Glioblastoma TCGA mesenchymal and IGS 23 tumors are identifiable by immunohistochemistry and have an immune-phenotype indicating a potential benefit from immunotherapy

Cristina Carrato¹, Francesc Alameda², Anna Esteve-Codina³, Estela Pineda⁴, Oriol Arpí⁵, Maria Martinez-García⁶, Mar Mallo⁷, Marta Gut⁸, Raquel Lopez-Martos¹, Sonia Del Barco⁹, Teresa Ribalta¹⁰, Jaume Capellades¹¹, Josep Puig¹², Oscar Gallego¹³, Carlos Mesia¹⁴, Ana M^a Muñoz-Marmol¹, Ivan Archilla¹⁰, Montserrat Arumí², Julie Marie Blanc¹, Beatriz Bellosillo², Silvia Menendez⁵, Anna Esteve¹⁵, Silvia Bague¹⁶, Ainhoa Hernandez¹⁵, Jordi Craven-Bartle¹⁷, Rafael Fuentes¹⁸, Noemí Vidal¹⁹, Iban Aldecoa^{10,20}, Nuria de la Iglesia²¹, Carmen Balana¹⁵

¹Pathology Department, Hospital Universitari Germans Trias i Pujol, Badalona, Spain.

²Pathology Department, Neuropathology Unit, Hospital del Mar, Institut Hospital del Mar d'Investigacions Mèdiques (IMIM), Barcelona, Spain.

³CNAG-CRG, Centre for Genomic Regulation, Barcelona Institute of Science and Technology, 08028 Barcelona, Spain.

⁴Medical Oncology, Hospital Clínic, Translational Genomics and Targeted Therapeutics in Solid Tumors, August Pi i Sunyer Biomedical Research Institute (IDIBAPS), Barcelona, Spain.

⁵Cancer Research Program, Institut Hospital del Mar d'Investigacions Mèdiques (IMIM), Barcelona, Spain.

⁶Medical Oncology, Hospital del Mar, Barcelona, Spain.

⁷Institut de Recerca Contra la Leucèmia Josep Carreras, Badalona, Spain.

⁸CNAG-CRG, Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Universitat Pompeu Fabra (UPF), Barcelona, Spain.

⁹Medical Oncology, Institut Català d'Oncologia (ICO) Girona, Hospital Josep Trueta, Girona, Spain.

¹⁰Pathology Department (Neuropathology), Hospital Clínic, Barcelona, Spain.

¹¹Radiology Department, Hospital del Mar, Barcelona, Spain.

¹²Radiology Department, Institut de Diagnòstic per la Imatge, Hospital Josep Trueta, Girona, Spain.

¹³Medical Oncology, Hospital de Sant Pau, Barcelona, Spain.

¹⁴Neuro-Oncology Unit & Medical Oncology Department, Institut Català d'Oncologia (ICO), Institut de Investigació Bellvitge (IDIBELL), L'Hospitalet, Barcelona, Spain.

¹⁵ Institut Català d'Oncologia (ICO) Badalona, Badalona Applied Research Group in Oncology (B-ARGO Group), Institut Investigació Germans Trias i Pujol (IGTP), Badalona, Spain.

¹⁶Pathology Department, Hospital de Sant Pau, Barcelona, Spain.

¹⁷Radiation Therapy Department, Hospital de Sant Pau, Barcelona, Spain.

¹⁸Radiation Therapy Department, Institut Català d'Oncologia (ICO), Girona, Spain.

¹⁹Pathology Department, Hospital de Bellvitge. Bellvitge, Spain.

²⁰Neurological Tissue Bank, Biobanc-Hospital Clínic-IDIBAPS, Barcelona, Spain.

²¹Glioma and Neural Stem Cell Group, Translational Genomics and Targeted Therapeutics in Solid Tumors Team, August Pi i Sunyer Biomedical Research Institute (IDIBAPS), Barcelona, Spain.

Running title: IHC & immune-phenotyping in GBM molecular subtypes

Correspondence to:

Carmen Balaña Quintero
Research Coordinator
President of the Internal Scientific Committee of the IGTP
Neuro-oncology and Sarcomas
Institut Català d'Oncologia
Ctra Canyet s/n
08916 Badalona
Spain
Tel: +34 93 497 89 25
Fax: +34 93 497 89 50
Email: cbalana@iconcologia.net
ORCID: 0000-0003-0771-0390

Conflict of interest. The authors report no potential conflict of interest.

Data availability

The datasets generated during the current study are available from the corresponding author on reasonable request. Molecular data underlying the findings described in the manuscript are fully available without restriction from the Bioproject Sequence Read Archive

<http://www.ncbi.nlm.nih.gov/bioproject/PRJNA613395>

Translational Relevance

FFPE samples from newly diagnosed, uniformly treated glioblastoma patients were submitted to RNA-Sequencing and tissue samples were included in TMAs, which allowed us to select immunohistochemistry antibodies based on the results of a differential gene expression analysis. As a result, we found that two previously unreported antibodies (SHC1 and TCIRG1) were able to identify the TCGA mesenchymal subtype, which overlapped molecularly and immunohistochemically with the IGS cluster 23. Moreover, mesenchymal tumors show an immune-phenotypical pattern that has been related to immunotherapy response in melanoma. These findings, together with those of our cell-type deconvolution analysis, suggest that these tumors may well be candidates for immunotherapy, especially with anti-M2 and/or dendritic cell vaccination.

ABSTRACT

Purpose: Molecular subtype classifications in glioblastoma may detect therapy sensitivities.

Immunohistochemistry would potentially allow the identification of molecular subtypes in routine clinical practice.

Experimental Design: Paraffin-embedded tumor samples of 124 uniformly treated, newly diagnosed glioblastoma patients were submitted to RNA-Sequencing, immunohistochemistry, and immune-phenotyping to identify differences in molecular subtypes associated with treatment sensitivities.

Results: We detected high molecular and immunohistochemical overlapping of the TCGA mesenchymal subtype with IGS cluster 23 and of the TCGA classical subtype with IGS cluster 18.

Immunohistochemical patterns, gene fusion profiles, and immune-phenotypes varied across subtypes.

Immunohistochemistry revealed that the TCGA classical subtype was identified by high expression of EGFR and low expression of PTEN, while the mesenchymal subtype was identified by low expression of SOX2 and high expression of two antibodies, SHC1 and TCIRG1, selected based on RNA differential transcriptomic expression. The proneural subtype was identified by frequent positive IDH1 expression and high Olig2 and Ki67 expression. Immune-phenotyping showed that mesenchymal and IGS 23 tumors exhibited a higher positive effector cell score, a higher negative suppressor cell score, and lower levels of immune checkpoint molecules. The cell-type deconvolution analysis revealed that these tumors are highly enriched in M2 macrophages, resting memory CD4⁺ T cells, and activated dendritic cells, indicating that they may be ideal candidates for immunotherapy, especially with anti-M2 and/or dendritic cell vaccination.

Conclusions: There is a subset of tumors, frequently classified as mesenchymal or IGS cluster 23, that may be identified with immunohistochemistry and could well be optimal candidates to immunotherapy.

Key words: glioblastoma, immune-phenotype, immunohistochemistry, SHC1, TCIRG1, mesenchymal

INTRODUCTION

For many years, the backbone of first-line therapy for glioblastoma has been maximal and safe surgery followed by radiation therapy with concomitant and adjuvant temozolomide, although recently the addition of tumor treating fields has shown some additional survival benefit.^{1,2} Prognosis is highly dependent of clinical and molecular factors, including *IDH* mutation status and *MGMT* promoter methylation.^{3,4}

Molecular subtyping of glioblastoma may help determine prognosis and identify the optimal therapy for specific populations, as it has been shown in other diseases as breast cancer.^{5,6} While several systems have been proposed for classifying glioblastoma⁷⁻²⁰, two classifications have been used in clinical trials to explore treatment sensitivities: The Cancer Genome Atlas (TCGA) and the Intrinsic Glioma Subtypes (IGS).²¹⁻²⁵ The latest TCGA classification identified three subtypes – classical, proneural, and mesenchymal – and suggested that there was intratumoral heterogeneity since some cases activated more than one transcriptional signature.¹³ The TCGA analysis also confirmed a different pattern of activation of the immune microenvironment, especially in mesenchymal tumors, which had an increased presence of macrophages/microglia, M1 and M2 macrophages, and neutrophil gene signatures.¹³ The IGS classification^{16,26} segregates gliomas of all histological grades into seven different clusters. Glioblastoma tumors are distributed mainly in clusters 18, 22, and 23 but can also be seen in other clusters, where prognosis is better. While the TCGA analyses were performed mainly in fresh-frozen specimens, the IGS used both fresh frozen and formalin-fixed, paraffin-embedded (FFPE) tissues. Parallelisms have been drawn between the two classifications; for example, samples assigned to IGS clusters 9 and 17 are generally assigned to the TCGA proneural subtype, those in IGS 18 to the classical subtype, and those in IGS 23 to the mesenchymal subtype.²³

The identification of molecular subtypes by immunohistochemistry has been explored in several studies, as it would allow these subtypes to be integrated in routine pathologic diagnosis²⁷⁻³¹ and be used in customizing treatment. In addition, recent findings on tumor-immune cell interaction and determinants of immunogenicity have highlighted the role of immunotherapies.³² The importance of gene fusions in solid tumors has been appreciated only recently, largely due to high-throughput technologies, such as RNA-Sequencing (RNA-Seq), and has led to important therapeutic implications.³³

We have explored the TCGA and IGS glioblastoma molecular subtypes by analyzing FFPE samples from a uniformly treated population of newly diagnosed glioblastoma patients. Our objectives were to characterize the subtypes by immunohistochemistry and to explore molecular alterations, such as immune-phenotypical patterns and gene fusions, that could identify different subtypes as candidates for selected therapies.

