UNIVERSITÀ
DEGLI STUDI
DI BRESCIA

DOTTORATO DI RICERCA IN

Scienze Biomediche e Medicina Traslazionale

Settore scientifico disciplinare MED/08

XXXIII ciclo

GLIOBLASTOMA WITH PRIMITIVE NEURONAL COMPONENT: IMMUNOPHENOTYPICAL AND MOLECULAR CHARACTERIZATION AND CONTRIBUTION OF EBF3 TO THE NEURONAL CELL COMMITMENT

DOTTORANDA

Dr.ssa Francesca Pagani

RELATORE

Prof. Pietro Luigi Poliani

COORDINATORE DEL DOTTORATO

Prof. Maurizio Memo

“Non est ad astra mollis

e terris vía”

Lucíus Annaeus Seneca

INDEX

PREFACE

"Glioblastoma with Primitive Neuronal Component" (GBM-PNC) is a rare histological variant of Glioblastoma, consistent of a mixed glial component with nodules of immature cells that display an early neuronal differentiation. Since GBM-PNC is a rare tumour, only few data regarding their characterization and their molecular alterations are available. We proposed to collect a consistent number of samples, in order to extensively characterise this rare Glioblastoma variant from a phenotypical and molecular point of view. In detail, the main focus was to find common or distinctive features for each of the two components, in order to shed a light on the origin of these tumours and on the evolution of the two components, whether common or subsequent. Moreover, we searched for a putative molecular alteration that could predict the development of GBM-PNC and distinguish them from the other GBM subtypes.

We then collected a cohort of 24 patients and deeply analysed these tumours, as detailed in the Part I of the present thesis work.

EBF3 is a transcription factor that, in the Central Nervous System, plays a key role in the process of neurogenesis and neuronal migration and in the differentiation towards the neuronal cell fate. Moreover, *EBF3* gene is frequently deleted or methylated in both primary and secondary glioblastomas, as well as in other tumour types, where epigenetic alterations are linked to tumour progression and metastatic ability, suggesting a potential role as tumour suppressor gene in the brain. In a previous paper, we demonstrated that in medulloblastoma, a malignant embryonal brain tumour, *EBF3* is highly expressed and acts as a major master regulator of neuronal differentiation. However, its sustained expression correlates with an immature phenotype, confirmed by lack of expression of mature markers of neuronal differentiation, and contributes to promote neoplastic progression, acting, according to specific cell contest, as an oncogene. For these reasons, we investigated EBF3 expression in GBN-PNC and, notably, we found that EBF3 was selectively expressed in the PNC and constantly negative in the GBM component.

We hypothesized that *EBF3* could have a major role in driving the divergent differentiation of the two components. In order to elucidate its function, we created a CRISPR/Cas9 EBF3 knockout model using a GBM stem cell line expressing EBF3. Concurrently we forced an EBF3 negative GBM stem cell line to overexpress *EBF3* by lentiviral transduction. These data are reported in Part II of the present thesis work.

ABSTRACT

PART I GLIOBLASTOMA WITH PRIMITIVE NEURONAL COMPONENT: IMMUNOPHENOTYPICAL AND MOLECULAR CHARACTERIZATION

"Glioblastoma with Primitive Neuronal Component" (GBM-PNC) is a rare histological variant of Glioblastoma, consistent of a mixed glial component with nodules of immature cells that display an early neuronal differentiation. We collected a cohort of samples (n=24) in order to deeply characterize these tumours. Patients with GBM-PNC have a lower median age at diagnosis (59.5 y/o) and a peculiar topographical distribution, with the temporal and frontal lobes being the most affected brain regions (75% of the cases). We performed immunohistochemical and molecular analysis, comprising Next Generation Sequencing with a GBM-customized panel and Genome Wide Methylation Analysis, on our sample cohort. Our findings confirm the dual phenotype of these tumours, since the GBM components show expression of glial-associated markers (GFAP, YAP1, CD44, Vimentin, EGFR) while the PNC components show immunoreactivity for neuronal/embryonal markers, such as Synaptophysin, NeuN. Noteworthy, c-Myc and n-Myc, a typical feature of the primary CNS-PNET, are basically expressed only in the PNC component. Both components show positivity for early stem cell markers, such as SOX1 and SOX2, and for markers linked to neurogenesis, such as DCX and βIIITubulin. Interestingly, we found a novel biomarker selectively expressed in the PNC and constantly negative in the GBM component: EBF3. EBF3 is a transcription factor usually deleted or methylated in both primary and secondary glioblastomas, but highly expressed in medulloblastoma, a malignant embryonal brain tumour, in which EBF3 acts as a major master regulator of neuronal differentiation. However, its sustained expression correlates with an immature phenotype. The functional role of this transcription factor will be investigated in Part II of the thesis.

Data show that our cohort is enriched in *IDH1* mutant tumours (16.7%) and the most represented transcriptional subtypes are the Classical and the Proneural subtypes, globally reaching the 66.7%, for the GBM component. Of note, the majority of the PNC components could not be classified (66.7%), supporting the evidence that this component is mainly constituted by cells with an undifferentiated phenotype. Notably, the presence or absence of molecular alterations, such as IDH1-R123H mutation, loss of ATRX and overexpression of p53, are common features in both components. As a matter of fact, methylation-based classification classified the two components from the same tumour under the same group: one as IDH-mutant GBM, the others as RTK I/II GBMs, classes that correspond to the Classical and Proneural transcriptional subgroups. Copy Number

Variation analysis showed for the GBM component the typical GBM profile with gain on chromosome 7 and loss on chromosome 10. Interestingly, the PNC component shows a higher degree of loss and gain of genetic material. Data are in line with the observation that the PNC component does not activate TERT and/or ALT as telomere maintenance mechanism, a feature that may contribute to tumour instability. NGS analysis with a custom panel designed to explore 75 genes that are relevant in GBM and PNC tumours revealed that all the samples harbour high frequency mutations shared between the two components on frequently GBM altered genes, such as *TP53*, *PTEN*, *IDH1*, *TERT*, *PDGFRA*, *MDM2/4*. Interestingly, all samples harbour private molecular alterations specific for each component, suggesting that GBM-PNC tumours may arise from a common ancestor characterized by a set of mutations/amplifications, while other genomic lesions are acquired after the emergence of either GBM and PNC components.

Interestingly, two out of six tumours analyzed by NGS show a high number of subclonal mutations such that we can hypothesize a hypermutator phenotype; mutations are more abundant in the PNC component, in line with the previous observation that this component is genetically unstable. According to the evidence that the *EBF3* gene is frequently deleted or methylated in glioblastoma, we found that in 5 samples out of 6 a copy of the *EBF3* gene is lost; however, albeit with an allelic loss, EBF3 is expressed in the PNC component, suggesting a specific transcriptional mechanism enabling EBF3 sustained expression.

Altogether, data indicate that GBM-PNCs take their origin from common immature progenitor cells. These cells are endowed with an intrinsic differentiation ability, as happens for neuroectodermic (NEC) progenitors in normal brain development, and may undergo a subsequent differentiation by the accumulation of additional molecular alterations, giving rise to the two components.

A possible genetic hallmark of the GBM-PNC tumours may be represented by mutations in the *RB1* gene. In fact, *RB1* gene is mutated or deleted in 5 tumours out of the 6 analyzed tumours (83.3%), while the only *RB1* wildtype tumour harbours a *CDK4* gene amplification that, from a functional point of view, may mimic the *RB1* loss. Mutations of *RB1* gene are related with the development of CNS-PNETs in children; these tumours display similar features of the PNC component. It has to be clarified how RB1 mutation and regulation of EBF3 expression may interact with each other within the GBM-PNC context.

PART II CONTRIBUTION OF EBF3 TO THE NEURONAL CELL COMMITMENT

In the first part of this work, we characterized a cohort of Glioblastomas with Primitive Neuronal Component. Immunohistochemistry analysis showed a glial profile for the GBM component, as expected, while the PNC component, not surprisingly, was usually negative for GBM typical markers. Notably, we found a novel biomarker selectively expressed in the PNC and constantly negative in the GBM component: EBF3. *EBF3* is a transcriptional factor belonging to a highly conserved four gene family, with an important role in neurogenesis and neuronal migration. EBF3 is frequently deleted or methylated in both primary and secondary glioblastomas, while in medulloblastoma it acts as a major master regulator of neuronal differentiation, even though its sustained expression correlates with an immature phenotype. Since EBF3 is highly expressed in the PNC component of the GBM-PNC, and is persistently negative in the GBM component, we hypothesized that *EBF3* could have a major role in driving the divergent differentiation of the two components. In order to elucidate its function, we developed an *in vitro* model using different Glioblastoma Stem like Cells (GSC). At first, we created a lentiviral vector targeting *EBF3* for CRISPR/Cas9 technology based genome editing. We thus selected a GSC line derived from a GBM-PNC (BT483) that constitutively expressed high levels of EBF3 and created 4 clonal *EBF3* knockout GBM-PNC GSCs lines. EBF3 overexpressing cells were obtained by transducing EBF3-negative L0512 GSCs with a lentiviral vector coding for mouse *ebf3* full-length cDNA.

EBF3-knockout clones showed a lower kinetic of expansion, as compared to the wildtype clones, while no differences were observed between mock and overexpressing clones. The same difference in self-renewal ability was observed when serial clonogenic assays were performed. We then investigated the expression of differentiation markers in transduced GSCs both under proliferative and differentiative conditions. Glial fibrillary acidic protein (GFAP), an astrocyte biomarker expressed by the glial component of GBM-PNC, appears to be expressed at higher levels in EBF3-knockout clones, at both protein and transcriptional level. Contrariwise, EBF3 overexpressing cells express a lower level of GFAP protein as compared to their mock control. Nestin is generally a recognized marker of undifferentiated CNS cells; it is more expressed in the GBM component of GBM-PNC, with a much lower expression in the PNC. In line with this observation, Nestin was expressed at higher level in EBF3-negative cell lines and at lower level in EBF3 expressing cell lines, as assessed by Western blot, RT-qPCR and IHC on cell blocks.

We then induced to terminally differentiate GSC cell lines through removal of mitogens and exposure to fetal bovine serum, in order to assess whether EBF3 could influence GSCs differentiation

ability. The differentiation process occurred efficiently, as demonstrated by the decrease of SOX2, a biomarker indicating an immature profile.

GFAP, reaches overall higher levels in all the clones, as compared to *TUBB3*, the gene coding for βIII-tubulin protein. Of note, EBF3 wildtype clones, constitutively expressing EBF3, and EBF3 overexpressing cells, whose expression is ectopically induced by viral transduction, showed lower fold increase of both *GFAP* and *TUBB3*, as compared to their matched GSC line. This observation suggests that EBF3 expression, both constitutive or induced, could lead to an impaired ability of differentiation.

Preliminary data from *in vivo* xenografts obtained with EBF3 expressing GBM-PNC derived GSCs show the growth of a biphasic tumour resembling the GBM-PNC, with neoplastic aggregates respectively positive or negative for GFAP and EBF3. Data suggest that GBM-PNC GSCs are endowed with the capacity to modulate EBF3 expression. On the contrary, xenografts from the corresponding EBF3-KO clones show an EBF3 negative tumour with a prominent glial aspect and a more widely distributed GFAP positivity.

RIASSUNTO

PARTE I GLIOBLASTOMA WITH PRIMITIVE NEURONAL COMPONENT: IMMUNOPHENOTYPICAL AND MOLECULAR CHARACTERIZATION

Il "Glioblastoma con componente primitiva neuronale" (GBM-PNC) è una variante istologica rara di Glioblastoma, e si caratterizza per avere una componente di natura gliale ed una componente primitiva neuronale immatura. Abbiamo raccolto una coorte di 24 campioni allo scopo di studiare in modo approfondito questi tumori. I pazienti affetti da GBM-PNC hanno un'età alla diagnosi più bassa rispetto ai GBM convenzionali (59,5); il tumore mostra un sito preferenziale di insorgenza a livello dell'area temporo-frontale (75% dei casi). L'analisi immunoistochimica ha confermato il fenotipo divergente di questi tumori, con la componente GBM positiva per i marcatori associati alla glia (GFAP, YAP1, CD44, Vimentin, EGFR), e la componente PNC positiva per i marcatori neuronali/embrionali, quali Sinaptofisina e NeuN. Degno di nota il fatto che c-Myc e n-Myc, marcatori tipicamente espressi dai tumori primitivi neuronali del Sistema Nervoso Centrale, sono espressi solo nella componente PNC. Ambedue le componenti mostrano immunoreattività per marcatori staminali precoci, quali SOX1 e SOX2, e per marcatori associati alla neurogenesi, quali DCX e βIIITubulina. Notevole l'identificazione di un nuovo marcatore espresso selettivamente dalla componente PNC e costantemente negativo nella GBM: EBF3. EBF3 è un fattore di trascrizione solitamente deleto o metilato nei GBM sia primari che secondari, ma espresso ad alti livelli nel medulloblastoma, un tumore cerebrale embrionale maligno in cui EBF3 ha la funzione di guidare il differenziamento neuronale. Tuttavia, la sua prolungata espressione correla con un fenotipo immaturo. Il ruolo funzionale di questo fattore di trascrizione verrà indagato approfonditamente nella Parte II del presente lavoro di tesi.

