, A. M., Werner-Wasik, M., Shih, H. A., et al. (2018). A randomized phase II study of everolimus in combination with chemoradiation in newly diagnosed glioblastoma: results of NRG Oncology RTOG 0913. *Neuro Oncol.* 20 (5), 666–673. doi:10.1093/neuonc/nox209

Chojak, R., Fares, J., Petrosyan, E., and Lesniak, M. S. (2023). Cellular senescence in glioma. *J. Neurooncol* 164, 11–29. doi:10.1007/s11060-023-04387-3

Choo, M., Mai, V. H., Kim, H. S., Kim, D. H., Ku, J. L., Lee, S. K., et al. (2023). Involvement of cell shape and lipid metabolism in glioblastoma resistance to temozolomide. *Acta Pharmacol. Sin.* 44 (3), 670–679. doi:10.1038/s41401-022-00984-6

Cibulskis, K., Lawrence, M. S., Carter, S. L., Sivachenko, A., Jaffe, D., Sougnez, C., et al. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* 31 (3), 213–219. doi:10.1038/nbt.2514

Colamaio, M., Tosti, N., Puca, F., Mari, A., Gattordo, R., Kuzay, Y., et al. (2016). HMGA1 silencing reduces stemness and temozolomide resistance in glioblastoma stem cells. *Expert Opin. Ther. Targets* 20 (10), 1169–1179. doi:10.1080/14728222.2016.1220543

Costello, M., Pugh, T. J., Fennell, T. J., Stewart, C., Lichtenstein, L., Meldrim, J. C., et al. (2013). Discovery and characterization of artifactual mutations in deep coverage targeted capture sequencing data due to oxidative DNA damage during sample preparation. *Nucleic Acids Res.* 41 (6), e67. doi:10.1093/nar/gks1443

Cox, D. R. (1972). Regression models and life-tables (with discussion). *J. R. Stat. Soc. Ser. B Methodol.* 34 (2), 187–220.

Eich, M., Roos, W. P., Nikolova, T., and Kaina, B. (2013). Contribution of ATM and ATR to the resistance of glioblastoma and malignant melanoma cells to the methylating anticancer drug temozolomide. *Mol. Cancer Ther.* 12 (11), 2529–2540. doi:10.1158/1535-7163.MCT-13-0136

Engeland, K. (2018). Cell cycle arrest through indirect transcriptional repression by p53: I have a DREAM. *Cell Death Differ.* 25 (1), 114–132. doi:10.1038/cdd.2017.172

Filippi-Chiela, E. C., Thome, M. P., Bueno e Silva, M. M., Pelegrini, A. L., Ledur, P. F., Garicochea, B., et al. (2013). Resveratrol abrogates the temozolomide-induced G2 arrest leading to mitotic catastrophe and reinforces the temozolomide-induced senescence in glioma cells. *BMC Cancer* 13, 147. doi:10.1186/1471-2407-13-147

Frattini, V., Trifonov, V., Chan, J. M., Castano, A., Lia, M., Abate, F., et al. (2013). The integrated landscape of driver genomic alterations in glioblastoma. *Nat. Genet.* 45 (10), 1141–1149. doi:10.1038/ng.2734

Gao, Z., Xu, J., Fan, Y., Zhang, Z., Wang, H., Qian, M., et al. (2022). ARPC1B promotes mesenchymal phenotype maintenance and radiotherapy resistance by blocking TRIM21-mediated degradation of IFI16 and HuR in glioma stem cells. *J. Exp. Clin. Cancer Res.* 41 (1), 323. doi:10.1186/s13046-022-02526-8

Gene Ontology, C. (2021). The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res.* 49 (D1), D325–D334. doi:10.1093/nar/gkaa1113

Golubovskaya, V. M., Huang, G., Ho, B., Yemma, M., Morrison, C. D., Lee, J., et al. (2013). Pharmacologic blockade of FAK autophosphorylation decreases human glioblastoma tumor growth and synergizes with temozolomide. *Mol. Cancer Ther.* 12 (2), 162–172. doi:10.1158/1535-7163.MCT-12-0701

