

Correlation between molecular features and genetic subtypes of Glioblastoma: critical analysis in 109 cases

Thais F Galatro^I, Paula Sola^I, Isabele F Moretti^I, Flavio K Miura^I, Sueli M Oba-Shinjo^I, Suely KN Marie^{II,III}, Antonio M Lerario^{III,IV}

^I Universidade de Sao Paulo, Faculdade de Medicina FMUSP, Department of Neurology, Laboratory of Molecular and Cellular Biology, LIM15, São Paulo, SP, Brazil.

^{II} University of São Paulo, Center for Studies of Cellular and Molecular Therapy (NAP-NETCEM-NUCEL), São Paulo, SP, Brazil.

^{III} University of Michigan, Division of Internal Medicine Metabolism, Endocrinology and Diabetes, Ann Arbor, MI, USA.

^{IV} Universidade de Sao Paulo, Faculdade de Medicina, Bioinformatics coordinator of SELA (NGS facility core), São Paulo, SP, BR.

OBJECTIVE: Glioblastoma, the most common and lethal brain tumor, is also one of the most defying forms of malignancies in terms of treatment. Integrated genomic analysis has searched deeper into the molecular architecture of GBM, revealing a new sub-classification and promising precision in the care for patients with specific alterations.

METHOD: Here, we present the classification of a Brazilian glioblastoma cohort into its main molecular subtypes. Using a high-throughput DNA sequencing procedure, we have classified this cohort into proneural, classical and mesenchymal sub-types. Next, we tested the possible use of the overexpression of the EGFR and CHI3L1 genes, detected through immunohistochemistry, for the identification of the classical and mesenchymal subtypes, respectively.

RESULTS: Our results demonstrate that genetic identification of the glioblastoma subtypes is not possible using single targeted mutations alone, particularly in the case of the Mesenchymal subtype. We also show that it is not possible to single out the mesenchymal cases through CHI3L1 expression.

CONCLUSION: Our data indicate that the Mesenchymal subtype, the most malignant of the glioblastomas, needs further and more thorough research to be ensure adequate identification.

KEYWORDS: glioblastoma, classification, DNA sequence analysis, CHI3L1, EGFR.

Galatro TF, Sola P, Moretti IF, Miura FK, Oba-Shinjo SM, Mariel SKN, Lerario AM. Correlation between molecular features and genetic subtypes of Glioblastoma: critical analysis in 109 cases. MedicalExpress. 2017 Oct; 4(5):M170505

Received for Publication on July 28,2017; First review on September 6, 2017; Accepted for publication on October 3, 2017; Online on October 20, 2017

E-mail: thaisgalatro@usp.br

INTRODUCTION

Glioblastoma multiforme (GBM) is the most common and most malignant form of brain tumor in adults. Glioblastomas belong to the glioma group of tumors. The World Health Organization¹ classifies gliomas according to their resemblance to their cell of origin, along with histological and molecular features. Glioblastomas, the most prevalent within the glioma group, are extremely aggressive grade IV tumors, exhibiting high mitotic rates, micro-vascular proliferation and necrosis; they also present the poorest prognosis, with a median survival of 15 months from the time of diagnosis.² Due to their invasive nature, complete surgical

resection is very difficult to achieve. The presence of residual tumor cells results in recurrence and malignant progression, albeit at different intervals.

Glioblastomas can be further divided into two subgroups: primary GBMs, which arise *de novo*, and secondary GBMs, which result from the progression of a lower grade tumor.³ These two clinical forms of GBM have different and extensively characterized molecular features.³ The past decade has seen the rise of high-throughput sequencing techniques that provide in-depth knowledge of molecular alterations in tumor cells. GBM has been one of the most molecularly profiled tumors by several groups, including The Cancer Genome Atlas (TCGA) Networks.⁴⁻¹⁰ These studies have singled out specific determinant mutations within the newly identified subtypes: proneural, neural, classical and mesenchymal. Proneural tumors present alterations

DOI: 10.5935/MedicalExpress.2017.05.05

in the IDH1, PDGFRA and TP53 genes; classical tumors display mainly the EGFR mutation/amplification and the presence of the EGFRvIII oncogenic variant; mesenchymal tumors show RB1 and NF1 mutations as their main features. The neural subtype presents overexpression of neural markers, but display no specific genetic alterations. Patient overall survival was also correlated with GBM molecular subtypes, with the mesenchymal subtype presenting the worst prognosis. However, the reproducibility, clinical relevance, and functional basis of these subclasses remain to be established.

In this study, we have developed a NGS-based gene panel to detect and analyze several genes commonly mutated in GBM. We classified a series of Brazilian GBM cases based on the somatic mutation signatures previously established for the three main subtypes: proneural, classical and mesenchymal. We validated our genetic findings through orthogonal methods and immunohistochemistry. In this multi-institution cohort of Brazilian patients, GBM subtype distribution and associated patient outcomes corroborate previously published results and further validates the clinical relevance of GBM subtyping. NGS-based approaches for GBM classification are fast, reliable, and therefore, may have value as a diagnostic tool that may add to the clinical decision-making process.

■ MATERIALS AND METHODS

Included patients and ethical statement. One hundred and nine GBM samples were obtained during therapeutic surgery of patients treated by the Neurosurgery Group of the Department of Neurology at Hospital das Clínicas at the School of Medicine of the University of São Paulo, in the period of 2000 to 2014. Written informed consent was obtained from all patients according to the ethical guidelines approved by the institutional Ethics Committee. (case # 0599/10). Complete patient information, along with the current WHO grading system¹ and clinical findings are presented in Supplemental Table 1, annexed to this article.

GBM sampling and diagnosis. GBM diagnosis was confirmed by neuropathologists from the Division of Pathological Anatomy of the same institution, according to the WHO grading system available at the time.¹¹ Samples were macrodissected and immediately snap-frozen in liquid nitrogen upon surgical removal. A 4µm-thick cryosection of each sample was analyzed under a light microscope after hematoxylin-eosin staining for assessment of cellular debris as well as necrotic and non-neoplastic areas; this was followed by removal from the frozen block by microdissection prior to DNA extractions.