I dati mostrano che la nostra coorte è arricchita per tumori mutati per il gene *IDH1* (16.7%) ed i gruppi trascrizionali più rappresentati nella componente GBM sono il Classico ed il Proneurale, che raggiungono globalmente il 66.7%. Da notare che la maggior parte delle componenti PNC non vengono classificate in nessun gruppo trascrizionale, supportando l'evidenza che questa componente sia costituita da cellule prevalentemente indifferenziate. Interessante il fatto che la presenza di eventuali alterazioni molecolari, come la mutazione IDH1-R123H, la perdita di ATRX e l'overespressione di p53, è condivisa da entrambe le componenti. In linea con questa osservazione, la classificazione basata sul profilo globale di metilazione di quattro campioni della nostra coorte ha classificato le due componenti dello stesso tumore nella stessa classe: un campione nella categoria *IDH-mutant* GBM, gli altri nella categoria RTK I/II GBMs, classe che corrisponde ai gruppi

trascrizionali Classico e Proneurale. L'analisi di *Copy Number Variation* ha mostrato per la componente GBM il tipico profilo associato ai glioblastomi, con *gain* sul cromosoma 7 e *loss* sul cromosoma 10. Tuttavia, a parità di profilo, la componente PNC mostra un grado maggiore di *gain/loss* di materiale genetico, dato in linea con l'osservazione che questa componente non attiva né TERT, né ALT come meccanismo di mantenimento dei telomeri, a supporto della maggiore instabilità genetica del tumore. L'analisi in *Next Generation Sequencing* con un pannello *custom* comprendente 75 geni rilevanti in GBM e PNC ha rivelato che tutti i campioni portano mutazioni ad alta frequenza condivise dalle due componenti su geni frequentemente alterati nei GBM, quali *TP53*, *PTEN*, *IDH1*, *TERT*, *PDGFRA*, *MDM2/4*. In aggiunta, tutti i campioni portano alterazioni molecolari private in ciascuna componente, suggerendo che questi tumori derivino da un precursore comune caratterizzato da un *set* di amplificazioni/mutazioni, mentre altre alterazioni genetiche sarebbero acquisite a seguito dell'emergere dei due subcloni riferibili alla componente GBM ed alla PNC. Due dei sei campioni analizzati in NGS hanno rivelato un elevato numero di mutazioni subclonali, tanto da far ipotizzare uno stato di ipermutazione; le mutazioni sono più numerose nella componente PNC, in linea con quanto precedentemente osservato in merito alla maggior instabilità di questa componente. In accordo con quanto frequentemente avviene nei GBM, una copia del gene *EBF3* è persa in 5 campioni su 6; ciononostante, EBF3 viene espresso dalla componente PNC, suggerendo un possibile meccanismo di regolazione trascrizionale.

Nel complesso, i dati sembrano indicare che i GBM-PNC derivino da un progenitore comune immaturo; queste cellule sono dotate di una capacità intrinseca di differenziamento, come accade per i progenitori neuroectodermici durante lo sviluppo fisiologico del cervello, e vadano poi incontro ad un successivo differenziamento divergente per accumulo di alterazioni molecolari aggiuntive, portando così allo sviluppo delle due componenti. Un possibile tratto peculiare di questi tumori potrebbe essere la presenza di mutazioni a carico del gene *RB1*. Infatti, questo gene è risultato essere mutato in 5 tumori sui 6 analizzati (83.3%), mentre il solo tumore *wildtype* porta un'amplificazione del gene *CDK4*, alterazione che, dal punto di vista funzionale, mima la perdita di *RB1*. Mutazioni del gene *RB1* sono frequentemente associate allo sviluppo di CNS-PNET in età pediatrica; questi tumori mostrano caratteristiche simili alla componente PNC dei GBM-PNC. Resta da chiarire come mutazioni di *RB1* e regolazione dell'espressione di EBF3 possano interagire all'interno del contesto tumorale dei GBM-PNC.

PART II CONTRIBUTION OF EBF3 TO THE NEURONAL CELL COMMITMENT

Nella prima parte del presente lavoro di tesi, abbiamo caratterizzato una coorte di Glioblastomi con Componente Primitiva Neuronale. L'analisi immunoistochimica ha rivelato un profilo gliale per la componente GBM, mentre la componente PNC, come atteso, è risultata tendenzialmente negativa per i marcatori gliali. Abbiamo però identificato un nuovo marcatore specifico per la componente PNC: EBF3. *EBF3* è un fattore di trascrizione con un importante ruolo nel processo di neurogenesi e migrazione neuronale. EBF3 è frequentemente deleto o metilato nei glioblastomi sia primari che secondari, ma espresso ad alti livelli nel medulloblastoma, un tumore cerebrale embrionale maligno in cui EBF3 ha la funzione di guidare il differenziamento neuronale. Tuttavia, la sua prolungata espressione correla con un fenotipo immaturo. Dal momento che EBF3 è espresso ad alti livelli nella componente PNC dei GBM-PNC, mentre è costantemente negativo nel GBM, abbiamo ipotizzato che possa avere un ruolo fondamentale nel guidare il differenziamento divergente delle due componenti. Allo scopo di indagare la sua funzione, abbiamo creato un modello *in vitro* utilizzando diverse linee di *Glioblastoma Stem like Cells* (GSC). Innanzitutto, abbiamo creato un vettore lentivirale codificante per il *genome editing* di EBF3 basato sulla tecnologia CRISPR/Cas9. Abbiamo quindi selezionato una linea di GSC derivate da GBM-PNC esprimenti costitutivamente alti livelli di EBF3 e creato 4 cloni *knockout*. Una linea over-esprimente EBF3 è stata poi creata mediante trasduzione di una linea di GSC EBF3-negativa con un vettore lentivirale codificante per il cDNA di *ebf3* murino.

I cloni *knockout* per EBF3 hanno mostrato una minore cinetica di espansione, paragonati ai cloni *wildtype*, mentre non ci sono differenze tra il clone over-esprimente ed il relativo controllo. La stessa differenza nella capacità di *self-renewal* è stata osservata nel saggio clonogenico. Abbiamo quindi indagato l'espressione di marcatori di differenziamento nelle GSC trasdotte sia in condizioni di coltura proliferative, che differenziative. La GFAP, un marcatore astrocitico espresso dalla componente gliale dei GBM-PNC, è risultato essere espresso ad alti livelli nei cloni *knockout*, sia a livello proteico che trascrizionale. Al contrario, le cellule over-esprimenti mostrano livelli più bassi di GFAP, se confrontate con il relativo controllo. La Nestina è un *marker* riconosciuto di cellule indifferenziate del Sistema Nervoso Centrale ed è più espresso nella componente GBM dei GBM-PNC. In linea con questa osservazione, la Nestina è risultata essere espressa a livelli più elevati nelle linee di GSCs negative per EBF3 ed a livelli più bassi nelle linee esprimenti EBF3 sia in *Western blot*, sia in RT-qPCR, sia in immunoistochimica su *cell blocks*.

Abbiamo quindi indotto le GSC a differenziare, rimuovendo i mitogeni ed aggiungendo siero fetale bovino al terreno di coltura, al fine di stabilire se EBF3 possa influenzare la capacità di differenziamento delle staminali.

GFAP, la cui espressione raggiunge livelli globalmente più elevati in tutte le linee, se paragonati a *TUBB3* (βIII-tubulina). I cloni *wildtype*, esprimenti costitutivamente EBF3, e le GSC over-esprimenti EBF3 ectopicamente, nostrano un ridotto *fold increase* sia di *GFAP* che di *TUBB3*, in confronto ai relativi controlli. Questa osservazione suggerisce che l'espressione sostenuta di EBF3 possa portare ad un blocco della capacità di differenziamento cellulare.

I dati preliminari dagli *xenografts* ottenuti dalle GSC derivate da GBM-PNC ed esprimenti EBF3 mostrano un tumore bi-fasico simile morfologicamente ad un GBM-PNC, con aggregati neoplastici rispettivamente positivi o negativi per GFAP ed EBF3. I dati suggeriscono che le GSC da GBM-PNC possano essere dotate di una capacità intrinseca di modulare l'espressione di EBF3. Al contrario, gli *xenografts* ottenuti dalle rispettive linee *knockout* mostrano tumori negativi per EBF3, con un prominente aspetto gliale ed una positività per GFAP più diffusa.

PART I

GLIOBLASTOMA WITH PRIMITIVE NEURONAL COMPONENT: IMMUNOPHENOTYPICAL AND MOLECULAR CHARACTERIZATION

INTRODUCTION

Glioblastoma

Glioblastoma (GBM), the most aggressive and common type of brain tumour, is relatively rare, with an incidence of 7.2 adults per 100,000 every year. Prognosis is poor, with a median overall survival (OS) of only 12-15 months (Dolecek, Propp et al. 2012), despite the current combined chemo- and radiotherapy treatment. The GBM histopathological hallmarks are the extreme invasiveness, anaplasia, nuclear pleomorphism, high mitotic activity, florid neoangiogenesis and necrosis (Stiver, Tan et al. 2004). In the recent WHO classification (Louis, Perry et al. 2016) glioblastomas are classified into: (I) IDH-wildtype, usually *de-novo* tumours in patients over 55 years; (II) IDH-mutant, usually secondary glioblastomas preferentially occurring in younger patients; (III) glioblastoma NOS, when the evaluation of IDH status is not possible (Figure 1). Within the IDH-wildtype glioblastoma group, different histopathological variants have been identified, including the giant cell glioblastoma and gliosarcoma and the newly recognized histotype variant, the epithelioid glioblastoma. In addition to these officially recognized histopathological variants, different GBM patterns have been recognized, as small cell glioblastoma, granular cell glioblastoma and the glioblastoma with primitive neuronal component.

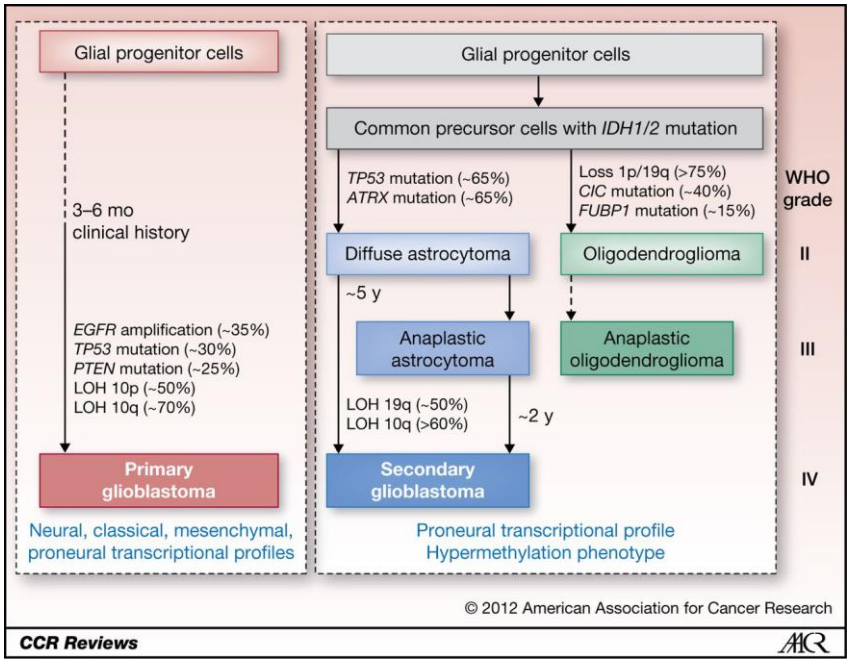


Figure 1. Genetic pathways to primary and secondary glioblastomas(Ohgaki and Kleihues 2013)

The standard therapeutic protocol for GBM is a multimodal approach that combines surgery, radiotherapy, and adjuvant chemotherapy with temozolomide (TMZ), a DNA alkylating agent (Stupp, Hegi et al. 2009). Even though it led to a significant survival improvement, it remains poorly effective. The locally infiltrative nature of the tumour often prevents complete surgical resection, radiations have devastating side-effects on the nervous system restrain, and the blood-brain barrier hampers the delivery of chemotherapeutics (Huse and E.C. 2010, Westphal and Lamszus 2011). Moreover, despite extensive efforts, this tumour is still devoid of targeted therapeutic options. Recent clinical trials with the anti-angiogenic agent bevacizumab (showing possible benefits in preliminary trials), or with inhibitors of tyrosine kinase receptors or signal transducers, failed to significantly prolong patients survival (Quant and Wen 2010, Tanaka, Louis et al. 2013, Thomas, Brennan et al. 2014). In order to develop more effective therapies, in the last decade, a broad investigation of genetic, molecular and cellular features of this tumour has been pursued.

Subclassification of glioblastoma based on genetic alterations and gene expression profiles into subtypes associated with distinct clinical features

As previously described, Glioblastoma comprises a large variety of morphological patterns and histological variants, some of which included in the most recent 2016 WHO classification of tumours of the central nervous system (Louis, Perry et al. 2016). The extensive intertumoural and intratumoural histological heterogeneity reflects an intrinsic genetic instability, to earn the attribute of “multiforme”. The advent of novel and powerful genomic technologies eventually provided the opportunity to recognize glioblastoma heterogeneity at the molecular level and identify subtypes associated with genetic features exploitable for a more accurate diagnosis and prognosis, and for the identification of prospective therapeutic targets (Huse, Phillips et al. 2011, Vitucci, Hayes et al. 2011).

In 2008, the first publication for GBM by The Cancer Genome Atlas Research Network (TCGA) reported the results of genomic and transcriptomic analysis of 206 GBMs, which included mutation sequencing of 600 genes in 91 of the sample (Network 2008). The publication mainly focused on biologically relevant alterations in three core pathways (Figure 2):

I. The Receptor Tyrosine Kinase-Ras-MAP Kinase/AKT pathway, controlling proliferation and survival, including EGF Receptor (*EGFR*, focally amplified and/or mutated in 57% cases), the Phosphatidylinositol-3 kinase inhibitor PTEN (homozygously inactivated in 41% of cases), and the Ras inhibitor NF1 (homozygously inactivated in 10% of cases);

II. the p53 pathway, controlling apoptosis and senescence, with *TP53* loss-of-function mutations observed in 28% of cases;

III. the RB pathway, controlling cell cycle progression, including the cell cycle inhibitors *CDKN2A* (p16/INK4A) and *CDKN2B*, which are alternatively homozygously deleted in about 60% of cases.

