Chromatin accessibility analysis identifies \textit{GSTM1} as a prognostic marker in human glioblastoma patients

Yin-Cheng Huang\textsuperscript{1,6†}, Joseph Chieh-Yu Lai\textsuperscript{2†}, Pei-Hua Peng\textsuperscript{3}, Kuo-Chen Wei\textsuperscript{1} and Kou-Juey Wu\textsuperscript{3,4,5*}

**Abstract**

**Background:** Glioblastoma (GBM) is a malignant human brain tumor that has an extremely poor prognosis. Classic mutations such as \textit{IDH} (isocitrate dehydrogenase) mutations, \textit{EGFR} (epidermal growth factor receptor) alternations, and \textit{MGMT} (O6-methylguanine-methyltransferase) promoter hypermethylation have been used to stratify patients and provide prognostic significance. Epigenetic perturbations have been demonstrated in glioblastoma tumorigenesis. Despite the genetic markers used in the management of glioblastoma patients, new biomarkers that could predict patient survival independent of known biomarkers remain to be identified.

**Methods:** ATAC-seq (assay for transposase accessible chromatin followed by sequencing) and RNA-seq have been used to profile chromatin accessible regions using glioblastoma patient samples with short-survival versus long-survival. Cell viability, cell cycle, and Western blot analysis were used to characterize the cellular phenotypes and identify signaling pathways.

**Results:** Analysis of chromatin accessibility by ATAC-seq coupled with RNA-seq methods identified the \textit{GSTM1} (glutathione S-transferase mu-1) gene, which featured higher chromatin accessibility in GBM tumors with short survival. \textit{GSTM1} was confirmed to be a significant prognostic marker to predict survival using a different GBM patient cohort. Knockdown of \textit{GSTM1} decreased cell viability, caused cell cycle arrest, and decreased the phosphorylation levels of the NF-\kappaB (nuclear factor kappa B) p65 subunit and STAT3 (signal transducer and activator of transcription 3) (pSer727).

**Conclusions:** This report demonstrates the use of ATAC-seq coupled with RNA-seq to identify \textit{GSTM1} as a prognostic marker of GBM patient survival. Activation of phosphorylation levels of NF-\kappaB p65 and STAT3 (pSer727) by \textit{GSTM1} is shown. Analysis of chromatin accessibility in patient samples could generate an independent biomarker that can be used to predict patient survival.

**Keywords:** Glioblastomas, Next generation sequencing, ATAC-seq, Chromatin accessibility, RNA-seq, \textit{GSTM1}, STAT3

**Introduction**

Glioblastoma (GBM) remains a devastating disease with a median survival of 13–16 months despite advances in surgery, radiation therapy, chemotherapy and targeted therapy [1–3]. The classifications of primary glioblastomas are based on classic mutations, including \textit{IDH} mutations, 1p/19q codeletion, \textit{MGMT} promoter methylation, \textit{G-CIMP} (glioma-CpG island methylator phenotype) methylation, \textit{TERT} (telomerase reverse transcriptase)
promoter mutations, EGFR genomic alterations/mutations, LOH (loss of heterozygosity) at 10q, p16/INK4A deletion, and PTEN (phosphatase and tensin homolog) mutations [4–8]. These classic mutations can be used to provide prognostic implications. For example, EGFR alterations and TERT promoter mutations are associated with poor prognosis, whereas IDH mutations and MGMT promoter methylation are associated with favorable prognosis in glioblastoma patients [6–9].

Epigenetic alterations or dysfunctions in glioblastoma involves three levels, including DNA methylation, histone modifications, and chromatin remodeling [10, 11]. Whole genome sequencing showed the frequent incidence of histone H3 mutations in pediatric glioma patient samples and aberrant expression of EZH2 (enhancer of zeste homolog 2) and KDMs (lysine demethylases) in GBM patient samples [7, 10, 11]. HDACs (histone deacetylases) overexpression have also been shown to confer therapeutic resistance and tumor growth in glioblastoma [11, 12]. Therefore, EZH2 inhibitors, KDM inhibitors, and HDAC inhibitors may serve as therapeutic agents for GBM patients, although some of the inhibitors did not have satisfying results [10–12].

ATAC-seq (assay for transposase-accessible chromatin followed by sequencing) coupled with RNA-seq could be used in the profiling of chromatin accessibility regions in different types of human cancers to generate prognostic markers [13, 14]. For example, utilizing ATAC-seq to explore the differences in chromatin accessibility between neural stem cells and tumor-specific/migratory stem cells, TEAD1 (TEA domain family member 1) has been identified as a regulator of migration in glioblastoma [15].