Targeted gene panel sequencing. Based on previous literature assessing GBM molecular sub classification,^{6,8,9} 40 genes were targeted for capture and deep sequencing. This

customized panel included genes such as NF1, RB1, EGFR, TP53, PTEN, IDH1 and PDGFRA, whose alterations have all been previously associated with the three main GBM subtypes: proneural, classical and mesenchymal. Using the SureDesign tool (Agilent Technologies Inc., USA), the targeted RNA capture enrichment baits were designed to include coding exon regions and 50 bp from both the 3' end and 5' end of the flanking intronic sequence. The purpose was to incorporate possible splicing site mutations in our analysis. A total of 973 regions from the 40 genes were targeted for a final capture size of 2.79 Mega base pairs (Mb).

A target-enrichment DNA library was constructed using the Agilent SureSelect XT Target Enrichment Kit (Agilent Technologies Inc.), following the recommended protocol. Two hundred nanograms of tumor DNA was sheared using a E220 focused-ultrasonicator (Covaris, USA) to generate DNA fragments with a mean peak around 150bp. Indexed and adaptor-ligated libraries were multiplexed and paired-end sequenced on Illumina NextSeq 500 platform (Illumina, USA). An additional library was built from peripheral blood DNA from all the patients pooled in equal amounts.

Bioinformatics analysis. Sequencing data was generated as 150-bp paired-end reads using the Illumina Next-Seq platform. Raw data was aligned to the hg38 assembly of the human genome using the BWA mapping software.¹² Aligned reads were coordinate-sorted with the bamsort tool from Biobambam2.¹³ Variant calling was performed simultaneously in all the tumor samples with the reference genome-free algorithm Platypus,¹⁴ and in the germline pool with freebayes, with the parameters properly set to call pooled data (pooled-discrete and ploidy parameter set to 132, which corresponds to the number of alleles present in the pool). The resulting VCFs file were annotated with SnpEff and SnpSIFT.¹⁵

Tissue microarray (TMA) construction and Immunohistochemistry. Two representative areas of each tumor were chosen by neuropathologists and marked both on HE sections and on the original paraffin block. Each of the 0.6 mm-diameter three cores of tumor tissue was extracted from the marked area of each donor block using an arraying machine (MTA-1, Beecher Instruments Inc., USA). The cores were inserted into a TMA recipient block in predetermined sites. Sections of 3µm-thickness were cut from the TMA block. A representative TMA section was initially stained with HE to assess the suitability of each core, and all other sections were paraffin coated and stored at -20°C until use.

For immunohistochemical detection, serial TMA sections were deparaffinized, rehydrated, treated for endogenous peroxidase blocking and subjected to antigen retrieval. Slides were immersed in 10 mM citrate buffer, pH 6.0 and incubated at 122°C for 3 min using an electric pressure cooker (BioCare Medical, USA). Specimens were

then blocked and further incubated with anti-human CHI3L1 (mouse monoclonal, clone AT4A3; Abcam, United Kingdom) and with anti-human EGFR (mouse monoclonal, clone 31G7; Thermo Fisher Scientific, USA) at 16-20°C for 16 hours. Development of the reaction was performed with a commercial kit (Novolink; Novocastra, United Kingdom) at room temperature, using diaminobenzidine and Harris hematoxylin for nuclear staining. Optimization using positive controls suggested by the manufacturer of each antibody was performed in order to obtain optimal dilution. Two observers (TFG and IFM) evaluated staining intensity of tissue sections independently. A semi-quantitative scoring system considering both intensity of staining and the respective percentage of stained cells was applied as follows: **score 0**: no cells stained; **score 1**: 10–25% cells stained; **score 2**: 26–50% cells stained; **score 3**: 51–75% cells stained; **score 4**: 76–100% cells stained. Only cases with positive cell staining with scores ≥ 2 were considered as positive. Digital photomicrographs of representative fields were captured and processed using Picasa 3 (Google, USA).

Statistical analysis. Survival data was assessed by comparing the Kaplan-Meier survival curves using the log-rank (Mantel Cox) test. A p-value < 0.05 was considered statistically significant. For this analysis, we included 77 cases with complete clinical follow-up. Calculations were performed using SPSS, version 23.0 (IBM, USA).

■ RESULTS

Molecular classification of GBM. We used the mutational profile to classify tumors into three major molecular subtypes: proneural, classical, and mesenchymal. We defined the proneural subtype by the presence of IDH1, PDGFRA, and TP53 mutations. The classical subtype was defined by EGFR alterations (amplification and mutations), PTEN mutations, and the presence of the oncogenic variant EGFRVIII. EGFR amplification results was based on a previous publication from our group.¹⁶ Finally, for the definition of the mesenchymal subtype, we used mutations in NF1 and RB1. For each patient, we used DNA from blood leukocytes for germline alteration subtraction. None of the found alteration were present in these germline controls. Aside from this control, we considered in this analysis a) known pathogenic mutations, such as the IDH1 R132H, b) variants that were not present in population databases (1000kg/ExAC), c) variants with predicted pathogenicity (nonsense or frameshift, or a missense mutation with high SIFT and polyphen scores), d) variants not present in our germline pool. All the tumors which did not carry a mutation in the genes described above and met the defined criteria were classified as “others”. Given the lack of a specific genetic profile defining the neural subtype, we were not able to identify this subclass in our dataset. The complete list of genetic alteration can be found in Supplemental Table 2.

The targeted NGS analysis was performed on 109 GBM samples. A 20x coverage was achieved in ~97% of the targeted regions of the GBM samples. The mean coverage of the germline pool was 6063.67x.

Following the proposed criteria, we were able to classify 89 of the 109 analyzed GBM samples within the three major subtypes. Genetic alterations typical of the classical subtype were the most prevalent, accounting for 45.9% of the classified cases. Samples with alterations inherent to the mesenchymal subtype corresponded to 19.3% of the cases; alterations particular to the proneural subtype contributed 16.5% of the cases. This information is diagrammatically shown in Figure 1. A preview of survival rates are also shown in this figure.

The overall survival of GBM patients according to our molecular classification is shown in Figure 2: patients classified as mesenchymal GBM had a shorter overall survival in comparison to other subgroups (medians of 8, 9, and 13 months for the mesenchymal, classical, and proneural, respectively; log-rank $p < 0.05$).

Correlation between molecular classification of GBM and immunohistochemical markers to differentiate between classical and mesenchymal subtypes.