METHODS

Patients and study design

This study was approved by the Institutional Review Board of the Hospital Germans Trias i Pujol and by the Ethics Committees of all the participating institutions and their Biobanks and was conducted in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments. All patients or their representatives gave their written informed consent to participate in this study.

From 2004 to 2014, the GLIOCAT project ³⁴ collected clinical data from 432 consecutive glioblastoma patients from six institutions, all of whom had received the standard first-line treatment (surgery followed by radiotherapy with concurrent and adjuvant temozolomide). The pathological diagnosis confirming glioblastoma was performed by several pathologists before patients were included in the project. Once selected for inclusion, each case was anonymized and given a number to identify it across all data. Pre- and post-surgical MRIs were uploaded on a shared platform linking each case to their clinical and molecular data. Demographic and treatment characteristics were recorded. DNA was extracted for the study of *MGMT* methylation status if it had not previously been assessed. TMAs were built for immunohistochemical analyses of all samples with sufficient tissue. RNA was extracted and assessed for quality before sending it for RNA-sequencing (RNA-Seq).

DNA extraction and assessment of MGMT methylation

DNA was extracted from two 15- μ m sections of FFPE tissue using the QIAamp DNA Mini Kit (QIAGEN GmbH, Hilden, Germany), following the manufacturer's protocol. In cases with less than 50% of tumor cells, the tumor tissue was macrodissected manually. Then 500 ng of extracted DNA was subjected to bisulfite treatment using the EZ DNA Methylation-Gold Kit (Zymo Research Corporation, Irvine, CA). DNA methylation patterns in the CpG island of the *MGMT* gene were determined by

methylation-specific PCR (MSP) using primers specific for either methylated or modified non-methylated DNA, as previously described.³⁵

Tissue microarray (TMA) construction and immunohistochemical analyses

Between two and four cores of each case were obtained, depending on the amount of tissue available.

High necrotic areas were not selected. TMAs were constructed using a Veridiam Tissue Array Instrument (El Cajon, California, USA), model VTA-100, using a 1-mm diameter needle.

Consecutive 4- μ m-thick sections were obtained and hematoxylin-eosin staining was done in sections 1, 20, and 40 in order to evaluate the persistence of the tumor at each spot.

Sixteen antibodies were preplanned, selected based on the literature (IDH1-R132H, p53, Olig2, Ki67, c-Met, p16, CD44, D2-40, nestin, SOX2, YKL40, EGFR, PDL1, H3K27M, H3.3 G34R and PTEN) and four (IDO1, TCIRG1, SHC1, RUNX3) were selected based on RNA-Seq results showing differential expression in the TCGA glioblastoma molecular subtypes: classical vs mesenchymal, classical vs proneural, and mesenchymal vs proneural. The differential expression analysis was performed as follows. Using the first 38 sequenced samples classified according to the SVM algorithm, we performed an unsupervised analysis with DESeq2 (RRID:SCR_015687)³⁶ (Supplementary Data File S1a-c and Fig. 1) and a further analysis between pairs (classical/mesenchymal, classical/proneural and mesenchymal/proneural) for the top 50 differentially expressed genes (Supplementary Fig. 1 A-C). Two of these 50 genes (TCIRG1 and SHC1) and an additional two not in the top 50 (RUNX3 and IDO1) were consistently related to the mesenchymal subtype. These four genes were selected for analysis by immunohistochemistry (Supplementary Table S1). Primary antibody characteristics, RRIDs, pretreatment conditions, dilutions and scoring systems used are shown in Supplementary Table S2. Immunoreactivity was scored on different cell sub-locations using a quantitative or semi-quantitative scale. Cores with less than 50% of tumor were considered non-evaluable, and samples with more than 50% of non-evaluable cores were considered non-informative.

Benchmark XT (Ventana Medical System Inc., Tucson, AZ), Bond Max (Leica Microsystems GmbH, Wetzlar, Germany), and Dako (Dako Cytomation, Glostrup, Denmark) automated immunolabelling systems were used for different antibodies.

RNA-Seq assessments

The highest-quality RNA samples were further anonymized and sent to the Centro Nacional de Análisis Genómico (CNAG-CRG, Barcelona, Spain) for analysis by RNA-Seq. Before performing RNA-Seq analysis in the formalin-fixed, paraffin-embedded (FFPE) samples, we compared results in four cases of paired FFPE and fresh frozen (FF) samples and confirmed the similarity of results in the two types of tissue samples.³⁴

RNA extraction and quality assessment

The RNA extraction of FFPE tumor specimens was performed on five 15µm-deep tissue cuts using the RNeasy FFPE Kit (Qiagen, Hilden, Germany) according to the manufacturer's recommendations. RNA quantity and purity were measured with the NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA) and the Qubit RNA HS Assay Kit (Invitrogen, Eugene, OR) used with the Qubit Fluorometer (Thermo Fisher Scientific). RNA was analyzed with the 2100 Bioanalyzer (Agilent, Waldbronn, Germany) and quality was assessed by Eukaryote Total RNA Nano Bioanalyzer assay (Agilent) according to the "DV200" standard – the percentage of RNA fragments >200 nucleotides – as recommended by Illumina (San Diego, CA) for FFPE samples.

FFPE samples were sequenced in different batches and we first selected the highest-quality RNAs from patients with the most complete clinical and radiological data. However, as many samples did not have RNA with the minimal quality requirements, we subsequently included cases without available MRIs and, finally, those without informative results from the *MGMT* methylation analyses.

RNA library construction and sequencing

The starting input material for the construction of the RNA library was DNA-free total RNA from FFPE samples, obtained using the KAPA RNA HyperPrep Kit with RiboErase (HMR) (Roche Kapa Biosystems, Wilmington, MA), according to the manufacturer's protocol. Briefly, the ribosomal RNA (rRNA) was depleted from 0.8-1µg of total RNA using RiboErase. rRNA-depleted RNA samples were fragmented in the presence of divalent ions at 65°C for 1 minute due to the low quality of the initial total RNA. Following RNA fragmentation (first and second strand synthesis), the Illumina-compatible bar-coded adapters were ligated at a concentration of 1.5 µM. Libraries were enriched with 15 cycles of PCR using KAPA HiFi HotStart ReadyMix (Kapa-Roche). The size and quality of the libraries were assessed in a High Sensitivity DNA Bioanalyzer assay (Agilent).

The libraries were sequenced on HiSeq2000 (Illumina) in paired-end mode with a read length of 2x76bp using TruSeq SBS Kit v4. Each sample was sequenced in a fraction of a sequencing v4 flow cell lane, following the manufacturer's protocol. Image analysis, base calling and quality scoring of the run were processed using Real Time Analysis (RTA 1.18.66.3) software and followed by the generation of FASTQ sequence files by CASAVA (RRID:SCR_001802).

RNA-Seq processing and data analysis

RNA-Seq paired-end reads were mapped against the human reference genome (GRCh38) using STAR (RRID:SCR_015899, version 2.5.2a)³⁷ with ENCODE parameters for long RNA. Genes were quantified using RSEM (RRID:SCR_013027, version 1.2.28)³⁸ and GENCODE, (RRID:SCR_014966, version 24). Mapping quality metrics were calculated with GEMtools (<http://gemtools.github.io/>) and RSeQC (RRID:SCR_005275).³⁹ Normalization and transformation of the expression data for plotting and analysis were done with the variance stabilizing transformation function (vst) of DESeq2 (RRID:SCR_015687).³⁶ Principal component analysis (PCA) was performed with the R prcomp function and the R ggplot2 library (RRID:SCR_014601).

Classification of glioblastoma molecular subtypes

The TCGA classification of glioblastoma molecular subtypes¹¹⁻¹³ was performed with the GlioVis portal,⁴⁰ which contains over 6500 tumor samples of approximately 50 expression datasets and facilitates classification with three algorithms: Support Vector Machine (SVM), K-nearest neighbor (KNN) and single sample Gene Set Enrichment Analysis (ssGSEA). G-CIMP status was predicted using the GlioVis SVM algorithm. The GlioVis glioblastoma TCGA cohort was selected as training dataset for both glioblastoma molecular subtype and G-CIMP predictions. The IGS classification of glioblastoma molecular subtypes was done with the R clusterRepro package^{13,16}, using the centroids for IGS0, IGS9, IGS16, IGS17, IGS18, IGS22 and IGS23 subtypes as described by Gravendeel.^{16 24,41}

Immune-phenotyping in glioblastoma molecular subtypes

We applied immune profiling algorithms to our dataset to characterize *in silico* the cellular composition of the intratumoral immune infiltrates. We used two different approaches. First, we used the ImmunoPhenogram tool (<https://tcia.at/tools/toolsMain>) to calculate the z-scores for MHC, CP molecules, EC, and SC.³² We then performed a cell-type deconvolution analysis with

CIBERSORT (mode: absolute)⁴² with the 22 immune cell type signature (LM22) as reference dataset, to qualitatively enumerate the immune cell types found within the tumors and differences among the subtypes and patient populations.