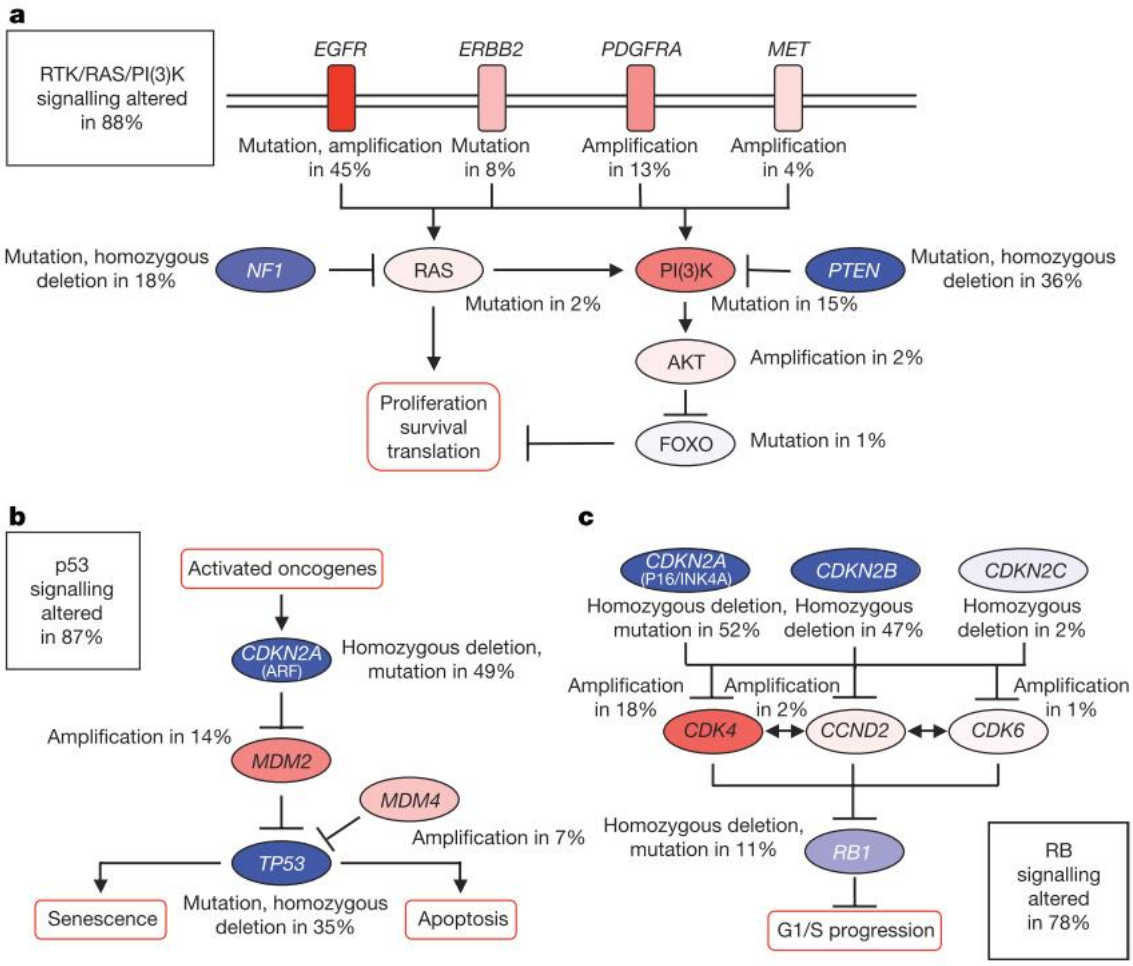


Figure 2 Frequent GBMs genetic alterations in three critical signalling pathways
Primary sequence alterations and significant copy number changes for components of the RTK/RAS/PI(3)K (a), p53 (b) and RB (c) signalling pathways. Red indicates activating genetic alterations, while blue indicates inactivating alterations (Adapted from The Cancer Genome Atlas Research Network 2008).

In 2013, the second publication for GBM by TCGA illustrated the landscape of somatic genomic alterations of more than 500 GBMs, especially emphasizing novel mutation and complex gene rearrangements in *EGFR*, *PDGFRA* and other signature receptors (Brennan, Verhaak et al. 2013). It was found that *TERT* promoter mutations correlated well with the increase of its mRNA expression in GBM. Overall, at least one gene alteration in each of the above three pathways is present in about 75% of patients (Huse and E.C. 2010).

It is currently accepted that GBMs can be classified according to their transcriptional profile in three major distinct molecular subgroups (Proneural, PN; Classical, CL; Mesenchymal, MES) associated with specific gene signatures and different prognostic and therapeutic implications (Phillips, Kharbanda et al. 2006, Verhaak, Hoadley et al. 2010, Wang, Hu et al. 2018). Expression of genes related to neural (e.g. *ASCL1*, *OLIG2*) and glioma-CpG island methylator phenotype (Noushmehr, Weisenberger et al. 2010) defines the PN subtype, frequently associated to *TP53* mutation. Interestingly, PN subgroup comprises *IDH1/2* mutated GBMs, associated with significantly better prognosis and younger age (Yan, Parsons et al. 2009). In contrast, CL and MES gene signatures are both associated with poor outcome and older age. CL subtype is characterized by high frequency of *EGFR* gene alterations (typically amplification and/or *EGFRvIII* mutation). *EGFR* amplification/overexpression is correlated with poor outcome (Verhaak, Hoadley et al. 2010) but this patients might strongly benefit from metronomic TMZ administration, with additional advantage if carrying *PTEN* loss (Cominelli, Grisanti et al. 2015). Homozygous deletion of 9p21.3 (*CDKN2A* locus), lack of *TP53* mutations and RB pathways alteration are also frequent (Verhaak, Hoadley et al. 2010). MES subtype is characterized by high expression of YKL40, MET and CD44, frequent deletion at 17q11.2 (containing NF1 gene), chromosomal aberrations in *CDK6*, *CDKN2A* and *RB1* and activation of the NF-kB pathway (Wang, Hu et al. 2018). Interestingly, recurrences are frequently associated with a phenotypic shift into MES subtype, suggesting that glioma progression may proceed from PN or CL into MES phenotype(Wood, Reis et al. 2016).

Recently, our group proposed an integrated molecular and immunohistochemical approach aimed at identifying GBM subtypes in routine paraffin-embedded material (Orzan, Pagani et al. 2020). The scoring system has been incorporated within the transcriptional status prediction algorithm (available at http://fisher.med.unibs.it:3838/GBMscore/) to make it easier to be applied in daily pathology practice (Figure 3). Nevertheless, the attempt to attribute a subgroup to every tumour highlighted the huge heterogeneous phenotype of some GBMs, particularly within the PN phenotype, that may not be assigned to a unique profile. This could be an impairment for an accurate diagnosis which is crucial for optimal management of patients. Researchers tried to overcome this limitation by using a different approach based on DNA methylation, whose profile has been shown to be highly robust and reproducible, even for poor quality samples (Hovestadt, Remke et al. 2013).

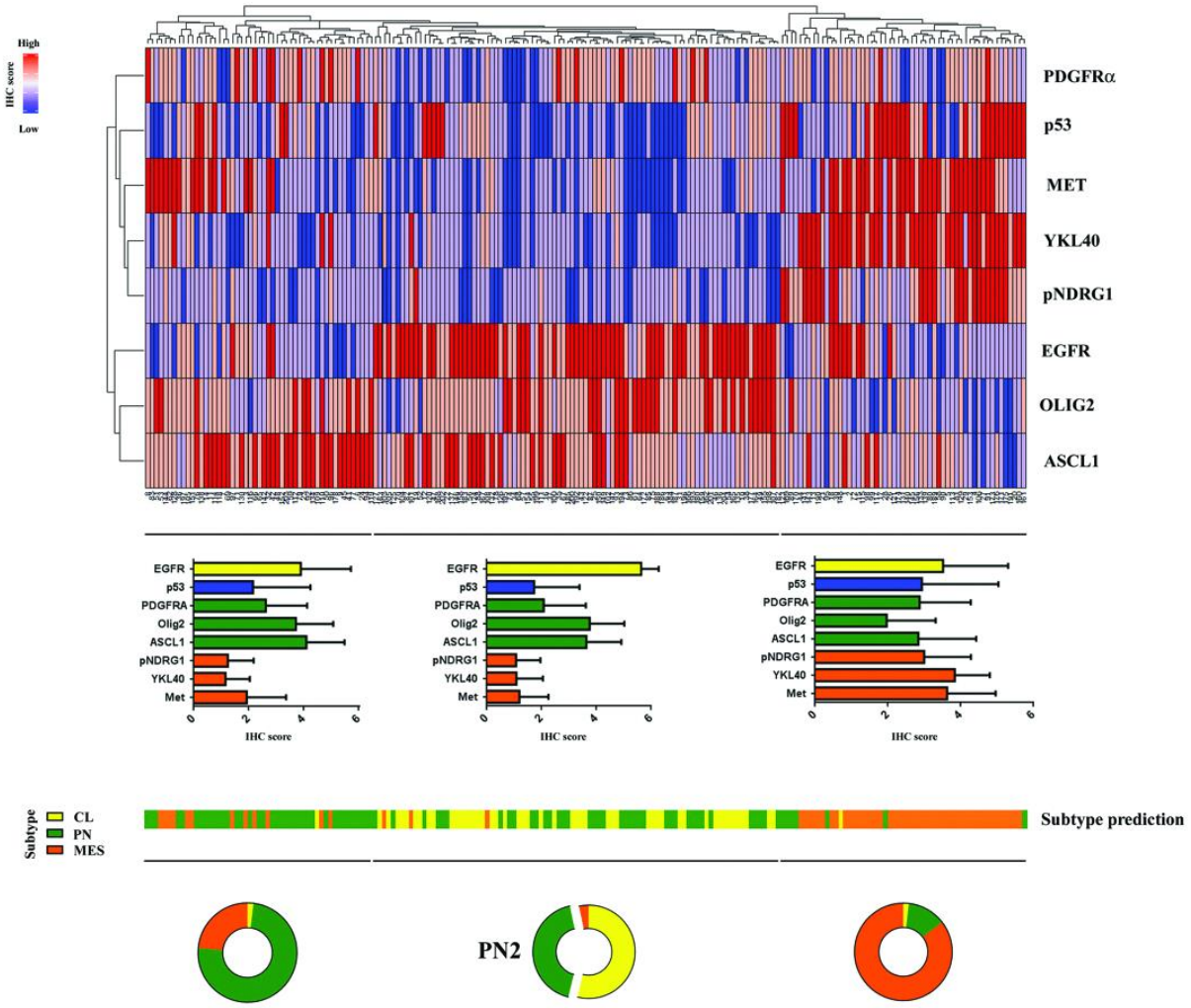


Figure 3. Transcriptional status prediction based on immunohistochemical profile
Dendrogram from cluster analysis based on combination of biomarkers expression shows distribution of GBMs in three major branches and their relative expression profiles within single group (plots). The lower string shows transcriptional subtypes assigned by the algorithm (available at http://fisher.med.unibs.it:3838/GBMscore/) based on combination of IHC and transcriptional data (Orzan, Pagani et al. 2020).

DNA methylation is one of the most studied epigenetic mechanisms (Kulis and Esteller 2010). It is a heritable, enzyme-induced modification to DNA structure without alteration of the specific sequence of the genome and, together with histone modifications, it can regulate the functioning of the genome by changing chromatin architecture. The molecular mechanism involves the covalent addition of a methyl group in cytosine within large clusters of CpG dinucleotides along the genome called CpG islands. Even though DNA methylation plays an essential role in normal biologic processes, aberrant patterns of methylation are observed in different types of tumours.

Methylation of the promoter regions of several genes, including tumour suppressor genes, results in their inactivation as well as hypomethylation can induce genomic instability and contribute to cell transformation. Hypermethylated promoters may serve as biomarkers and, since DNA methylation is reversible, it is extremely interesting for therapy approaches. As a matter of fact, in glioblastomas,

the epigenetic silencing of the *MGMT* (O6-methylguanine-DNA methyltransferase) DNA-repair gene by promoter methylation compromises DNA repair and has been associated with longer survival in patients with glioblastoma who receive the alkylating agent temozolomide (Hegi, Diserens et al. 2005). In 2012 Sturm and colleagues used an integrative analysis of the methylation profile and gene expression data and demonstrated that it was possible to subclassifiy GBMs into 6 groups, mostly indistinguishable by their histological appearance, but correlating with molecular-genetic alterations as well as key clinical variables such as patient age and tumour location (Sturm, Witt et al. 2012). As schematised in Figure 4, the methylation class "*IDH glioma, subclass high grade astrocytoma*" is mainly comprised of *IDH* mutant glioblastoma and anaplastic astrocytoma. This class universally harbours mutations of either *IDH1* or *IDH2* and the associated glioma CpG island methylator (CIMP) phenotype. The methylation class "*glioblastoma, IDH wildtype, H3.3 K27 mutant*" comprises *IDH* wildtype juvenile tumours that harbour mutations of codon 27 of the H3.3 gene (*H3F3A*). The methylation class "*glioblastoma, IDH wildtype, H3.3 G34 mutant*" comprises *IDH* wildtype juvenile tumours that harbour mutations of codon 34 of the H3.3 gene (*H3F3A*). The methylation class "*glioblastoma, IDH wildtype, subclass RTK I*" is composed of *IDH* wildtype GBMs whose expression profiles often resemble the TCGA Proneural subgroup. Recurrent chromosomal alterations are gain of chromosome 7 with or without *EGFR* amplification (>80%), loss of 9p21 (*CDKN2A/B*; >50%) and chromosome 10 loss (>70%). The methylation class "*glioblastoma, IDH wildtype, subclass RTK II*" comprised the *IDH* wildtype GBMs that correspond to the Classical subgroup.

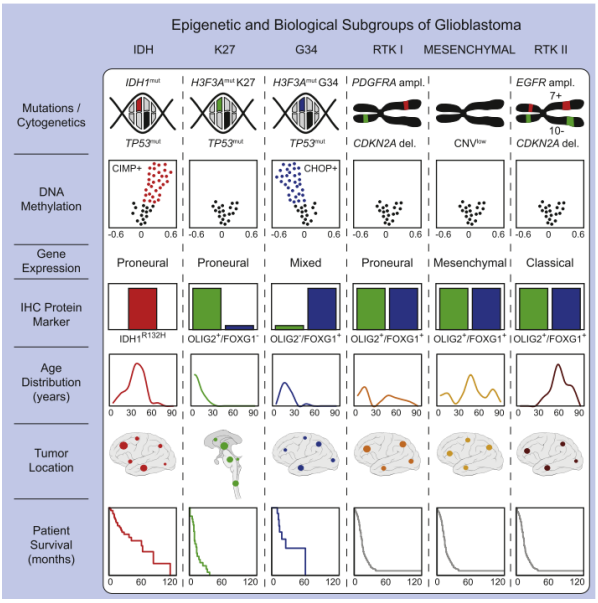


Figure 4. Summary of key molecular features of GBM subgroups and biological correlations
Simplified schematic representation of key genetic and epigenetic findings in six GBM subgroups as identified by methylation profiling and correlations with clinical patient data (Sturm, Witt et al. 2012).

