Identifying potential diagnostic biomarkers in human glioblastoma using integrative transcriptomic analysis

Andres Turizo-Smith[a,*] • Daniel Garcia-Niño[b] • Liliana Lopez-Kleine[c]

[a,] Departamento de Patología, Facultad de Medicina, Universidad Nacional de Colombia, Bogotá, Colombia. andturizosm@unal.edu.co*
[b] Departamento de Matemáticas, Facultad de Ciencias, Universidad Nacional de Colombia, Bogotá, Colombia.
[c] Departamento de Estadística, Facultad de Ciencias, Universidad Nacional de Colombia, Bogotá, Colombia.

Resumen

El glioblastoma (GB) es el tumor cerebral maligno más común y agresivo en humanos y presenta un pronóstico desfavorable. Los tratamientos existentes han tenido un éxito limitado en la prolongación de la supervivencia. La heterogeneidad tumoral y la caracterización precisa del GB siguen siendo un desafío, y actualmente no se dispone de biomarcadores efectivos y específicos de la enfermedad. Por lo tanto, la identificación y comprensión de moléculas clave responsables del fenotipo maligno del GB permitirán generar nuevos blancos terapéuticos potenciales.

El objetivo de este estudio es identificar genes diferencialmente expresados entre pacientes y controles, con el fin de reconocer firmas genéticas que permitan diferenciar los tipos de GB. Reanalizamos cinco conjuntos de datos transcriptómicos de GB disponibles públicamente, comparando controles y pacientes para la detección de genes diferencialmente expresados (DEG). Además, realizamos análisis de enriquecimiento y supervivencia, y confrontamos los resultados con la literatura científica para proponer biomarcadores diagnósticos y pronósticos potenciales.

Identificamos tres grupos de DEG que afectan diversas características que favorecen el desarrollo tumoral, así como cuatro DEG que no habían sido previamente reportados en GB. Los genes identificados representan nuevos biomarcadores potenciales para el diagnóstico o la mejora de tratamientos, y también permiten obtener nuevos conocimientos sobre los mecanismos moleculares implicados y una mejor comprensión de los subtipos de GB.

Los DEG clave son posibles blancos para el diagnóstico, la subtipificación, el pronóstico y el tratamiento, con el objetivo de mejorar la terapia y la supervivencia en GB. Los hallazgos principales indican que PDPN está asociado con la linfangiogénesis y la metástasis, lo que lo convierte en un marcador atractivo para la clasificación del GB, mientras que la glicosilación alterada sugiere un papel en la progresión del cáncer, haciendo de NEBL un buen biomarcador potencial de progresión.

Palabras Claves: *Astrocitoma, Clasificación, Biología Computacional, Diagnóstico, Transcriptoma.*

Abstract

Glioblastoma (GB) is the most common and aggressive malignant brain tumor in humans and has a bad prognosis. Existing treatments have had limited success in prolonging survival. Tumor heterogeneity and accurate characterization of GB remain challenging, and effective disease-specific biomarkers are currently unavailable. Therefore, identifying and understanding of key molecules responsible for the malignant phenotype of GB will allow the generation of new potential therapeutic targets.

The aim of this study is to identify differentially expressed genes between patients and controls to recognize genetic signatures that will allow differentiation of GB types. We re-analyzed five publicly available GB transcriptomic datasets, comparing controls and patients for the detection of differentially expressed genes (DEG). Additionally, we performed enrichment and survival analysis, and confronted the results with scientific literature to propose potential diagnostic and prognostic biomarkers.

We identified three groups of DEGs that affect several characteristics that favor tumor development and four DEGs that had not been reported earlier for GB. The identified genes are potential new biomarkers for diagnosis or to improve treatments, and they also allow gaining new insights into understanding molecular mechanisms and a better understanding of GB subtypes.

Key DEGs are potential targets for diagnosis, subtyping, prognosis, and therapy to improve GB treatment and survival. Main findings are that PDPN is associated with lymphangiogenesis and metastasis, making it an attractive marker for GB classification as well as the altered glycosylation indicating a role in cancer progression making of NEBL a good potential progression biomarker.

Key Words: *Astrocytoma, Classification, Computational Biology, Diagnosis, Transcriptome.*

Introduction

Glioblastoma (GB) is the most common and lethal variant of glioma. Mean survival is only 14-15 months with 10 % survival probability of 5 years (Gallego, 2015). Improving the survival time of GB patients remains a challenging management problem. Patients with GB have a dramatically poor prognosis due to its cellular and molecular heterogeneity. For that reason, it is essential to find molecular targets that can serve as biomarkers and possible druggable targets, that should lead to a new classification system for patient stratification, prognosis prediction and selection of appropriate therapies that could provide therapeutic opportunities for this deadly neoplasm. This tumor is also flexible and adaptative to different adverse conditions, such as nutrient deprivation and other multiple interactions between the tumor microenvironment and immune cells that may drive and maintain its development. The identification of potential biomarkers in such a complex disease requires the integration of several data sources and a data analysis pipeline ensuring correct pre-processing, filtering and selection of core genes that could serve the above purpose.

Considering the glioblastoma characteristics, we searched for gene expression data of glioblastoma patients and healthy controls, with the aim to identify a gene signature or some biomarkers that are potentially good candidates for diagnosis, subtyping, improving treatment and, also to gain new insights to understand the underlying molecular mechanisms in GB. To achieve this goal, we identified differentially expressed genes in five publicly available GB datasets comparing controls and patients. This identification helped us to propose a gene signature for the differentiation of GB types. Further, we conducted an enrichment analysis of the differentially expressed genes (DEGs) lists. This allowed the confirmation of previously known processes in GB and, also, the identification of new crucial genes that had not previously been reported in GB. We selected the principal potential biomarkers through a thorough literature review and survival analysis to propose them for diagnosis or as future drug targets.

Our main finding is that the key DEGs can be classified into three groups: Genes that code for proteins of the collagen family (COL1A1, COL4A2, COL6A1), genes that code for transcription factors (HOXA10, HOXD10, HOXA5) and genes that code for proteins related to angiogenesis and/or cellular remodeling and a higher invasive potential (PDPN, VEGFA, CHI3L1). Moreover, we found four DEGs that have not been reported earlier as related to GB and could have a high potential of being valuable biomarkers and be possible druggable objectives. They have been reported to be implicated in carcinogenesis

or progression in other cancers but not directly related to GB (HS3ST4, GCNT4, NEBL, ST6GALNAC1).

Materials and methods

We selected five data sets from the NCBI GEO database that compare GB vs. regular brain tissue expression. Each data set was pre-processed and analyzed separately. Pre-processing and quality control were conducted: selecting random samples from each condition (GB/normal) to balance samples, gene filtering, and normalization to prevent mean-variance dependence. The results of these steps were confirmed using descriptive statistics.

Transcriptomic analysis.

Further, we identified DEGs using Bioconductor specialized packages (Carlson, 2017) version 3.16 in R version 4.2.1 and conducted an enrichment analysis for common DEGs in DAVID tool . The direction of differential expression was reported by fold-change outputs. Main DEGs were used for survival analysis.

The Genome Expression Omnibus database of NCBI (https://www.ncbi.nlm.nih.gov/gds) was searched for gene expression data of glioblastoma patients and healthy controls (without previous therapy). The datasets listed in Table 1 met the criteria. Most of these datasets had chip-seq, non-coding RNA or even methylation data that were not used. All datasets were accessed in the second semester of 2021.

Table 1. Data sets from NCBI with accession numbers, description of content and samples used in our investigation after random selection.