Gene fusions in glioblastoma molecular subtypes

We then screened for fusions that could characterize each subtype. For fusion detection, RNA-Seq reads were mapped against the reference human genome with STAR-fusion version 0.7.0 with default options using GENCODE version 24. Previously described fusions were selected as final candidates to describe subtypes.⁴³

Statistical analyses

Categorical variables were compared with the χ^2 or Fisher's exact test. Clinical data were merged with molecular and immunohistochemical results for statistical analyses. Immunohistochemical values were presented as positive or negative, median values with 25-75% percentiles or in quartiles if the median was near 0. Differences in protein expression among subtypes were compared with the Kruskal-Wallis test. The Spearman coefficient was used to estimate the correlation between gene and protein expression. Analyses were performed with R software v3.4.2 and SPSS v24 (IBM, RRID:SCR_002865).

RESULTS

Tumor samples were obtained from 329 of the 432 glioblastoma patients registered in the GLIOCAT project.³⁴ After multiple RNA extractions from each sample, 357 RNA libraries were prepared and RNA-Seq results were obtained for 151 tumor samples. Twenty-seven of these samples were not included in subsequent analyses: 18 because the RNA was degraded and/or contaminated and nine because they were obtained at second surgeries at recurrence. The remaining 124 samples were included in further analyses. In addition, 17 tissue microarrays (TMAs) were prepared and 280 samples were analyzed by immunohistochemistry. One-hundred and thirteen patients had both RNA-Seq results and immunohistochemical data. (Supplementary Table 3).

RNA-Seq characterization of glioblastoma molecular subtypes

Based on the RNA-Seq results, the tumor samples were classified into the three TCGA molecular subtypes¹³, using three GlioVis algorithms SVM, KNN and ssGSEA (Supplementary Data File S1d).⁴⁰ There was agreement among all three algorithms on the subtype classification of 82 samples, which were

considered a separate subgroup in further analyses (Supplementary Table 3). Samples were also classified according to the IGS molecular subtypes and information about TCGA subtype, IGS subtype, *MGMT* status and presence or absence of fusions is shown in Fig. 2. Most tumors classified as IGS cluster 18 aligned with the TCGA classical subtype and those classified as IGS cluster 23 aligned with the TCGA mesenchymal subtype, especially for those samples that called the same subtype group by all three algorithms used. IGS 9 and IGS 17 aligned mostly with proneural tumors.

Characterization of TCGA molecular subtypes by immunohistochemistry

We tested 20 antibodies to analyze samples classified in the TCGA molecular subtypes by each of the three algorithms (Supplementary Table S1e). No patients had histone H3F3A mutations. Five of the preplanned antibodies (EGFR, SOX2, IDH1-R132H, Olig2, and Ki67) showed statistically significant different patterns of expression among subtypes, which confirmed parallelisms between the classical and the IGS 18 and the mesenchymal and the IGS 23 subtypes (Table 1). The differential expression study provided information about the genes that were expressed differentially in the molecular subtypes. Of the 50 genes with the greatest differential expression (Supplementary Data File S1a-c and Supplementary Fig. 1), the mesenchymal subtype showed the highest significant adjusted p-values to be differentiated from the other subtypes. We selected four of the genes for immunohistochemical analysis (Supplementary Table S1). A correlation was observed between gene and protein expression for all the antibodies selected (Supplementary Fig. 2).

Two of these genes (TCIRG1 and SHC1) clearly identified the mesenchymal subtype, as classified by all three algorithms, as well as the IGS 23 cluster. In summary, immunohistochemistry revealed that the TCGA classical subtype was identified by high expression of EGFR and low expression of PTEN, while the mesenchymal subtype and IGS 23 cluster were identified by low expression of SOX2 and high expression of SHC1 and TCIRG1. In addition, the proneural subtype was identified by frequent positive expression of IDH1 and high expression of Olig2 and Ki67. Most of these patterns of expression were shared by IGS 18 and IGS 23 tumors (Table 1, Fig. 3 and Supplementary Fig. S3). Supplementary Fig. S4 shows representative staining for each of these antibodies.

Immune-phenotyping of TCGA molecular subtypes and IGS clusters

We then explored differences in immune cell infiltration among the TCGA subtypes (Fig. 4) and found no significant differences in immunophenoscores. Nevertheless, when we broke down the scores into major

histocompatibility complex (MHC), immune checkpoint (CP), effector cells (EC), and suppressor cells (SC), we found that mesenchymal, independently of the algorithm used, and IGS 23 tumors had a higher positive EC score and a more negative SC score than classical or proneural tumors. Mesenchymal and IGS 23 tumors also had a more negative CP score indicating that they express lower levels of immune checkpoint molecules.

The cell-type deconvolution analysis showed that mesenchymal tumors were highly enriched in M2 macrophages, resting memory CD4⁺ T cells, and activated dendritic cells and that IGS 23 tumors showed the same pattern (Fig. 5).

Gene fusions in TCGA and IGS clusters

Among the 124 samples with informative RNA-Seq results, 25 (20.1%) had gene fusions (Fig. 1). Fusions were more prevalent in the classical subtype: 12 samples with fusions (63.1%) were classified as classical by all three methods, while six were classified as classical by one or two methods. Fifteen of the 25 samples with fusions (60%) were classified as IGS 18. Four samples (3.2%) harbored more than one fusion, mainly EGFR with other genes.

The most frequent gene involved in fusions was EGFR, which occurred in 14 samples (11.2%), most of which were IGS 18 and/or TCGA classical tumors (Supplementary data file S1f).

DISCUSSION

We have examined the TCGA¹³ and IGS¹⁶ molecular subtypes in FFPE tumor samples from a uniformly treated cohort of 124 newly diagnosed glioblastoma patients. We explored the feasibility of classifying tumors using immunohistochemistry, which can easily be integrated into clinical practice, and we analyzed biological differences, such as immune-phenotyping patterns and gene fusions, that could identify different subtypes as candidates for selected therapies.

We found that samples were classified differently depending on the Gliovis algorithm⁴⁰ used for the TCGA subtypes. Nevertheless, we confirmed previous observations²³ of a high overlapping between the TCGA mesenchymal subtype and the IGS cluster 23 and between the TCGA classical subtype and the IGS cluster 18, which was especially evident for those samples that called only one subtype. We also identified an alignment of IGS clusters 9 and 17 with the proneural subtype.

Several previous groups have explored the use of immunohistochemistry to predict glioblastoma subtypes²⁷⁻³¹ and have proposed different immunohistochemical panels for identifying the TCGA subtypes but none analyzed their cases based on results at the RNA level. By basing our immunohistochemical analyses on the differential expression observed in our RNA-Seq data, we were able to select antibodies directed to specific genes.

Our immunohistochemistry results showed that the molecular subtypes had different patterns of protein expression. Certain proteins were under- or overexpressed in each particular subtype regardless of the algorithm used to classify the sample. In line with previous reports^{27-29,31}, the classical subtype was consistently identified by immunohistochemistry by overexpression of EGFR, as has previously been described at the molecular level.^{11,12} The IGS cluster 18 showed a similar immunohistochemical pattern. In addition, gene fusions were predominant in these subtypes and most of the EGFR fusions were detected in tumors classified as classical by at least one of the algorithms. Proneural tumors were clearly identified by high expression of Olig2 and Ki67 and frequent IDH1-R132H mutations. Both IDH1 and Olig2 have previously been related to the proneural subtype.^{10,11,44}

Importantly, the differential expression analysis allowed us to identify some genes that were highly and differentially expressed in tumors mainly classified into the mesenchymal subtype by most of the algorithms.. Immunohistochemistry targeting two of these genes, *SHC1* and *TCIRG1*, was able to identify these tumors. The same immunohistochemical pattern was observed in tumors classified as IGS cluster 23. To the best of our knowledge, neither SHC1 nor TCIRG1 antibodies have previously been described as a marker of the mesenchymal subtype. *SHC1* is a RAS-pathway protein that interacts with *EGFR* and receptors with protein tyrosine kinase activity; it has been related to angiogenesis and cytokine-mediated signaling pathways, among other functions (<https://www.uniprot.org/uniprot/P29353>). *TCIRG1* is involved in several biological processes, including B-cell differentiation, cellular response to cytokine stimulus, immunoglobulin production and mediated immune response, inflammatory response, interferon-gamma secretion, memory T cell activation and differentiation, and T-helper 1 cell activation (<https://www.uniprot.org/uniprot/A0A024R5E5>).⁴⁵

In line with previous observations, immune-phenotyping also revealed differences according to glioblastoma molecular subtype. Immunophenoscores were not uniformly different across the subtypes but we found that mesenchymal tumors – and the overlapping IGS cluster 23 – are more infiltrated by

immune effector cells and contain a lower number of suppressor cells. This pattern has been related to immunotherapy response in melanoma.³² Moreover, as previously reported¹³, these tumors are enriched in M2 macrophages, resting memory CD4+ T cells, and activated dendritic cells. Therefore, this group of tumors, which are frequently classified as mesenchymal and IGS 23, seem to present a stronger host anti-tumor response, making them potentially ideal candidates for immunotherapy approaches, such as anti-M2 and/or dendritic vaccination.⁴⁶⁻⁴⁹ The association of increased levels of M2 macrophages in mesenchymal tumors has been described by Wang et al.,¹³ but in contrast with their findings, we could not confirm an increased expression of dendritic cells in the classical subtype.