Concomitant with the molecular classification of GBMs, we also assessed the protein expression levels of CHI3L1 (YKL-40) and EGFR. CHI3L1 has been previously associated with the mesenchymal subtype of GBM, while EGFR overexpression is a known characteristic of the classical subtype. Immunohistochemistry staining was performed in a subset of our molecularly classified samples, comprising a cohort of 40 cases (8 mesenchymal, 9 proneural, 20 classical and 3 others). Figure 3 shows the results and Figure 4 depicts examples of stained tissue microarray sections. While only one of the mesenchymal samples showed negative levels of CHI3L1, cases positive for CHI3L1 (scoring above 2) comprised samples of the three subtypes (6 mesenchymal, 6 proneural and 10 classical cases were positive for CHI3L1). For EGFR, all proneural and 6 mesenchymal cases proved to be negative for this marker. Despite presenting genetic characteristics of classical subtypes, 4 of these cases did not present positivity for EGFR immunostaining. A complete listing of immunohistochemistry score can be found in Supplemental Table 2.

■ DISCUSSION

The advent of next generation sequencing and large-scale molecular analysis over the past decade revealed that molecular alterations predict patients' responses to treatment, overall survival and clinical outcome. A new light has been shed on the high level of GBM heterogeneity and new sub classifications have emerged. For GBM, several studies have singled out specific determinant mutations of the main, newly identified subtypes: proneural, classical

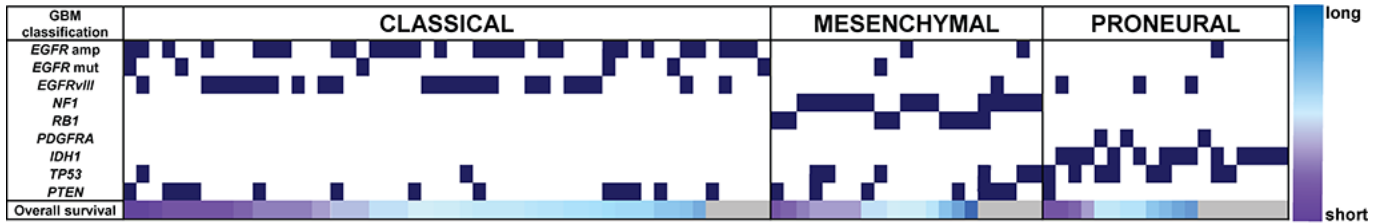


Figure 1 - Distribution of the molecularly classified GBM cases. Depiction of the main genetic alterations of the cases molecularly classified in our GBM cohort. Classification was achieved using a customized gene panel, followed by exome sequencing (refer to text for details). The color gradient at the bottom depicts patient overall survival (blue indicates long; purple short survival as shown in the color gradient at the left); gray represents cases without a complete clinical follow-up. amp: amplification; mut: mutation.

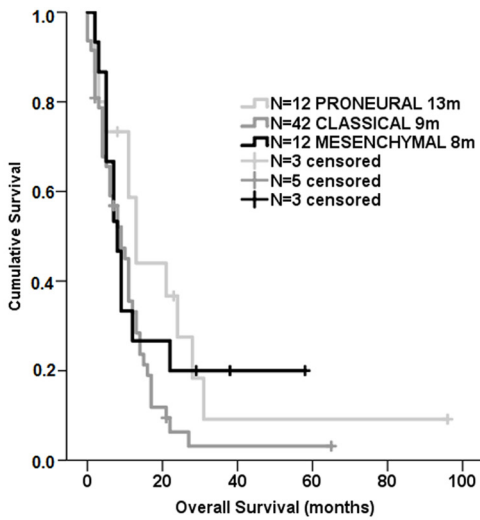


Figure 2: Survival curve of the molecularly classified GBM cases. Kaplan-Meier of 77 GBM cases detailing median survival (log rank test, $p=0.128$) times according to our molecular classification. Patients alive at the time of the last clinical follow-up were censored from this analysis. m, months.

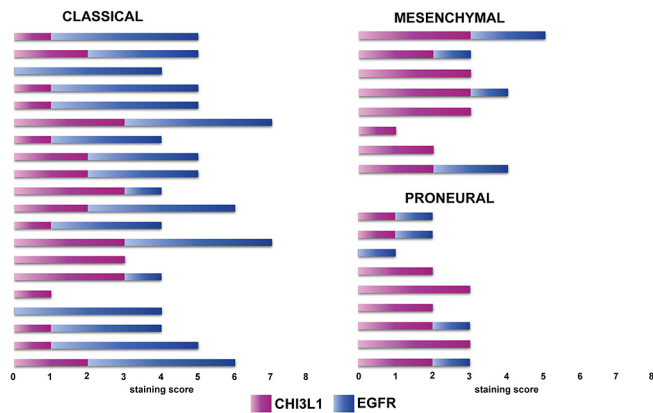


Figure 3: Scoring of immunohistochemical analysis for CHI3L1 and EGFR in GBM cases. Each GBM sample, previously classified by the molecular approach, was scored for the classical (EGFR, blue) and mesenchymal (CHI3L1, pink) immunohistochemical markers. Scoring ranged from 0-4, in which score 0: no cells stained; score 1: 10–25% cells stained; score 2: 26–50% cells stained; score 3: 51–75% cells stained; score 4: 76–100% cells stained. EGFR staining was prevalent in classical samples, although classical cases did not present significant staining of the marker, while other mesenchymal samples presented comparable levels of EGFR. CHI3L1 was present in high levels in cases from all three GBM subtypes.