As for RTK I, recurrent chromosomal alterations are gain of chromosome 7 with or without *EGFR* amplification (>90%), loss of 9p21 (*CDKN2A/B*; >70%) and chromosome 10 loss (>90%). The methylation class "*glioblastoma, IDH wildtype, subclass mesenchymal*" is comprised of GBMs that resembles the Mesenchymal TCGA subtype. Recurrent chromosomal alterations are gain of chromosome 7 with or without *EGFR* amplification (>80%), loss of 9p21 (*CDKN2A/B*; >60%) and chromosome 10 loss (>90%), as well as alterations of *NF1*.Ceccarelli et al. made a deeper analysis and correlated DNA methylation clusters with telomere length and maintenance, as well as with other biomarkers and clinical features, in order to elucidate the mechanism of progression from low grade gliomas to glioblastoma (Ceccarelli, Barthel et al. 2016). Furthermore, this approach was applied to other malignant brain tumours, the primitive neuroectodermal tumours of the central nervous system (CNS-PNETs), that often show a difficult diagnosis, and allowed to classify most of them into well-defined CNS tumours entity (Sturm, Orr et al. 2016). Data collected from this extensive studies led to a comprehensive approach aimed to classify all CNS tumours on the basis of the DNA-methylation profile (Capper, Jones et al. 2018). A collaboration between Heidelberg University and Hospital, the German Cancer Research Center (DKFZ) consortium and the German Consortium for Translational Cancer Research allowed to develop a platform, available on line, where users can upload their data and obtain the classification of their samples (https://www.molecularneuropathology.org/mnp).

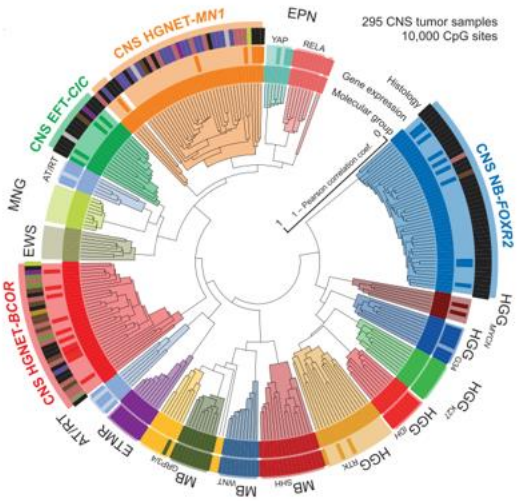


Figure 5. Identification of New CNS Tumour Entities across Histologies
DNA methylation patterns of CNS-PNET and CNS tumours of varying histology. Molecular subgroup assignment by DNA methylation (inner circle) or gene expression patterns (middle circle) correspond to subgroup labels. Original tumour histology (outer circle) is depicted for tumours from new molecular CNS tumour entities by colored bars as indicated (Sturm, Orr et al. 2016).

Hypermutation in human cancer

Mutations in cancer genes can be inherited, spontaneous, or be acquired over time during tumour evolution or by therapeutic effect, but the degree to which mutation rates influence cancer initiation and development and/or enable therapeutic resistance is still under debate. There are emerging evidence that, for many cancer types, outliers with a much higher mutation burden (hypermutation) do exist (Campbell, Light et al. 2017).

Systematic retrospective studies, such as those by The Cancer Genome Atlas (TCGA) and International Cancer Genome Consortium (ICGC), have revealed that mutation loads can differ by several orders of magnitude, with a wide variety of tumour types, such as melanoma, lung, stomach, colorectal, endometrial, and cervical cancers, displaying high mutation loads consistent with hypermutation, which may generate drivers of malignancy, even though it is not clear and most are typically classified as “variants of unknown significance”.

Statistical approaches for extracting information from cancer genomes mutation datasets have provided the ability to assess sources and consequences of mutations during cancer development by analysing mutational patterns. Apparent irregularities in the distribution of mutation types and position are compared to the null hypothesis of random mutation spectrum. These informations are matched against mechanistic knowledge about the chemistry of a mutagenic factor and genetic systems expected to repair the resulting DNA lesions, allowing statistical analysis of cancer mutation spectra (Roberts and Gordenin 2014).

Specific genomic features susceptible of a high rate of mutations have been identified, such as rearrangement breakpoints, replication timing, transcription levels, and chromatin organization, that can in some cases be grouped in mutation clusters.

Moreover, some specific mutation signature have been paired with their source, both extrinsic, as UV light (Pfeifer, You et al. 2005) or tobacco carcinogens (Pleasance, Stephens et al. 2010), and intrinsic, as the dysregulation of apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like (APOBEC) family members (Roberts, Lawrence et al. 2013).

The integrity of the genome is guaranteed by the DNA damage repair machinery. Defects in these machinery by mutations that compromise proofreading, performed by the major replicative enzymes Polε and Polδ1, or DNA mismatch repair, are associated with hypermutation in colorectal, endometrial, and other cancers (Muzny and Network 2012, Network 2012, Kandoth, Schultz et al. 2013). DNA replication repair mutations are also found in cancer predisposition syndromes, such as Lynch syndrome. In particular, dysfunction of the DNA mismatch repair (MMR), the mechanisms

that corrects improper nucleotide pairings that arise from replication errors, is associated with a hypermutator phenotype and Microsatellite instability (MSI) (Baretti and Le 2018).

MLH1, MutS protein homologue 2 (MSH2), MutS homologue 6 (MSH6) and PMS1 homologue 2 (PMS2) are the main proteins involved in this MMR system, and they interact as heterodimers; the complexes are ultimately responsible for the recognition of mismatches and insertion–deletion loops (Genschel, Littman et al. , Genschel, Littman et al. 1998) and subsequent recruitment of the MLH1/PMS2 complex will degrade the mutated stretch and initiates resynthesis.

In glioblastomas, the use of temozolomide has been associated with improved overall survival of the patients, specifically those with transcriptional silencing of the *MGMT* gene. However, the are several studies showing that a fraction of TMZ-treated patients acquired a hypermutator phenotype (Johnson, Mazor et al. 2014, Wang, Cazzato et al. 2016) as direct result of TMZ-induced mutagenesis due to inactivation of the DNA mismatch repair pathway (MSH6, MSH2, MHS4, MSH5, PMS1, PMS2, MLH1, and MLH2). The vast majority of the new mutations were G:C>A:T transitions, the signature of TMZ-induced mutagenesis. This mechanism is important both in the evolution of the tumour and for the development of therapy-resistance in the recurrences. Interestingly, Sa and colleagues identified a rare subset of pre-treatment adult glioma patients with a *de novo* hypermutator phenotype. TMZ-naïve hypermutated tumours lacked somatic mutation of *IDH1* and *MGMT* promoter methylation, and harboured both germline and somatic dysregulation of MMR encoding genes. Patients with TMZ-naïve hypermutagenesis demonstrated high incidence of cancer-development history in their immediate family members suggesting that germline dysfunction of the MMR pathway could potentially pose hereditary risk to genetic predisposition of carcinogenesis in gliomas (Sa, Choi et al. 2019).

Tumour heterogeneity reflects distinct genetic alterations and gene expression profiles within the same tumour bulk

Molecular and morphological classification suffers from the limitation caused by the attempt to categorize samples that are composed by different populations of cells with heterogeneous genetic lesions and different phenotypes. However, the extent of intra-tumour heterogeneity is still poorly understood.

A recent paper from Neftel and colleagues demonstrates, using an integrative approach, that heterogeneity in GBM is driven by four cellular states present at the same moment in the tumour, influenced by both genetics and microenvironment factors. *In vivo* single-cell tracing supports plasticity between these four states and the relative frequency of each state varies between GBM

samples, influencing the global aspect of the bulk tumour profile and its categorization (Neftel, Laffy et al. 2019).

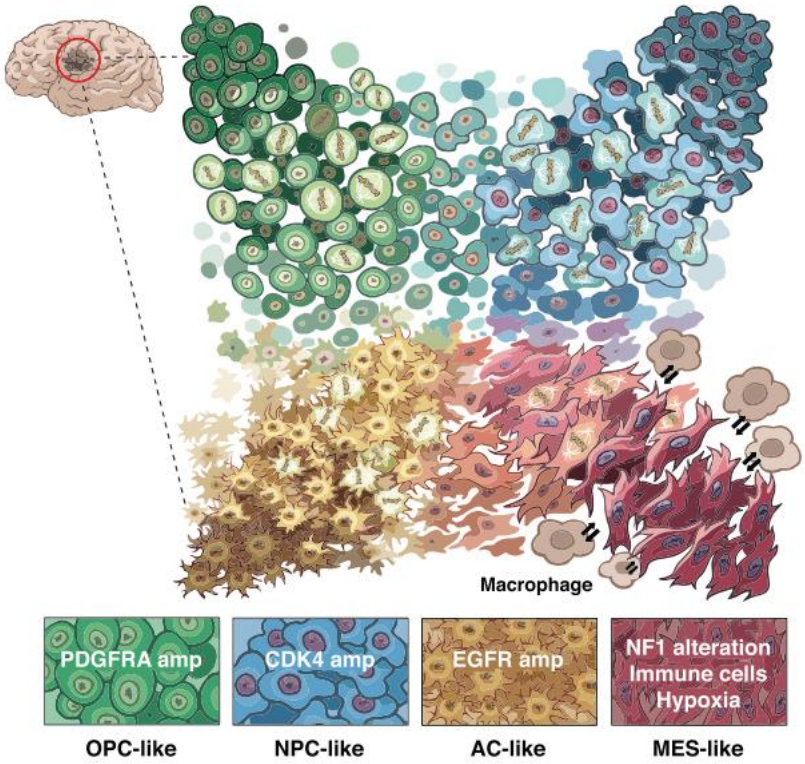


Figure 6. Model for the cellular states of glioblastoma and their genetic and micro-environmental determinants
Malignant cells in glioblastoma exist in four main cellular states that recapitulate distinct neural cell types; they are influenced by the tumour microenvironment, and exhibit plasticity. Each state is characterized by copy number amplifications of the *CDK4*, *EGFR* and *PDGFRA* loci and by mutations in the *NF1* locus (Neftel, Laffy et al. 2019).

The root of glioblastoma heterogeneity lies in the fact that these tumours contain a subpopulation of hierarchical organized cancer cell population led by cancer stem-like cells.

Brain tissue was historically considered a quiescent tissue, which does not undergo cell turnover, implying the absence of stem cells. This concept was overwhelmed after the discovery that neurogenesis persists in the adult age in various organisms, including humans (Lie, Song et al. 2004). The main neurogenetic region in the adult mammalian brain was then identified in the subventricular zone (Doetsch, Caillé et al. 1999).Consequently, the concept of stem cells has been extended to brain tumours and the subventricular zone is likely the major source of glioma cells (Sanai, Alvarez-Buylla et al. 2005), along with other minor neurogenic regions within the brain.

The first evidence of the existence of cells with stem-like properties in human brain tumours was reported in the paper of Ignatova and colleagues (Ignatova, Kukekov et al. 2002). By applying the same protocol used for neural stem cells, Galli and colleagues isolated clonogenic, neurosphere-

forming progenitors from adult human GBM. These cells were endowed with self-renewal ability and long-term propagation, are responsible for tumour-initiating activity and the generation of a heterogeneous cell population including non-stem cells that lack tumourigenic potential (Galli, Binda et al. 2004). These evidence support the fact that these cells likely represent brain tumour stem cells and are commonly, even though improperly, mentioned as Glioblastoma Stem Cells (GSCs). Since then, a bulk of paper was published, but the characterization of GSCs is still challenging and the identification of the brain tumours "cell of origin" remains controversial. Indeed, there is no surface marker that can unequivocally identify GSCs. Dirks and colleagues identified CD133+ cells as the tumour initiating subset *in vivo* and sorting for this antigen could enrich for neurosphere forming cells (Singh, Hawkins et al. 2004). Unfortunately this marker, as well as others (CD44, CD15), is not completely sensitive or specific for Glioblastoma stem cells (Beier, Hau et al. 2007, Gimple, Bhargava et al. 2019).

Despite their elusive nature, the *in vitro* model represented by neurospheres enriched of stem-like cells made possible the study of some features related to different GBM subtypes. For instance, tumour-initiating cells expressing EGFR, a marker associated with GBM Classical subtype that plays a fundamental role in gliomagenesis, displayed the most malignant functional and molecular phenotype among multiple GBM cell populations (Mazzoleni, Politi et al. 2010). In neurospheres, *EGFR* amplification and expression is mutually exclusive with the oncogene *MET*, a marker associated with mesenchymal and proneural subtype (De Bacco, Casanova et al. 2012). *MET* identifies a subpopulation of GSCs with high clonogenic potential and long-term self-renewal ability *in vitro* and enhanced growth kinetics *in vivo*, and promotes radioresistance (De Bacco, D'Ambrosio et al. 2016). *ASCL1*, a gene classifier for the proneural transcriptional subgroup, promotes the acquisition of a PN phenotype in GSCs by inducing a glial-to-neuronal switch and concomitantly represses mesenchymal features by directly downregulating the expression *NDRG1* (Narayanan, Gagliardi et al. 2018).

After remarkable advances in the study of GSC, it is now accepted that GSC population also presents a notable heterogeneity, which is reflected in the GBM tissue heterogeneity (Piccirillo, Combi et al. 2009, Chen, Nishimura et al. 2010). Additional studies have demonstrated that distinct GSC clones, even from the same tumour, could display variability in gene expression profile and functional properties, which could underlie different therapeutic sensitivity (Stieber, Golebiewska et al. 2014). Conventional therapies for GBM can lead to eradication or reduction of the tumour bulk, but almost all GBMs recur, suggesting that a subset of cells, presumably attributable to the stem cell bulk, are

responsible for sustaining long-term tumour growth and recurrence (Chen, Li et al. 2012), possibly due to their chemoresistance ability (Liu, Yuan et al. 2006).

A recent paper of Orzan and colleagues on GSCs from primary GBMs and recurrences showed that the complex heterogeneity observed in whole-tissue GBMs are reflected at the stem cell level, since primary and recurrent GSCs likely derive from a putative common ancestor by divergent genetic evolution and are positively selected by chemoradiotherapy (Orzan, De Bacco et al. 2017).

Investigating the genetic features of the cancer stem cell population seems to be the most suitable approach to achieve a deeper understanding of the pathogenesis of this tumour, to explain treatment failure, and to set up new therapeutic strategies.