Our study has some limitations, including its retrospective nature and the use of FFPE specimens for the TCGA subtyping. Importantly, however, this is the first study based on RNA-Seq analysis of FFPE samples in a sizeable cohort ($n = 124$) of newly diagnosed, uniformly treated glioblastoma patients with available paired specimens for immunohistochemical analysis. The subtypes differed in regard to protein expression, gene fusions, and immune-phenotype and could be identified by protein expression studied by immunohistochemistry. Our results lead us to postulate therapeutic implications that warrant further investigation in larger series and clinical trials. Glioblastoma is a heterogeneous disease with different biologic and metabolic pathways and there is a subset of tumors that are frequently classified as mesenchymal or IGS cluster 23. These tumors may be identified with immunohistochemistry in routine clinical practice by high TCIRG1 and SHC1 expression and low SOX2 expression and could well be optimal candidates for immunotherapy.

Acknowledgements

This study was funded by the Fundació La Marató TV3 (665/C/2013)

(<http://www.ccma.cat/tv3/marato/projectes-financats/2012/231/>), PIs: Carmen Balana, Francesc Alameda and Nuria de la Iglesia. Anna Esteve Codina is funded by ISCIII/MINECO (PT17/0009/0019) and co-funded by FEDER. Silvia Menéndez is supported by the Health Department of the Generalitat de Catalunya (PERIS SLT006/17/00040), Carmen Balana is funded by ISCIII/AES (INT19/00032).

The authors thank the biobanks of the participating institutions: Fundació Institut Mar d'Investigacions Mèdiques, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol (IGTP), Fundació Institut de Recerca de l'Hospital de la Santa Creu i St. Pau (IIB Sant Pau), and the Xarxa de Bancs de Tumors sponsored by Pla Director

d'Oncologia de Catalunya (XBTC). We also thank Natalia Garcia-Balaña for her role as data manager and in coordinating the group. We are grateful for the collaboration of all the professionals of the Glioma Catalonia (GLIOCAT) Group.

Other Clinical Investigators of the GLIOCAT Group that contributed to this study:

Marc Dabad¹, Avelina Tortosa², Teresa Pujol³, Laura Oleaga³, Cristian de Quintana-Schmidt⁴, Miquel Gil-Gil⁵, Izaskun Valduvieto⁶, Anna Martinez-Cardús⁷, Marta Domenech⁷, Anna Estival⁷, Salvador Villa⁸, Angels Camins⁹, Jordi Marruecos¹⁰, Sira Domenech¹¹

¹CNAG-CRG, Centre for Genomic Regulation, Barcelona Institute of Science and Technology, 08028 Barcelona, Spain.

²Laboratori de Químio-resistència i Càncer, School of Medicine and Health Sciences, University of Barcelona, Department of Fundamental Care and Medical-Surgical Nursing, Bellvitge Biomedical Research Institute (IDIBELL), Bellvitge, Spain.

³Radiology Department, Hospital Clínic, Barcelona, Spain.

⁴Neurosurgery Department, Hospital de Sant Pau, Barcelona, Spain.

⁵Neuro-Oncology Unit & Medical Oncology Department, Institut Català d'Oncologia (ICO), Institut de Investigació Bellvitge (IDIBELL), L'Hospitalet, Barcelona, Spain.

⁶Radiation Therapy Department, Hospital Clínic, Barcelona, Spain.

⁷ Institut Català d'Oncologia (ICO) Badalona, Badalona Applied Research Group in Oncology (B-ARGO Group), Institut Investigació Germans Trias i Pujol (IGTP), Badalona, Spain.

⁸Radiation Oncology Department, Institut Català d'Oncologia (ICO), Badalona, Spain.

⁹Radiology Department, Institut de Diagnòstic per la Imatge, Hospital de Bellvitge, L'Hospitalet de Llobregat, Barcelona, Spain

¹⁰Radiation Oncology Department, Institut Català d'Oncologia (ICO), Girona, Spain.

¹¹Radiology Department, Institut de Diagnòstic per la Imatge, Hospital Germans Trias i Pujol, Badalona, Spain.

REFERENCES

1. Stupp R, Mason WP, van den Bent MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med*. 2005; 352(10):987-996.
2. Stupp R, Taillibert S, Kanner A, et al. Effect of Tumor-Treating Fields Plus Maintenance Temozolomide vs Maintenance Temozolomide Alone on Survival in Patients With Glioblastoma: A Randomized Clinical Trial. *JAMA*. 2017; 318(23):2306-2316.
3. Hegi ME, Diserens AC, Gorlia T, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med*. 2005; 352(10):997-1003.

4. Bleeker FE, Atai NA, Lamba S, et al. The prognostic IDH1(R132) mutation is associated with reduced NADP+-dependent IDH activity in glioblastoma. *Acta Neuropathol.* 2010; 119(4):487-494.
5. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature.* 2000; 406(6797):747-752.
6. Cancer Genome Atlas N. Comprehensive molecular portraits of human breast tumours. *Nature.* 2012; 490(7418):61-70.
7. Li A, Walling J, Ahn S, et al. Unsupervised analysis of transcriptomic profiles reveals six glioma subtypes. *Cancer Res.* 2009; 69(5):2091-2099.
8. Freije WA, Castro-Vargas FE, Fang Z, et al. Gene expression profiling of gliomas strongly predicts survival. *Cancer Res.* 2004; 64(18):6503-6510.
9. Nigro JM, Misra A, Zhang L, et al. Integrated array-comparative genomic hybridization and expression array profiles identify clinically relevant molecular subtypes of glioblastoma. *Cancer Res.* 2005; 65(5):1678-1686.
10. Phillips HS, Kharbanda S, Chen R, et al. Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. *Cancer Cell.* 2006; 9(3):157-173.
11. Verhaak RG, Hoadley KA, Purdom E, et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell.* 2010; 17(1):98-110.
12. Brennan CW, Verhaak RG, McKenna A, et al. The somatic genomic landscape of glioblastoma. *Cell.* 2013; 155(2):462-477.
13. Wang Q, Hu B, Hu X, et al. Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. *Cancer Cell.* 2017; 32(1):42-56 e46.
14. Nutt CL, Mani DR, Betensky RA, et al. Gene expression-based classification of malignant gliomas correlates better with survival than histological classification. *Cancer Res.* 2003; 63(7):1602-1607.

15. Godard S, Getz G, Delorenzi M, et al. Classification of human astrocytic gliomas on the basis of gene expression: a correlated group of genes with angiogenic activity emerges as a strong predictor of subtypes. *Cancer Res.* 2003; 63(20):6613-6625.
16. Gravendeel LA, Kouwenhoven MC, Gevaert O, et al. Intrinsic gene expression profiles of gliomas are a better predictor of survival than histology. *Cancer Res.* 2009; 69(23):9065-9072.
17. Karcher S, Steiner HH, Ahmadi R, et al. Different angiogenic phenotypes in primary and secondary glioblastomas. *Int J Cancer.* 2006; 118(9):2182-2189.
18. Liang Y, Diehn M, Watson N, et al. Gene expression profiling reveals molecularly and clinically distinct subtypes of glioblastoma multiforme. *Proc Natl Acad Sci U S A.* 2005; 102(16):5814-5819.
19. Rickman DS, Bobek MP, Misek DE, et al. Distinctive molecular profiles of high-grade and low-grade gliomas based on oligonucleotide microarray analysis. *Cancer Res.* 2001; 61(18):6885-6891.
20. Sallinen SL, Sallinen PK, Haapasalo HK, et al. Identification of differentially expressed genes in human gliomas by DNA microarray and tissue chip techniques. *Cancer Res.* 2000; 60(23):6617-6622.
21. Cloughesy T, Finocchiaro G, Belda-Iniesta C, et al. Randomized, Double-Blind, Placebo-Controlled, Multicenter Phase II Study of Onartuzumab Plus Bevacizumab Versus Placebo Plus Bevacizumab in Patients With Recurrent Glioblastoma: Efficacy, Safety, and Hepatocyte Growth Factor and O-6-Methylguanine-DNA Methyltransferase Biomarker Analyses. *Journal of Clinical Oncology.* 2017; 35(3):343-+.
22. Sandmann T, Bourgon R, Garcia J, et al. Patients With Proneural Glioblastoma May Derive Overall Survival Benefit From the Addition of Bevacizumab to First-Line Radiotherapy and Temozolomide: Retrospective Analysis of the AVAglio Trial. *J Clin Oncol.* 2015; 33(25):2735-2744.
23. Erdem-Eraslan L, Gravendeel LA, de Rooi J, et al. Intrinsic molecular subtypes of glioma are prognostic and predict benefit from adjuvant procarbazine, lomustine, and vincristine chemotherapy in combination with other prognostic factors in anaplastic oligodendroglial brain tumors: a report from EORTC study 26951. *J Clin Oncol.* 2013; 31(3):328-336.