GEO accession number	Description	Samples used (before - after random selection)[a]
GSE119834 (D. W. Huang et al., 2009; Mack et al., 2019; Sherman et al., 2022)	RNAseq from GB samples, glioblastoma stem cells (GSC) and neural stem cells (NSC). Fragments Per Kilobase Million (FPKM) transformation.	*Before*: 45 GB; 44 GSC; 9 NSC. 19471 genes. *After*: 9 GB; 9 NSC. 19048 genes.
GSE145645 (Xu et al., 2021)	RNAseq from GB biopsies, normal brain tissues (NB), and cell line counterparts. FPKM-lg2 transformation.	*Before*: 32 GB; 3 NB. 19774 genes. *After*: 4 GB; 3 NB. 17963 genes.
GSE147352 (T. Huang et al., 2021)	RNAseq from GB biopsies, lower grade gliomas (LGG) and NB.	*Before*: 85 GB; 18 LGG; 15 NB. 35149 genes. *After*: 17 GB; 15 NB. 32661 genes.
GSE151352	RNAseq of normal/tumor tissue pairs of GB patients (same patient). rlog-normalized count data.	*Before*: 12 GB; 12 NB. 33245 genes. *After*: 12 GB; 12 NB. 20565 genes.
GSE159851 (Schaffenrath et al., 2021)	RNAseq from endothelial cells isolated from primary and secondary brain tumors. GB, Adenocarcinoma brain metastasis (BM) and NB.	*Before*: 5 GB; 6 NB; 6 BM. 20336 genes. *After*: 5 GB; 6 NB. 19174 genes.

[a]Filtered by quality analysis based on descriptive statistics analysis.

Ethical Considerations

This study is based solely on the re-analysis of publicly available transcriptomic data, which were generated under the ethical guidelines and approvals of the original studies. Since the data used are anonymized and publicly available, this research falls under the category of secondary research, which, according to international and national standards, does not require new ethical approvals. In Colombia, Resolution 8430 of 1993 requires ethical approval for studies involving direct human participation or biological samples, but exempts studies using anonymized, publicly available data. Law 1751 of 2015 emphasizes the protection of dignity and privacy, while Decree 1374 of 2013 regulates the use of data in scientific research, ensuring that anonymized public data usage does not infringe on patient's rights. International guidelines, such as the Declaration of Helsinki, also confirm that secondary research using anonymized data

usually does not require further ethical review.

Summarization
When needed, summarization was performed to a specific reference human genome for each data set to ensure analysis of genes using gene symbols as an identifier. For GSE147352, GSE151352, and GSE159851, we used *Genome comprehensive annotation for the Human* package available in Bioconductor.

Data preprocessing
Data quality used descriptive statistics (boxplots, correlations, density plots, and Euclidean distance dendrogram) to identify atypical or non-comparable samples. Filters for non- informative row data were conducted: zero expression, non-variability between samples (zero variation coefficient). Some datasets needed a resampling because of an excessive difference between the number of cases (GB) and controls (NB) or because of a high sample number to balance conditions and controls. Finally, to correct mean-variance dependence Variance Stabilizing Normalization (VSN) (Huber et al., 2002) was applied for data without previous normalization.

Significance Analysis of Microarrays (SAM) (Schwender, 2017) was the method used for identifying DEGs in datasets for continuous data (microarrays, FPKM transformations, etc.). Differential gene expression analysis based on the negative binomial distribution (DESeq2) (Love et al., 2014) was used for identifying DEGs for raw count data. Both methods were used with a false discovery rate (FDR) < 0.05 (Benjamini & Hochberg, 1995) as control of multiplicity effects and error rate for many hypothesis tests.

Enrichment analysis
For common DEGs in the five data sets, enrichment search was conducted using DAVID tools (https://david.ncifcrf.gov/, 2021 version). Based on the enrichment terms we grouped them into the main higher category

and searched for their relationship with processes related to glioblastoma onset, development, and metastasis that could be potentially useful for monitoring the disease. Moreover, a reduced number of genes with coherent differential expression was chosen for further analysis to identify potential biomarkers.

Survival analysis
We used the GEPIA tool (Tang et al., 2017) for the survival analysis of the main common DEGs and saved the p-value of the hazard ratio test (Spruance et al., 2004) to select the best potential biomarkers. GEPIA was selected because it allows for the analysis of disease-free survival and overall survival.

Gene expression validation using TCGA
To validate our findings and confirm the results on a larger dataset, we used UALCAN (Chandrashekar et al., 2017), a comprehensive web resource for analyzing cancer omics data. UALCAN provides easy access to publicly available cancer transcriptome data from The Cancer Genome Atlas (TCGA) project. We used this tool to examine the expression levels of our candidate biomarker genes in glioblastoma samples from the TCGA database. The web resource was consulted in the 2022 version (Chandrashekar et al., 2022).

Results

Differentially expressed genes and functional enrichment
Statistical analysis of the five GB datasets allowed us the identification of thirteen commonly deregulated genes identified as critical genes related to the presence of tumor processes (Table 2). Their importance was further investigated through annotation and enrichment, the interaction of the protein products, and survival (Table 3).

Survival analysis and Gene expression validation using TCGA
The main finding of this study, taking together all steps, is that the DEGs can be classified into three groups. These groups are related to molecular subtypes, therapy response, closed protein interactions between them, and overall patient survival. The genes belonging to these three groups have been reported earlier (Table 3). Kaplan-Meier analysis shows Disease-Free Survival (DFS) and Overall Survival (OS) for the five significant genes identified using the GEPIA tool. The analysis included genes where either or both survival metrics were statistically significant.

Notably, for COL6A1, both DFS and OS were significant, suggesting their potential as a biomarker. In other cases, DFS was significant but not OS for PDPN and NEBL, while OS was significant but not DFS for VEGFA and HOXD10 in glioblastoma.

Table 2: Log fold change values and functional classification of the thirteen key genes identified by differential expression analysis of five GB datasets.

	log2 Fold Change by data set[a]					
DEGs	GSE119834	GSE145645	GSE147352	GSE151352	GSE159851	Functional group
COL1A1	-1,18	3,79	7,18	0,44	3,09	Cellular differentiation factors, extracellular matrix related, epithelial-mesenchymal transition.
COL4A2	-0,52	2,00	3,82	0,30	3,43	Cellular differentiation factors, angiogenesis, extracellular matrix related, epithelial-mesenchymal transition.
COL6A1	-0,51	0,87	1,86	0,19	2,26	Cellular differentiation factors, epithelial-mesenchymal transition.
HOXA10	-1,84	7,20	10,21	0,79	2,99	Epithelial-mesenchymal transition.
HOXD10	-1,44	6,99	11,92	0,83	7,69	Epithelial-mesenchymal transition.
HOXA5	-11,00	5,65	9,85	0,82	4,89	Epithelial-mesenchymal transition.
PDPN	-11,00	1,57	3,91	0,34	1,51	Epithelial-mesenchymal transition.
VEGFA	0,29	1,70	3,76	0,37	3,03	Cellular differentiation factors, angiogenesis, extracellular matrix related, epithelial-mesenchymal transition.
CHI3L1	-11,00	1,44	4,97	0,28	3,40	Cellular differentiation factors, epithelial-mesenchymal transition.
HS3ST4	-6,48	-2,22	-3,48	-1,16	-2,11	Epithelial-mesenchymal transition.
GCNT4	-2,60	-1,45	-3,50	-3,06	-2,82	Epithelial-mesenchymal transition.
NEBL	-11,00	-0,33	-1,49	-0,53	-2,42	Extracellular matrix related.
ST6GALNAC1	-6,86	-1,73	-1,29	-1,36	-3,33	Epithelial-mesenchymal transition.

[a]Showing a qualitative expression by color. Red for down-regulated and green for up-regulated genes. For additional information see Supplementary Table.

Moreover, we found four DEGs that have not been reported earlier and could have a high potential for being useful biomarkers and possible druggable objectives. These DEGS were HS3ST4, GCNT4, NEBL, and ST6GALNAC1 which have been reported to be implicated in carcinogenesis or progression in other cancers. Global references related to cancer are listed in Table 3.