Glutathione S-transferase mu-1 (GSTM1) is a member of the glutathione-S transferase (GST) family, which detoxifies by conjugating reduced glutathione (GSH) to target proteins in different organs [16]. GSTs could play a role in drug resistance through direct detoxification or involvement in the MAP (mitogen-activated protein) kinase pathway [17]. GSTM1 is one of the most abundant proteins in astrocytes [18]. In contrast to the association between GSTM1 polymorphism and tumorigenesis in other tumor types, there is no association between GSTM1 polymorphism and adult glioma [19]. Deletion of GSTM1 has been shown to be associated with pediatric glioma [20]. In mechanistic studies, GSTM1 has been shown to promote inflammatory signaling through activation of NF-kB and production of GM-CSF (granulocyte–macrophage colony stimulating factor) and CCL2 (chemokine [C–C motif] ligands 2) in astrocytes during brain inflammation [21]. GSTM1 also modulates allergen-induced NF-kB activation in asthmatic airway epithelium [22]. Although NF-kB signaling has been shown to promote gliomagenesis [23], the role of GSTM1 in enhancing NF-kB activity in glioblastoma remains to be determined. Regarding other signaling pathways involved in GBM, STAT3 has been shown to be a molecular hub and its role in GBM tumor progression has been demonstrated [24, 25]. However, the role of GSTM1 in activating STAT3 signaling also remains to be determined.

To compare the chromatin accessibility landscape between short-term (<0.5 year) and long-term (>2 year) surviving GBM sample types for the discovery of novel biomarkers, ATAC-seq coupled with RNA-seq from these two different GBM sample types was performed. Among the genes upregulated in GBM samples with short survival, GSTM1 was selected for further analysis due to its high hazard ratio and possible role in regulating signaling pathways (NF-kB, STAT3) involved in gliomagenesis [23–25]. Our results confirmed GSTM1 as a biomarker that could predict the survival of GBM patients. The STAT3 signaling pathway was also shown to be regulated by GSTM1. This report demonstrates that analysis of the chromatin accessibility landscape can lead to the discovery of a useful biomarker that predicts the survival of GBM patients.

Material and methods
Clinical samples collection
The patient samples were collected from tissues cryo-preserved by liquid nitrogen from the tissue bank of Chang Gung Memorial Hospital at Linkou and they were confirmed WHO grade IV glioblastoma samples (Table 1). The first set of 10 glioblastoma tissue (5 short-term survival and 5 long-term survival patients) were obtained for gene analysis (IRB consent, CGMH-201900138B0). These samples were selected based on the survival duration. From our archives, the short-term survival was defined as less than 0.5 years and long-term survival was defined as longer than 2 years. Patients with short-term or long-term survival reflected the chemoresistant (CR) or chemosensitive (CS) status, respectively. For abbreviation, the terminology of “short-survival” and “long-survival” will be used. The median age was 58 (ranged from 41 to 72, mean = 58.3). Seven patients were male and three were female. The tissue was stored in liquid nitrogen immediately after surgical resection. The DNA/RNA extraction and library construction were performed by standard protocols. For sequencing, we selected 2 normal tissues and 4 tumor tissues of patients with long survival and 4 tumor tissues of patients with short survival for RNA-seq; 4 normal tissues, 4 tumor tissues of patients with long survival and 4 tumor tissues of patients with long survival for ATAC-seq (the detailed sequencing information was shown in Table 1). The second cohort was 125 paraffin-embedded grade IV glioblastoma samples.
which were also obtained from the tissue bank of Chang Gung Memorial Hospital under IRB approval (CGMH-201900138B0). It was an independent cohort and the tumor tissues were collected consecutively from the neurosurgical department archives without any selection. These samples were used for immunohistochemistry (IHC) staining.

**ATAC-seq on frozen tissue**

Omni-ATAC was performed for gathering tissue ATAC-seq data by following the previous study [14]. Frozen tissue was thawed into 1 ml iced cold HB buffer (320 mM sucrose, 0.1 mM EDTA, 0.1% NP40, 5 mM CaCl2, 3 mM Mg(Ac)2, 10 mM Tris pH 7.8, 1 × protease inhibitors (Roche, cOmplete), and 167 μM β-mercaptoethanol in ultrapure water) inside 2 ml Dounce homogenizer. The thawed frozen tissue was homogenized with pestle A (loose) for 10 times stroke, followed by filtering homogenized sample with 100 μm nylon mesh. The filtered sample was homogenized again with pestle B (tight) for 20 times stroke. After releasing the nuclei by douncing, 400 ul of homogenized tissue was mixed with 400 ul of 50% iodixanol (inside 1 × homogenization buffer) followed by pipetting, making the final 25% gradient solution with a total volume of 800 ul of homogenized tissue. 29% iodixanol and 35% iodixanol solution was prepared and layered underneath the 25% gradient solution that contained homogenized tissue. The sample was centrifuged in swing-bucket 3000 g for 30 min at 4 °C to form a nuclei band that stayed at the interface between 29 and 35% iodixanol solutions. 300 μl of solution containing nuclei was collapsed followed by gathering 5 × 10⁴ counted nuclei into 1 ml ATAC-RSB buffer (10 mM Tris–HCl pH 7.4, 10 mM NaCl, and 3 mM MgCl₂ in ultrapure water) with 0.1% Tween-20. The nuclei were pelleted by centrifugation at 500 g for 5 min at 4 °C. After removing the supernatant, transposase mixture (25 μl 2 × TD buffer, 2.5 μl transposase (100 nM final), 16.5 μl PBS, 0.5 μl 1% digitonin, 0.5 μl 10% Tween-20, 5 μl ultrapure water) was directly added into nuclear pellet followed by pipetting 6 times. Transposase reaction was performed at Thermo-Mixer 1000 rpm for 30 min at 37 °C. Transposed DNA was collapsed followed by Zymo DNA Clean and Concentrator. Library was constructed by PCR amplification with NEB Next High-Fidelity PCR mix according to previous method [26]. The amplified library was purified by 1.8X Ampure XP beads clean-up through following the manufacture’s protocol. Constructed library was quality-controlled by Tapstation to avoid primer dimer, then sequenced by NextSeq 550 with 75 bp paired-end sequence.