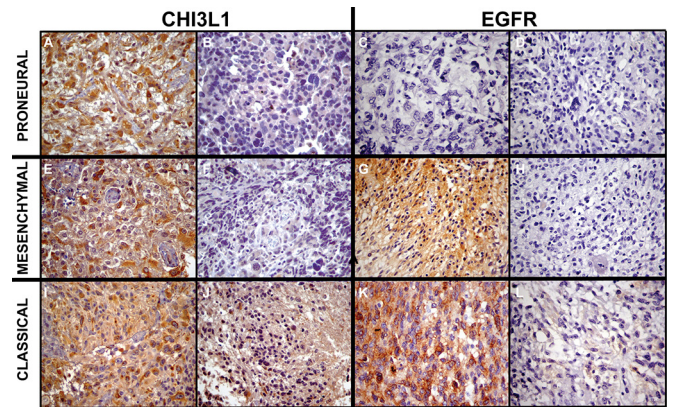


Figure 4: Representative immunohistochemical sections for the evaluation of CHI3L1 and EGFR in molecularly classified GBM samples. Proneural (A-D), mesenchymal (E-H) and classical (I-L) GBM representative cases were stained for CHI3L1 and EGFR, whose overexpression has been associated with mesenchymal and classical subtypes, respectively. CHI3L1 expression was widespread among GBM samples, and not exclusive of mesenchymal cases. Both proneural (A) and classical (I) samples displayed high levels of this marker. Despite being negative in proneural cases (D), EGFR overexpression failed to differentiate between mesenchymal and classical cases, with both groups showing positive and negative cases. The reaction was performed in paraffin embedded tissue sections with a commercial polymer kit (Novolink; Novocastra, UK), using diaminobenzidine as developer and Harris hematoxylin for nuclear counterstaining. 400x magnification for all images. Scale bar 10 μ m.

and mesenchymal.⁴⁻¹⁰ Their potential to aid the choice of better treatment options is a step forward toward precision medicine.

We have performed a somatic mutation analysis in our GBM cohort utilizing a customized gene panel that comprises all the coding regions and splicing regions the most commonly mutated genes described in GBM. Our NGS panel allowed us to classify approximately ~85% of our samples according to molecular subtype using a single assay; this would have been very laborious and time-consuming using traditional methods, such as Sanger sequencing.¹⁷ The results we obtained, regarding the percentage of each GBM subtype, as well as their association with patient overall survival, is in accordance with what has been published so far.⁵⁻⁹

Since the first documentation of the molecular subclasses of GBMs (available by 2010), these have proved

to be fundamental for new discoveries that may ultimately lead to better clinical approaches and precision medicine. For instance, Davis et al¹⁸ used the classification to compare genetic alterations in cultured brain-tumor initiating cells. This model system is important to study treatment options and GBM biology. Natash et al¹⁹ explored the role of oncostatin, a cytokine present in the microenvironment of GBM and its signaling pathway, which is associated with poor prognosis. They related such features to the aggressive nature in the mesenchymal GBM subtype. Chen and Xu²⁰ have recently developed an algorithm that matches FDA approved drugs to the molecular subtype of GBMs, based on the genetic alterations of each subtype. Thus, it is important to specifically differentiate classical from mesenchymal GBMs: classical GBMs are amenable to treatment by specific tyrosine kinase inhibitors and adjuvant antibodies;²¹ in contrast, mesenchymal GBMs have only one potentially useful drug, currently introduced as a phase 1 clinical trial (clinical trial identifier: NCT02272270). Therefore, new practical proposals, more suitable for major diagnostic centers, have been looked for. The exemplified studies highlight the usefulness of GBM classification in improving current knowledge, biological understanding and diagnosis and treatment options for this tumor.

However, and because the basis for such classification is the genetic approach, requiring the implementation of molecular biology techniques, other more feasible methodologies have been searched for. A recent study revealed that MRI derived quantitative volumetric tumor phenotype features only moderately predict the molecular GBM subtypes, suggesting that subtypes do not generally alter the size composition of tumor areas.²² Therefore, the more doable proteomic IHC-based approach, still based on the reported molecular findings, has been the preferred technique for patient classification.

Both proneural and classical GBM subtypes present genetic point mutations or alterations (IDH1 R132H and EGFRvIII, respectively) that have made it possible to develop, over recent years, specific antibodies to achieve a cheaper and more suitable approach in major diagnostic centers.^{23,24} Nonetheless, the alteration of EGFRvIII was found in only 28.4% of our GBM cohort, while classical cases corresponded to 45.9% of all cases. For those extra cases, EGFR protein overexpression still excluded 4 out of 23 samples with a classical genetic subtype.

The mesenchymal subtype presents a high level of heterogeneity of alterations within the usually mutated genes of this subtype (NF1 and RB1). CHI3L1 (also known as YKL-40) is a secreted lectin, related to the regulation of hypoxia-induced injury response and has been previously associated with the mesenchymal GBM.^{6,8} Previous studies have proposed CHI3L1 as an additional marker to be used as a diagnosis tool through an immunohistochemical approach for the sub classification of GBMs.^{25,26} Yet, our

results indicate that the higher expression of CHI3L1 is not an event exclusive of the mesenchymal cases, because high percentages of both proneural and classical samples displayed equivalent high levels.

■ CONCLUSION

Our results indicate the need for a genetic approach to further classify GBMs, so that a higher number of patients can profit from the precision provided by such classification in terms of treatment options.

■ ACKNOWLEDGEMENTS

The authors sincerely thank doctors and residents from the Discipline of Neurosurgery of the Department of Neurology at Hospital das Clínicas of School of Medicine, University of São Paulo, from the Neurosurgery group at Sírío Libanês Hospital and, particularly, Drs André M Bianco and Egmond Alves, for the therapeutic and diagnostic procedures of all patients included in this study. We also thank the physicians and technicians at the Division of Pathological Anatomy of the same institutions, particularly Dr Sergio Rosemberg, for the histological classification of tumor samples and tissue section processing. The authors thank São Paulo Research Foundation (FAPESP), grants #2013/07704-3, #2013/06315-3, #2013/02162-8, #2014-50137-5 and #2016/15652-1, CAPES-NUFFIC (062/15), Conselho Nacional de Pesquisa (CNPq 305730/2015-0) and Faculdade de Medicina FMUSP for financial support.

■ AUTHOR CONTRIBUTION

AML, SKNM and SMOS conceived the study. AML performed the bioinformatics analysis. TFG and PS performed the experiments. SKNM and FKM followed-up the patients. SKNM and AML provided supervision. AML, SKNM and TFG wrote the manuscript. All authors contributed to the editing of the paper.

■ CONFLICT OF INTEREST

The authors declare no competing interests.