Glioblastoma with Primitive Neuronal Component

"Glioblastoma with Primitive Neuronal Component" (GBM-PNC) has been now officially recognized as a peculiar histological pattern in the 2016 WHO classification (Louis, Perry et al. 2016). Previously referred in the literature as "Glioblastoma with PNET-like component", is a rare variant, representing approximately 0.5% of all GBMs. It is described as a tumour that "usually comprised of a diffuse astrocytoma of any grade (or oligodendroglioma in rare cases) that has well-demarcated nodules containing primitive cells that display neuronal differentiation […] and that sometimes has *MYC* or *MYCN* amplification" (Figure 7). The PNC component consists of hypercellular nodules of primitive appearing cells with high nuclear-to-cytoplasmic ratios, markedly hyperchromatic nuclei, and high mitotic rates, with evidence of neuronal differentiation, as gain of Synaptophysin positivity and loss of GFAP expression; they often display Homer Wright rosettes and anaplasia, as also seen in the large cell/anaplastic variant of medulloblastomas (Perry, Miller et al. 2009). In contrast, glioma-associated alterations involve both components, 10q loss (50%) being the most common.

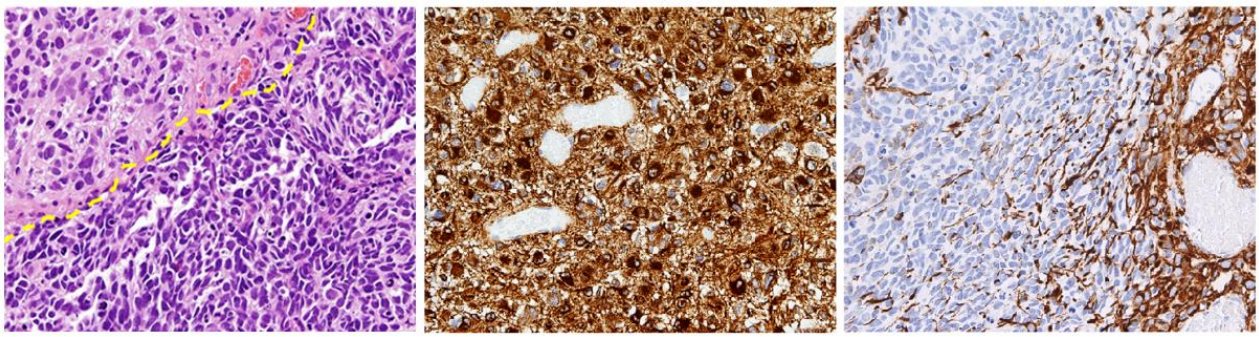


Figure 7. Glioblastoma with Primitive Neuronal Component
On the left, H&E staining showing a sample of GBM-PNC combining an area of high grade GBM (upper left of the yellow line) with a hypercellular nodule with primitive neuronal aspect; central image shows positivity to GFAP IHC of the glial component, while right image shows negativity of the PNC component for GFAP (40X original magnification).

These tumours may display OLIG2 expression, a feature that contrasts with primary CNS-PNETs, the main differential diagnostic, and they show higher frequency of *IDH1* mutation (Joseph, Phillips et al. 2013). The latter is consistent with the hypothesis of a secondary origin underlying high-grade glioma with secondary development of PNET-like foci and possible explanations are: (I) neuronal/neuroblastic metaplasia; and (II) nodular expansion of the cancer stem cell population (Perry, Miller et al. 2009).

When encountered, they pose a diagnostic challenge and are frequently classified as pure Glioblastomas, since the PNC component may be a minority of the tumour bulk, making it difficult to place them into a single diagnostic category.

Although large clinical studies are lacking, a variety of case reports have been reported (Wharton, Whittle et al. 2001, Ishizawa, Kan-nuki et al. 2002, Dulai, Bosanko et al. 2004, Kaplan and Perry 2007, Yao, Qi et al. 2015).

Recent studies (Perry, Miller et al. 2009, Song, Andrew Allen et al. 2011) revealed a significantly increased propensity for cerebrospinal fluid dissemination, as for CNS-PNETs, as compared to conventional GBMs and a possible benefit from the CNS-PNET platinum-based chemotherapy upfront or after standard failure of the conventional GBM therapy (irradiation and temozolomide). Song et al. evaluated 10 patients with GBM-PNC, with 3 patients who died having a median survival time of 17 months and 2 who were alive on follow-up at 15 and 31 months, suggesting a possible better prognosis compared to conventional GBMs. In a larger multi-institutional series of 53 cases, Perry et al. did not report any significant differences in survival for GBM-PNC compared with conventional GBM. No cases of solid organ metastasis have been reported, apart from one case of extracranial lung metastases (Tamai, Kinoshita et al. 2019). Given the differences in biology, clinical course, staging and potential implications for treatment, identification of PNET-like foci within GBM may be crucial. For this purpose, imaging could be of some help, since it was reported a substantially reduced apparent diffusion coefficient (ADC) values on Magnetic Resonance Imaging (MRI) (Ali, Joseph et al. 2014).

Since GBM-PNC is a rare tumour, only few data regarding their genetic alterations are available and very few cases, if any, were included in the TCGA.

In 2018, Xu and Li investigated the difference of mRNA expression profiles between GBM-PNC and conventional GBM using Human Cancer Drug Targets PCR Arrays. The main finding was the high expression of CDK4 and EGFR in the primitive neuronal component and in the glial component of

GBM-PNC respectively, suggesting that the combination of CDK4/6 inhibitor and targeted therapy against EGFR might be potential effective therapeutic regimen for GBM-PNC (Xu and Li 2018).

A more recent study of the same group reported whole exome sequencing for 11 GBM-PNC cases and analysed common mutations. Nine out of 11 had *TP53* mutations, while eight out of 11 cases showed mutations in PTEN–PI3K pathway, with an incidence much higher if compared to TCGA data for GBM (Xu, Zheng et al. 2019).

An additional contribution to the understanding of GBM-PNC biology comes from a recent work of our group about the telomere maintenance mechanism (TMM) in juvenile and adult brain tumours (Idilli, Pagani et al. 2020). The up-regulation of TMM is a common feature of cancer cells and a hallmark of cancer. In adult GBMs mutations within the promoter region of the telomerase catalytic subunit, TERT, leading to its overexpression, occur in 55–83% of the patients (Mangerel, Price et al. 2014, Sturm, Bender et al. 2014). In contrast, the Alternative Telomeres Lengthening mechanism (ALT) is a non-canonical mechanism of telomere maintenance developed by cancer cells with no-functional telomerase based on homologous recombination. It is found mostly in tumours with a mesenchymal origin (sarcomas) and in a subset of malignant pediatric brain tumours (Heaphy, Subhawong et al. 2011), including High Grade Gliomas and Primitive Neuroectodermal Tumours (PNET). In adult brain tumours ALT develops in approximately 15% of cases, and is associated with IDH1 mutations and better prognosis (McDonald, McDonnell et al. 2010). Our group analysed the TMM in a cohort of tumours: as reported, ALT is preferentially activated by CNS-PNETs, while GBMs express high levels of TERT. Surprisingly, the GBM component of the analysed GBM-PNC samples showed activation of both mechanisms assessed by the presence of TERT expression, as well C-Circles and PML bodies as markers of ALT. These markers were mostly confined to the GBM component, whereas the PNC component showed absence of TERT expression and no signs of ALT activation (Idilli, Pagani et al. 2020).

MATERIALS AND METHODS

Patient cohort

This study was conducted in compliance with the Declaration of Helsinki and with policies approved by the Ethics Board of Spedali Civili di Brescia, University of Brescia for retrospective and exclusively observational study on archival material obtained for diagnostic purpose and patient consent was not needed (Delibera del Garante n. 52 del 24/7/2008 and DL 193/2003). Sixteen patients with newly diagnosed pathologically confirmed GBM with PNC component were retrieved from the Institutional database of the Department of Pathology (Spedali Civili of Brescia) from 2007 to 2019. The other eight cases were collected from other hospitals: Candiolo Institute, Cremona, Padova, Vicenza and Cosenza. Histological diagnosis was revised according to the recent World Health Organization criteria (Louis, Perry et al. 2016) and formalin-fixed paraffin-embedded representative sections for each lesion were selected based on adequate tissue preservation, as assayed by hematoxylin and eosin (H&E) staining. Information regarding clinical features, treatment and outcome were collected from the medical records.

Immunostaining on paraffin-embedded sections

Briefly, 2-µm-thick paraffin sections were obtained from Formalin-Fixed Paraffin-Embedded samples (FFPE). Sections were de-waxed, re-hydrated and endogenous peroxidase activity blocked with 0.3% H_2O_2 in methanol for 20 minutes. Antigen retrieval (when necessary) was performed using a microwave or a thermostatic bath in either 1.0 mM EDTA buffer (pH 8.0) or 1 mM Citrate buffer (pH 6.0). Sections were then washed in TBS (pH 7.4) and incubated for one hour or overnight in the specific primary antibody diluted in TBS 1% bovine serum albumin.

The following primary antibodies have been used: rabbit polyclonal anti-GFAP (1:1500, Dako), mouse monoclonal anti-Vimentin (1:100, Biocare Medical), mouse monoclonal anti-EGFR (clone E30, 1:20, Dako), mouse monoclonal anti-YAP1 (clone 63.7, 1:50, Santa Cruz Biotechnology), mouse monoclonal anti-CD44 (clone 156-3C11, 1:200, Thermo Scientific), rabbit polyclonal anti-Olig2 (1:600, Millipore), mouse monoclonal anti-Synaptophysin (clone DAKSYNAP, 1:50, Dako), mouse monoclonal anti-EBF3 (clone 8D6, 1:3000, Abnova), mouse monoclonal anti-NeuN (clone A60, 1:1500, Millipore), rabbit polyclonal anti-SOX1 (1:300, Acris Antibodies), mouse monoclonal anti-SOX2 (clone 245610, 1:100, R&D Systems), rabbit polyclonal anti-βIII-tubulin (1:1000, Biolegend), goat polyclonal anti-DCX (1:200, Santa Cruz Biotechnology), mouse monoclonal anti-Nestin (clone 10C2, 1:2000, Millipore), mouse monoclonal anti-IDH1R132H (clone H09, 1:100, Dianova), rabbit polyclonal anti-ATRX (1:100, Sigma Aldrich), mouse monoclonal anti-P53 (clone DO-7, 1.100, Thermo Scientific), mouse monoclonal anti-TERT (clone 2C4, 1:100, NovusBiologicals), rabbit polyclonal anti-TRF2 (1:200, NovusBiologicals), mouse monoclonal anti-PML bodies (clone PG-M3, 1:100, Santa Cruz Biotechnology), rabbit monoclonal anti-c-Myc (clone Y69, 1:50, Abcam), rabbit polyclonal anti-n-

Myc (1:5000, Novus Biological), mouse monoclonal anti-MASH1/ASCL1 (clone 24B72D11.1, 1:50, BD Pharmingen), rabbit polyclonal anti-PDGFRα (1:50, Thermo Scientific), rabbit polyclonal anti-YKL40 (1:50, Quidel Corporation), rabbit polyclonal anti-MET(C12) (1:50, Santa Cruz Biotechnology), rabbit monoclonal anti-phosphoNDRG1 (clone D98G11, 1:100, Cell Signaling).

The reaction was revealed by using Novolink Polymer (Leica Microsystems) or Dako EnVision+Dual Link System Peroxidase (Dako Cytomation) followed by DAB and slides counterstained with Hematoxylin.

For double immunostainings, after completing the first immune reaction, second antibodies was applied; chromogen reaction was developed with Ferangi Blue Chromogen Kits (Biocare Medical) and nuclei were counterstained with Hematoxylin (Leica Microsystems).

Markers expression was semi-quantitatively scored on representative tumour regions based on both percentage [score ranges: 0 (0-5%), 1 (6-29%), 2 (30-69%), 3 (≥70%)] and intensity (score ranges: 0, no expression; 1, weak; 2, moderate; 3, high) of immunoreactive neoplastic cells with a combined cumulative score ranging from 0 to 6. Images were then acquired with Olympus DP70 camera mounted on Olympus Bx60 microscope using AnalySIS imaging software (Soft Imaging System GmbH).

For the classification of GBMs according to their transcriptional profile, the following panel of biomarkers was analysed: p53, EGFR, ASCL1, OLIG2, PDGFRα, pNDRG1, YKL40, MET. The panel was previously published (Orzan, Pagani et al. 2020) and gene classifiers were selected because for being mostly inversely correlated and differentially expressed between different GBMs, allowing for discrimination and clustering of the samples (Figure 8). IHC scores of the separate components were loaded on the algorithm available at http://fisher.med.unibs.it:3838/GBMscore/ for the transcriptional status prediction.

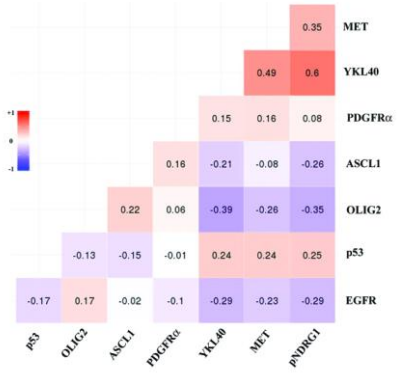


Figure 8. Selection of biomarkers for transcriptional status prediction based on immunohistochemical profile
Correlation index between the expression gene classifiers. Selected biomarkers were found to be mostly inversely correlated and differentially expressed between different GBMs, allowing for discrimination and clustering.

IDH Pyrosequencing assay

Genomic DNA was isolated from five 5-µm-thick paraffin tissue sections. At least 1 slice was stained with Hematoxylin and Eosin to check for the percentage of tumour cells and adequacy of the sample. DNA extraction was performed by QIAamp DNA FFPE tissue (Qiagen) according to the manufacturer's protocol. The "IDH1/2 status" kit (Diatech) was used to identify the main variants in codon 132 of *IDH1* gene, exon 2

(R132H, R132L, R132C, R132G, R132S) and in codon 172 of *IDH2* gene, exon 4 (R172K). Pyrosequencing was carried out on a PyroMark system (Qiagen-Diatech) according to the manufacturer's protocol based on the "sequencing by synthesis" principle. Pyrograms outputs were analysed by the Pyromark Q24 software (Qiagen) using the Allele Quantification software to determine the percentage of mutant versus wildtype alleles according to percentage relative peak height.