**RNA extraction from frozen tissue and RNA-seq**

Frozen tissue was homogenized by MagNA Lyser Green Beads (03358941001, Roche). After homogenization, RNA was extracted by RNeasy according to the manufacturer’s protocol. RNA RIN score was checked by RNA screen tape with Tapstation (Agilent). RNA-seq library was constructed using the KAPA RiboMinus kit under stranded library construction. Library was sequenced by Illumina Nextseq 550 150 bp paired-end sequence.

**Sequence data analysis**

ATAC-seq data were aligned to hg38 by bowtie2 [27]. After the alignment, files were filtered according to the mapping quality and the PCR duplicates were removed.
Chromatin accessible regions were called by MACS [28]. Differential chromatin accessibility was performed by diffbind, R package [29]. RNA-seq data were aligned to hg38 by hisat2 [30]. After the alignment, files were filtered according to the mapping quality, and the abundance of gene expression by ht-seq were counted. Differential gene analysis was performed by edgeR [31]. For further comprehensive analysis, peak sets from diffbind results were annotated by ChIPSeeker, a R package with Gencode v31 [32]. Promoter was defined to be -3000 to 500 bp surrounding the transcription start sites of each transcript. Under the consideration of accessible region located in the promoter that correlated with gene activation, we selected promoter peak sets that correlated with the differential results and linked the accessible region to gene expression by the coordinate of each gene.

Quality control of clinical sample
ATAC-seq library were quality-controlled by tapstation before sequencing. TSS enrichment score were calculated for data filtering (TSS enrichment score >1) and principal component analysis (PCA) was performed during the bioinformatics processing. Before RNA-seq library construction, extracted RNAs were quality-controlled by RNA RIN value with tapstation. Samples with RIN value above 4 were selected for library construction followed by sequencing.

Immunohistochemistry (IHC) staining
The IHC staining of paraffin-embedded tissues of the independent cohort that contained 125 paraffin-embedded GBM samples was performed as previously described [33]. The tissue (4–5 μM in thickness) was mounted onto slides. The tissue was deparaffinized and rehydrated. Antigen retrieval was performed with citrate buffer and boiled at 100 °C for 10 min. The primary antibody was diluted as suggested (GSTM1, 1H4F2, Novus Biologicals, NBP2-22186) at 4 °C overnight and washed for 3 times. Then secondary antibody was incubated for 1 h at room temperature. After another 3 washes, DAB (3,3’-diaminobenzidine) chromogenic was applied, followed by hematoxylin counterstain (Thermo-Fisher Scientific Inc. UltraVision Quanto Detection System, TL-125-QHD).

IHC scoring
A scoring system for IHC was performed as described, from 0 to 3+ [34]. To further categorize the heterogeneous expression of DAB chromogen in the tissue, the scoring system of 0–1+ was defined as low-expression and 2+ to 3+ as high-expression. The personnel who scored the level of expression was an experienced neurosurgeon with extensive laboratory experience.

Survival analysis
A Kaplan–Meier survival analysis, non-parametric statistics, was used to estimate the survival functions of high and low expression of target genes. Other factors, such as sex, age, MGMT (O6-methylguanine-methyltransferase) status and IDH1 (isocitrate dehydrogenase 1) mutation was analyzed with log-rank and Wilcoxon test. The person who scored the expression of IHC staining was blind to the survival data of the patients.

Expression of GSTM1 by quantitative RT-PCR (qRT-PCR)
The quantitative real-time PCR (qRT-PCR) was used to quantify the expression of GSTM1 with SYBR-Green system. Primers for GSTM1 and β-actin was designed by Primer 3 (GSTM1; forward: AGCCGGCCATGTTTTGCAGGAA, reverse: TTCTCCAGCCCTCAAGCGG). Briefly, the cDNA triplicate experiments were amplified at 95 °C for 1 min, followed by denaturing, annealing and extension cycles in SYBR Green master mix (Invitrogen) with LightCycler 480 (Roche Molecular Systems, Inc). The expression level is calculated as ΔCT = CT(target gene) – CT(a reference gene).

siRNA knockdown
The siRNA that targeted GSTM1 was obtained from Origene (SR301988, OriGene Tech, Rockville, MD, USA) with a scrambled non-silencing oligonucleotide as a control. U87 cells were seeded in 6 and 12-well plate with DMEM (10%FBS). The cells were transfected with siRNA (100 pmol) at a confluent density of 70–80% using Lipofectamine 2000 (Thermo Fisher Scientific, Inc.) according to the protocol. The efficiency of knockdown is validated by qRT-PCR.