Gronborg, M., Kristiansen, T. Z., Iwahori, A., Chang, R., Reddy, R., Sato, N., et al. (2006). Biomarker discovery from pancreatic cancer secretome using a differential proteomic approach. *Mol. Cell Proteomics* 5 (1), 157–171. doi:10.1074/mcp.M500178-MCP200

Gusev, Y., Bhuvaneshwar, K., Song, L., Zenklusen, J. C., Fine, H., and Madhavan, S. (2018). The REMBRANDT study, a large collection of genomic data from brain cancer patients. *Sci. Data* 5, 180158. doi:10.1038/sdata.2018.158

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009). The WEKA data mining software: an update. *ACM SIGKDD Explor. Newsl.* 11 (1), 10–18. doi:10.1145/1656274.1656278

Han, K., Ren, M., Wick, W., Abrey, L., Das, A., Jin, J., et al. (2014). Progression-free survival as a surrogate endpoint for overall survival in glioblastoma: a literature-based meta-analysis from 91 trials. *Neuro Oncol.* 16 (5), 696–706. doi:10.1093/neuonc/not236

He, S. W., Liang, Y. L., Zhang, Y., Liu, X., Gong, S., Ye, M. L., et al. (2023). LINC00173 facilitates tumor progression by stimulating RAB1B-mediated PA2G4 and SDF4 secretion in nasopharyngeal carcinoma. *Mol. Oncol.* 17 (3), 518–533. doi:10.1002/1878-0261.13375

Hegi, M. E., Diserens, A. C., Gorlia, T., Hamou, M. F., de Tribolet, N., Weller, M., et al. (2005). MGMT gene silencing and benefit from temozolomide in glioblastoma. *N. Engl. J. Med.* 352 (10), 997–1003. doi:10.1056/NEJMoa043331

Hirose, Y., Berger, M. S., and Pieper, R. O. (2001). p53 effects both the duration of G2/M arrest and the fate of temozolomide-treated human glioblastoma cells. *Cancer Res.* 61 (5), 1957–1963.

Hirose, Y., Katayama, M., Mirzoeva, O. K., Berger, M. S., and Pieper, R. O. (2005). Akt activation suppresses Chk2-mediated, methylating agent-induced G2 arrest and protects from temozolomide-induced mitotic catastrophe and cellular senescence. *Cancer Res.* 65 (11), 4861–4869. doi:10.1158/0008-5472.CAN-04-2633

Hofree, M., Shen, J. P., Carter, H., Gross, A., and Ideker, T. (2013). Network-based stratification of tumor mutations. *Nat. Methods* 10 (11), 1108–1115. doi:10.1038/nmeth.2651

Huang, G., Ho, B., Conroy, J., Liu, S., Qiang, H., and Golubovskaya, V. (2014). The microarray gene profiling analysis of glioblastoma cancer cells reveals genes affected by FAK inhibitor Y15 and combination of Y15 and temozolomide. *Anticancer Agents Med. Chem.* 14 (1), 9–17. doi:10.2174/18715206113139990141

Huang, J. K., Jia, T., Carlin, D. E., and Ideker, T. (2018). pyNBS: a Python implementation for network-based stratification of tumor mutations. *Bioinformatics* 34 (16), 2859–2861. doi:10.1093/bioinformatics/bty186

Indeglia, A., Leung, J. C., Miller, S. A., Leu, J. I., Dougherty, J. F., Clarke, N. L., et al. (2023). An african-specific variant of TP53 reveals PADI4 as a regulator of p53-

mediated tumor suppression. *Cancer Discov.* 13 (7), 1696–1719. doi:10.1158/2159-8290.CD-22-1315

Ji, H., Greening, D. W., Kapp, E. A., Moritz, R. L., and Simpson, R. J. (2009). Secretome-based proteomics reveals sulindac-modulated proteins released from colon cancer cells. *Proteomics Clin. Appl.* 3 (4), 433–451. doi:10.1002/prca.200800077