24. Erdem-Eraslan L, van den Bent MJ, Hoogstrate Y, et al. Identification of Patients with Recurrent Glioblastoma Who May Benefit from Combined Bevacizumab and CCNU Therapy: A Report from the BELOB Trial. *Cancer Res.* 2016; 76(3):525-534.
25. Gao Y, Weenink B, van den Bent MJ, et al. Expression-based intrinsic glioma subtypes are prognostic in low-grade gliomas of the EORTC22033-26033 clinical trial. *European journal of cancer (Oxford, England : 1990).* 2018; 94:168-178.
26. Gravendeel LA, de Rooi JJ, Eilers PH, van den Bent MJ, Sillevs Smitt PA, French PJ. Gene expression profiles of gliomas in formalin-fixed paraffin-embedded material. *Br J Cancer.* 2012; 106(3):538-545.
27. Conroy S, Krut FA, Joseph JV, et al. Subclassification of newly diagnosed glioblastomas through an immunohistochemical approach. *PLoS One.* 2014; 9(12):e115687.
28. Popova SN, Bergqvist M, Dimberg A, et al. Subtyping of gliomas of various WHO grades by the application of immunohistochemistry. *Histopathology.* 2014; 64(3):365-379.
29. Nagy A, Garzuly F, Padanyi G, et al. Molecular Subgroups of Glioblastoma- an Assessment by Immunohistochemical Markers. *Pathol Oncol Res.* 2019; 25(1):21-31.
30. Motomura K, Natsume A, Watanabe R, et al. Immunohistochemical analysis-based proteomic subclassification of newly diagnosed glioblastomas. *Cancer Sci.* 2012; 103(10):1871-1879.
31. Le Mercier M, Hastir D, Moles Lopez X, et al. A simplified approach for the molecular classification of glioblastomas. *PloS one.* 2012; 7(9):e45475.
32. Charoentong P, Finotello F, Angelova M, et al. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep.* 2017; 18(1):248-262.
33. Nikanjam M, Okamura R, Barkauskas DA, Kurzrock R. Targeting fusions for improved outcomes in oncology treatment. *Cancer.* 2019.
34. Esteve-Codina A, Arpi O, Martinez-Garcia M, et al. A Comparison of RNA-Seq Results from Paired Formalin-Fixed Paraffin-Embedded and Fresh-Frozen Glioblastoma Tissue Samples. *PloS one.* 2017; 12(1):e0170632.
35. Esteller M, Hamilton SR, Burger PC, Baylin SB, Herman JG. Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is a common event in primary human neoplasia. *Cancer Res.* 1999; 59(4):793-797.

36. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014; 15(12):550.
37. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013; 29(1):15-21.
38. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics.* 2011; 12:323.
39. Wang L, Wang S, Li W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics.* 2012; 28(16):2184-2185.
40. Bowman RL, Wang Q, Carro A, Verhaak RG, Squatrito M. GlioVis data portal for visualization and analysis of brain tumor expression datasets. *Neuro-oncology.* 2017; 19(1):139-141.
41. Gao Y, Weenink B, van den Bent MJ, et al. Expression-based intrinsic glioma subtypes are prognostic in low-grade gliomas of the EORTC22033-26033 clinical trial. *Eur J Cancer.* 2018; 94:168-178.
42. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods.* 2015; 12:453.
43. Stransky N, Cerami E, Schalm S, Kim JL, Lengauer C. The landscape of kinase fusions in cancer. *Nature communications.* 2014; 5:4846.
44. Noushmehr H, Weisenberger DJ, Diefes K, et al. Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell.* 2010; 17(5):510-522.
45. Bulwin GC, Heinemann T, Bugge V, et al. TIRC7 inhibits T cell proliferation by modulation of CTLA-4 expression. *Journal of immunology (Baltimore, Md. : 1950).* 2006; 177(10):6833-6841.
46. Gregoire H, Roncali L, Rousseau A, et al. Targeting Tumor Associated Macrophages to Overcome Conventional Treatment Resistance in Glioblastoma. *Front Pharmacol.* 2020; 11:368.
47. Wen PY, Reardon DA, Armstrong TS, et al. A Randomized Double-Blind Placebo-Controlled Phase II Trial of Dendritic Cell Vaccine ICT-107 in Newly Diagnosed Patients with Glioblastoma. *Clin Cancer Res.* 2019; 25(19):5799-5807.
48. Pombo Antunes AR, Scheyltjens I, Duerinck J, Neyns B, Movahedi K, Van Ginderachter JA. Understanding the glioblastoma immune microenvironment as basis for the development of new immunotherapeutic strategies. *Elife.* 2020; 9.

49. Prins RM, Soto H, Konkankit V, et al. Gene expression profile correlates with T-cell infiltration and relative survival in glioblastoma patients vaccinated with dendritic cell immunotherapy. *Clin Cancer Res.* 2011; 17(6):1603-1615.

FIGURE LEGENDS

Fig. 1 Heatmap showing the scaled expression of the top differentially expressed genes between the glioblastoma subtypes (SVM classification).

Fig. 2 Classification of tumor samples by molecular subtype (classical, mesenchymal, proneural) according to the KNN, SVM and ssGSEA algorithms, indicating the correspondence of IGS clusters, G-CIMP status, *MGMT* methylation status, and gene fusions with each subtype. There is a high overlap between mesenchymal and IGS 23 tumors, especially for those that are classified as mesenchymal by all three algorithms used for TCGA.

MET, methylated *MGMT*; UNMET, unmethylated *MGMT*

Fig. 3 SHC1 and TCIRG1 (A) RNA expression and (B) IHC protein expression by IGS (first row) and TCGA subtype (second row: KNN, third row: SVM, fourth row: ssGSEA). Significant differences (Kruskall Wallis test) were seen between mesenchymal and classical or proneural tumors but not between classical and proneural tumors.

Fig. 4 Immune-phenotyping results of tumors in each of the IGS and the three TCGA molecular subtypes as classified by each of the three GlioVis algorithms (K-nearest neighbor [KNN], single sample Gene Set Enrichment Analysis [ssGSEA], and Support Vector Machine [SVM]): (A) immunophenoscores (IPS); (B) major histocompatibility complex (MHC); (C) effector cells (EC); (D) suppressor cells (SC); (E) immune checkpoint (CP).

Fig. 5 Comparison of resting memory CD4⁺ T cells, activated dendritic cells and M2 macrophage enrichment among IGS clusters and TCGA molecular subtypes as classified by each of the GlioVis algorithms (KNN, SVM and ssGSEA).

Table 1 Summary of *P*-values for association between immunohistochemical results and glioblastoma molecular subtypes as identified by GlioVis algorithms

Protein	SVM	KNN	ssGSEA	All three algorithms
Classical				
High EGFR ^a	<0.001	<0.001	<0.001	<0.001
Low PTEN ^b	0.074	0.003	0.025	0.001
Mesenchymal				
High TCIRG1 ^a	0.013	0.015	0.002	0.010
High SHC1 ^a	<0.001	<0.001	<0.001	<0.001
Low SOX2 ^a	<0.001	0.001	0.003	0.002
Proneural				
IDH1-R132H positive ^a	0.04	0.003	0.004	0.027
High Olig2 ^a	<0.001	<0.001	<0.001	<0.001
High Ki67 ^a	<0.001	0.03	0.001	0.003

^aSignificant association between expression levels and molecular subtype regardless of the algorithm used for subtype classification based on RNA-Seq results.

^bCommon trend towards association between expression levels and molecular subtype significant in at least two algorithms. IGS 18 cluster shared with the classical subtype a higher expression of EGFR and IGS 23 cluster shared with the mesenchymal subtype a high expression of TCIRG1 and SHC1. Supplementary Data file S1e shows details of all immunohistochemical results with corresponding *P*-values. Fig. 3 and Supplementary Fig. S3 show a graphic representation of these results.

Supplementary Fig S2. Spearman correlation between gene and protein expression for the four antibodies selected according to RNA-Seq differential expression.

Supplementary Fig. S3. IHC protein expression of IDH1, EGFR (intensity 3 as high versus 0-2), high intensity, PTEN, OLIG2, Ki67 and SOX2 by IGS (first row) and TCGA subtype (second row: KNN, third row: SVM, fourth row: ssGSEA). Significant differences were seen between mesenchymal and classical or proneural tumors. See Supplementary Data File S1e for results simplified in Table 1.