CORRELAÇÃO ENTRE AS CARACTERÍSTICAS MOLECULARES E OS SUBTIPOS GENÉTICOS DOS GLIOBLASTOMAS: ANÁLISE CRÍTICA DE 109 CASOS

OBJETIVO: O glioblastoma (GBM), o tumor cerebral mais comum e mais letal, é também um dos tipos de tumores de mais difícil tratamento. Análises genômicas integradas têm contribuído para um melhor entendimento

da arquitetura molecular dos GBMs, revelando uma nova subclassificação com a promessa de precisão no tratamento de pacientes com alterações específicas. Neste estudo, nós apresentamos a classificação de uma casuística brasileira de GBMs dentro dos principais subtipos do tumor.

MÉTODOS: Usando sequenciamento de DNA em larga escala, foi possível classificar os tumores em proneural, clássico e mesenquimal. Em seguida, testamos o possível uso da hiperexpressão de EGFR e CHI3L1 para a identificação dos subtipos clássico e mesenquimal, respectivamente.

RESULTADOS: Nossos resultados deixam claro que a identificação genética dos subtipos moleculares de GBM não é possível utilizando-se apenas um único tipo de mutação, em particular nos casos de GBMs mesenquimais. Da mesma forma, não é possível distinguir os casos mesenquimais apenas com a expressão de CHI3L1.

CONCLUSÃO: Nossos dados indicam que o subtipo mesenquimal, o mais maligno dos GBMs, necessita de uma análise mais aprofundada para sua identificação.

PALAVRAS-CHAVE: glioblastoma, classificação, análise de sequência de DNA, CHI3L1, EGFR

■ REFERENCES

- Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol.* 2016;131(6):803–20. DOI:10.1007/s00401-016-1545-1.
- Wen PY, Kesari S. Malignant gliomas in adults. *N Engl J Med.* 2008;359(5):492–507. DOI:10.1056/NEJMra0708126.
- Ohgaki H, Kleihues P. The definition of primary and secondary glioblastoma. *Clin Cancer Res Off J Am Assoc Cancer Res.* 2013;19(4):764–72. DOI:10.1158/1078-0432.CCR-12-3002.
- Cancer Genome Atlas Research Network, Brat DJ, Verhaak RGW, Aldape KD, Yung WKA, Salama SR, et al. Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. *N Engl J Med.* 2015(26);372(26):2481–98. DOI:10.1056/NEJMoa1402121.
- Ceccarelli M, Barthel FP, Malta TM, Sabedot TS, Salama SR, Murray BA, et al. Molecular Profiling Reveals Biologically Discrete Subsets and Pathways of Progression in Diffuse Glioma. *Cell.* 2016;164(3):550–63. DOI:10.1016/j.cell.2015.12.028.
- Verhaak RGW, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, 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. DOI:10.1016/j.ccr.2009.12.020.
- Stieber D, Golebiewska A, Evers L, Lenkiewicz E, Brons NHC, Nicot N, et al. Glioblastomas are composed of genetically divergent clones with distinct tumorigenic potential and variable stem cell-associated phenotypes. *Acta Neuropathol.* 2014;127(2):203–19. DOI:10.1007/s00401-013-1196-4.
- Phillips HS, Kharbanda S, Chen R, Forrester WF, Soriano RH, Wu TD, 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):157–73. DOI:10.1016/j.ccr.2006.02.019.
- Brennan CW, Verhaak RGW, McKenna A, Campos B, Nounshmehr H, Salama SR, et al. The somatic genomic landscape of glioblastoma. *Cell.* 2013;155(2):462–77. DOI:10.1016/j.cell.2013.09.034.
- Parsons DW, Jones S, Zhang X, Lin JC-H, Leary RJ, Angenendt P, et al. An integrated genomic analysis of human glioblastoma multiforme. *Science.* 2008;321(5897):1807–12. DOI:10.1126/science.1164382.
- Louis DN, Ohgaki H, Wiestler OD, Cavenee WK, Burger PC, Jouvet A, et al. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol.* 2007 Aug;114(2):97–109. DOI:10.1007/s00401-007-0243-4
- Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinform Oxf Engl.* 2009;25(14):1754–60. DOI:10.1093/bioinformatics/btp324.
- Tischler G, Leonard S. biobambam: tools for read pair collation based algorithms on BAM files. *Source Code Biol Med.* 2014;9:13. DOI:10.1186/1751-0473-9-13.
- Platypus, a reference genome-free algorithm that rapidly calls variants in clinical sequencing data. *SciBX Sci-Bus Exch.* 2014;7. DOI:10.1038/scibx.2014.936.
- Cingolani P, Patel VM, Coon M, Nguyen T, Land SJ, Ruden DM, et al. Using *Drosophila melanogaster* as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. *Front Genet.* 2012;3:35. DOI:10.3389/fgene.2012.00035.
- Carvalho PO, Uno M, Oba-Shinjo SM, Rosemberg S, Wakamatsu A, da Silva CC, et al. Activation of EGFR signaling from pilocytic astrocytomas to glioblastomas. *Int J Biol Markers.* 2014;29(2):e120-128. DOI:10.5301/IJBM.5000045.
- Le Gallo M, Lozy F, Bell DW. Next-Generation Sequencing. *Adv Exp Med Biol.* 2017;943:119–48. DOI:10.1007/978-3-319-43139-0_5.
- Davis B, Shen Y, Poon CC, Luchman HA, Stechishin OD, Pontifex CS, et al. Comparative genomic and genetic analysis of glioblastoma-derived brain tumor-initiating cells and their parent tumors. *Neuro-Oncol.* 2016;18(3):350–60. DOI:10.1093/neuonc/nov143.
- Natesh K, Bhosale D, Desai A, Chandrika G, Pujari R, Jagtap J, et al. Oncostatin-M differentially regulates mesenchymal and proneural signature genes in gliomas via STAT3 signaling. *Neoplasia.* 2015;17(2):225–37. DOI:10.1016/j.neo.2015.01.001.
- Chen Y, Xu R. Drug repurposing for glioblastoma based on molecular subtypes. *J Biomed Inform.* 2016;64:131–8. DOI:10.1016/j.jbi.2016.09.019.
- Padfield E, Ellis HP, Kurian KM. Current Therapeutic Advances Targeting EGFR and EGFRvIII in Glioblastoma. *Front Oncol.* 2015;5:5. DOI:10.3389/fonc.2015.00005.
- Grossmann P, Gutman DA, Dunn WD, Holder CA, Aerts HJWL. Imaging-genomics reveals driving pathways of MRI derived volumetric tumor phenotype features in Glioblastoma. *BMC Cancer.* 2016;16:611. DOI:10.1186/s12885-016-2659-5.
- Kato Y. Specific monoclonal antibodies against IDH1/2 mutations as diagnostic tools for gliomas. *Brain Tumor Pathol.* 2015;32(1):3–11. DOI:10.1007/s10014-014-0202-4.
- Gupta P, Han S-Y, Holgado-Madruga M, Mitra SS, Li G, Nitta RT, et al. Development of an EGFRvIII specific recombinant antibody. *BMC Biotechnol.* 2010;10:72. DOI:10.1186/1472-6750-10-72.
- Conroy S, Kruyt FAE, Joseph JV, Balasubramanian V, Bhat KP, Wagemakers M, et al. Subclassification of Newly Diagnosed Glioblastomas through an Immunohistochemical Approach. *PLoS ONE* 2014;9(12):e115687. DOI:10.1371/journal.pone.0115687.
- Joseph JV, Conroy S, Pavlov K, Sontakke P, Tomar T, Eggen-Meijer E, et al. Hypoxia enhances migration and invasion in glioblastoma by promoting a mesenchymal shift mediated by the HIF1 α -ZEB1 axis. *Cancer Lett.* 2015;359(1):107–16. DOI:10.1016/j.canlet.2015.01.010.