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alfoldi, J., Wang, Q., et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 581 (7809), 434–443. doi:10.1038/s41586-020-2308-7

Karran, P., Macpherson, P., Ceccotti, S., Dogliotti, E., Griffin, S., and Bignami, M. (1993). O6-methylguanine residues elicit DNA repair synthesis by human cell extracts. *J. Biol. Chem.* 268 (21), 15878–15886. doi:10.1016/s0021-9258(18)82335-3

Kebir, S., Ullrich, V., Berger, P., Dobersalske, C., Langer, S., Rauschenbach, L., et al. (2023). A sequential targeting strategy interrupts AKT-driven subclone-mediated progression in glioblastoma. *Clin. Cancer Res.* 29 (2), 488–500. doi:10.1158/1078-0432.CCR-22-0611

Kendall, M., and Gibbons, J. D. (1990). *Rank correlation methods.* 5th ed. New York: NY: Oxford University Press, 8–10.

Khasraw, M., Fujita, Y., Lee-Chang, C., Balyasnikova, I. V., Najem, H., and Heimberger, A. B. (2022). New approaches to glioblastoma. *Annu. Rev. Med.* 73, 279–292. doi:10.1146/annurev-med-042420-102102

Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., and Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Comput. Struct. Biotechnol. J.* 13, 8–17. doi:10.1016/j.csbj.2014.11.005

Kristensen, L. S., Michaelsen, S. R., Dyrbye, H., Aslan, D., Grunnet, K., Christensen, I. J., et al. (2016). Assessment of quantitative and allelic MGMT methylation patterns as a prognostic marker in glioblastoma. *J. Neuropathol. Exp. Neurol.* 75 (3), 246–255. doi:10.1093/jnen/nlv024

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25 (14), 1754–1760. doi:10.1093/bioinformatics/btp324

Li, P., Nijhawan, D., Budihardjo, I., Srinivasula, S. M., Ahmad, M., Alnemri, E. S., et al. (1997). Cytochrome c and dATP-dependent formation of Apaf-1/caspase-9 complex initiates an apoptotic protease cascade. *Cell* 91 (4), 479–489. doi:10.1016/s0092-8674(00)80434-1

Liu, M., Qu, Y., Teng, X., Xing, Y., Li, D., Li, C., et al. (2019). PADI4-mediated epithelial-mesenchymal transition in lung cancer cells. *Mol. Med. Rep.* 19 (4), 3087–3094. doi:10.3892/mmr.2019.9968

Luo, J., Li, Z., Zhu, H., Wang, C., Zheng, W., He, Y., et al. (2016). A novel role of cab45-G in mediating cell migration in cancer cells. *Int. J. Biol. Sci.* 12 (6), 677–687. doi:10.7150/ijbs.11037

Malley, D. S., Hamoudi, R. A., Kocialkowski, S., Pearson, D. M., Collins, V. P., and Ichimura, K. (2011). A distinct region of the MGMT CpG island critical for transcriptional regulation is preferentially methylated in glioblastoma cells and xenografts. *Acta Neuropathol.* 121 (5), 651–661. doi:10.1007/s00401-011-0803-5

Margolin-Miller, Y., Yanichkin, N., Shichrur, K., Toledano, H., Ohali, A., Tzaridis, T., et al. (2017). Prognostic relevance of miR-124-3p and its target TP53INP1 in pediatric ependymoma. *Genes Chromosom. Cancer* 56 (8), 639–650. doi:10.1002/gcc.22467

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20 (9), 1297–1303. doi:10.1101/gr.107524.110

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R., Thormann, A., et al. (2016). The ensembl variant effect predictor. *Genome Biol.* 17 (1), 122. doi:10.1186/s13059-016-0974-4

Mi, H., Muruganujan, A., Ebert, D., Huang, X., and Thomas, P. D. (2019a). PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res.* 47 (D1), D419–D426. doi:10.1093/nar/gky1038

Mi, H., Muruganujan, A., Huang, X., Ebert, D., Mills, C., Guo, X., et al. (2019b). Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nat. Protoc.* 14 (3), 703–721. doi:10.1038/s41596-019-0128-8