New potential biomarkers of therapeutic targets in GB

NEBL

NEBL is a protein-encoding gene belonging to the nebulin family. The proteins of this family play an essential role in cell adhesion and the architecture of the actin filament in the cell (Pappas et al., 2011). Upregulation of the NEBL gene has been reported in several high-grade and metastatic stage

Table 3: Thirteen key genes identified by differential expression analysis of five transcriptomic GB datasets, results of their survival analysis for each of them and summary of the reported implication of these genes in cancer.

Gene symbol	Disease free survival	Overall survival	Description
COL1A1	p(HR)=0.24	p(HR)=0.14	COL1A1, COL4A2, and COL6A1 genes are protein-coding genes that can encode the different types of collagen and belong to the collagen family (Chandrashekar et al., 2017).
COL4A2	p(HR)=0.11	p(HR)=0.35	The COL1A1 expression level in GB is much higher than in LGG (Chandrashekar et al., 2022).
COL6A1	p(HR)=0.03*	p(HR)=0.0058*	The COL4A2 gene had a significantly high-expression level in anaplastic astrocytoma and glioblastoma, and its expression level was closely related to glioma malignancy (Chandrashekar et al., 2017, 2022; Jiang et al., 2020). COL6A1 is a highly expressed tumor biomarker, including GB, with low levels in most normal tissues (S. Sun et al., 2018).
HOXA10	p(HR)=0.32	p(HR)=0.17	*Transcription Factor (TF)* A high expression of HOXA10 has been observed in head and neck squamous cell carcinoma (HNSCC) which correlates with poor prognosis (Choi et al., 2018).
HOXD10	p(HR)=0.03*	p(HR)=0.03*	HOXD10 may play different or even opposite roles at different stages of GB onset and development. For patients with GB, HOXD10 may be a valid predictor of prognosis (Feng et al., 2021; Y. Li et al., 2021; Turtoi et al., 2014), but it is an unfavorable prognostic marker in renal cancer (see proteinatlas.org).
HOXA5	p(HR)=0.49	p(HR)=0.49	HOXA5 is seen to be dysregulated in several tumor types, including cervical cancer and breast cancer, suggesting that HOXA5 may be an important tumor suppressor (Hussain et al., 2020).
PDPN	(HR)=0.0086*	p(HR)=0.35	*Angiogenesis/remodeling* It was observed that PDPN could be considered as a possible biomarker of stem cells derived from glioma, which confers resistance to ionizing radiation and would serve as a prognostic marker on patient outcomes (Sulman et al., 2008).
VEGFA	p(HR)=0.17	p(HR)=0.17	CHI3L1 and PDPN are increased in GB tumors staining for markers associated with the mesenchymal gene expression pattern (Wood et al., 2016).
CHI3L1	p(HR)=0.16	p(HR)=0.16	VEGF and PDPN have been identified as angiogenesis and/or lymphangiogenesis regulators and might be essential to restrict tumor growth, progression, and metastasis (Belfort-Mattos et al., 2016). CHI3L1 up-regulated VEGF expression in GB; they synergistically promote endothelial cell angiogenesis (Francescone et al., 2011; Shao, 2013). Combination therapies, including anti-CHI3L1 and other traditional antiangiogenic agents together with chemotherapy/radiotherapy could be an interesting approach to treat GB (Francescone et al., 2011).
HS3ST4	p(HR)=0.52	p(HR)=0.52	*New Posible Genes implicated in GB* HS3ST4 is involved in post-synthetic modification of heparan sulfate proteoglycan (HSPG); tumors of different histotypes, including breast, lung, brain, pancreas, skin, and colorectal cancer, are characterized by profound alterations in the fine structure of proteoglycans leading to uncontrolled proliferation, immune escape, metastasis and differentiation (Knelson et al., 2014; Zizza et al., 2019).
GCNT4	p(HR)=0.51	p(HR)=0.99	GCNT4 is a prognostic marker in renal cancer (see proteinatlas.org); GCNT4 mediates O-Glycan biosynthesis in mucin-type biosynthesis (*GCNT4*, 2024; *GCNT4 Gene - Glucosaminyl (N-Acetyl) Transferase 4*, 2025). Alterations of O-glycans, such as increased expression of Tn antigens, are commonly detected in cancer cells (Gao, 2013). Additionally, expression of Tn antigen has been described in human glioblastoma cell lines (Dusoswa et al., 2020) and the developing
NEBL	p(HR)=0.049*	p(HR)=0.99	
ST6GALNAC1	p(HR)=0.42	p(HR)=0.072	

mouse brain but not in healthy human brain tissues (Dusoswa et al., 2020).
Nebulette (NEBL) overexpression increases cell migration (Mamizadeh et al., 2021). It could be a favorable prognostic marker in renal cancer and an unfavorable one in urothelial cancer (see proteinatlas.org).
High expression of NEBL may be beneficial for the prognosis of glioma (Liu et al., 2021).
ST6N-acetylgalactosaminide alpha-2,6-sialyltransferase 1 (ST6GALNAC1) appears to be a favorable prognostic marker in head and neck cancer (see proteinatlas.org), overexpression of ST6GALNAC1 in gastric, breast, prostate and ovarian cancer cell lines and tissues has been directly associated with poor prognosis (Hugonnet et al., 2021), it has been observed that other members of sialyltransferases family could be related with glioma progression (Cuello et al., 2020; Iwasawa et al., 2018; Suzuki et al., 2005).

*Statistical significance with p-value < 0.05

cancers (Mamizadeh et al., 2021), where NEBL overexpression increases cell migration (Hosseini et al., 2018). Currently there is no information on a clear correlation and significance of the expression of NEBL and gliomas beyond a study that proposes that NEBL is associated with tubuline (Hosseini et al., 2018) to a greater degree than with actin in the cytoskeleton of glioma cells (Dunina-Barkovskaya, 2013).

ST6GALNAC1

The ST6GALNAC1 gene encodes a homonymous protein that is a sialyltransferase expressed in the Golgi apparatus and transfers sialic acid to the O-linked sugar chain of the spleen backbone, receptor protein and produces the sialyl-Tn antigen (STn antigen). The STn antigen is overexpressed in some adenocarcinomas, including colon, gastric, pancreatic, breast, prostate, and ovarian adenocarcinomas, but has limited or no expression in normal organs (Marcos et al., 2004). The functions of the STn antigen are thought to be related to cell-to-cell attachment and cell migration, but recent studies have suggested associations with cancer aggressiveness and poor prognosis (Ferreira et al., 2013; Ozaki et al., 2012).

Since ST6GALNAC1-positive cancers are associated with poor prognosis, targeting ST6GALNAC1 and STn antigen could be an attractive novel treatment to prevent metastasis and recurrence of adenocarcinomas, including colorectal cancer (Ogawa et al., 2017). In the case of gliomas and GB, it is unknown what their correlation and significance could be, since there are currently no studies in this regard. The confirmation of the significance of this gene was not established based on the TCGA data and therefore, more than the other genes, which significance was validated using TCGA data, it needs to be validated experimentally.

HS3ST4

Heparan sulfate (HS) is a highly sulfated glycosaminoglycan found on the cell surface and in the extracellular matrix. It is involved in cell-cell and cell-matrix communications and regulates the binding of a large number of ligands, resulting in a variety of physiological and pathological effects, such as in embryonic development, cell growth and differentiation, homeostasis, inflammatory response, tumor growth and microbial infection (Hellec et al., 2018). Altered expression of HS- modifying enzymes has been frequently observed in cancer. Consequently, dysregulation of the HS biosynthetic machinery results in dramatic changes in HS structure, affecting a variety of fundamental cellular processes involved in tumorigenesis and cancer progression, including proliferation, migration, apoptosis and immune evasion (Denys & Allain, 2019).

It has been observed that the increase in the expression of HS isoenzymes such as HS3ST4 may involve a mechanism that allows tumor cells to modulate the activation of natural killer (NK) cells and thus prevent their elimination, thus promoting evasion of the immune system by tumor cells (Denys & Allain, 2019; Hellec et al., 2018). Although HS3ST4 is abundantly expressed in the cerebral cortex and cerebellum (Denys & Allain, 2019) to

date there are no studies that correlate its expression in gliomas.