Evaluation of MGMT promoter methylation status

Evaluation of O6-methylguanine DNA methyltransferase (*MGMT*) promoter methylation status was performed by using "MGMT plus" kit (Diatech). DNA was extracted with QIAamp DNA FFPE tissue (Qiagen) and subsequently subjected to polymerase chain reaction (PCR) amplification with a forward primer and a biotinylated reverse primer using the "MGMT PLUS" kit (Diatech Pharmacogenetics), according to manufacturer's instructions. Pyrosequencing methylation assay was performed in order to evaluate 10 CpG sites in the following regions: chr 10: 131,265,507–131,265,556 using sequencing primer of MGMT Kit (Diatech Pharmacogenetics).

Fluorescence In Situ Hybridization (FISH)

EGFR FISH was performed with the Spectrum Orange locus-specific (excitation 547nm- emission 572nm) identifier EGFR probe specific for the *EGFR* locus (7p12) DNA probe "Chromosome Enumeration probe" (CEP)7 specific for centromeric region of human chromosome 7 (7p11.1-q11.1, locusD7z1) labelled with fluorophore spectrum green (excitation 497nm, emission 524nm). For the FISH slides were de-waxed, de-hydrated in 100% ethanol and dried at 45°C. Sections were then pre-treated by immersion in HCl 0,2N for 20 minutes and incubated in sodium thiocyanate 1M at 80°C for 30 minutes. Slides were then placed in a 0,3% pepsin solution at 37°C for 40 minutes and air dried. Probe (LSI EGFR Spectrum Orange/CEP7 Spectrum Green, Abbott Molecular) was applied to the dried slides that were sealed under a coverslip with tuber cement before incubation at 73°C for 5 minutes to denature the DNA followed by hybridization in a humidified chamber at 37°C overnight. Sections were then incubated in post-hybridization solution (2x55c with 0,3% NP-40) at 72°C for 2 minutes, air dried and counterstained with DAPI/Antifade Solution. Evaluation and scoring were determined according to criteria as previously described (Cominelli, Grisanti et al. 2015). Images have been acquired by Nikon Eclipse 90i microscope using Genikon imaging software v 3.4.8.

DNA methylation-based CNS tumour classification using a comprehensive machine learning

We obtained DNA of the two separated components from 4 samples of our cohort. Using H&E-stained slides as a guide, the GBM or PNC cancer cell-enriched areas in unstained 20 µm tissue sections were scraped from the glass using the sharp edge of a sterilized razor. The tissue slices were collected into 1.5 ml and sent to Diagenode for DNA extraction and quality control. The Illumina Infinium MethylationEPIC array BeadChip (850K) was carried out by the Epigenomic Services from Diagenode (Cat nr. G02090000). Raw .idat files were

uploaded on the site https://www.molecularneuropathology.org/mnp , where the data were compared to methylation data of a reference cohort comprising over 2800 neuropathological tumours of almost all known entities (currently over 80 tumour classes or subclasses). After the upload, data are automatically normalized, and classified by Random Forest classification. The calibrated score accepted for the attribution of a methylation class is 0.84; however, for diagnostic samples with low cell content, calibrated scores between 0.5 and 0.8 would be accepted. Genome-wide DNA methylation array data can also be used to perform analysis of copy-number variations (CNV). Areas with high copy-number ratios correspond to areas with a gain of chromosomal material (trisomies, larger sub-chromosomal gains), while areas with low copy-number ratios represent lost DNA (deletion). The baseline (value: 0) is the line where the median absolute deviation to all data points is minimal. Thus, the baseline is close to the predominant copy-number state of a sample. Values that deflect from the baseline of a score higher than the log2 value of 0.4 are considered as gain or loss of genetic material. CNV analysis provides a good overview of gross structural alterations in the tumour genome. High-level amplifications (for example EGFR) and homozygous deletions are usually clear when present. In particular, the algorithm examines 29 genes, listed in Table 1, that are known to be altered in CNS tumours.

chr	start	end	name
chr1	204480507	204527248	MDM4
chr2	16075560	16087129	MYCN
chr2	121549867	121750229	GLI2
chr4	1739326	1808410	FGFR3/TACC3
chr4	55090264	55164412	PDGFRA
chr5	1253287	1300162	TERT
chr6	135497453	135540310	MYB
chr7	55081725	55275031	EGFR
chr7	92234235	92468231	CDK6
chr7	116307459	116438440	MET
chr7	138552722	140487384	KIAA1549/BRAF
chr8	38279316	38688695	FGFR1/TACC1
chr8	67474411	67530480	MYBL1
chr8	128743315	128753680	MYC
chr9	21967751	22014312	CDKN2A/B
chr9	98205264	98275831	PTCH1
chr10	89618195	89731687	PTEN
chr10	131260454	131565783	MGMT
chr11	69450873	69469242	CCND1
chr12	4377902	4414522	CCND2
chr12	58141510	58151230	CDK4
chr12	69196952	69239324	MDM2
chr13	48872883	49056026	RB1
chr17	7571720	7595868	TP53
chr17	29416945	29704695	NF1
chr17	58672544	58743640	PPM1D
chr19	54164928	54265684	C19MC
chr22	24124150	24176705	SMARCB1
chr22	29994545	30094589	NF2

Table 1: List of 29 CNS tumour relevant genes analysed for CNV by the Heidelberg algorithm (Capper, Stichel et al. 2018)

A summary of the CNV plots with an overview of typical chromosomal aberrations found in the main GBMs DNA methylation classes identified by Sturm (Sturm, Witt et al. 2012) are shown in Figure 9 (Capper, Stichel et al. 2018). The DNA methylation-based CNS tumour classifier does not rely on the CNV pattern of a given tumour for classification. However, since losses or gains of chromosomal regions are of high diagnostic impact in some instances, it is highly recommendable an approach that contemplates an integrated diagnosis of morphological and molecular genetic findings. The results from the CNV analysis can be considered as independent from results of the methylation classifier, and both readouts can independently contribute to the final diagnostic interpretation.

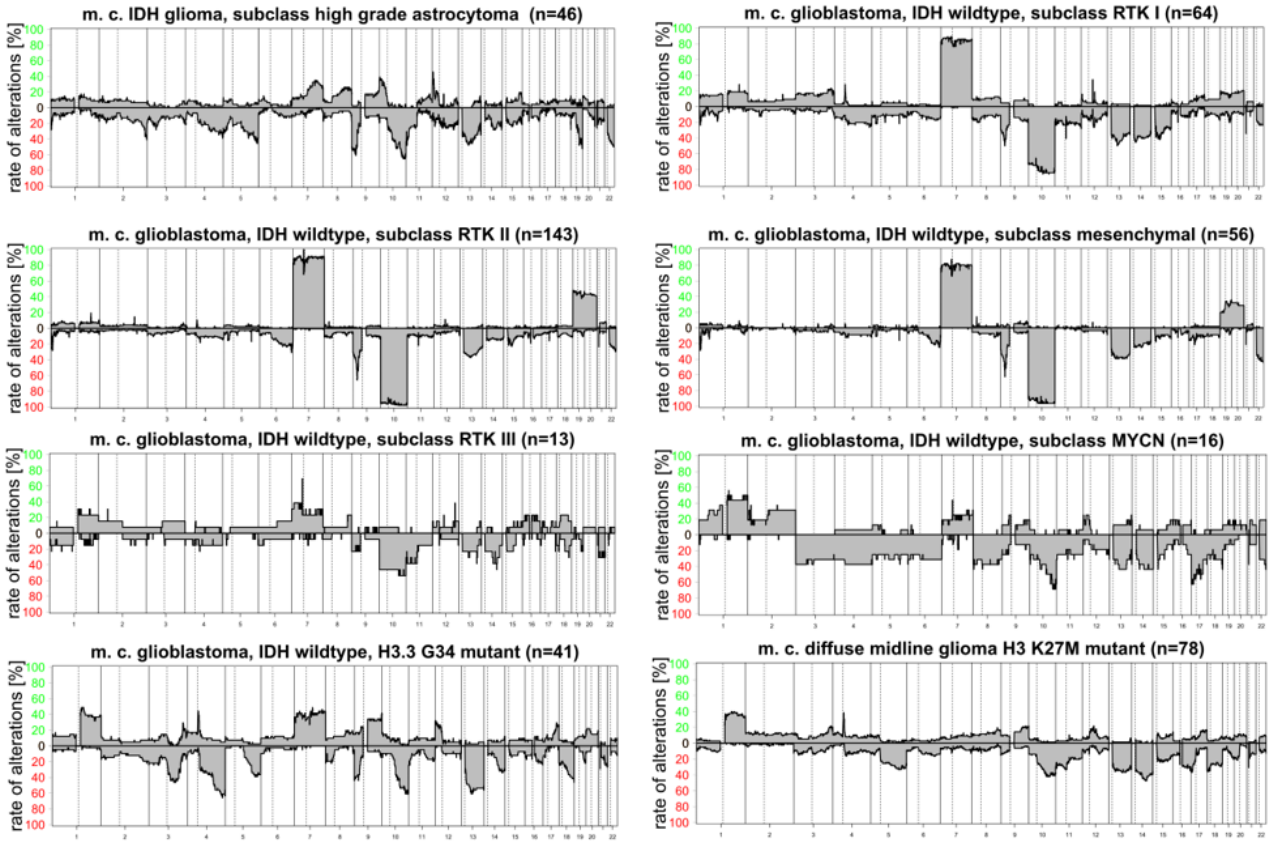


Figure 9. Summary CNV plots of GBMs methylation classes
An overview of the typical chromosomal aberrations found in the main DNA methylation classes of GBMs and diffuse midline K27M glioma is shown (Capper, Stichel et al. 2018). The p-arm (left) and the q-arm (right) separated by a dotted line. Gains/amplifications represent positive, losses negative deviations from the baseline.

Next Generation Sequencing

We obtained DNA of the two separated components from 6 samples of our cohort. Using H&E-stained slides as a guide, the GBM or PNC cancer cell-enriched areas in unstained 20 µm tissue sections were scraped from the glass using the sharp edge of a sterilized razor. The tissue slices were collected into 1.5 ml and DNA was extracted using QIAamp DNA FFPE tissue (Qiagen) according to the manufacturer's protocol. DNA isolated FFPE samples was fragmented using the M220 Focused-ultrasonicator (Covaris®). The preparation of the NGS libraries started with End-Repair and A-Tailing reactions, followed by adapter ligation (NXSEQ Ampfree Low DNA Library kit, Lucigen®; TruSeq™ DNA Single Indexes, Illumina). Then, the target of interest has been

captured and enriched using the GBM-custom panel probes (xGen Predesigned Gene Capture Pools and xGen Hybridization and Wash kit, IDT®) following manufacterer's protocol. Final libraries were sequenced on MiSeq sequencer (Illumina®).

We used a panel of genes designed in the laboratory of Candiolo Cancer Institute in collaboration with the University of Torino to detect changes in 75 target genes whose alterations are known to be relevant for glioblastoma tumours (listed in Table 2). Custom-panel displays a genomic size as little as 196 Kbps. Genetic analysis was performed as reported in Corti et al. and Crisafulli et al. (Corti, Bartolini et al. 2019, Crisafulli, Mussolin et al. 2019). In detail: the pipeline uses automated and semi-automated methodologies, *ad hoc* developed and other open-sources freely available software. Sequences generated by Illumina MiSeq were preprocessed to remove all bases in the read with a Phred quality score less than 30. Reads were mapped to the human reference, assembly hg19, using BWA-mem algorithm. PCR duplicates were removed using the RMDUP command of SAMtools package. Custom pipeline for genetic analysis was used in order to call somatic variations when supported by at least 1% allelic frequency and 5% Fisher's Test significance level, according to previously published methods(Corti, Bartolini et al. 2019, Crisafulli, Mussolin et al. 2019). Insertions and deletions (indels) were identified using Pindel software and were annotated by custom scripts printing out gene information, number of normal and mutated reads, the allelic frequencies and the variation effect. Each of these entries was associated with the corresponding number of occurrences in the COSMIC database. Copy Number Variation (CNV) was obtained by calculating the ratio of median gene depth to the median depth of whole exome. For each gene, copy number variation is reported as 2* tumour/normal CNV. The circular binary segmentation (CBS) algorithm, as implemented in the DNA copy R module, was used to cluster all the gene copy-number alterations. In the copy number analysis, a panel of 10 PBMCs was used as normal reference.

Statistical analysis

Values were expressed as mean ± standard deviation (SD). In vitro experiments were repeated at least three times. Statistical analyses were performed using GraphPad Prism Software (GraphPad Software Inc.). Unpaired two-sided Student's t –test was used. P -value < 0.05 was considered to be significant.

* $p<0.05$; ** $p<0.01$; *** $p<0.001$; **** $p<0.0001$.

GBM NGS Custom Panel								
Chr.	**Gene**	**Region**	**Chr.**	**Gene**	**Region**	**Chr.**	**Gene**	**Region**
chr2	*ACVR*	all cds	chr1	*H3F3A*	all cds	chr4	*PDGFRA*	all cds
chr14	*AKT1*	all cds	chr6	*HIST1H3B*	all cds	chr3	*PIK3CA*	all cds
chr19	*AKT2*	all cds	chr2	*IDH1*	all cds	chr3	*PIK3CB*	all cds
chr1	*AKT3*	all cds	chr15	*IDH2*	all cds	chr5	*PIK3R1*	all cds
chrX	*ATRX*	all cds	chr2	*IRS1*	all cds	chr1	*PIK3R3*	all cds
chr7	*BRAF*	all cds	chr12	*MDM2*	all cds	chr3	*PIK3R4*	all cds
chr12	*CDK4*	all cds	chr1	*MDM4*	all cds	chr2	*PMS1*	all cds
chr7	*CDK6*	all cds	chr7	*MET*	all cds	chr7	*PMS2*	all cds
chr9	*CDKN2A*	all cds	chr10	*MGMT*	all cds	chr19	*POLD1*	all cds
chr9	*CDKN2B*	all cds	chr3	*MLH1*	all cds	chr7	*POLD2*	all cds
chr19	*CIC*	all cds	chr14	*MLH3*	all cds	chr11	*POLD3*	all cds
chr6	*DAXX*	all cds	chr2	*MSH2*	all cds	chr11	*POLD4*	all cds
chr2	*DNMT3A*	all cds	chr5	*MSH3*	all cds	chr12	*POLE*	all cds
chr10	*EBF3*	all cds	chr1	*MSH4*	all cds	chr15	*POLG*	all cds
chr7	*EGFR*	all cds	chr6	*MSH5*	all cds	chr10	*PTEN*	all cds
chr17	*ERBB2*	all cds	chr2	*MSH6*	all cds	chr13	*RB1*	all cds
chr12	*ERBB3*	all cds	chr8	*MYC*	all cds	chr4	*RFC1*	all cds
chr2	*ERBB4*	all cds	chr2	*MYCN*	all cds	chr7	*RFC2*	all cds
chr1	*EXO1*	all cds	chr17	*NF1*	all cds	chr13	*RFC3*	all cds
chr1	*EXO5*	all cds	chr22	*NF2*	all cds	chr3	*RFC4*	all cds
chr8	*FGFR1*	all cds	chr9	*NOTCH1*	all cds	chr12	*RFC5*	all cds
chr10	*FGFR2*	all cds	chr1	*NTRK1*	all cds	chr19	*SMARCA4*	all cds
chr4	*FGFR3*	all cds	chr9	*NTRK2*	all cds	chr22	*SMARCB1*	all cds
chr5	*FGFR4*	all cds	chr15	*NTRK3*	all cds	chr5	*TERT*	all cds
chr1	*FUBP1*	all cds	chr20	*PCNA*	all cds	chr17	*TP53*	all cds

Table 2. Genes analysed with the NGS custom panel

RESULTS

Patient cohort and clinical-pathological features

The study was conducted on 24 diagnosed GBM with PNC component upon revision of histological diagnosis.