Mineo, M., Ricklefs, F., Rooj, A. K., Lyons, S. M., Ivanov, P., Ansari, K. I., et al. (2016). The long non-coding RNA HIF1A-AS2 facilitates the maintenance of mesenchymal glioblastoma stem-like cells in hypoxic niches. *Cell Rep.* 15 (11), 2500–2509. doi:10.1016/j.celrep.2016.05.018

Murat, A., Migliavacca, E., Gorlia, T., Lambiv, W. L., Shay, T., Hamou, M. F., et al. (2008). Stem cell-related "self-renewal" signature and high epidermal growth factor receptor expression associated with resistance to concomitant chemoradiotherapy in glioblastoma. *J. Clin. Oncol.* 26 (18), 3015–3024. doi:10.1200/JCO.2007.15.7164

Neira, J. L., Araujo-Abad, S., Camara-Artigas, A., Rizzuti, B., Abian, O., Giudici, A. M., et al. (2022). Biochemical and biophysical characterization of PADI4 supports its involvement in cancer. *Arch. Biochem. Biophys.* 717, 109125. doi:10.1016/j.abb.2022.109125

Newlands, E. S., Stevens, M. F., Wedge, S. R., Wheelhouse, R. T., and Brock, C. (1997). Temozolomide: a review of its discovery, chemical properties, pre-clinical development and clinical trials. *Cancer Treat. Rev.* 23 (1), 35–61. doi:10.1016/s0305-7372(97)90019-0

Nie, E., Jin, X., Wu, W., Yu, T., Zhou, X., Zhi, T., et al. (2016). BACH1 promotes temozolomide resistance in glioblastoma through antagonizing the function of p53. *Sci. Rep.* 6, 39743. doi:10.1038/srep39743

Oh, S., Geistlinger, L., Ramos, M., Morgan, M., Waldron, L., and Riester, M. (2020). Reliable analysis of clinical tumor-only whole-exome sequencing data. *JCO Clin. Cancer Inf.* 4, 321–335. doi:10.1200/CCI.19.00130

Pearson, J. R. D., Cuzzubbo, S., McArthur, S., Durrant, L. G., Adhikaree, J., Tinsley, C. J., et al. (2020). Immune escape in glioblastoma multiforme and the adaptation of immunotherapies for treatment. *Front. Immunol.* 11, 582106. doi:10.3389/fimmu.2020.582106

Porta, C., Paglino, C., and Mosca, A. (2014). Targeting PI3K/Akt/mTOR signaling in cancer. *Front. Oncol.* 4, 64. doi:10.3389/fonc.2014.00064

Puca, R., Nardinocchi, L., Sacchi, A., Rechavi, G., Givol, D., and D'Orazi, G. (2009). HIPK2 modulates p53 activity towards pro-apoptotic transcription. *Mol. Cancer* 8, 85. doi:10.1186/1476-4598-8-85

R-Core, T. (2021). *A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Salam, R., Saliou, A., Bielle, F., Bertrand, M., Antoniewski, C., Carpentier, C., et al. (2023). Cellular senescence in malignant cells promotes tumor progression in mouse and patient Glioblastoma. *Nat. Commun.* 14 (1), 441. doi:10.1038/s41467-023-36124-9

Sase, T., Arito, M., Onodera, H., Omoteyama, K., Kurokawa, M. S., Kagami, Y., et al. (2017). Hypoxia-induced production of peptidylarginine deiminases and citrullinated proteins in malignant glioma cells. *Biochem. Biophys. Res. Commun.* 482 (1), 50–56. doi:10.1016/j.bbrc.2016.10.154

Schober, P., Boer, C., and Schwarte, L. A. (2018). Correlation coefficients: appropriate use and interpretation. *Anesth. Analg.* 126 (5), 1763–1768. doi:10.1213/ANE.0000000000002864