GCNT4

Members of the glucosaminyl (N-acetyl) transferase (GCNT) family are critical mediators in the synthesis, branching, and oligomerization of the mucin backbone (H. Sun et al., 2020). Altered glycosylation is a hallmark of cancer (H. Sun et al., 2020). Members of the GCNT family, including GCNT2, GCNT3, and GCNT4, have been previously identified as being associated with multiple human malignancies, influencing cancer genesis by regulating cell growth and apoptosis in pancreatic, prostate, and colon cancer (Hu et al., 2021; Zeng et al., 2022). However, the relationship between the expression levels of GCNT and GC family members has not been thoroughly investigated, including their expression and significance in gliomas.

Discussion

Cancer is a group of diseases characterized by abnormal cell growth and the potential to spread. Both environmental and genetic factors are crucial in cancer development. At the cellular level, oncogenes promote uncontrolled growth, whereas tumor suppressor genes protect against malignancy. Interestingly, some genes, such as the NOTCH receptors—key components of the evolutionarily conserved Notch signaling pathway—can function paradoxically as both oncogenes and tumor suppressors, depending on the context (Aster et al., 2017; Shen et al., 2018; Soussi & Wiman, 2015; Yip et al., 2010).

Transcriptomic studies help classify tumors into subtypes, aiding therapeutic response and clinical outcomes. The study's analysis aligns with previous classifications, including Verhaak et al.'s proneural, neural, classic, and mesenchymal GB types (Verhaak et al., 2010). It also considers potential contamination and integrates signaling process-based classifications (Kim et al., 2021; Wang et al., 2017). Combining transcriptional analysis results with these elements harmonized classifications, revealing proneural and mesenchymal characteristics, which may result from PMT, a mechanism similar to EMT in other cancers. EMT involves carcinoma cells improving their invasive capacity, losing cell polarity, and acquiring a mesenchymal phenotype, promoting metastasis (Behnan et al., 2019; Garofano et al., 2021; Kim et al., 2021; Wang et al., 2017) (Figure 1a-1b).

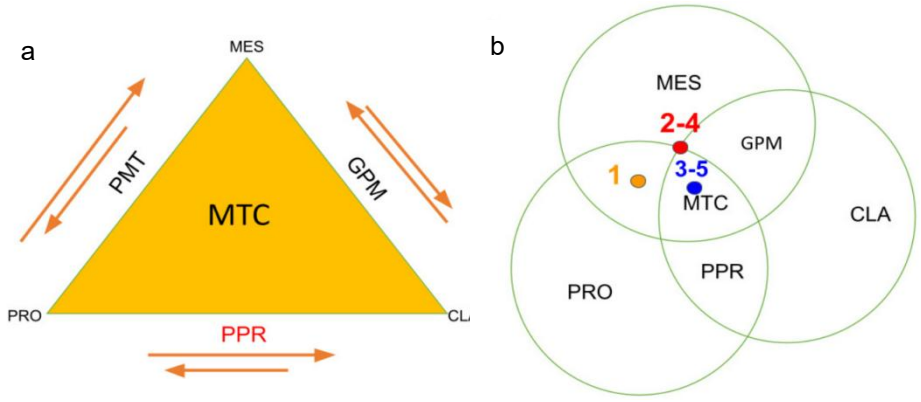

Fig 1 a-b: Proposed transcriptomic profiling and signaling pathway-based integration classification in GB. **(a)** Describes at each vertex the main transcriptomic types described to date, each edge represents subtypes based on analyzes of signaling pathways and transitions that have been observed experimentally between the PMT analogous to what happens in other types of cancer with EMT so that the tumors can manage to prosper. **(b)** Venn diagram between the transcriptomic types, subtypes based on signaling routes and clusters analyzed in this document. PRO: Proneural, MES: Mesenchymal CLA: Classical, GPM: Glycolytic/Plurimetabolic, MTC: Mitochondrial, NEU: Neuronal, PPR: PRoliferative/PRogenitor, PMT: Proneural-mesenchymal transition, 1-5 Cluster of GB patient samples analyzed.

A key unresolved question is whether PMT arises intrinsically or is induced by the tumor microenvironment (TME). Some evidence points to PMT being triggered by external factors like chemotherapy and radiotherapy (Lau et al., 2015; Lu et al., 2012; Segerman et al., 2016; Yang et al., 2021), while other studies highlight internal master regulators such as STAT3, C/EBPβ, TAZ, and NF-κB as key drivers (Yang et al., 2021; Zhang et al., 2020). Additionally, members of the HOX family of transcription factors have been implicated in modulating carcinogenesis, with context-dependent effects that can either promote or suppress tumor progression (Feng et al., 2021; Q. Li et al., 2014; Shah et al., 2012).

Within this context, several molecular markers have emerged as potential tools for classification and therapy. PDPN, a mucin-like transmembrane protein associated with lymphangiogenesis and metastasis, is widely expressed across tumors. Its interaction with receptor type C lectin-like 2 (CLEC-2) promotes platelet aggregation and metastasis, while also modulating cytoskeletal dynamics to enhance migration, invasion, and angiogenesis independently of VEGF pathways (Grau et al., 2015; Modrek et al., 2020). Given its expression patterns and the results of GEPIA analyses, PDPN represents a promising marker for GB classification.

Similarly, NEBL, which encodes a cytoskeletal matrix protein, is overexpressed in various high-grade and metastatic cancers (Cóser et al., 2010; Hosseini et al., 2018). Although it has not been previously linked to GB, GEPIA data suggest a potential association with progression-free survival, making it a candidate for further study and potential therapeutic targeting.

Another hallmark of cancer is altered glycosylation, involving the dysregulation of sialyltransferases and heparan sulfate (HS)–modifying enzymes (Vajaria & Patel, 2017). High-grade gliomas often exhibit elevated levels of terminal sialoglycans and sialyltransferases (Hugonnet et al., 2021). In this study, we observed aberrant expression of glycosylation-related genes, including ST6GALNAC1, GCNT4, and HS3ST4. While their specific roles in GB remain unclear, their dysregulation suggests involvement in key processes such as proliferation and migration (Denys & Allain, 2019). Notably, ST6GALNAC2 was enriched in E2F and MYC target pathways, and E2F7 has been shown to activate EZH2 transcription, which in turn induces mTOR signaling, a crucial pathway in glioblastoma progression (Ahmad et al., 2024).

In addition to glycosylation-related pathways, remodeling of the extracellular matrix (ECM) plays a crucial role in glioma progression. This dynamic process facilitates the activation and migration of endothelial cells, thereby promoting tumor angiogenesis. Within this context, members of the collagen gene superfamily such as COL1A1, COL4A2, and COL6A1, have been implicated in the angiogenic cascade and may serve as potential biomarkers or therapeutic targets for GB classification and treatment (Pan et al., 2020).

However, interpreting gene expression data related to ECM components and other molecular features requires consideration of the inherent biological complexity of glioblastoma. Differences in gene expression trends across datasets can be attributed to several factors, including biological variability, inclusion of different cell types, technical differences, and sample composition. GB's heterogeneity, maintained by GSCs, contributes to its resistance to treatment and presents challenges and insights for targeted therapies. Finally, but not least importantly, the observed differences in gene expression trends, particularly the significant variance between the GSE119834 dataset and the other four datasets, could be attributed to several factors. Biological variability is a primary factor, as the GSE119834 dataset includes glioblastoma stem cells (GSC) and neural stem cells (NSC) in addition to glioblastoma (GB) samples, while the other datasets focus primarily on GB and normal brain tissues. This inclusion of different cell types might influence gene expression profiles and lead to distinct trends. Furthermore, the multiform and progressive nature of GB may lead to the evolution of molecular profiles in other datasets towards those

observed in the GSE119834 dataset. Glioblastoma tumors exhibit significant heterogeneity at both the single-cell and spatial levels, driven by distinct populations of cells with specific transcriptional signatures and the unique microenvironments created by hypoxia gradients. This heterogeneity, as noted by (Ou et al., 2020), is maintained by GSCs with concordant genomic mutations, contributing to the tumor's chemo- and radioresistance. Given the dynamic nature of GB, the molecular profiles observed in the other datasets may transition or evolve towards those seen in the GSE119834 dataset, which captures the complexity of GB more comprehensively. This potential for molecular evolution highlights the importance of considering PMT and other transition mechanisms in the study of GB.