Cell growth assay
WST-1, a cell growth assay (TAAR-WBF9, Tools Biotech. Co. Ltd.), was performed to assess the cell viability under knockdown of target gene. In brief, 4,000 cells were plated in 96-well plates and 10 μl of WST-1 was added to each well after knockdown of target gene for 48 h. The relative cell density was obtained with a ELISA reader (absorbance 450 nm, Infinite M200, Tecan Group Ltd., Switzerland) 2 h later after a gentle shake. The optical density (OD) was calculated as: inhibition ratio (%) = (1 − OD value of the experimental group/OD of the control group) × 100.

Cell cycle analysis
The glioma U87 cells undergoing GSTM1 knockdown for 48 h were harvested and washed in phosphate-buffered saline, along with scrambled siRNA-treated control cells. The cells were fixed in ethanol and treated
with RNase prior to propidium iodide (PI) staining (50 ug/ml). According to the standard analysis protocol, the PI histogram was demonstrated with 5000 cell counts (BD LSRFortessa flow cytometer). FlowJo with Watson Pragmatic analysis was used to analyze the DNA contents of cells.

**Western blot analysis**
The cell lysate was denatured and the protein amount were quantified before loading into the polyacrylamide gel. After adequate separation of protein contents by electrophoresis, the gel was transferred to a PVDF membrane. After blocking with 5% skim-milk buffer, overnight incubation with the primary antibody (diluted as suggested in the individual data sheet), followed by a 1:10000X dilution of horseradish peroxidase–conjugated anti–rabbit or anti-mouse antibody (GeneTex, Hsinchu, Taiwan) for one hour, was performed. Signals were detected by using Clarity Western ECL blotting Substrate (Bio-Rad Laboratories, Hercules, CA, US).

**The list of antibodies used**
GSTM1(1H4F2) (IHC and western): Novus Biologicals, LLC. (NBP2-22186); STAT3(C-20): Santa Cruz Biotechnology, Inc. (sc-482); Phospho-STAT3 (pSer727) Cell Signaling Technology, Inc. (#9134); p38 MAPK: Cell Signaling Technology, Inc. (#9212); Phospho-p38 MAPK (Thr180/Tyr182): Cell Signaling Technology, Inc. (#9211); NF-κB p65 (C22B4): Cell Signaling Technology, Inc. (#4764); NF-kB phospho-p65 (Ser536)(93HI): Cell Signaling Technology, Inc. (#3033); β-Actin Antibody (C4): Santa Cruz Biotechnology, Inc. (sc-47778).

**Results**

**Analysis of ATAC-seq and RNA-seq datasets in GBM patient samples**
To identify prognostic biomarkers of glioblastoma (GBM) based on chromatin accessibility and differential gene expression, we performed ATAC-seq and RNA-seq on tumor samples collected from GBM patients with short-term (< 0.5 years) and long-term (> 2 years) survival (2 normal tissues and 4 tumor tissues from each of the long- and short-term survival patients were analyzed by
RNA-seq and ATAC-seq, while 2 additional normal tissues were analyzed by ATAC-seq (see the “Methods” section and Table 1 for details) (Fig. 1a). Comparing the chromatin accessibility between these two sets of samples (see “Methods” section), we identified 2,957 regions with significantly increased accessibility (cut off by 1.5-fold change, \(p \text{ value} < 0.05\) in short-survival (chemoresistant; CR) vs. long-survival (chemosensitive; CS) patient samples (Fig. 1b). Among the 2,957 regions, we only focused on 1,053 regions that also exhibited increased accessibility when compared with normal brain tissues (Fig. 1c). To understand the functional implications of these regions, we annotated their genomic features. Among the 1053 genomic regions with increased accessibility, 14.05% were located in the promoter regions of 143 genes (Fig. 1d). KEGG (Kyoto Encyclopedia of Genes and Genomes) enrichment analysis of these 143 genes revealed various pathways, such as the hedgehog signaling pathway (Fig. 1e). In addition to the analysis of ATAC-seq datasets, we also applied differential gene expression analysis by RNA-seq and identified 4,003 and 630 upregulated genes whose expression levels were increased in tumor vs. normal and short-survival vs. long-survival patient samples, respectively (Additional file 1: Fig. S1a). KEGG analysis of increased gene expression in tumor tissues (vs. normal tissues) and short-survival (vs. long-survival) tumor samples showed the pathways of cell cycle and MAPK signaling in these two groups (Additional file 1: Fig. S1b, c). We ranked the fold change of gene expression with significant induction of chromatin accessibility, and the results showed that the expression of most genes was increased with the induction of chromatin accessibility (Additional file 1: Fig. S1d).

High chromatin accessibility and gene expression in tumor samples from patients with short survival identified GSTM1 as a putative biomarker