Schober, P., and Vetter, T. R. (2018). Survival analysis and interpretation of time-to-event data: the tortoise and the hare. *Anesth. Analg.* 127 (3), 792–798. doi:10.1213/ANE.0000000000003653

Schwarzenbach, C., Tatsch, L., Brandstetter Vilar, J., Rasenberger, B., Beltzig, L., Kaina, B., et al. (2021). Targeting c-IAP1, c-IAP2, and bcl-2 eliminates senescent glioblastoma cells following temozolomide treatment. *Cancers (Basel)* 13 (14), 3585. doi:10.3390/cancers13143585

Seillier, M., Peuget, S., Gayet, O., Gauthier, C., N'Guessan, P., Monte, M., et al. (2012). TP53INP1, a tumor suppressor, interacts with LC3 and ATG8-family proteins through the LC3-interacting region (LIR) and promotes autophagy-dependent cell death. *Cell Death Differ.* 19 (9), 1525–1535. doi:10.1038/cdd.2012.30

Singh, N., Miner, A., Hennis, L., and Mittal, S. (2021). Mechanisms of temozolomide resistance in glioblastoma - a comprehensive review. *Cancer Drug Resist* 4 (1), 17–43. doi:10.20517/cdr.2020.79

Stupp, R., Hegi, M. E., Gorlia, T., Erridge, S. C., Perry, J., Hong, Y. K., et al. (2014). Cilengitide combined with standard treatment for patients with newly diagnosed glioblastoma with methylated MGMT promoter (CENTRIC EORTC 26071-22072 study): a multicentre, randomised, open-label, phase 3 trial. *Lancet Oncol.* 15 (10), 1100–1108. doi:10.1016/S1470-2045(14)70379-1

Stupp, R., Mason, W. P., van den Bent, M. J., Weller, M., Fisher, B., Taphoorn, M. J., et al. (2005). Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N. Engl. J. Med.* 352 (10), 987–996. doi:10.1056/NEJMoa043330

Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., et al. (2023). The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* 51 (D1), D638–D646. doi:10.1093/nar/gkac1000

Tanikawa, C., Ueda, K., Nakagawa, H., Yoshida, N., Nakamura, Y., and Matsuda, K. (2009). Regulation of protein Citrullination through p53/PADI4 network in DNA damage response. *Cancer Res.* 69 (22), 8761–8769. doi:10.1158/0008-5472.CAN-09-2280

Tate, J. G., Bamford, S., Jubb, H. C., Sondka, Z., Beare, D. M., Bindal, N., et al. (2019). COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Res.* 47 (D1), D941–D947. doi:10.1093/nar/gky1015

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 58 (1), 267–288. doi:10.1111/j.2517-6161.1996.tb02080.x

Tomasini, R., Samir, A. A., Carrier, A., Isnardon, D., Cecchinelli, B., Soddu, S., et al. (2003). TP53INP1s and homeodomain-interacting protein kinase-2 (HIPK2) are partners in regulating p53 activity. *J. Biol. Chem.* 278 (39), 37722–37729. doi:10.1074/jbc.M301979200

Tomita, Y., Marchenko, N., Erster, S., Nemajerova, A., Dehner, A., Klein, C., et al. (2006). WT p53, but not tumor-derived mutants, bind to Bcl2 via the DNA binding domain and induce mitochondrial permeabilization. *J. Biol. Chem.* 281 (13), 8600–8606. doi:10.1074/jbc.M507611200

Tusher, V. G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. U. S. A.* 98 (9), 5116–5121. doi:10.1073/pnas.091062498

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinforma.* 43 (1110), 1–11. doi:10.1002/0471250953.bi1110s43

van Solinge, T. S., Nieland, L., Chiocca, E. A., and Broekman, M. L. D. (2022). Advances in local therapy for glioblastoma - taking the fight to the tumour. *Nat. Rev. Neurol.* 18 (4), 221–236. doi:10.1038/s41582-022-00621-0

Venkataramani, V., Yang, Y., Schubert, M. C., Reyhan, E., Tetzlaff, S. K., Wissmann, N., et al. (2022). Glioblastoma hijacks neuronal mechanisms for brain invasion. *Cell* 185 (16), 2899–2917.e31. doi:10.1016/j.cell.2022.06.054