The integration of data from UALCAN, TCGA, and GEPIA in our study has provided a comprehensive view of gene expression in glioblastoma. Each dataset has its strengths: UALCAN offers an accessible interface for survival analysis and gene expression validation, TCGA provides a rich, detailed dataset that allows for in-depth genomic and transcriptomic analysis, and GEPIA bridges the gap between TCGA cancer data and GTEx normal tissue data, offering a broader context for gene expression analysis. Despite these strengths, the discrepancies observed in the gene expression trends across these datasets can be attributed to several factors:

- Sample Composition and Heterogeneity: The composition of samples in each dataset can differ significantly. For instance, TCGA and UALCAN primarily focus on tumor samples, while GEPIA includes normal tissue data from GTEx, which may affect comparative analyses.
- Technical Variations: Differences in sequencing technologies, data processing pipelines, and normalization methods across these platforms can lead to variability in gene expression measurements.
- Biological Variability: Intrinsic biological differences, such as tumor heterogeneity, different subtypes, and the tumor microenvironment, can contribute to inconsistent gene expression patterns observed in different datasets.

Our study demonstrates that integrating multiple datasets can enhance the reliability of biomarker identification and provide a more nuanced view of the molecular underpinnings of glioblastoma.

Conclusions

Glioblastoma is one of the most aggressive and lethal human brain tumors. The high invasiveness, the propensity to disperse throughout the brain parenchyma, and the elevated vascularity and necrosis make these tumors extremely recidivist, resulting in a short patient median survival even after surgical resection and chemoradiotherapy (D'Alessio et al., 2019). There is a growing body of evidence demonstrating the existence of PMT, which suggests that first-line therapy for a primary disease may not work effectively for recurrent tumors due to this situation (Kim et al., 2021). In order to contribute to the understanding, diagnosis and treatment of glioblastoma, transcriptomic studies for tumor classification are crucial for classifying tumors into subtypes, aiding in therapeutic response and clinical outcomes. The classifications align with prior studies and integrate signaling processes. Proneural-Mesenchymal Transition is analogous to EMT in other cancers and is significant in GB. PMT may be intrinsic or induced by external factors like chemotherapy, radiotherapy, or master regulators such as STAT3 and NF-KB.

Regarding the potential biomarkers we identified through the integrative transcriptomic analysis and further confirmation, the main findings are that PDPN is associated with lymphangiogenesis and metastasis, making it an attractive marker for GB classification. Collagen Superfamily genes like COL6A1 are significant in tumor angiogenesis and prognosis. NEBL, a cytoskeletal matrix protein, is also proposed as a potential target for GB therapy. The altered Glycosylation represented by sialyltransferases and HS modifying enzymes show altered expressions, indicating a role in cancer progression that could be used for therapeutic decisions.

Moreover, it is important to mention that differences in gene expression trends across datasets are due to biological variability, inclusion of different cell types, technical differences, and sample composition. GB's heterogeneity contributes to its treatment resistance and presents challenges and insights for targeted therapies and findings need to be validated in different populations and ancestries. These findings underscore the importance of considering various genetic, molecular, and environmental factors in developing effective cancer treatments, particularly for GB. It seems clear that to combat GB there must be a change from classic cytotoxic chemoradiotherapy to more targeted therapies, considering the high degree of heterogeneity and molecular complexity of the tumor. GB subtyping is a key process to impact a patient's survival by providing targeted therapies.

Authorship contribution statement

Andres David Turizo-Smith: Conceptualization, Validation, Investigation, Writing - original draft. **Daniel Garcia-Niño:** Software, Formal analysis, Resources, Data curation, Visualization. **Liliana Lopez-Kleine:** Methodology, Writing - review & editing, Supervision, Project administration. All authors approved this study.

Disclosure of interest

The authors report there are no competing interests to declare.

Ethical approval

It does not apply as the data were obtained from publicly available databases.

Funding

No funding was received for conducting this study.

Data availability

All data sets are available in The Genome Expression Omnibus database from NCBI by GEO accession number. Statistical code and analysis are available from D.E.G. upon reasonable request via email.

References

Ahmad, M. S., Braoudaki, M., & Siddiqui, S. S. (2024). Differential expression of ST6GALNAC1 and ST6GALNAC2 and their clinical relevance to colorectal cancer progression. *PLOS ONE*, *19*(9), e0311212. https://doi.org/10.1371/journal.pone.0311212

Aster, J. C., Pear, W. S., & Blacklow, S. C. (2017). The Varied Roles of Notch in Cancer. *Annual Review of Pathology: Mechanisms of Disease*, *12*(Volume 12, 2017), 245–275. https://doi.org/10.1146/annurev-pathol-052016-100127

Behnan, J., Finocchiaro, G., & Hanna, G. (2019). The landscape of the mesenchymal signature in brain tumours. *Brain*, *142*(4), 847–866. https://doi.org/10.1093/brain/awz044

Belfort-Mattos, P. N., Focchi, G. R. D. A., Ribalta, J. C. L., Megale De Lima, T., Nogueira Carvalho, C. R., Kesselring Tso, F., & De Góis Speck, N. M. (2016). Immunohistochemical Expression of VEGF and Podoplanin in Uterine Cervical Squamous Intraepithelial Lesions. *Disease Markers*, *2016*, 1–8. https://doi.org/10.1155/2016/8293196

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Carlson, M. (2017). *Org.Hs.eg.db* [Computer software]. Bioconductor. https://doi.org/10.18129/B9.BIOC.ORG.HS.EG.DB

Chandrashekar, D. S., Bashel, B., Balasubramanya, S. A. H., Creighton, C. J., Ponce-Rodriguez, I., Chakravarthi, B. V. S. K., & Varambally, S. (2017). UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia*, *19*(8), 649–658. https://doi.org/10.1016/j.neo.2017.05.002

Chandrashekar, D. S., Karthikeyan, S. K., Korla, P. K., Patel, H., Shovon, A. R., Athar, M., Netto, G. J., Qin, Z. S., Kumar, S., Manne, U., Creighton, C. J., & Varambally, S. (2022). UALCAN: An update to the integrated cancer data analysis platform. *Neoplasia*, *25*, 18–27. https://doi.org/10.1016/j.neo.2022.01.001

Choi, S. W., Shin, H., Sa, J. K., Cho, H. J., Koo, H., Kong, D., Seol, H. J., & Nam, D. (2018). Identification of transcriptome signature for predicting clinical response to bevacizumab in recurrent glioblastoma. *Cancer Medicine*, *7*(5), 1774–1783. https://doi.org/10.1002/cam4.1439

Cóser, V. M., Meyer, C., Basegio, R., Menezes, J., Marschalek, R., & Pombo-de-Oliveira, M. S. (2010). Nebulette is the second member of the nebulin family fused to the MLL gene in infant leukemia. *Cancer Genetics and Cytogenetics*, *198*(2), 151–154. https://doi.org/10.1016/j.cancergencyto.2009.12.013

Cuello, H. A., Ferreira, G. M., Gulino, C. A., Toledo, A. G., Segatori, V. I., & Gabri, M. R. (2020). Terminally sialylated and fucosylated complex N-glycans are involved in the malignant behavior of high-grade glioma. *Oncotarget*, *11*(52), 4822–4835. https://doi.org/10.18632/oncotarget.27850