ANNEX

Table 1 - Patient data

Diagnosis	Case	Sample ID	2016 WHO grading [†]	Gender [†]	Age [‡]	Date of birth	Date of surgery	Overall Survival [§]	CHI3L1	EGFR
GBM	1	35	GBM IDH wt	M	53	25/11/47	24/11/00	45m		
GBM	2	74	GBM IDH wt	M	74	25/10/26	19/3/01	17m	2	4
GBM	3	175	GBM IDH wt	F	56	7/4/45	26/10/01	5m		
GBM	4	194	GBM IDH wt	M	71	29/05/30	30/11/01	22m	1	4
GBM	5	204	GBM IDH wt	F	70	1/1/30	21/12/01	7m		
GBM	6	208	GBM IDH wt	M	62	9/8/39	4/1/02	6m	3	2
GBM	7	256	GBM IDH wt	M	41	25/11/60	21/03/02	2m		
GBM	8	269	GBM IDH wt	F	65	28/02/37	12/4/02	0	3	2
GBM	9	274	GBM IDH wt	F	47	20/12/54	19/04/02	17m		
GBM	10	297	GBM IDH wt	F	78	14/06/23	23/05/02	11m	0	4
GBM	11	317	GBM IDH wt	F	71	22/12/30	24/06/02	2m		
GBM	12	356	GBM IDH wt	M	74	17/06/28	30/08/02	11m	1	0
GBM	13	370	GBM IDH wt	M	45	12/7/57	27/09/02	13m		
GBM	14	384	GBM IDH wt	M	45	27/04/57	25/10/02	14m		
GBM	15	391	GBM IDH wt	F	54	16/03/48	7/11/02	12m		
GBM	16	397	GBM IDH mut	F	58	8/12/43	20/11/02	5m	2	1
GBM	17	405	GBM IDH wt	M	74	12/7/57	10/12/02			
GBM	18	427	GBM IDH wt	F	51	11/11/51	23/01/03	5m		
GBM	19	450	GBM IDH wt	F	61	26/08/41	6/3/03	14m		
GBM	20	458	GBM IDH wt	M	62	23/04/42	16/03/03	3m	3	1
GBM	21	485	GBM IDH wt	M	67	18/02/36	1/5/03	6m	3	0
GBM	22	496	GBM IDH wt	F	57	30/10/45	22/05/03	8m	3	4
GBM	23	498	GBM IDH wt	F	17	6/2/86	23/05/03	4m	1	3
GBM	24	503	GBM IDH wt	M	63	17/07/39	2/6/03	2m		
GBM	25	510	GBM IDH wt	M	56	11/10/46	10/6/03	23m		
GBM	26	522	GBM IDH mut	M	48	20/05/55	27/06/03	24m	3	0
GBM	27	524	GBM IDH wt	F	59	5/11/43	30/06/03	8m	2	4
GBM	28	547	GBM IDH mut	M	71	15/08/31	31/07/03	13m	2	1
GBM	29	555	GBM IDH wt	M	57	20/07/46	14/08/03	18m		
GBM	30	592	GBM IDH wt	F	40	1/10/62	24/09/03		2	0
GBM	31	629	GBM IDH wt	M	64	25/06/38	11/11/03	14m	3	1
GBM	32	632	GBM IDH mut	F	41	15/08/62	13/11/03	28m	3	0
GBM	33	638	GBM IDH wt	M	55	15/08/62	13/11/03	2m		
GBM	34	640	GBM IDH wt	F	58	12/12/44	21/11/03	5m		
GBM	35	642	GBM IDH wt	M	42	14/04/62	21/11/03	17m	2	3
GBM	36	663	GBM IDH wt	M	66	11/12/37	5/12/03	7m		
GBM	37	684	GBM IDH wt	F	56	15/12/47	12/2/04	1m		
GBM	38	687	GBM IDH wt	M	45	7/3/58	16/02/04	9m		
GBM	39	698	GBM IDH wt	M	58	25/09/45	27/02/04	7m	2	2
GBM	40	724	GBM IDH wt	F	62	1/2/42	8/4/04		2	3
GBM	41	743	GBM IDH wt	M	52	9/6/52	6/5/04	2m	1	3
GBM	42	750	GBM IDH wt	M	51	23/09/52	13/05/04	6m		