Villalonga-Planells, R., Coll-Mulet, L., Martinez-Soler, F., Castano, E., Acebes, J. J., Gimenez-Bonafe, P., et al. (2011). Activation of p53 by nutlin-3a induces apoptosis and cellular senescence in human glioblastoma multiforme. *PLoS One* 6 (4), e18588. doi:10.1371/journal.pone.0018588

Wang, Y., Li, P., Wang, S., Hu, J., Chen, X. A., Wu, J., et al. (2012). Anticancer peptidylarginine deiminase (PAD) inhibitors regulate the autophagy flux and the mammalian target of rapamycin complex 1 activity. *J. Biol. Chem.* 287 (31), 25941–25953. doi:10.1074/jbc.M112.375725

Weller, M., Felsberg, J., Hartmann, C., Berger, H., Steinbach, J. P., Schramm, J., et al. (2009). Molecular predictors of progression-free and overall survival in patients with newly diagnosed glioblastoma: a prospective translational study of the German Glioma Network. *J. Clin. Oncol.* 27 (34), 5743–5750. doi:10.1200/JCO.2009.23.0805

Wen, P. Y., Macdonald, D. R., Reardon, D. A., Cloughesy, T. F., Sorensen, A. G., Galanis, E., et al. (2010). Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. *J. Clin. Oncol.* 28 (11), 1963–1972. doi:10.1200/JCO.2009.26.3541

Wen, P. Y., Weller, M., Lee, E. Q., Alexander, B. M., Barnholtz-Sloan, J. S., Barthel, F. P., et al. (2020). Glioblastoma in adults: a Society for Neuro-Oncology (SNO) and European Society of Neuro-Oncology (EANO) consensus review on current management and future directions. *Neuro Oncol.* 22 (8), 1073–1113. doi:10.1093/neuonc/noaa106

Wick, W., Weller, M., van den Bent, M., Sanson, M., Weiler, M., von Deimling, A., et al. (2014). MGMT testing--the challenges for biomarker-based glioma treatment. *Nat. Rev. Neurol.* 10 (7), 372–385. doi:10.1038/nrneurol.2014.100

Wiedemeyer, W. R., Dunn, I. F., Quayle, S. N., Zhang, J., Chheda, M. G., Dunn, G. P., et al. (2010). Pattern of retinoblastoma pathway inactivation dictates response to CDK4/6 inhibition in GBM. *Proc. Natl. Acad. Sci. U. S. A.* 107 (25), 11501–11506. doi:10.1073/pnas.1001613107

Willmott, C. J., and Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* 30 (1), 79–82. doi:10.3354/cr030079

Yang, H., Yuan, L., Ibaragi, S., Li, S., Shapiro, R., Vanli, N., et al. (2022). Angiogenin and plexin-B2 axis promotes glioblastoma progression by enhancing invasion, vascular association, proliferation and survival. *Br. J. Cancer* 127 (3), 422–435. doi:10.1038/s41416-022-01814-6

Yip, S., Miao, J., Cahill, D. P., Iafrate, A. J., Aldape, K., Nutt, C. L., et al. (2009). MSH6 mutations arise in glioblastomas during temozolomide therapy and mediate temozolomide resistance. *Clin. Cancer Res.* 15 (14), 4622–4629. doi:10.1158/1078-0432.CCR-08-3012

Yuan, F., Cong, Z., Cai, X., Zhu, J., Yuan, L., Wang, Y., et al. (2022). BACH1 as a potential target for immunotherapy in glioblastomas. *Int. Immunopharmacol.* 103, 108451. doi:10.1016/j.intimp.2021.108451

Zhang, Z., Hernandez, K., Savage, J., Li, S., Miller, D., Agrawal, S., et al. (2021). Uniform genomic data analysis in the NCI genomic data commons. *Nat. Commun.* 12 (1), 1226. doi:10.1038/s41467-021-21254-9