D'Alessio, A., Proietti, G., Sica, G., & Scicchitano, B. M. (2019). Pathological and Molecular Features of Glioblastoma and Its Peritumoral Tissue. *Cancers*, *11*(4), 469. https://doi.org/10.3390/cancers11040469

Denys, A., & Allain, F. (2019). The Emerging Roles of Heparan Sulfate 3-O-Sulfotransferases in Cancer. *Frontiers in Oncology*, *9*, 507. https://doi.org/10.3389/fonc.2019.00507

Dunina-Barkovskaya, L. V. (2013). *Effect of desmin peptides and the characterization of nebulin family proteins in glial cells*. https://hdl.handle.net/1969.1/157610

Dusoswa, S. A., Verhoeff, J., Abels, E., Méndez-Huergo, S. P., Croci, D. O., Kuijper, L. H., De Miguel, E., Wouters, V. M. C. J., Best, M. G., Rodriguez, E., Cornelissen, L. A. M., Van Vliet, S. J., Wesseling, P., Breakefield, X. O., Noske, D. P., Würdinger, T., Broekman, M. L. D., Rabinovich, G. A., Van Kooyk, Y., & Garcia-Vallejo, J. J. (2020). Glioblastomas exploit truncated O linked glycans for local and distant immune modulation via the macrophage galactose-type lectin. *Proceedings of the National Academy of Sciences*, *117*(7), 3693–3703.https://doi.org/10.1073/pnas.1907921117

Feng, Y., Zhang, T., Wang, Y., Xie, M., Ji, X., Luo, X., Huang, W., & Xia, L. (2021). Homeobox Genes in Cancers: From Carcinogenesis to Recent Therapeutic Intervention. *Frontiers in Oncology*, *11*, 770428. https://doi.org/10.3389/fonc.2021.770428

Ferreira, J. A., Videira, P. A., Lima, L., Pereira, S., Silva, M., Carrascal, M., Severino, P. F., Fernandes, E., Almeida, A., Costa, C., Vitorino, R., Amaro, T., Oliveira, M. J., Reis, C. A., Dall'Olio, F., Amado, F., & Santos, L. L. (2013). Overexpression of tumour-ssociated carbohydrate antigen sialyl-Tn in advanced bladder tumours. *Molecular Oncology*, *7*(3), 719–731. https://doi.org/10.1016/j.molonc.2013.03.001

Francescone, R. A., Scully, S., Faibish, M., Taylor, S. L., Oh, D., Moral, L., Yan, W., Bentley, B., & Shao, R. (2011). Role of YKL-40 in the Angiogenesis, Radioresistance, and Progression of Glioblastoma. *Journal of Biological Chemistry*, *286*(17), 15332–15343. https://doi.org/10.1074/jbc.M110.212514

Gallego, O. (2015). Nonsurgical Treatment of Recurrent Glioblastoma. *Current Oncology*, *22*(4), 273–281. https://doi.org/10.3747/co.22.2436

Gao, Y. (2013). *Glycosylation of Prostate Cancer Cell Lines and Glycosyltransferase Characterization and Inhibition*. https://qspace.library.queensu.ca/items/0caefc 81-f156-4a58-a8a4-7acfbe9a4c03

Garofano, L., Migliozzi, S., Oh, Y. T., D'Angelo, F., Najac, R. D., Ko, A., Frangaj, B., Caruso, F. P., Yu, K., Yuan, J., Zhao, W., Di Stefano, A. L., Bielle, F., Jiang, T., Sims, P., Suvà, M. L., Tang, F., Su, X.-D., Ceccarelli, M., … Iavarone, A. (2021). Pathway-based classification of glioblastoma uncovers a mitochondrial

subtype with therapeutic vulnerabilities. *Nature Cancer*, *2*(2), 141–156. https://doi.org/10.1038/s43018-020-00159-4

GCNT4. (2024). The Human Protein Atlas. https://www.proteinatlas.org/ENSG00000017692 8-GCNT4

GCNT4 Gene—Glucosaminyl (N-Acetyl) Transferase 4. (2025). The Human Gene Database. https://www.genecards.org/cgi-bin/carddisp.pl?gene=GCNT4

Grau, S. J., Trillsch, F., Tonn, J.-C., Goldbrunner, R. H., Noessner, E., Nelson, P. J., & von Luettichau, I. (2015). Podoplanin increases migration and angiogenesis in malignant glioma. *International Journal of Clinical and Experimental Pathology*, *8*(7), 8663–8670.

Hellec, C., Delos, M., Carpentier, M., Denys, A., & Allain, F. (2018). The heparan sulfate 3-O-sulfotransferases (HS3ST) 2, 3B and 4 enhance proliferation and survival in breast cancer MDA-MB-231 cells. *PLOS ONE*, *13*(3), e0194676. https://doi.org/10.1371/journal.pone.0194676

Hosseini, S. M., Mahjoubi, F., Majidzadeh, T., Khaje-Hosseini, F., & Haghipanah, M. (2018). Nebulette Expression Is Associated with Lymph Node Metastasis in Patients with Colorectal Cancer. *Middle East Journal of Digestive Diseases*, *10*(3), 174–179. https://doi.org/10.15171/mejdd.2018.107

Hu, W., Zheng, X., Liu, J., Zhang, M., Liang, Y., & Song, M. (2021). MicroRNA MiR-130a-3p promotes gastric cancer by targeting Glucosaminyl N-acetyl transferase 4 (GCNT4) to regulate the TGF-β1/SMAD3 pathway. *Bioengineered*, *12*(2), 11634–11647. https://doi.org/10.1080/21655979.2021.1995099

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, *4*(1), 44–57. https://doi.org/10.1038/nprot.2008.211

Huang, T., Yang, Y., Song, X., Wan, X., Wu, B., Sastry, N., Horbinski, C. M., Zeng, C., Tiek, D., Goenka, A., Liu, F., Brennan, C. W., Kessler, J. A., Stupp, R., Nakano, I., Sulman, E. P., Nishikawa, R., James, C. D., Zhang, W., … Cheng, S.-Y. (2021). PRMT6 methylation of RCC1 regulates mitosis, tumorigenicity, and radiation response of glioblastoma stem cells. *Molecular Cell*, *81*(6), 1276-1291.e9. https://doi.org/10.1016/j.molcel.2021.01.015

Huber, W., Von Heydebreck, A., Sültmann, H., Poustka, A., & Vingron, M. (2002). Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics*, *18*(suppl_1), S96–S104. https://doi.org/10.1093/bioinformatics/18.suppl_1.S96

Hugonnet, M., Singh, P., Haas, Q., & Von Gunten, S. (2021). The Distinct Roles of Sialyltransferases in Cancer Biology and Onco-Immunology. *Frontiers in Immunology*, *12*, 799861. https://doi.org/10.3389/fimmu.2021.799861

Hussain, I., Deb, P., Chini, A., Obaid, M., Bhan, A., Ansari, K. I., Mishra, B. P., Bobzean, S. A., Udden, S. M. N., Alluri, P. G., Das, H. K., Brothers, R. M., Perrotti, L. I., & Mandal, S. S. (2020). HOXA5 Expression Is Elevated in Breast Cancer and Is Transcriptionally Regulated by Estradiol. *Frontiers in Genetics*, *11*, 592436. https://doi.org/10.3389/fgene.2020.592436

Iwasawa, T., Zhang, P., Ohkawa, Y., Momota, H., Wakabayashi, T., Ohmi, Y., Bhuiyan, R., Furukawa, K., & Furukawa, K. (2018). Enhancement of malignant properties of human glioma cells by ganglioside GD3/GD2. *International Journal of Oncology*. https://doi.org/10.3892/ijo.2018.4266

Jiang, Y., He, J., Guo, Y., Tao, H., Pu, F., &

Li, Y. (2020). Identification of genes related to low-grade glioma progression and prognosis based on integrated transcriptome analysis. *Journal of Cellular Biochemistry*, *121*(5–6), 3099–3111. https://doi.org/10.1002/jcb.29577