Supplementary Fig. S4. Representative images comparing high versus low immunohistochemical expression of **(A,B)** EGFR, **(C,D)** PTEN, **(E,F)** SOX-2, **(G,H)** SHC1, **(I,J)** TCIRG1, **(K,L)** IDH1R132H (K:positive, L: negative), **(M,N)** Olig2, and **(O,P)** Ki67

Supplementary Table S1. Statistically significant adjusted *P*-values for differential expression of genes in the pairwise comparison of molecular subtypes. Two of these genes were selected for immunohistochemical analysis for the identification of the mesenchymal subtype.

Gene	Ensembl ID	Classical vs. Mesenchymal	Classical vs. Proneural	Mesenchymal vs. Proneural	Overexpression in Mesenchymal Subtype
<i>SHC1</i>	ENSG00000160691	0.003	NS	1.35E-08	yes
<i>TCIRG1</i>	ENSG00000110719	0.0008	NS	1.18E-08	yes
<i>IDO1</i>	ENSG00000131203	0.0039	NS	0.0015	no
<i>RUNX3</i>	ENSG00000020633	0.0137	NS	1.88E-05	no

Supplementary Table S2. Key resource identifiers (RRIDs), antibodies, methods, and scoring systems used for immunohistochemistry analyses.

Gene	Antibody	Chromosome	SOURCE/RRID	Clone	Antigen Retrieval	Dilution	Cellular Location	Scoring System
Isocitrate dehydrogenase (NADP(+)) 1, cytosolic ^a	IDH1-R132H	2q33	(Dianova Cat# DIA-H09, RRID:AB_2335716)	H-09	Heat CC1	1/40	Cyt	Positive vs negative expression
Tumor protein p53	p53	17p13	(Ventana Medical Systems Cat# 790-2912, RRID:AB_2335971)	DO-7	Citrate buffer pH 7.3	Prediluted	Nu	Percentage of positive tumoral cells and positive (>10% of intense staining in tumoral cells) vs negative expression
Marker of proliferation Ki-67	Ki67	10q25	(Ventana Medical Systems Cat# 790-4286, RRID:AB_2631262)	30/09	Heat CC1	Prediluted	Nu	Percentage of positive tumoral cells
Cyclin-dependent kinase inhibitor 2A, multiple tumor suppressor 1	p16	9p21,3	(Ventana Medical Systems Cat #790-4430, RRID: NA)	E6H4	Heat CC1	Prediluted	Cyt/Nuc	Semiquantitative scale of positivity (0: negative; 1: <5%; 2: 5-25%; 3: 25-50%; 4: 50-75%; 5: >75%)
Epidermal growth factor receptor	EGFR	7q11,2	(Ventana Medical Systems Cat# 790-4347, RRID:AB_2617183)	HER1	Proteinase K	Prediluted	Mem/Cyt	Scale of intensity (0: negative; 1, 2, 3, increasing intensity)
Podoplanin	D2-40	1p36,21	(Cell Marque Cat# 322M-18, RRID:AB_1160805)	D2-40	Heat CC1	Prediluted	Cyt	Scale of intensity (0: negative; 1, 2, 3, increasing intensity)
CD44 (CD44R, CSPG8, HCELL, IN, MC56, MDU2, MDU3, MIC4, Pgp1)	CD44	11p13	Roche Cat#790-4547, RRID:NA)	SP37	Heat	Prediluted	Mem/Cyt	H-Score (intensity x percentage of positive tumoral cells)
SRY-box transcription factor 2	SOX2	3q26,3	(Millipore Cat# MAB4343, RRID:AB_827493)	SP76	Heat	1/100	Nu	H-Score (intensity x percentage of positive tumoral cells)
Nestin	Nestin	1q23,1	(Abcam Cat# ab22035, RRID:AB_446723)	SP103	PtLink Low	1/100	Mem/Cyt	Percentage of positive tumoral cells

MET proto-oncogene, receptor tyrosine kinase	c-Met	7q31.2	Roche Cat#790-4430, RRID:NA)	SP44	Heat CC1	Prediluted	Mem/Cyt	Percentage of positive tumoral cells
Oligodendrocyte transcription factor 2	Olig2	21q22	(Abcam Cat# ab56643, RRID:AB_944547)	EP-112	Heat	1/100	Nu	Percentage of positive tumoral cells
Chitinase 3 like 1	YKL-40	1q32.1	(Abcam Cat# ab77528, RRID:AB_2040911)	polyclonal	Heat	1/100	Mem/Cyt	H-Score (intensity x percentage of positive tumoral cells)
Programmed cell death 1 ligand 1 precursor	PD-L1	9p24.1	(Roche Cat# 790-4905, RRID:AB_2819099)	SP263	Heat CC1	Prediluted	Mem	Percentage of positive tumoral cells
Anti-Histone H3.3	H3K27M	1q41.12	(Millipore Cat# ABE419, RRID:AB_2728728)	polyclonal	HIER + Citrate buffer pH 6	1/500	Nu	Positive vs negative expression
Anti-Histone H3.3 G34R	H3.3 G34R	1q42.12	(RevMab Biosciences Cat# 31-1120-00, RRID:AB_2716433)	RM240	HIER + Citrate buffer pH 6	1/200	Nu	Positive vs negative expression
Indoleamine 2,3-dioxygenase 1	IDO1	8p11	(Sigma-Aldrich Cat# HPA027772, RRID:AB_1846222)	polyclonal	PtLink High	1/100	Nu/Cyt	Percentage of positive tumoral cells
Phosphatase and tensin homolog	PTEN	10q23.31	(Agilent Cat# M3627, RRID:AB_2174185)	6H2.1	Tris Edta pH9	1/200	Nu/Cyt	High (strong staining intensity in >50% of tumoral cells) vs low expression
T cell immune regulator 1, ATPase H ⁺ transporting V0 subunit a3	TCIRG1	11q13	(Sigma-Aldrich Cat# HPA038742, RRID:AB_10795138)	polyclonal	PTLink low	1/400	Mem/Cyt	H-Score (intensity x percentage of positive tumoral cells)
SHC adaptor protein 1	SHC1	10q21.3	(Sigma-Aldrich Cat# HPA001844, RRID:AB_1856826)	polyclonal	PtLink Low	1/100	Nu/Cyt	Percentage of positive tumoral cells
Runt related transcription factor 3	RUNX3	1p36	(Abcam Cat# ab40278, RRID:AB_883789)	R35-G4	PTLink low	1/200	Nu/Cyt	Percentage of positive tumoral cells

^a Six doubtful cases were sequenced for IDH1, two of which showed a mutation.

Supplementary Table S3. Patient characteristics and molecular subtypes

	All Patients with RNA-Seq Results <i>n</i> = 124 <i>n</i> (%)	All Patients with both RNA-Seq and IHC Results <i>n</i> = 113 <i>n</i> (%)	All Patients with IHC Results <i>n</i> = 280 <i>n</i> (%)
Age, yrs			
median (range)	64.0 (55.0-71.0)	64.0 [55.0-71.0]	62.0 (53.0-69.0)
≤65 yrs	70 (56.5)	64 (56.6)	186 (66.4)
≤50 yrs	15 (12.1)	49 (43.4)	47 (16.8)
Sex			
Male	76 (61.3)	68 (60.2)	164 (58.6)
Female	48 (38.7)	45 (39.8)	116 (41.4)
Glioblastoma type			
Primary	120 (96.8)	109 (96.5)	273 (97.5)
Secondary	4 (3.23)	4 (3.54)	7 (2.50)
Extent of surgery			
Total resection	27 (23.9)	27 (26.2)	74 (28.1)
Partial resection	71 (62.8)	65 (63.1)	171 (65.0)
Biopsy	15 (13.3)	11 (10.7)	18 (6.8)
MGMT methylation status			
Methylated	64 (51.6)	60 (53.1)	132 (47.1)
Unmethylated	52 (41.9)	49 (43.4)	138 (49.3)
Unknown	8 (6.5)	4 (3.54)	10 (3.6)
IDH1 status			
Wild-type	107 (86.3)	107 (94.7)	264 (94.3)
Mutated	5 (4.0)	5 (4.4)	11 (3.9)
Unknown	12 (9.7)	1 (0.9)	5 (1.8)
G-CIMP+	6 (4.8)	6 (5.3)	6 (2.1)
Unknown	0 (0.0)	0 (94.7)	167 (59.6)
TCGA subtype by SVM			
Classical	46 (37.1)	40 (35.4)	40 (14.3)
Mesenchymal	40 (32.3)	36 (31.9)	36 (12.9)
Proneural	38 (30.6)	37 (32.7)	37 (13.2)
Unknown	0 (0.0)	0 (0.0)	167 (59.6)
TCGA subtype by KNN			
Classical	80 (64.5)	70 (61.9)	70 (25.0)
Mesenchymal	14 (11.3)	14 (12.4)	14 (5.0)
Proneural	30 (24.2)	29 (25.7)	29 (10.4)
Unknown	0 (0.0)	0 (0.0)	167 (59.6)
TCGA subtype by ssGSEA			
Classical	53 (42.7)	45 (39.8)	45 (16.1)
Mesenchymal	32 (25.8)	30 (26.5)	30 (10.7)
Proneural	39 (31.5)	38 (33.6)	38 (13.6)
Unknown	0 (0.0)	0 (0.0)	167 (59.6)
TCGA subtype by SVM+KNN+ssGSEA (<i>n</i> = 82)			
Classical	42 (51.2)	36 (48.0)	36 (12.9)
Mesenchymal	13 (15.9)	13 (17.3)	13 (4.6)
Proneural	27 (32.9)	26 (34.7)	26 (9.3)
Unknown	0 (0.0)	0 (0.0)	167 (59.6)

IGS classification			
0	6 (4.8)	5 (4.4)	5 (1.8)
9	14 (11.3)	14 (12.4)	14 (5.0)
16	2 (1.6)	2 (1.8)	2 (0.7)
17	12 (9.6)	11(9.7)	11 (3.9)
18	68 (54.8)	59 (62.2)	59 (21.1)
22	4 (3.2)	4(3.5)	4 (1.4)
23	18 (14.5)	18 (15.9)	18 (6.4)
Unknown	0 (0.0)	0 (0.0)	167 (59.6)

IHC, immunohistochemistry.