To compare the variations of the genomic signature between samples, we performed principal component analysis (PCA) on the ATAC-seq and RNA-seq datasets. The results showed that the tumor samples (short-survival, long survival) were grouped together compared to the normal tissue group (Additional file 1: Fig. S2a). After analyzing ATAC-seq and RNA-seq datasets separately, we combined the analyses of these two datasets by defining the accessible regions on the promoter (Additional file 1: Fig. S2b) and presenting scatter plots between chromatin accessibility on the promoter and gene expression (Additional file 1: Fig. S2c). We found that there were positive correlations between accessible regions and gene expression in both tumor vs. normal brain tissue and short-survival vs. long-survival patients (Additional file 1: Fig. S2c). To consolidate the results from these two datasets, we set the cut off of log2(2) fold change in the ATAC-seq and RNA-seq datasets and the Venn diagrams of overlapping subsets showed that 11 overlapping genes contained higher chromatin accessibility in the promoter regions and higher gene expression both in tumor vs. normal and short-survival vs. long-survival patients (Fig. 2a). A web-based TCGA (The Cancer Genome Atlas) data analysis was applied to these 11 genes, and the glutathione S-transferase mu-1 (GSTM1) gene was identified as the most significant gene predicting poor survival in clinical patients (Table 2). Heatmap analysis of gene expression and chromatin accessibility showed that GSTM1 had the highest chromatin accessibility in tumor samples with short survival times (Fig. 2b, c). The ATAC-sequencing tracks from the genome datasets (short-survival, long-survival, normal tissue) showed high chromatin-accessible regions in the promoter of GSTM1 in certain short-survival tumor samples (Fig. 2d), consistent with the genomic dataset analysis. The ATAC sequencing tracks from a different genomic locus were shown as a control (Fig. 2d). Positive correlations between chromatin accessibility and gene expression were found in both the promoter and the gene body in GSTM1 (Additional file 1: Fig. S2d).

Verification of GSTM1 as a prognostic marker in glioblastoma from a different patient cohort and analysis of TCGA dataset

To further verify the clinical significance of GSTM1 using a different cohort of grade IV glioblastoma patients (\(N=125\)), IHC staining of GSTM1 was performed (Fig. 3a). High levels of GSTM1 expression (IHC staining: ++ to ++++) compared to low GSTM1 expression (IHC staining: – to +) (see “Methods” section) were defined to compare the GSTM1 IHC staining in GBM samples (Fig. 3a). To measure the extent of GSTM1 expression in the tumor region, direct counting of GSTM1-positive cells in 50 regions from the tumor vs. normal brain region was performed (high power field, 400X) (Additional file 1: Fig. S3). The statistics of high GSTM1 staining were shown in tumor tissues versus normal tissues (Fig. 3b). Seventy-four out of the 125 patients had high levels of GSTM1 expression (++ to ++++) compared to 51 patients who had low GSTM1 expression (– to +) (Fig. 3c). To examine the independent prognostic value of GSTM1, further analysis showed that the expression of GSTM1 was independent of other factors, including sex, age, MGMT promoter hypermethylation, or IDH1 mutation (Fig. 3c). Furthermore, the expression of GSTM1 was significantly correlated with overall survival in this GBM cohort (Fig. 3d, \(p = 0.0055\) and 0.0088 by log-rank and Wilcoxon test, respectively), indicating that GSTM1 could serve as...
a biomarker to predict GBM patients’ clinical outcome. Finally, GSTM1 as a prognostic marker in GBM was corroborated through analysis of TCGA dataset (Additional file 1: Fig. S4a). Interestingly, a search for a downregulated gene biomarker in short-survival patients showed that ZNF727 expression was repressed...
and its gene tracks were decreased in short-survival samples (Additional file 1: Fig. S4b–d). TCGA analysis also showed that low expression of ZNF727 could predict poor overall survival in GBM patients (Additional file 1: Fig. S4e).

Knockdown of GSTM1 decreased cell viability and caused cell cycle arrest

To further examine the role of GSTM1 in regulating tumor cell phenotypes, GSTM1 was knocked down using an siRNA approach in U87 glioma cells. The expression of GSTM1 was knocked down by 96.8±1.3% as demonstrated by RT–PCR after 48 h (Fig. 4a). Knockdown of GSTM1 decreased cell viability (down to 67±3% compared to scrambled siRNA control) at 48 h (Fig. 4b). Further analysis of the cell cycle profile showed that knockdown of GSTM1 caused more tumor cells to be arrested at G0/G1 phase (Fig. 4c, d). The above results indicate that knockdown of GSTM1 reduces tumor cell viability and causes cell cycle arrest.

Regulation of the NF-κB and STAT3 signaling pathways by GSTM1

Since GSTM1 has been shown to regulate NF-κB signaling [21, 22], we examined whether the NF-κB signaling pathway may be regulated by GSTM1 in U87 cells. Knockdown of GSTM1 in U87 cells decreased the phosphorylated NF-κB p65 (phospho-p65) levels but not the phosphorylated MAP kinase (phospho-p38) levels (Fig. 5a). Since STAT3 has been shown to be a molecular hub and plays a significant role in GBM tumor progression [24, 25], we further examined whether knockdown of GSTM1 significantly decreased phosphorylated STAT3 (STAT3-pSer727) levels. The results showed that knockdown of GSTM1 decreased the phosphorylated STAT3 (STAT3-pSer727) levels, but not the total STAT3 levels (Fig. 5a). This result indicates that GSTM1 regulates both the STAT3 pathway by GSTM1 is direct or indirect remains to be determined.
**Discussion**