Kim, Y., Varn, F. S., Park, S.-H., Yoon, B. W., Park, H. R., Lee, C., Verhaak, R. G. W., & Paek, S. H. (2021). Perspective of mesenchymal transformation in glioblastoma. *Acta Neuropathologica Communications*, *9*(1), 50. https://doi.org/10.1186/s40478-021-01151-4

Knelson, E. H., Nee, J. C., & Blobe, G. C. (2014). Heparan sulfate signaling in cancer. *Trends in Biochemical Sciences*, *39*(6), 277–288. https://doi.org/10.1016/j.tibs.2014.03.001

Lau, J., Ilkhanizadeh, S., Wang, S., Miroshnikova, Y. A., Salvatierra, N. A., Wong, R. A., Schmidt, C., Weaver, V. M., Weiss, W. A., & Persson, A. I. (2015). STAT3 Blockade Inhibits Radiation-Induced Malignant Progression in Glioma. *Cancer Research*, *75*(20), 4302–4311. https://doi.org/10.1158/0008-5472.CAN-14-3331

Li, Q., Ding, C., Chen, C., Zhang, Z., Xiao, H., Xie, F., Lei, L., Chen, Y., Mao, B., Jiang, M., Li, J., Wang, D., & Wang, G. (2014). miR-224 promotion of cell migration and invasion by targeting Homeobox D 10 gene in human hepatocellular carcinoma. *Journal of Gastroenterology and Hepatology*, *29*(4), 835–842. https://doi.org/10.1111/jgh.12429

Li, Y., Ma, K., Xie, Q., Zhang, X., Zhang, X., Chen, K., Kong, L., & Qian, R. (2021). Identification of HOXD10 as a Marker of Poor Prognosis in Glioblastoma Multiforme. *OncoTargets and Therapy*, *Volume 14*, 5183–5195. https://doi.org/10.2147/OTT.S336225

Liu, Z., Lian, X., Zhang, X., Zhu, Y., Zhang, W., Wang, J., Wang, H., Liu, B., Ren, Z., Zhang, M., Liu, M., & Gao, Y. (2021). ESPL1 Is a Novel Prognostic Biomarker Associated With the Malignant Features of Glioma. *Frontiers in Genetics*, *12*, 666106. https://doi.org/10.3389/fgene.2021.666106

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Lu, K. V., Chang, J. P., Parachoniak, C. A., Pandika, M. M., Aghi, M. K., Meyronet, D., Isachenko, N., Fouse, S. D., Phillips, J. J., Cheresh, D. A., Park, M., & Bergers, G. (2012). VEGF Inhibits Tumor Cell Invasion and Mesenchymal Transition through a MET/VEGFR2 Complex. *Cancer Cell*, *22*(1), 21–35. https://doi.org/10.1016/j.ccr.2012.05.037

Mack, S. C., Singh, I., Wang, X., Hirsch, R., Wu, Q., Villagomez, R., Bernatchez, J. A., Zhu, Z., Gimple, R. C., Kim, L. J. Y., Morton, A., Lai, S., Qiu, Z., Prager, B. C., Bertrand, K. C., Mah, C., Zhou, W., Lee, C., Barnett, G. H., … Rich, J. N. (2019). Chromatin landscapes reveal developmentally encoded transcriptional states that define human glioblastoma. *Journal of Experimental Medicine*, *216*(5), 1071–1090. https://doi.org/10.1084/jem.20190196

Mamizadeh, Z., Kalani, M. R., Parsania, M., Soltan Dallal, M. M., & Moradi, A. (2021). NEBL and AKT1 maybe new targets to eliminate the colorectal cancer cells resistance to oncolytic effect of vesicular stomatitis virus M-protein. *Molecular Therapy - Oncolytics*, *23*, 593–601. https://doi.org/10.1016/j.omto.2021.11.013

Marcos, N. T., Pinho, S., Grandela, C., Cruz, A., Samyn-Petit, B., Harduin-Lepers, A., Almeida, R., Silva, F., Morais, V., Costa, J., Kihlberg, J., Clausen, H., & Reis, C. A. (2004). Role of the Human ST6GalNAc-I and ST6GalNAc-II in the Synthesis of the Cancer-Associated Sialyl-Tn Antigen. *Cancer Research*, *64*(19), 7050–7057.

https://doi.org/10.1158/0008-5472.CAN-04-1921

Modrek, A. S., Eskilsson, E., Ezhilarasan, R., Wang, Q., Goodman, L. D., Bhat, K., Le, T., Barthel, F. P., Tang, M., Yang, J., Long, L., Gumin, J., Lang, F., Verhaak, R., Aldape, K., & Sulman, E. P. (2020). PDPN+ Tumor Initiating, Treatment Resistant Glioblastoma Cells Promote Radiation Resistance Via PRC2. *International Journal of Radiation Oncology, Biology, Physics*, *108*(3), e510. https://doi.org/10.1016/j.ijrobp.2020.07.1606

Ogawa, T., Hirohashi, Y., Murai, A., ishidate, T., Okita, K., Wang, L., Ikehara, Y., atoyoshi, T., Usui, A., Kubo, T., Nakastugawa, M., Kanaseki, T., Tsukahara, T., Kutomi, G., Furuhata, T., Hirata, K., Sato, N., Mizuguchi, T., Takemasa, I., & Torigoe, T. (2017). ST6GALNAC1 plays important roles in enhancing cancer stem phenotypes of colorectal cancer via the Akt pathway. *Oncotarget*, *8*(68), 112550–112564. https://doi.org/10.18632/oncotarget.22545

Ou, A., Yung, W. K. A., & Majd, N. (2020). Molecular Mechanisms of Treatment Resistance in Glioblastoma. *International Journal of Molecular Sciences*, *22*(1), 351. https://doi.org/10.3390/ijms22010351

Ozaki, H., Matsuzaki, H., Ando, H., Kaji, H., Nakanishi, H., Ikehara, Y., & Narimatsu, H. (2012). Enhancement of metastatic ability by ectopic expression of ST6GalNAcI on a gastric cancer cell line in a mouse model. *Clinical & Experimental Metastasis*, *29*(3), 229–238. https://doi.org/10.1007/s10585-011-9445-1

Pan, Y., Wang, S., Yang, B., Jiang, Z., Lenahan, C., Wang, J., Zhang, J., & Shao, A. (2020). Transcriptome analyses reveal molecular mechanisms underlying phenotypic differences among transcriptional subtypes of glioblastoma. *Journal of Cellular and Molecular Medicine*, *24*(7), 3901–3916. https://doi.org/10.1111/jcmm.14976

Pappas, C. T., Bliss, K. T., Zieseniss, A., & Gregorio, C. C. (2011). The Nebulin family: An actin support group. *Trends in Cell Biology*, *21*(1), 29–37. https://doi.org/10.1016/j.tcb.2010.09.005

Schaffenrath, J., Wyss, T., He, L., Rushing, E. J., Delorenzi, M., Vasella, F., Regli, L., Neidert, M. C., & Keller, A. (2021). Blood-brain barrier alterations in human brain tumors revealed by genome-wide transcriptomic profiling. *Neuro-Oncology*, *23*(12), 2095–2106. https://doi.org/10.1093/neuonc/noab022

Schwender, H. (2017). *Siggenes* [Computer software]. Bioconductor. https://doi.org/10.18129/B9.BIOC.SIGGENES

Segerman, A., Niklasson, M., Haglund, C., Bergström, T., Jarvius, M., Xie, Y., Westermark, A., Sönmez, D., Hermansson, A., Kastemar, M., Naimaie-Ali, Z., Nyberg, F., Berglund, M., Sundström, M., Hesselager, G., Uhrbom, L., Gustafsson, M., Larsson, R., Fryknäs, M., … Westermark, B. (2016). Clonal Variation in Drug and Radiation Response among Glioma-Initiating Cells Is Linked to Proneural-Mesenchymal Transition. *Cell Reports*, *17*(11), 2994–3009. https://doi.org/10.1016/j.celrep.2016.11.056