Clinical information with complete follow-up was available for 19 cases. Median age at diagnosis was 59.5 years (range: 25 – 82) and female/male ratio was 1:1.2 (female = 11, male = 13). The preferential sites are frontal (9 out of 24; 37.5%) and temporal (9 out of 24; 37.5%), with a lower incidence in the parietal zone (4 out of 24; 16.7%). 6 patients developed short-term recurrences, while 3 patients showed cranio-spinal dissemination; 1 patient developed extra-CNS multiple metastasis to liver and bones.

Clinical features, including KPS and recursive partitioning analysis (RPA) classification are summarized in Table 3. In 16 patients (66.7%), surgery resulted in complete removal with no residual enhancement seen on postoperative CT or MRI scans. Partial removal occurred in 2 patients (8.3%). Two patients (8.3%) did not receive any therapy due to sudden worsening of clinical conditions. A total of 5 patients (20.8%) received radiotherapy or chemotherapy alone, whereas radiotherapy and concomitant and sequential temozolomide was the most frequently adopted schedule: 50.0% of the cohort, up to 63.2% if we consider only patients for whom clinical information are available (12 out of 19).

MGMT promoter methylation is present in 62.5% of the samples (15 out of 24). Of note, for 6 tumours we could analyse the DNA of the two components separately and the methylation status was the same for both. Overall survival information for our cohort was available for 19 patients out of 24. One patient deceased of post-surgical complications. Two patients were alive at the time of the analysis. In summary, the OS for our cohort is 15.5±7.4 months, in line with the OS of GBMs as reported in literature (Dolecek, Propp et al. 2012).

5 patients of the 19 for whom clinical information was available reported that at least one parent previously deceased for neoplastic pathology. One patient was previously diagnosed with a well differentiated liposarcoma with amplification of the gene *MDM2*; we found the same molecular alteration in the GBM-PNC (described in details further). It would be interesting to study in depth the correlation between GBM-PNC and a potential genetic or familial predisposition.

Characteristic	(No. 24) No. (%)
Gender	
female	11 (45.8)
male	13 (54.2)
Age, years (median 59.5 years)	
< 65	14 (58.3)
≥ 65	10 (41.7)
KPS	
< 70	4 (16.7)
≥ 70	16 (66.7)
NA	4 (16.7)
RTOG RPA classes	
IV	14 (58.3)
V	1 (4.2)
VI	5 (20.8)
NA	4 (16.7)
Site	
Frontal	9 (37.5)
Parietal	4 (16.7)
Temporal	9 (37.5)
Others	2 (8.3)
Hemispheres	
Right	10 (41.7)
Left	13 (54.2)
Bilateral	1 (4.2)
Extent of surgery	
Complete resection	16 (66.7)
Partial resection	2 (8.3)
NA	6 (25.0)
Therapy	
Combined radio-chemotherapy	12 (50.0)
Radiotherapy only	4 (16.7)
Chemotherapy only	1 (4.2)
None	2 (8.3)
NA	5 (20.8)
MGMT promoter status	
methylated	15 (62.5)
unmethylated	9 (37.5)

Table 3: Clinical feature of the studied cohort

Immunohistochemical profile

We performed the characterization of our sample cohort by immunohistochemistry, analysing the GBM and the PNC components separately.

At first, we evaluated the expression of markers associated with glial or neuronal differentiation. As shown in Figure 10, some markers were expressed only in the glial component, even though at variable levels. These markers are GFAP, Vimentin, EGFR, YAP1, CD44, biomarkers known to be linked to the glial phenotype and usually expressed in GBMs. On the contrary, the neuronal marker Neu-N, associated to a mature neuronal phenotype, was detected, albeit at low levels, only within the PNC component, while EBF3, expressed in neuronal progenitors, was expressed at a higher level exclusively in the PNC component. OLIG2, marker of oligodendroglial lineage, usually expressed in diffuse gliomas, was expressed in the GBM component as well as in the PNC component, even though at a lower level. Synaptophysin, a broad-spectrum neuroendocrine marker, is expressed mainly in the PNC component, with a low expression in the GBM component. The analysis of EGFR expression in our sample cohort is noteworthy. As shown in the plot in Figure 10, EGFR expression was mostly confined to the GBM component, with a medium score of 4.3±1.9, while in PNC component is extremely low (0.2±1.0). In 14 out of 24 samples, EGFR IHC score was 5 or 6.

We then analysed the expression of markers known to be expressed during neurogenesis in cells at different maturative stages. As illustrated in the histogram in Figure 11, both components display an immature profile and the expression of the analysed markers is comparable, as demonstrated by the levels of Sox1, Sox2, βIII-tubulin and DCX. Interestingly, Synaptophysin was expressed mainly in the PNC portion, with a low expression in the GBM, as previously reported. On the contrary, Nestin, generally recognized as a marker of undifferentiated CNS cells, is more expressed in the GBM portion, with a much lower expression in the PNC.

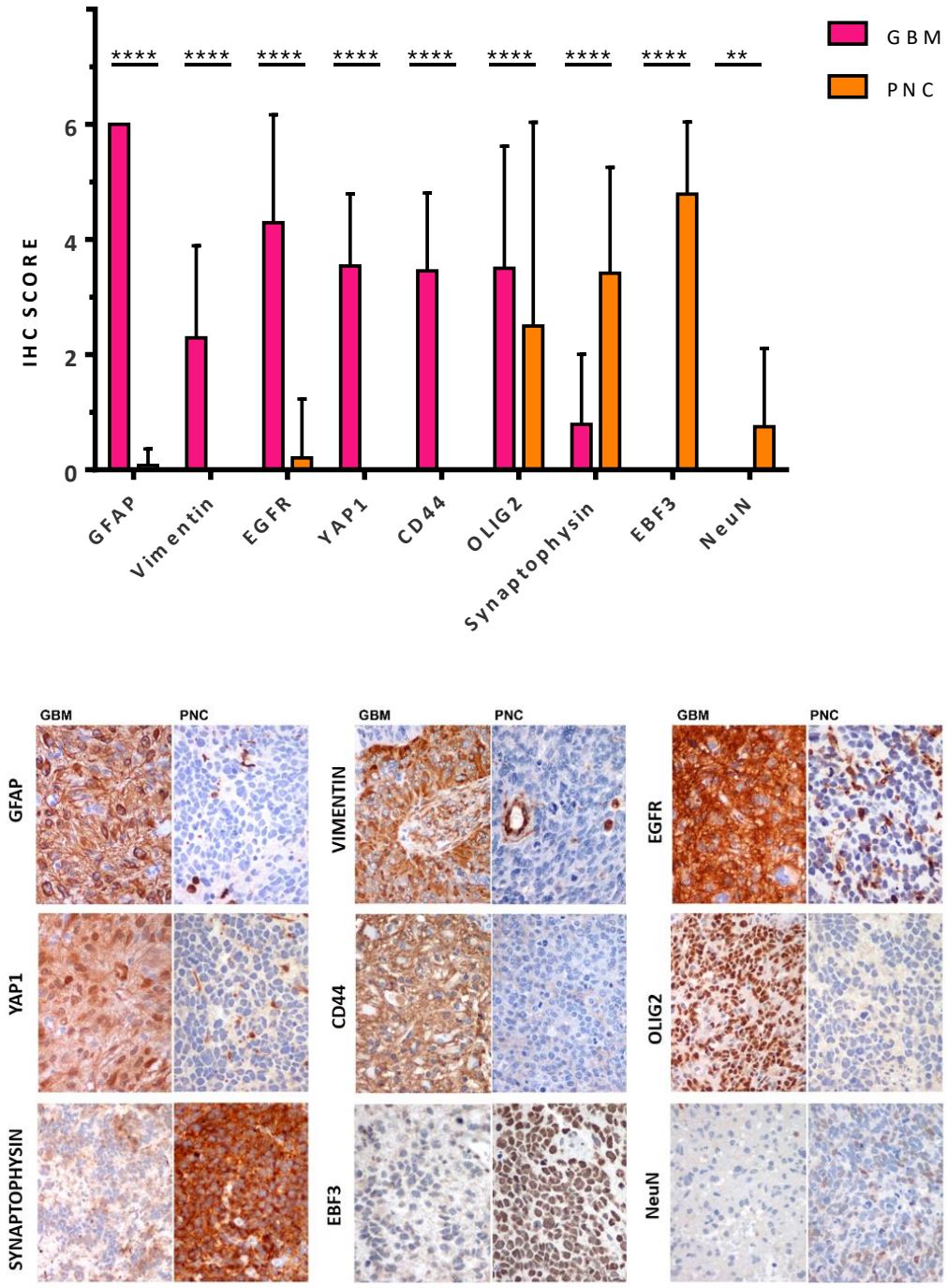


Figure 10. Analysis of glial and neuronal markers expression in GBM and PNC components
The histogram shows the mean value of markers expression (quantitative analysis performed by applying a "score" from 1 to 6 as described in the “Materials and Methods” section) in the two components of GBM-PNC. ** p<0.01; **** p<0.0001, Student’s t test. The panels below show representative images of markers expression in the two distinct components of GBM-PNC tumours (40X original magnification).

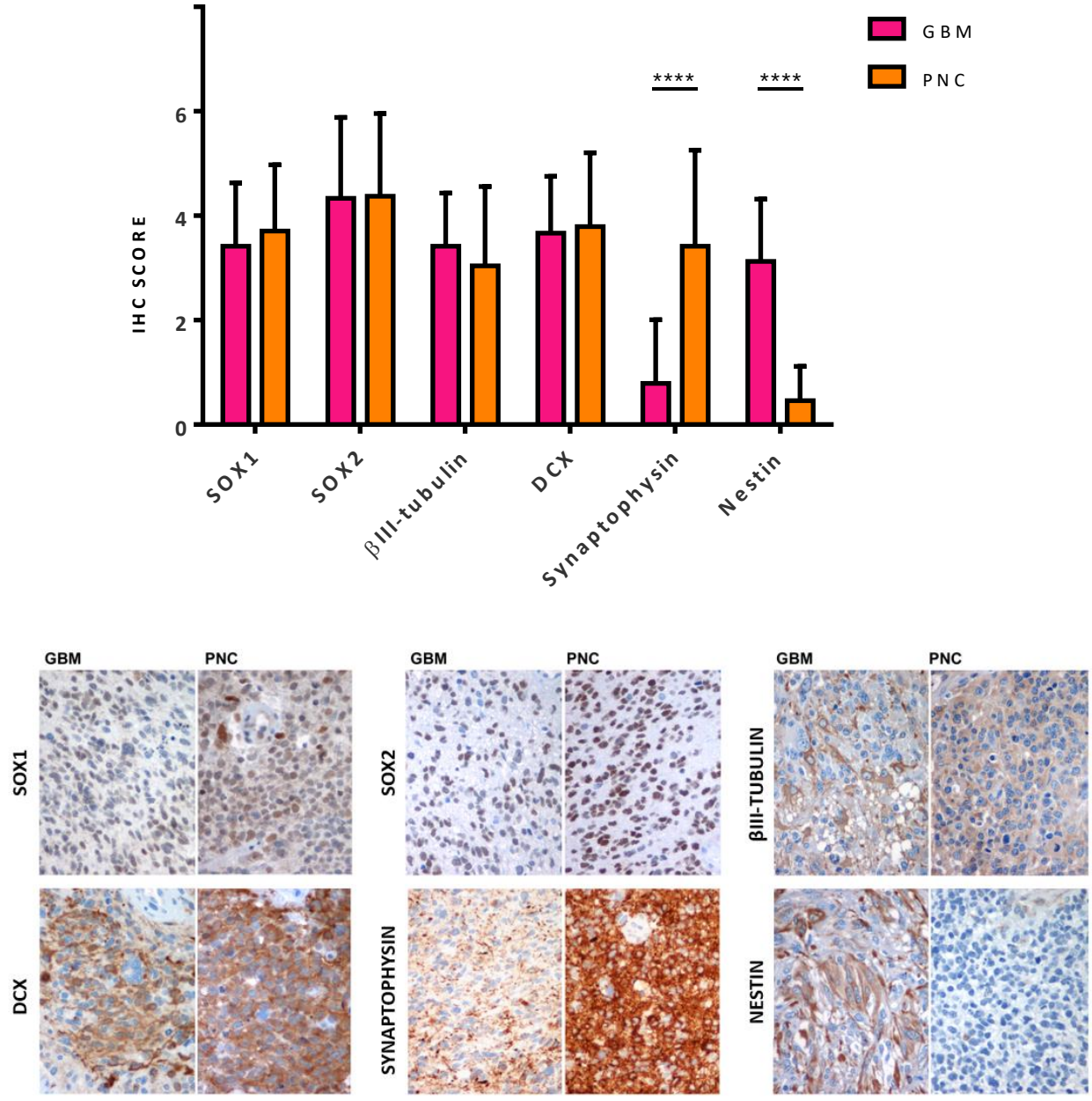


Figure 11. Analysis of neurogenesis markers expression in GBM and PNC components
The histogram shows the mean value of markers expression (quantitative analysis performed by applying a "score" from 1 to 6 as described in the "Materials and Methods" section) in the two components of GBM-PNC. **** $p<0.0001$, Student's t test. The panels below show representative images of markers expression in the two distinct components of GBM-PNC tumours (40X original magnification).