Classic genetic analyses in glioblastoma have shown that IDH1 mutation, MGMT promoter methylation, 1p/19q codeletion, G-CIMP methylation, BRAFV600E mutation, and H3K27 mutation are associated with glioblastoma overall survival [6–9]. However, biomarkers independent of these known biomarkers to predict the survival between short-survival and long-survival GBM patients are unknown [35]. In this report, we selected samples from short-survival (survival < 0.5 years) and long-survival (survival > 2 years) patients followed by ATAC-seq and RNA-seq analyses and identified GSTM1 as a prognostic marker of GBM, which was confirmed in a different cohort and through TCGA analysis (Additional file 1: Fig. S4a). Since GSTM1 promoter accessibility and increased expression was only detected in 2 out of 4 short survival patient samples (Fig. 2d), further search for other markers should be mandatory. Following this rationale, similar analysis showed that low expression of ZNF727 could predict poor overall survival in GBM patients (Additional file 1: Fig. S4b–e). These results suggest that our approach is able to effectively identify prognostic markers that could predict GBM patients’ clinical outcomes.

GSTM1 methylation has been identified in an integrative analysis of molecular aberrations in GBM genomes [36]. Interestingly, GSTM1 methylation and CD40 methylation were coupled under the G-CIMP subtype of GBM [7]. Since G-CIMP methylation has been linked to a favorable prognosis in GBM patients [7], methylation of GSTM1 may contribute to a favorable prognosis. Based on the ability of GSTM1 to trigger NF-kB and

---

**Fig. 4** Viability and cell cycle analysis using siRNA-mediated knockdown of GSTM1 showed that knockdown of GSTM1 decreased cell viability and caused cell cycle arrest in U87 cells. **a** Knockdown of GSTM1 showed a significant reduction of GSTM1 mRNA expression by RT-PCR assays. Data from three independent experiments were expressed as mean ± s.d. (s.d.: standard deviation). The asterisk (*) indicated statistical significance (P < 0.05) between GSTM1 siRNA and scrambled siRNA groups. **b** Knockdown of GSTM1 decreased the percentage of viable U87 cells down to 84.3%, compared to control knockdown cells at 48 h. Data from three independent experiments were expressed as mean ± s.d. The asterisk (*) indicated statistical significance (P < 0.05) between GSTM1 siRNA and scrambled siRNA groups. **c** The cell cycle profile showed that knockdown of GSTM1 caused G0/G1 cell cycle arrest after GSTM1 was knocked down at 48 h. **d** The percentage of G0/G1 cells was significantly increased in GSTM1 siRNA knockdown U87 cells compared to scrambled siRNA U87 cells. Data from three independent experiments were expressed as mean ± s.d. A bar graph containing different cell cycle phases (G0/G1, S, G2/M) was shown on the right.
STAT3 signaling, our results also support the linkage of GSTM1 overexpression to poor prognosis of GBM patients. In addition to IHC staining of GSTM1 in GBM samples, whether serum levels of GSTM1 could be used as a prognostic marker of GBM patients remains to be determined.

The role of GSTM1 in tumor cell phenotypes has been shown: knockdown of GSTM1 decreased tumor cell viability and caused cell cycle arrest at G0/G1 phase. The possible pathways regulated by GSTM1 are the NF-kB and STAT3 pathways. In vitro results support the clinical correlation. Our result is also supported by another glioblastoma cohort in which STAT3-pSer727 levels were correlated with clinical outcome [37]. GSTM1 may conjugate reduced glutathione (GSH) to target proteins to further activate NF-kB and STAT3 signaling. Further searches for putative protein targets by GSTM1 should reveal the molecular mechanism of GSTM1-mediated activation of STAT3 and NF-kB pathways. A model of GSTM1-activated kinase pathways is depicted in Fig. 5b.

In conclusion, using ATAC-seq coupled with RNA-seq to analyze datasets from short-survival versus long-survival glioblastoma and tumor versus normal tissues, GSTM1 was shown to be a biomarker capable of predicting GBM patient survival. Genomic approaches will facilitate the discovery of new biomarkers for glioblastoma patients.

**Abbreviations**

GBM: Glioblastoma; IDH1: Isocitrate dehydrogenase 1; EGFR: Epidermal growth factor receptor; MGMT: O6-methylguanine-methyltransferase; G-CIMP: Glioma CpG island methylator phenotype; ATAC-seq: Assay for transposase accessible chromatin followed by sequencing; GST: Glutathione S-transferase; GSTM1: Glutathione S-transferase mu-1; LOH: Loss of heterozygosity; KDM: Lysine demethylase; HDAC: Histone deacetylase; RT-PCR: Real-time PCR; CR: Chemo-resistant; CS: Chemo-sensitive; KEGG: Kyoto encyclopedia of genes and genomes; IHC: Immunohistochemistry.

**Supplementary Information**

The online version contains supplementary material available at [https://doi.org/10.1186/s13148-021-01181-8](https://doi.org/10.1186/s13148-021-01181-8).

**Additional file 1: Fig. S1.** Analysis of RNA-seq datasets from GBM patient samples revealed the number of upregulated genes and the possibly pathways involved. A model summarizes that GSTM1 overexpression induces the STAT3 and NF-kB pathways. The possible mechanism may go through GSH-conjugated targets that subsequently induce the phosphorylation of STAT3 and/or NF-kB p65.