Shah, C. A., Bei, L., Wang, H., Platanias, L. C., & Eklund, E. A. (2012). HoxA10 Protein Regulates Transcription of Gene Encoding Fibroblast Growth Factor 2 (FGF2) in Myeloid Cells. *Journal of Biological Chemistry*, *287*(22), 18230–18248. https://doi.org/10.1074/jbc.M111.328401

Shao, R. (2013). YKL-40 acts as an angiogenic factor to promote tumor angiogenesis. *Frontiers in Physiology*, *4*. https://doi.org/10.3389/fphys.2013.00122

Shen, L., Shi, Q., & Wang, W. (2018). Double

agents: Genes with both oncogenic and tumor-suppressor functions. *Oncogenesis*, *7*(3), 25. https://doi.org/10.1038/s41389-018-0034-x

Sherman, B. T., Hao, M., Qiu, J., Jiao, X., Baseler, M. W., Lane, H. C., Imamichi, T., & Chang, W. (2022). DAVID: A web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Research*, *50*(W1), W216–W221. https://doi.org/10.1093/nar/gkac194

Soussi, T., & Wiman, K. G. (2015). TP53: An oncogene in disguise. *Cell Death & Differentiation*, *22*(8), 1239–1249. https://doi.org/10.1038/cdd.2015.53

Spruance, S. L., Reid, J. E., Grace, M., & Samore, M. (2004). Hazard Ratio in Clinical Trials. *Antimicrobial Agents and Chemotherapy*, *48*(8), 2787–2792. https://doi.org/10.1128/AAC.48.8.2787-2792.2004

Sulman, E., Goodman, L., Le, T., Pelloski, C., Bhat, K., Gumin, J., Lang, F., Colman, H., & Aldape, K. (2008). A novel marker of glioma stem cells that is prognostic for treatment esponse and patient outcome. *Cancer Research*, *68*(9_Supplement), 3777.

Sun, H., Chang, J., Ye, M., Weng, W., Zhang, M., Ni, S., Tan, C., Huang, D., Wang, L., Du, X., Xu, M., & Sheng, W. (2020). GCNT4 is Associated with Prognosis and Suppress Cell Proliferation in Gastric Cancer. *OncoTargets and Therapy*, *Volume 13*, 8601–8613. https://doi.org/10.2147/OTT.S248997

Sun, S., Wang, Y., Wu, Y., Gao, Y., Li, Q., Abdulrahman, A., Liu, X., Ji, G., Gao, J., Li, L., Wan, F., Li, Y., & Gao, D. (2018). Identification of COL1A1 as an invasion-related gene in malignant astrocytoma. *International Journal of Oncology*. https://doi.org/10.3892/ijo.2018.4568

Suzuki, M., Suzuki, M., Nakayama, J., Suzuki, A., Angata, K., Chen, S., Sakai, K., Hagihara, K., Yamaguchi, Y., & Fukuda, M. (2005). Polysialic acid facilitates tumor invasion by glioma cells. *Glycobiology*, *15*(9), 887–894. https://doi.org/10.1093/glycob/cwi071

Tang, Z., Li, C., Kang, B., Gao, G., Li, C., & Zhang, Z. (2017). GEPIA: A web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Research*, *45*(W1), W98–W102. https://doi.org/10.1093/nar/gkx247

Turtoi, A., Blomme, A., Bianchi, E., Maris, P., Vannozzi, R., Naccarato, A. G., Delvenne, P., De Pauw, E., Bevilacqua, G., & Castronovo, V. (2014). Accessibilome of Human Glioblastoma: Collagen-VI-alpha-1 Is a New Target and a Marker of Poor Outcome. *Journal of Proteome Research*, *13*(12), 5660–5669. https://doi.org/10.1021/pr500657w

Vajaria, B. N., & Patel, P. S. (2017). Glycosylation: A hallmark of cancer? *Glycoconjugate Journal*, *34*(2), 147–156. https://doi.org/10.1007/s10719-016-9755-2

Verhaak, R. G. W., Hoadley, K. A., Purdom, E., Wang, V., Qi, Y., Wilkerson, M. D., Miller, C. R., Ding, L., Golub, T., Mesirov, J. P., Alexe, G., Lawrence, M., O'Kelly, M., Tamayo, P., Weir, B. A., Gabriel, S., Winckler, W., Gupta, S., Jakkula, L., … Hayes, D. N. (2010). Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell*, *17*(1), 98–110. https://doi.org/10.1016/j.ccr.2009.12.020

Wang, Q., Hu, B., Hu, X., Kim, H., Squatrito, M., Scarpace, L., deCarvalho, A. C., Lyu, S., Li, P., Li, Y., Barthel, F., Cho, H. J., Lin, Y.-H., Satani, N., Martinez-Ledesma, E., Zheng, S., Chang,

E., Sauvé, C.-E. G., Olar, A., … Verhaak, R. G. W. (2017). Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. *Cancer Cell*, *32*(1), 42-56.e6. https://doi.org/10.1016/j.ccell.2017.06.003

Wood, M. D., Reis, G. F., Reuss, D. E., & Phillips, J. J. (2016). Protein Analysis of Glioblastoma Primary and Posttreatment Pairs Suggests a Mesenchymal Shift at Recurrence. *Journal of Neuropathology & Experimental Neurology*, *75*(10), 925–935. https://doi.org/10.1093/jnen/nlw068

Xu, L., Chen, Y., Huang, Y., Sandanaraj, E., Yu, J. S., Lin, R. Y.-T., Dakle, P., Ke, X.-Y., Chong, Y. K., Koh, L., Mayakonda, A., Nacro, K., Hill, J., Huang, M.-L., Gery, S., Lim, S. W., Huang, Z., Xu, Y., Chen, J., … Koeffler, H. P. (2021). Topography of transcriptionally active chromatin in glioblastoma. *Science Advances*, *7*(18), eabd4676. https://doi.org/10.1126/sciadv.abd4676

Yang, M., Chen, X., Zhang, J., Xiong, E., Wang, Q., Fang, W., Li, L., Fei, F., & Gong, A. (2021). ME2 Promotes Proneural–Mesenchymal Transition and Lipogenesis in Glioblastoma. *Frontiers in Oncology*, *11*, 715593. https://doi.org/10.3389/fonc.2021.715593

Yip, S.-C., Saha, S., & Chernoff, J. (2010). PTP1B: A double agent in metabolism and oncogenesis. *Trends in Biochemical Sciences*, *35*(8), 442–449. https://doi.org/10.1016/j.tibs.2010.03.004

Zeng, L., Liao, Q., Zeng, X., Ye, J., Yang, X., Zhu, S., Tang, H., Liu, G., Cui, W., Ma, S., & Cui, S. (2022). Noncoding RNAs and hyperthermic intraperitoneal chemotherapy in advanced gastric cancer. *Bioengineered*, *13*(2), 2623–2638. https://doi.org/10.1080/21655979.2021.2021348

Zhang, G., Tanaka, S., Jiapaer, S., Sabit, H., Tamai, S., Kinoshita, M., & Nakada, M. (2020). RBPJ contributes to the malignancy of glioblastoma and induction of proneural-mesenchymal transition via IL-6-STAT3 pathway. *Cancer Science*, *111*(11), 4166–4176. https://doi.org/10.1111/cas.14642

Zizza, P., Dinami, R., Porru, M., Cingolani, C., Salvati, E., Rizzo, A., D'Angelo, C., Petti, E., Amoreo, C. A., Mottolese, M., Sperduti, I., Chambery, A., Russo, R., Ostano, P., Chiorino, G., Blandino, G., Sacconi, A., Cherfils-Vicini, J., eonetti, C., … Biroccio, A. (2019). TRF2 positively regulates SULF2 expression increasing VEGF-A release and activity in tumor microenvironment. *Nucleic Acids Research*, *47*(7), 3365–3382. https://doi.org/10.1093/nar/gkz041