Continuation Table 1

GBM	43	792	GBM IDH wt	M	35	21/12/69	16/07/04	5m	2	0
GBM	44	795	GBM IDH wt	M	28	15/06/76	22/08/04	11m	3	4
GBM	45	852	GBM IDH wt	M	60	19/08/44	22/10/04	13m	2	0
GBM	46	854	GBM IDH wt	M	46	17/06/58	27/10/04		1	4
GBM	47	875	GBM IDH wt	M	35	21/12/69	23/11/04		1	0
GBM	48	879	GBM IDH wt	M	61	27/07/43	2/12/04	3m	3	1
GBM	49	881	GBM IDH wt	M	49	25/01/55	7/12/04	4m	1	4
GBM	50	884	GBM IDH wt	F	52	15/01/52	10/12/04	27m	0	4
GBM	51	885	GBM IDH wt	F	86	24/04/18	10/12/04	2m		
GBM	52	891	GBM IDH wt	M	57	25/10/47	21/12/04	7m		
GBM	53	901	GBM IDH mut	M	16	7/11/88	5/1/05	2m	0	1
GBM	54	903	GBM IDH wt	M	55	8/2/45	11/1/05	22m	3	1
GBM	55	925	GBM IDH wt	M	40	8/9/64	25/02/05	12m		
GBM	56	930	GBM IDH mut	M	26	1/10/78	3/3/05	31m		
GBM	57	1002	GBM IDH wt	M	40	29/01/65	15/07/05	2m	3	2
GBM	58	1003	GBM IDH wt	F	68	25/03/37	17/07/05	4m	2	3
GBM	59	1007	GBM IDH mut	F	28	18/06/77	22/07/05	3m	1	1
GBM	60	1009	GBM IDH wt	F	38	25/10/66	25/07/05	21m	1	1
GBM	61	1070	GBM IDH wt	M	72	9/3/33	29/11/05	8m		
GBM	62	1074	GBM IDH wt	M	32	25/07/73	3/12/05		3	0
GBM	63	1084	GBM IDH wt	M	54	18/04/51	13/01/06	5m	2	1
GBM	64	1091	GBM IDH wt	M	55	5/9/49	3/2/06	21m		
GBM	65	1103	GBM IDH wt	M	54	18/04/1951	3/3/06		3	2
GBM	66	1118	GBM IDH wt	F	61	10/7/59	18/04/06	2m		
GBM	67	1122	GBM IDH wt	M	68	26/07/37	1/5/06	5m		
GBM	68	1123	GBM IDH wt	M	53	29/10/53	2/5/06	11m		
GBM	69	1124	GBM IDH wt	M	63	19/03/43	5/5/06	4m	1	4
GBM	70	1133	GBM IDH wt	M	52	28/12/53	26/05/06	9m		
GBM	71	1144	GBM IDH wt	M	76	15/08/29	25/06/06	2m		
GBM	72	1161	GBM IDH wt	M	39	13/01/67	26/07/06	13m		
GBM	73	1162	GBM IDH wt	F	68	20/05/38	31/07/06	15m		
GBM	74	1169	GBM IDH wt	F	56	13/03/50	9/8/06	2m		
GBM	75	1190	GBM IDH wt	F	58	16/09/48	24/10/06	9m		
GBM	76	1194	GBM IDH mut	M	26	1/10/78	3/11/06			
GBM	77	1199	GBM IDH mut	M	30	6/11/75	24/11/06	11m		
GBM	78	1205	GBM IDH wt	M	69	13/12/36	5/12/06		3	3
GBM	79	1212	GBM IDH mut	M	31	6/11/75	5/1/07			
GBM	80	1232	GBM IDH wt	M	58	7/4/48	16/02/07	12m		
GBM	81	1237	GBM IDH wt	M	59	18/05/47	27/02/07	7m		
GBM	82	1243	GBM IDH wt	M	47	22/09/59	13/03/07	10m		
GBM	83	1250	GBM IDH wt	M	63	11/1/44	30/03/07	7m		
GBM	84	1252	ND	M	45	12/2/61	4/9/07	11m		
GBM	85	1272	GBM IDH wt	M	70	5/5/37	18/05/07	6m		
GBM	86	1274	GBM IDH wt	M	58	29/05/1948	25/05/07	13m		
GBM	87	1282	GBM IDH wt	M	56	29/01/1951	6/7/07	9m		

Continuation Table 1

GBM	88	1295	GBM IDH wt	F	60		3/8/07	7m
GBM	89	1299	GBM IDH wt	M	77	29/10/1930	17/08/07	4m
GBM	90	1303	GBM IDH wt	M	49	31/08/1957	7/9/07	5m
GBM	91	1310	GBM IDH wt	M	48	14/11/1958	5/10/07	16m
GBM	92	1315	GBM IDH wt	M	68	31/07/1939	19/10/07	11m
GBM	93	1318	GBM IDH wt	M	58	12/11/49	26/10/07	21m
GBM	94	1319	GBM IDH wt	F	76	25/03/1931	27/10/07	3m
GBM	95	1331	GBM IDH wt	M	66	29/08/1941	21/05/08	0
GBM	96	1332	ND	M	46	1/14/62	5/28/08	7m
GBM	97	1335	GBM IDH mut	F	53	6/2/55	6/20/08	23m
GBM	98	1368	GBM IDH mut	F	29	1/21/86	3/14/09	
GBM	99	1380	GBM IDH wt	M	58	18/12/1950	29/04/09	29m
GBM	100	1386	GBM IDH wt	F	66	20/01/1943	20/05/09	2m
GBM	101	1398	GBM IDH wt	M	78	6/7/30	17/06/09	7m
GBM	102	1416	GBM IDH wt	M	32	8/11/77	6/8/09	7m
GBM	103	1417	GBM IDH wt	M	36	29/10/71	7/8/09	4m
GBM	104	1503	GBM IDH wt	F	49	16/12/66	18/02/14	
GBM	105	1507	GBM IDH wt	M	45	2/11/70	23/04/14	
GBM	106	1509	GBM IDH wt	M	71	28/10/42	9/6/14	
GBM	107	1518	GBM IDH wt	M	55	30/01/1959	22/10/14	
GBM	108	1520	GBM IDH wt	F	79	29/10/1935	2/12/14	
GBM	109	1522	GBM IDH wt	F	53	3/5/61	5/12/14	

† M, male. F, female, ‡ Age at diagnosis, years, § m, months, ¶ Classificaion according to the 2016 WHO grading, based on IDH1 mutational status. Wt, wild-type; mut, mutated; ND, non-determined. last update - dec/2014