S1e Immunohistochemical results of samples classified in the three TCGA molecular subtypes [according to the GlioVis algorithms KNN, SVM, and ssGSEA] and KGS clusters. P-values for association between immunohistochemical results and molecular subtypes																											
IHC	All [N=113]	CLASSIFICATION BY KNN				CLASSIFICATION BY ssGSEA				CLASSIFICATION BY SVM				CLASSIFICATION BY ALL THREE ALGORITHMS				CLASSIFICATION BY KGS CLUSTERS									
		Classical [N=70]	Mesenchymal [N=14]	Pronеural [N=29]	P	Classical [N=45]	Mesenchymal [N=30]	Pronеural [N=38]	P	Classical [N=40]	Mesenchymal [N=36]	Pronеural [N=37]	P	Classical [N=36]	Mesenchymal [N=13]	Pronеural [N=26]	P	0 [N=5]	9 [N=14]	16 [N=2]	17 [N=11]	18 [N=59]	22 [N=4]	23 [N=18]	P		
IDH1-R132H	Positive	5 [4.99%]	0 [0.00%]	0 [0.00%]	5 [17.2%]	0 [0.00%]	0 [0.00%]	5 [13.9%]	0.003	1 [2.5%]	0 [0.00%]	4 [11.4%]	0.04	0 [0.00%]	0 [0.00%]	4 [15.4%]	0.012	0 [0.00%]	2 [14.3%]	0 [0.00%]	3 [27.3%]	0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]		
	Negative	107 [95.4%]	69 [100%]	14 [100%]	24 [82.8%]	45 [100%]	30 [100%]	32 [86.5%]		39 [97.5%]	36 [100%]	32 [88.9%]		36 [100%]	13 [100%]	22 [84.6%]		5 [100%]	12 [85.7%]	1 [100%]	8 [72.7%]	57 [100%]	4 [100%]	17 [100%]	0.006		
	p53	2.75 [0.50;7.00]	3.00 [0.94;3.8]	3.00 [1.00;7.00]	1.70 [0.30;5.75]	0.664	2.50 [1.00;5.00]	3.00 [0.30;7.00]	2.25 [0.30;21.6]	0.93	1.88 [0.35;4.94]	3.00 [1.00;6.17]	3.25 [0.30;21.8]	0.296	2.00 [1.00;5.00]	3.30 [1.00;9.00]	1.88 [0.35;9.69]	0.886	0.00 [0.00;0.25]	3.88 [1.00;30.0]	0.15 [0.08;0.22]	1.50 [0.28;14.5]	3.00 [1.00;6.30]	9.90 [4.41;14.1]	3.00 [1.00;7.00]	0.200	
	Negative c10	85 [75.2%]	53 [75.7%]	10 [71.4%]	22 [75.9%]		38 [84.4%]	22 [73.3%]	25 [65.8%]		33 [82.5%]	29 [80.6%]	23 [62.2%]		29 [80.6%]	9 [69.2%]	19 [73.1%]		4 [80.0%]	8 [57.1%]	2 [100%]	8 [72.7%]	46 [80.7%]	4 [100%]	13 [76.5%]		
Ki67	Positive	25 [22.1%]	15 [21.4%]	3 [21.4%]	7 [24.1%]	0.742	5 [11.1%]	7 [23.3%]	13 [34.2%]	0.062	5 [12.5%]	6 [16.7%]	14 [37.8%]	0.038	5 [13.9%]	3 [23.1%]	7 [26.9%]	0.434	1 [20.0%]	6 [42.9%]	0 [0.00%]	3 [27.3%]	11 [19.3%]	0 [0.00%]	4 [23.5%]	0.541	
	Negative	4.62 [2.06;7.94]	4.75 [2.38;7.00]	2.50 [1.76;4.55]	7.50 [3.00;14.0]	0.03	4.15 [1.60;6.55]	2.88 [2.00;5.19]	7.88 [4.25;14.1]	0.001	3.50 [1.25;5.88]	3.62 [2.19;6.25]	7.75 [4.62;14.2]	<0.001	4.30 [1.45;6.25]	2.75 [1.80;3.50]	7.88 [4.50;14.2]	0.003	1.00 [0.00;3.25]	11.8 [7.89;17.6]	0.30 [0.30;0.30]	4.62 [2.26;7.38]	4.75 [2.31;7.00]	12.9 [7.55;21.9]	2.60 [1.85;3.45]	<0.001	
	p16	1.00 [0.00;4.00]	0.50 [0.00;4.00]	0.50 [0.00;1.75]	1.00 [1.00;5.00]	0.144	0.00 [0.00;2.00]	1.00 [0.00;4.00]	1.00 [0.00;5.00]	0.001	0.00 [0.00;3.25]	1.00 [0.00;3.25]	1.00 [0.00;5.00]	0.19	0.00 [0.00;4.00]	1.00 [0.00;2.00]	1.00 [1.00;5.00]	0.162	0.00 [0.00;0.00]	2.00 [1.00;5.00]	2.50 [1.25;3.75]	1.00 [0.50;2.50]	1.00 [0.00;4.00]	0.00 [0.00;1.25]	0.50 [0.00;1.75]	0.106	
	Negative																										
EGFR	0	2 [1.79%]	2 [2.90%]	0 [0.00%]	0 [0.00%]		0 [0.00%]	2 [6.67%]	0 [0.00%]		0 [0.00%]	2 [5.56%]	0 [0.00%]		0 [0.00%]	0 [0.00%]	0 [0.00%]		0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]	1 [1.69%]	0 [0.00%]	1 [5.56%]		
	1	10 [8.93%]	3 [4.35%]	3 [21.4%]	4 [13.8%]		0 [0.00%]	6 [20.0%]	4 [10.8%]		0 [0.00%]	6 [16.7%]	4 [11.1%]		0 [0.00%]	3 [23.1%]	4 [15.4%]		0 [0.00%]	2 [14.3%]	0 [0.00%]	1 [9.09%]	3 [5.08%]	1 [25.0%]	3 [16.7%]		
	2	28 [25.0%]	8 [13.0%]	8 [57.1%]	11 [37.9%]		4 [8.89%]	11 [36.7%]	13 [35.1%]		3 [7.50%]	13 [36.1%]	12 [33.3%]		3 [8.33%]	7 [53.8%]	10 [38.5%]		0 [0.00%]	7 [50.0%]	1 [100%]	3 [27.3%]	8 [13.6%]	1 [25.0%]	8 [44.4%]		
	3	7 [64.3%]	55 [79.7%]	3 [21.4%]	14 [48.3%]	<0.001	41 [91.1%]	11 [36.7%]	20 [54.1%]	<0.001	37 [92.5%]	15 [41.7%]	20 [55.6%]	<0.001	33 [91.7%]	3 [23.1%]	12 [46.2%]	<0.001	5 [100%]	5 [35.7%]	0 [0.00%]	7 [63.6%]	47 [79.7%]	2 [50.0%]	6 [33.3%]	0.096	
O2-40	0	20 [19.2%]	10 [15.2%]	1 [9.09%]	9 [33.3%]		7 [16.7%]	3 [11.1%]	10 [28.6%]		7 [18.9%]	5 [15.2%]	8 [23.5%]		6 [18.2%]	0 [0.00%]	8 [33.3%]		1 [20.0%]	4 [30.8%]	0 [0.00%]	4 [40.0%]	10 [17.9%]	0 [0.00%]	1 [6.67%]		
	1	30 [28.8%]	18 [27.3%]	4 [36.4%]	8 [29.6%]		8 [19.0%]	10 [37.0%]	12 [34.3%]		5 [13.5%]	12 [36.4%]	13 [38.2%]		4 [12.1%]	4 [40.0%]	7 [29.2%]		3 [60.0%]	4 [30.8%]	1 [100%]	1 [10.0%]	13 [23.2%]	2 [50.0%]	6 [40.0%]		
	2	41 [39.4%]	27 [40.9%]	5 [45.5%]	9 [33.3%]		19 [45.2%]	11 [40.7%]	11 [31.4%]		17 [45.9%]	9 [33.4%]	10 [29.4%]		16 [48.5%]	5 [50.0%]	8 [33.3%]		1 [20.0%]	5 [38.5%]	0 [0.00%]	4 [40.0%]	23 [41.1%]	2 [50.0%]	6 [40.0%]		
	3	13 [12.5%]	11 [16.7%]	1 [9.09%]	1 [3.70%]	0.33	8 [19.0%]	3 [11.1%]	2 [5.71%]	0.201	8 [21.6%]	2 [6.00%]	3 [8.82%]	0.1	7 [21.2%]	1 [10.0%]	1 [4.17%]	0.065	0 [0.00%]	0 [0.00%]	0 [0.00%]	1 [10.0%]	10 [17.9%]	0 [0.00%]	2 [13.3%]	0.534	