---

**Fig. 5** Western blot analysis showed that knockdown of GSTM1 by siRNA decreased the phosphorylated levels of NF-kB p65 and STAT3 (pSer727) and a model to summarize the results in this report was presented. **a** Western blot analysis showed that knockdown of GSTM1 in U87 cells at 48 h decreased the phosphorylated NF-kB p65 (phospho-p65) and phosphorylated STAT3 (pSer727) levels, but not the phosphorylated MAP kinase levels. The average levels of proteins from densitometer scanning normalized by β-actin were shown in the bottom of the protein bands. Bar graph showed the levels of different proteins that were quantified and normalized with β-actin by Western blot analysis (triplicate results were shown in Fig. S5a, Additional file 1: Fig. S5b). Data from three independent experiments are expressed as mean ± s.d. The asterisk (*) indicated statistical significance (P < 0.05) between GSTM1 siRNA and scrambled siRNA groups. **b** A model summarizes that GSTM1 overexpression induces the STAT3 and NF-kB pathways. The possible mechanism may go through GSH-conjugated targets that subsequently induce the phosphorylation of STAT3 and/or NF-kB p65.
Acknowledgements
We thank Dr. Chia-Lin Wei (Jackson Laboratory) for comments on the manuscript. We would like to acknowledge the contributions of research assistants (You-jie Chen, Wei-Ting Lin) for patient sample handling and bioinformatics analysis.

Authors’ contributions
Y.C.H. and K.J.W. processed patient samples, clinical data analysis, performed IHC staining; P.H.P. performed Western blot analysis and cell line experiments; J.C.Y.L. performed bioinformatics analysis; K.J.W. wrote the manuscript with the help form Y.C.H. and P.H.P. All authors read and approved the final manuscript.

Funding
This work was supported in part to K.J.W. by Ministry of Science and Technology Summit and Frontier Grants (MOST 108-2321-B-182A-005, MOST 109-2326-B-182A-002, MOST 110-2326-B-182A-004), Chang Gung Memorial Hospital (OMRPG30012, NMRPG36192, CORPGRG30023, NMRPG36072, CORPGRG30261, CORPGRG30662, NMRPG36073, NMRPG36193, CORPGRG30233, OMRPG30013); to Y.C.H. by Chang Gung Memorial Hospital (NMRPG36141, 36142); and to P.H.P. by Ministry of Science and Technology (MOST 109-2320-B-182A-022, MOST 110-2320-B-182A-019), Chang Gung Memorial Hospital (NMRPG30511, NRRPG3L0151).

Availability of data and materials
The datasets used for this report are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
The study conducted in this report was approved by the ethics committee of Chang Gung Memorial Hospital at Linkou (IRB consent, CGMH-20190013880). Written informed consent was obtained from all participants.

Consent for publications
Not applicable.

Competing interests
The authors declared that they have no conflict of interests.

Author details
1 Department of Neurosurgery, Chang Gung Memorial Hospital at Linkou, Taoyuan 333, Taiwan. 2 Institute of Biomedical Science, China Medical University, Taichung 404, Taiwan. 3 Cancer Genome Research Center, Chang Gung Memorial Hospital at Linkou, Guishan District, Taoyuan 333, Taiwan. 4 Institute of Cellular and Organismic Biology, Academia Sinica, Taipei 115, Taiwan. 5 Graduate Institute of Clinical Medical Sciences, Chang Gung University, Taoyuan 333, Taiwan. 6 Department of Medicine, Chang Gung University, Taoyuan 333, Taiwan.

Received: 28 April 2021 Accepted: 4 October 2021
Published online: 03 November 2021

References
1. Surawicz TS, Davis F, Freels S, Laws ER Jr, Menck HR. Brain tumor survival: results from the national cancer data base. J Neurooncol. 1998;40(2):151–60.
2. Laws ER, Parney IF, Huang W, Anderson F, Morris AM, Asher A, et al. Survival following surgery and prognostic factors for recently diagnosed malignant glioma: data from the glioma outcomes project. J Neurosurg. 2003;99(3):467–73.
3. Gramatzki D, Roth P, Rushing EJ, Weller J, Andratschke N, Hofer S, et al. Bevacizumab may improve quality of life, but not overall survival in glioblastoma: an epidemiological study. Ann Oncol. 2018;29(6):1431–6.
4. Brennan CW, Verhaak RG, McKenna A, Campo B, Noushmehr H, Salama SR, et al. The somatic genomic landscape of glioblastoma. Cell. 2013;155(2):462–77.
5. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 world health organization classification of tumors of the central nervous system: a summary. Acta Neuropathol. 2016;131(6):803–20.
6. Bell EH, Pugh SL, McElroy JP, Gilbert MR, Mehta M, Klimovitch AC, et al. Molecular-based recursive partitioning analysis model for glioblastoma in the temozolomide era: a correlative analysis based on NRG oncology RTOG 0525. JAMA Oncol. 2017;3(6):784–92.
7. Aquafresca R, Huhn R, Bangalore S, Salam A, Bhanot P, Spath D, et al. Survival of glioma patients: a correlative analysis. Acta Neuropathol. 2018;135(4):467–79.
8. Ramos AD, Magge RS, Ramakrishna R. Molecular pathogenesis and emerging treatment for glioblastoma. World Neurosurg. 2018;116:495–504.
9. Leu S, von Felten S, Frank S, Vassella E, Vajta I, Taylor E, et al. IDH/MGMT-driven molecular classification of low-grade glioma is a strong predictor for long-term survival. Neuro Oncol. 2013;15(4):469–79.
10. Romani M, Pistillo MR, Bignell G. Epigenetic targeting of glioblastoma. Front Oncol. 2018;8:448.
11. Suter RK, Rodriguez-Blanco J, Ayad NG. Epigenetic pathways and plasticity in brain tumors. Neurobiol Dis. 2020;145:105060.
12. Yang WB, Hsu CC, Hsu TI, Liou JP, Chang KY, Chen PY, et al. Increased activation of HDAC1/2/6 and Sp1 underlies therapeutic resistance and tumor growth in glioblastoma. Neuro Oncol. 2020;22(10):1439–51.

13. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013;10(12):1213–8.

14. Corces MR, Granja JM, Shams S, Louie BH, Seoane JA, Zhou W, et al. The chromatin accessibility landscape of primary human cancers. Science. 2018;362(6413):eaav1698.

15. Tomé-García J, Erfani P, Nudelman G, Tsankov AM, Katsyv I, Tejero R, et al. Analysis of chromatin accessibility uncovers TEAD1 as a regulator of migration in human glioblastoma. Nat Commun. 2018;9(1):4020.

16. Strange RC, Spiteri MA, Ramachandran S, Fryer AA. Glutathione-S-transferase family of enzymes. Mutat Res. 2001;481(1–2):21–6.

17. Townsend DM, Tew KD. The role of glutathione-S-transferase in anti-cancer drug resistance. Oncogene. 2003;22(47):7369–75.

18. Sharma K, Schmitt S, Bergner CG, Tyanova S, Kannaiyan N, Manrique-Hoyos N, et al. Cell type- and brain region-resolved mouse brain proteome. Nat Neurosci. 2015;18(12):1819–31.

19. Lai R, Crevier L, Thabane L. Genetic polymorphisms of glutathione s-transferases and the risk of adult brain tumors: a meta-analysis. Cancer Epidemiol Biomark Prev. 2005;14(7):1784–90.

20. Salnikova LE, Belopolskaya OB, Zelinskaya NI, Rubanovich AV. The potential effect of gender in CYP1A1 and GSTM1 genotype-specific association with pediatric brain tumor. Tumour Biol. 2013;34(5):2709–19.

21. Kano SI, Choi EY, Dohi E, Agarwal S, Chang DJ, Wilson AM, et al. Analysis of chromatin accessibility genome-wide. Curr Protoc Mol Biol. 2015;109:21–91.

22. Polosukhin VV, Polosukhin IV, Hoskins A, Han W, Abdolrasulnia R, Blackwell TS, et al. Glutathione S-transferase M1 modulates allergen-induced NF-κB activation in asthmatic airway epithelium. Allergy. 2014;69(12):1666–72.

23. Cahill KE, Morshed RA, Yamini B. Nuclear factor-κB in glioblastoma: insights into regulators and targeted therapy. Neuro Oncol. 2015;17(3):329–39.

24. Ouedraogo ZG, Biau J, Kemeny J-L, Morel L, Verelle P, Chautard E. Role of STAT3 in genesis and progression of human malignant gliomas. Mol Neurobiol. 2017;54(8):5780–97.

25. Brantley EC, Benveniste EN. Signal transducer and activator of transcription-3: a molecular hub for signaling pathways in gliomas. Mol Cancer Res. 2008;6(5):675–84.

26. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: a method for assaying chromatin accessibility genome-wide. Curr Protoc Mol Biol. 2015;109:21–91.

27. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357–9.

28. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9(9):R137.

29. Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature. 2012;481(7381):389–93.

30. Kim D, Langmead B, Salzberg SL, HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12(4):357–60.

31. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139–40.

32. Yu G, Wang LG, He QY. ChIPseeker: an R/bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics. 2015;31(14):2382–3.

33. Yang M-H, Chiang W-C, Chou T-Y, Chang S-Y, Chen P-M, Teng S-C, et al. Increased NBS1 expression is a marker of aggressive head and neck cancer and overexpression of NBS1 contributes to transformation. Clin Cancer Res. 2006;12(2):507–15.

34. Fedchenko N, Reifenrath J. Different approaches for interpretation and reporting of immunohistochemistry analysis results in the bone tissue - a review. Diagn Pathol. 2014;9:221.

35. Mazaris P, Hong X, Altshuler D, Schultz L, Poisson LM, Jain R, et al. Key determinants of short-term and long-term glioblastoma survival: a 14-year retrospective study of patients from the Hermelin Brain Tumor Center at Henry Ford Hospital. Clin Neurol Neurosurg. 2014;120:103–12.

36. Sintupisut N, Liu PL, Yeang CH. An integrative characterization of recurrent molecular aberrations in glioblastoma genomes. Nucleic Acids Res. 2013;41(19):8803–21.

37. Lin GS, Chen YP, Lin ZX, Wang XF, Zheng ZQ, Chen L. STAT3 serine 727 phosphorylation influences clinical outcome in glioblastoma. Int J Clin Exp Pathol. 2014;7(6):3141–9.

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