Table 2 - Classificatory genetic alterations in GBMs

Diagnosis	Case [‡]	GBM subtype	EGFR amplification	EGFR mutation	EGFRvIII	NF1 mutation	RB1 mutation or loss	PDGFRA mutation	IDH1 mutation	TP53 mutation	PTEN mutation
GBM	35	MESENCHYMAL					loss				
GBM	74	CLASSICAL	+								
GBM	175	MESENCHYMAL				c.1630.T					
GBM	194	CLASSICAL	+		+						
GBM	204	OTHERS									
GBM	208	OTHERS									
GBM	256	CLASSICAL	+		+						
GBM	269	CLASSICAL	+		+					c.375G>T	
GBM	274	CLASSICAL	+								c.260_322dup
GBM	297	CLASSICAL	+		+						c.388C>T
GBM	317	OTHERS									
GBM	356	CLASSICAL	+								
GBM	370	OTHERS									
GBM	384	OTHERS									
GBM	391	MESENCHYMAL				c.1882delT					
GBM	397	PRONEURAL							c.395G>A		
GBM	405	OTHERS									

Continuation Table 2

GBM	427	CLASSICAL			+				
GBM	450	CLASSICAL			+				
GBM	458	CLASSICAL			+				
GBM	485	CLASSICAL		c.1874C>T					
GBM	496	CLASSICAL			+				
GBM	498	CLASSICAL			+				
GBM	503	PRONEURAL					c.764_766 delTCA	c.51_53delAGA	
GBM	510	OTHERS							
GBM	522	PRONEURAL				c.395G>A	c.298C>T/ c.742C>T		
GBM	524	CLASSICAL	+					c.287C>T	
GBM	547	PRONEURAL			+	c.395G>A			
GBM	555	OTHERS							
GBM	592	PRONEURAL				c.863A>G			
GBM	629	CLASSICAL			+				
GBM	632	PRONEURAL				c.395G>A			
GBM	638	CLASSICAL			+				
GBM	640	OTHERS							
GBM*	642	CLASSICAL						c.210 +1G>A	
GBM	663	OTHERS							
GBM*	684	CLASSICAL							
GBM	687	CLASSICAL			+				
GBM	698	MESENCHYMAL				IVS45- -1G>C (splice site)		c.271G>T	
GBM	724	CLASSICAL	+						
GBM	743	CLASSICAL	+					c.335T>C	
GBM	750	CLASSICAL	+		+			c.407G>A	
GBM	792	MESENCHYMAL				c.6514_6515delGA	c.742C>T/ c.473G>A		
GBM	795	CLASSICAL	+						
GBM	852	PRONEURAL				c.1087A>T			
GBM	854	CLASSICAL	+		+				
GBM	875	MESENCHYMAL				c.1882delT	c.210_211insA		
GBM	879	MESENCHYMAL				loss			
GBM	881	CLASSICAL	+		+			c.388C>T	
GBM	884	CLASSICAL	+						
GBM	885	OTHERS							
GBM	891	OTHERS					c.524G>A		
GBM	901	PRONEURAL			+		c.395G>A		
GBM	903	MESENCHYMAL				c.12_13del/ c.45_53del			
GBM	925	CLASSICAL			+				
GBM	930	PRONEURAL			+		c.395G>A		
GBM	1002	OTHERS							
GBM	1003	CLASSICAL	+		+				
GBM	1007	PRONEURAL					c.395G>A	c.206delG	

Continuation Table 2

GBM	1009	PRONEURAL						c.733G>A
GBM	1070	MESENCHYMAL				loss		
GBM	1074	MESENCHYMAL			c.60+1G>T	c.2211+1G>T	c.184G>T	c.407G>A
GBM	1084	MESENCHYMAL			c.6852_6855del		c.210_211insA	
GBM	1091	CLASSICAL						c.158_160delTAG
GBM	1103	MESENCHYMAL		+	c.6852_6855del			c.465T>G
GBM	1118	CLASSICAL						c.355G>T
GBM	1122	MESENCHYMAL			c.499_502del			
GBM	1123	CLASSICAL	+	+				
GBM	1124	CLASSICAL						
GBM	1133	CLASSICAL	+	+				
GBM	1144	CLASSICAL		+				
GBM	1161	CLASSICAL	+	+				
GBM	1162	CLASSICAL		+				
GBM	1169	MESENCHYMAL				loss		c.355_356insGG
GBM	1190	MESENCHYMAL			c.7190-11_7190-8del (splice)			
GBM	1194	PRONEURAL					c.395G>A	
GBM	1199	PRONEURAL					c.395G>A	c.658T>C
GBM	1205	OTHERS						
GBM	1212	PRONEURAL	+				c.395G>A	c.658T>C
GBM	1232	OTHERS						
GBM	1237	OTHERS						
GBM	1243	CLASSICAL		+				c.1043T>C
GBM	1250	CLASSICAL	+					
GBM	1252	OTHERS						
GBM	1272	CLASSICAL	+					
GBM	1274	CLASSICAL	+					
GBM	1282	MESENCHYMAL	+		c.3739_3742del			
GBM	1295	MESENCHYMAL			c.323G>A	loss		c.210_211insA
GBM	1299	CLASSICAL	+					
GBM	1303	OTHERS						
GBM	1310	CLASSICAL	+		c.1859G>A			c.464A>G
GBM	1315	PRONEURAL				c.368-3C>T (splice)	c.97-6C>T (splice)	
GBM	1318	CLASSICAL			c.2240_2254del			
GBM	1319	OTHERS						
GBM	1331	CLASSICAL	+		c.685A>T			c.641dupA

Continuation Table 2

GBM	1332	OTHERS					
GBM	1335	PRONEURAL				c.395G>A	
GBM	1368	PRONEURAL				c.395G>A	
GBM	1380	MESENCHYMAL			c.1475 delA		c.822G>A
GBM	1386	CLASSICAL		c.760T>A			c.517C>T
GBM	1398	CLASSICAL	+				
GBM	1416	CLASSICAL	+				
GBM	1417	MESENCHYMAL			c.4742_ 4749del		
GBM	1503	PRONEURAL				loss	
GBM	1507	MESENCHYMAL	+		c.1546C>T	c.451C>T	
GBM	1509	CLASSICAL					c.115G>A
GBM	1518	CLASSICAL	+				
GBM	1520	CLASSICAL		c.1280G >T			
GBM	1522	MESENCHYMAL			c.1924C>T		c.765dupA