CD44	220 [154.281]	243 [185.289]	190 [163.300]	175 [75.0;250]	0.011	229 [189.289]	250 [160.300]	183 [87.6;250]	0.035	237 [189.275]	213 [149.300]	193 [95.2;271]	0.147	237 [194.275]	198 [178.300]	179 [52.5;244]	0.024	102 [50.0;285]	170 [52.5;244]	250 [250;250]	175 [124.201]	245 [191.300]	282 [246.289]	183 [115;273]	0.037		
	SOX2	187 [97.0;252]	190 [99.9;242]	69.9 [31.6;150]	224 [147;283]	0.001	200 [99.9;256]	127 [76.7;182]	220 [147;281]	0.003	200 [101;255]	125 [73.2;187]	220 [150;282]	<0.001	200 [101;258]	80.2 [30.8;161]	227 [150;292]	0.002	128 [79.5;175]	271 [167;292]	80.1 [80.1;80.1]	200 [46;225]	194 [91.9;250]	264 [208;300]	125 [31.6;168]	0.001	
	Nestin	40.0 [20.0;57.0]	45.5 [31.5;61.3]	23.0 [14.9;43.2]	28.7 [13.5;40.0]	0.004	46.3 [31.5;62.5]	30.9 [20.0;50.7]	36.0 [15.0;47.5]	0.06	45.0 [19.8;60.6]	39.1 [20.5;59.2]	36.4 [19.1;48.1]	0.341	45.0 [30.6;62.5]	24.3 [17.5;46.7]	35.5 [16.4;40.0]	0.076	28.7 [28.3;56.7]	33.8 [9.85;40.0]	13.5 [13.5;13.5]	32.5 [11.7;41.8]	45.0 [32.0;60.6]	50.0 [39.2;67.5]	23.0 [18.1;47.3]	0.077	
	C-Met																										
>25%	105 [96.3%]	64 [95.5%]	12 [92.3%]	29 [100%]		42 [97.3%]	27 [93.1%]	36 [97.3%]		37 [97.4%]	33 [94.3%]	35 [97.2%]		33 [97.1%]	11 [91.7%]	26 [100%]	0.308	5 [100%]	14 [100%]	1 [100%]	11 [100%]	54 [94.7%]	4 [100%]	16 [94.1%]			
	Olig2	4 [3.67]	3 [4.48%]	1 [7.69%]	0 [0.00%]	0.675	1 [2.33%]	2 [6.90%]	1 [2.70%]	0.259	1 [2.63%]	2 [5.71%]	1 [2.78%]	0.381	1 [2.94%]	1 [8.33%]	0 [0.00%]		0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]	3 [4.26]	0 [0.00%]	1 [5.88%]	0.929	
	>51%																										
	H3K27M negative	24.8 [8.57;52.5]	24.1 [10.0;45.0]	3.00 [0.00;5.00]	52.5 [23.8;65.0]	<0.001	30.6 [15.0;47.5]	3.85 [2.38;19.0]	46.6 [21.0;61.2]	<0.001	30.0 [12.5;55.0]	10.0 [3.00;24.6]	48.8 [22.5;62.5]	<0.001	35.0 [15.0;55.9]	2.75 [0.00;4.25]	52.5 [24.6;66.2]	<0.001	21.7 [16.0;67.0]	53.8 [52.5;69.2]	30.0 [30.0;30.0]	21.1 [11.2;32.1]	25.9 [12.5;47.5]	41.0 [23.7;54.3]	3.00 [1.50;4.25]	<0.001	
<25%	56 [51.9%]	35 [53.0%]	12 [92.3%]	9 [31.0%]		20 [45.2%]	23 [82.1%]	13 [36.1%]		18 [46.2%]	26 [76.5%]	12 [34.3%]		14 [40.0%]	11 [91.7%]	7 [26.9%]		3 [60.0%]	1 [7.14%]	0 [0.00%]	7 [70.0%]	29 [50.0%]	1 [25.0%]	15 [93.8%]			
	>25%	25 [23.1%]	17 [25.8%]	1 [7.69%]	5 [17.2%]		13 [29.5%]	4 [14.3%]	6 [16.7%]		10 [25.6%]	7 [20.6%]	6 [17.1%]		10 [28.6%]	1 [8.33%]	4 [15.4%]		0 [0.00%]	2 [14.3%]	1 [100%]	2 [20.0%]	16 [27.6%]	1 [25.0%]	1 [6.25%]	0.011	
	>51%	2 [1.85%]	2 [3.03%]	0 [0.00%]	0 [0.00%]		1 [2.33%]	1 [3.45%]	0 [0.00%]		0 [0.00%]	1 [2.86%]	1 [2.86%]		0 [0.00%]	0 [0.00%]	0 [0.00%]		0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]		
	>75%	4 [3.70%]	2 [3.03%]	0 [0.00%]	2 [6.90%]	0.003	1 [2.27%]	1 [3.57%]	2 [5.56%]	<0.001	1 [2.56%]	1 [2.94%]	2 [5.71%]	<0.001	1 [2.86%]	0 [0.00%]	2 [7.69%]	0.003	0 [0.00%]	2 [14.3%]	0 [0.00%]	2 [20.0%]	2 [3.45%]	0 [0.00%]	0 [0.00%]		
YKL-40	5.00 [0.00;30.0]	5.25 [0.00;24.1]	33.3 [4.90;80.8]	2.20 [0.00;31.7]	0.107	4.50 [0.00;17.3]	20.4 [4.40;35.6]	2.10 [0.00;31.7]	0.105	4.70 [0.00;21.3]	13.3 [3.30;41.7]	0.30 [0.00;23.8]	0.023	3.30 [0.00;19.3]	20.0 [4.85;40.2]	1.73 [0.00;30.8]	0.188	19.2 [15.0;28.3]	0.15 [0.00;26.2]	0.00 [0.00;0.00]	5.00 [1.18;21.6]	4.70 [0.00;25.4]	0.00 [0.00;0.08]	26.7 [5.85;47.1]	0.033		
	PD-L1																										
	<25%	100 [92.6%]	60 [90.9%]	11 [84.6%]	29 [100%]		39 [90.7%]	25 [86.2%]	36 [100%]		36 [94.7%]	30 [85.7%]	34 [97.1%]		32 [94.1%]	10 [83.3%]	26 [100%]		5 [100%]	14 [100%]	1 [100%]	11 [100%]	51 [89.5%]	3 [100%]	15 [88.2%]		
	>50%	2 [1.85%]	2 [3.03%]	0 [0.00%]	0 [0.00%]		1 [2.33%]	1 [3.45%]	0 [0.00%]		0 [0.00%]	1 [2.86%]	1 [2.86%]		0 [0.00%]	0 [0.00%]	0 [0.00%]		0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]	2 [3.51%]	0 [0.00%]	0 [0.00%]		
H3K3634R negative	51-75%	0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]		0 [0.00%]	0 [0.00%]	0 [0.00%]		0 [0.00%]	0 [0.00%]	0 [0.00%]		0 [0.00%]	0 [0.00%]	0 [0.00%]		0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]		
	>75%	6 [5.56%]	4 [6.06%]	2 [15.4%]	0 [0.00%]	0.275	3 [6.98%]	3 [10.3%]	0 [0.00%]		2 [5.26%]	4 [11.4%]	0 [0.00%]	0.155	2 [5.88%]	2 [16.7%]	0 [0.00%]	0.115	0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]	4 [7.02%]	0 [0.00%]	2 [11.8%]	0.923	
	H3K3634R negative	113 [100%]	70 [100%]	14 [100%]	29 [100%]	NA	45 [100%]	30 [100%]	38 [100%]	NA	40 [100%]	36 [100%]	37 [100%]	NA	36 [100%]	13 [100%]	26 [100%]	NA	5 [100%]	14 [100%]	2 [100%]	11 [100%]	59 [100%]	4 [100%]	18 [100%]	NA	
	H3K3634R negative	70 [100%]	70 [100%]	14 [100%]	29 [100%]	NA	45 [100%]	30 [100%]	38 [100%]	NA	40 [100%]	36 [100%]	37 [100%]	NA	36 [100%]	13 [100%]	26 [100%]	NA	5 [100%]	14 [100%]	2 [100%]	11 [100%]	59 [100%]	4 [100%]	18 [100%]	NA	
PTEN	low expression																										