Evaluation of molecular alterations of diagnostic value

Afterward, we evaluated the expression of some markers of diagnostic importance for the identification of mutations or alterations specific for CNS tumours (Figure 12).

We analysed the presence of the mutation on *IDH1* gene using an antibody that recognizes the most frequent mutation on the protein: the missense substitution R132H. *IDH* mutations are the earliest

detectable genetic alteration in precursor low-grade diffuse astrocytomas and discriminates between primary and secondary glioblastomas (Louis, Perry et al. 2016). Subsequently we confirmed the mutation at the gene level, for positive samples, using pyrosequencing; this technique is also able to detect the most frequent mutations in *IDH2* gene. In our cohort of samples, 4 samples out of 24 (16.7%) were positive for the IDH1-R132H mutation. The mutation was shared between the GBM and the PNC portions of the tumours. This percentage is consistently higher if compared to the incidence of *IDH1* mutations reported in literature for high grade gliomas, that is between 4 and 12% (Parsons, Jones et al. 2008, Cominelli, Grisanti et al. 2015).

Subsequently, we evaluated the overexpression of p53. This protein, as previously outlined, regulates the cell cycle and functions as a tumour suppressor; mutations of this protein or its pathway are observed in 28% of GBMs (Network 2008). For diagnostic purposes, the overexpression of p53 protein by IHC is generally considered to be associated to a mutation, usually missense, in the *TP53* gene (Takami, Yoshida et al. 2015). We considered as "overexpressing" only the samples with a IHC score of 5-6. In our cohort, 14 cases out of 24 are p53 overexpressing (58,3%) and are likely mutant for *TP53* gene. Also this feature is shared among the two components of the tumours.

We then evaluated ATRX expression. ATRX is a chromatin remodeling protein; its status has been shown to correlate with patient age, tumour histopathology, and prognosis. For instance, *ATRX* mutations confer a better progression free and overall survival in low grade glioma harbouring *IDH* mutations, but, in GBMs, ATRX deficiency has been shown to impair non-homologous end joining (Haase, Garcia-Fabiani et al. 2018). For the evaluation of ATRX expression, we considered the gene as mutated in case of loss of expression assessed by IHC. With the exception of one sample, not evaluable for technical reasons, we found 6 samples out of 23 as negative for ATRX expression (26.1%). The presence or loss of ATRX expression is shared in both GBM and PNC components of the tumours.

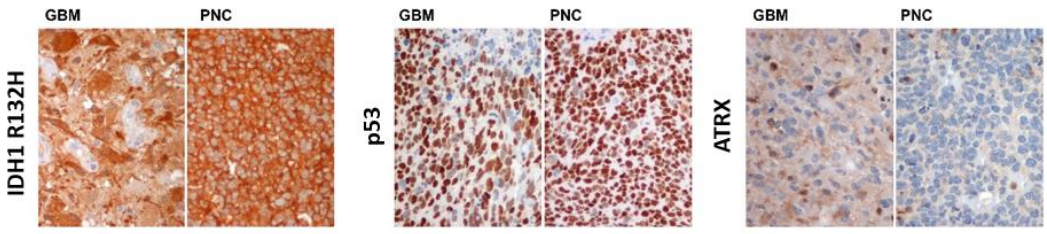
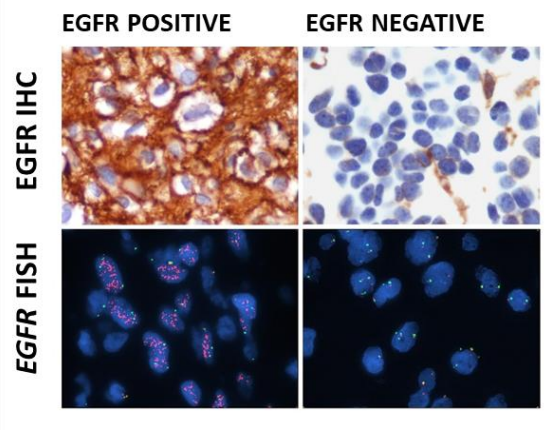


Figure 12. Molecular alterations of diagnostic value in GBM-PNC
Representative images of molecular alterations assessed by IHC in the two distinct components of GBM-PNC tumours: IDH1-R132H mutated protein; p53 overexpression; loss of ATRX expression (40X original magnification).

As previously shown in Figure 10, EGFR expression was mostly confined to the GBM component, with a medium score of 4.3±1.9, while in the PNC component is extremely low (0.2±1.0). In 15 out of 24 samples, EGFR IHC score was 5 or 6. As previously reported in Cominelli et al, EGFR overexpression strictly correlates with gene amplification (Cominelli, Grisanti et al. 2015). We thus selected for FISH analysis EGFR overexpressing samples (n=15; EGFR IHC score 5 or 6). As expected, almost all of the cases with IHC score 6 (10 out of 11) showed a clear gene amplification in the GBM components. Taking into consideration the PNC components, 4 samples out of 10 showed a diploid gene copy number; 4 samples showed gene amplification, while in 2 samples chromosome 7 polysomy was identified. In one sample out of 11, chromosome 7 polysomy was detected in both components. Regarding samples with IHC score 5 (n=4), amplification occurred in one case in both components and one case only in the glial component. One sample showed a chromosome 7 polysomy in both components; one case had a normal diploid status.



IHC SCORE	# of samples	FISH GBM	FISH PNC
• 5	1	A	A
	1	A	NA
	1	POL	POL
	1	NA	NA
• 6	4	A	A
	4	A	NA
	1	POL	POL
	2	A	POL

Figure 13. Evaluation of the expression and molecular status of EGFR
The panel shows representative images of one EGFR overexpressing sample as assessed by IHC (upper left picture) and the corresponding FISH analysis showing *EGFR* gene amplification status (lower left picture). On the right, one negative (upper) and not amplified (lower) sample is depicted (60X original magnification). On the right, a brief summary of FISH results of the analysed samples. (A=amplified; NA=not amplified; POL=polysomic).

The heterogeneity of the results and the discrepancy between IHC and FISH analyses could be due to the difficulty to clearly distinguish the two components in fluorescence microscopy, particularly for the sample where the two components are tightly intermixed. This issue is worth to be deeply clarified, however present data suggest that EGFR expression may be modulated by a specific transcriptional mechanism, independently from the presence or absence of gene amplification. In conclusion, these data suggest that some molecular alterations (*IDH1*, *TP53*, *ATRX*) are common for both components, suggesting their possible role in early oncogenic stages. Conversely, other mutations are specific for each component and may contribute to the development of the different phenotypes.

Evaluation of biomarkers of diagnostic value

The up-regulation of a telomere maintenance mechanism (TMM) is a common feature of cancer cells and a hallmark of cancer. In adult glioblastomas, overexpression of TERT due to mutations within the promoter region occur in more than 50% of the patients, while ALT develops in approximately 15% of cases, in association with *IDH1* mutations. In a recent study on the TMM in brain tumours, our group assessed that the presence of TERT, as well ALT, was confined mostly to the GBM component, whereas the PNC component showed an absence of TERT expression and no signs of ALT activation (Idilli, Pagani et al. 2020). In order to confirm this data, we performed IHC for TERT expression and a double IHC for the assessment of TRF2 and PML bodies co-localization, as a marker of ALT activation, on our GBM-PMC cohort of samples; we evaluated the two components separately. For TERT expression, we evaluated the intensity of the signal in immunoreactive neoplastic cells with a score ranging from 0 to 3 (0, no expression; 1, weak; 2, moderate; 3, high). As depicted in Figure 14, the GBM components of our samples are mostly positive for TERT expression (0=1; 1=9; 2=12; 3=2), with only 4 samples out of 24 negative; the PNC components are all negative for TERT expression.

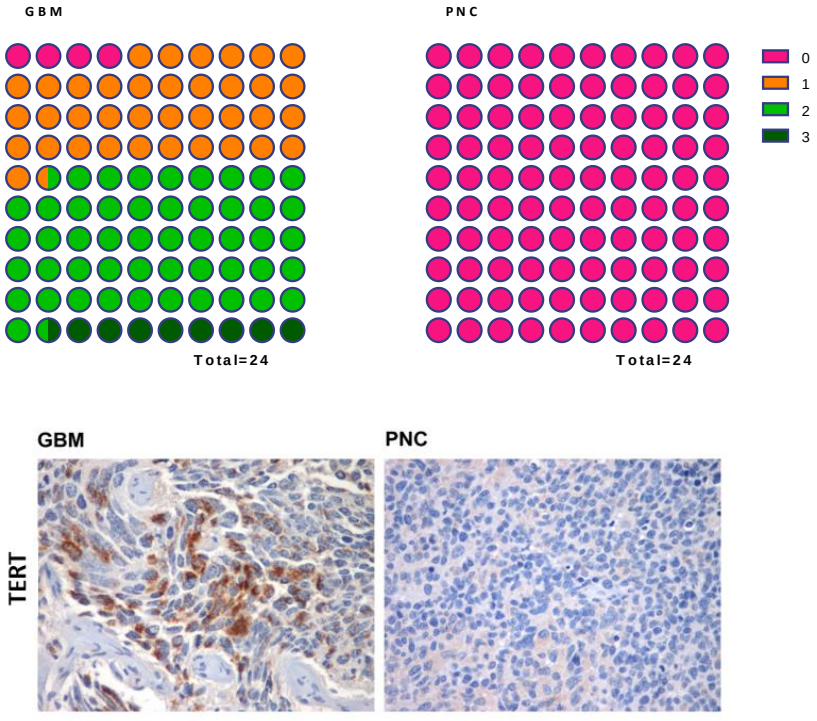


Figure 14. Evaluation of TERT expression in GBM-PNC
Upper plots show the number of samples assigned to each score; on the right, the evaluation of the GMB components, on the left, the evaluation of the PNC components. The difference is statistically significant: $p<0.0001$, Student's t test. Lower panels show representative images of the glial component of a GBM PNC tumour positive for TERT (score=3) (left image); the primitive neuronal component of a GBM PNC tumour negative for TERT (score=0) (right image) (40X original magnification).

For the evaluation of the co-localization of the PML bodies and TRF2, as a marker of ALT activation, on our samples, we attributed a score of "+" in case all of the cells displayed co-localization of the signals; "+/-" in case the co-localization was present only in some areas of the sample; "-" in case of absence of co-localization. Two samples were not adequate and were excluded from the analysis. As shown in Figure 15, the co-localization of PML bodies and TRF2 was present at least in some areas of the tissue in the GBM components of 21 out of 22 tumours ("+/-"=4; "+"= 17). Only 1 sample out of 22 showed co-localization of PML bodies and TRF2 in the PNC component, while the other 21 were negative.

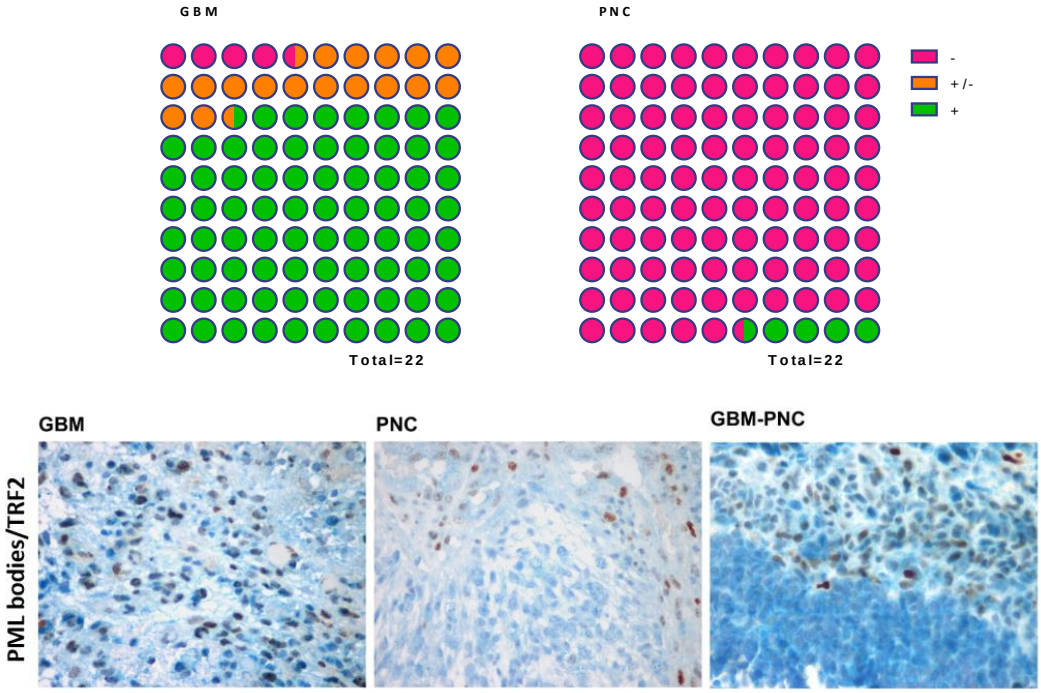


Figure 15. Evaluation of ALT activation in GBM-PNC
Upper plots show the number of samples assigned to each score; on the right, the evaluation of the GMB components, on the left, the evaluation of the PNC components. The difference is statistically significant: **** p<0.0001, Student's t test. Lower panels show representative images of the glial component of a GBM PNC tumour with co-localization of PML bodies and TRF2 (left image); the primitive neuronal component of a GBM PNC tumour without co-localization of PML bodies and TRF2 (central image); comparison between glial and primitive neuronal components of a GBM PNC (right image) (40X original magnification).

In summary, in line with our previous data, the presence of TERT or the presence of ALT activation markers was confined mostly to the GBM component, whereas the PNC component showed absence of any of the two TMM mechanisms.

c-Myc and n-Myc are transcription factors typically expressed in highly immature CNS tumours with neuronal differentiation (i.e. medulloblastoma). The immunohistochemical analysis of these markers in our GMB-PNC cohort revealed that the expression of both is mostly confined to the PNC component (mean scores: c-Myc 2.3±1.9 and n-Myc 2.7±2.1), with low levels in the GBM (mean

scores: c-Myc 0.6±0.9 and n-Myc 0.5±0.8) (Figure 16). Of note, c-Myc expression and n-Myc expression are inversely correlated.

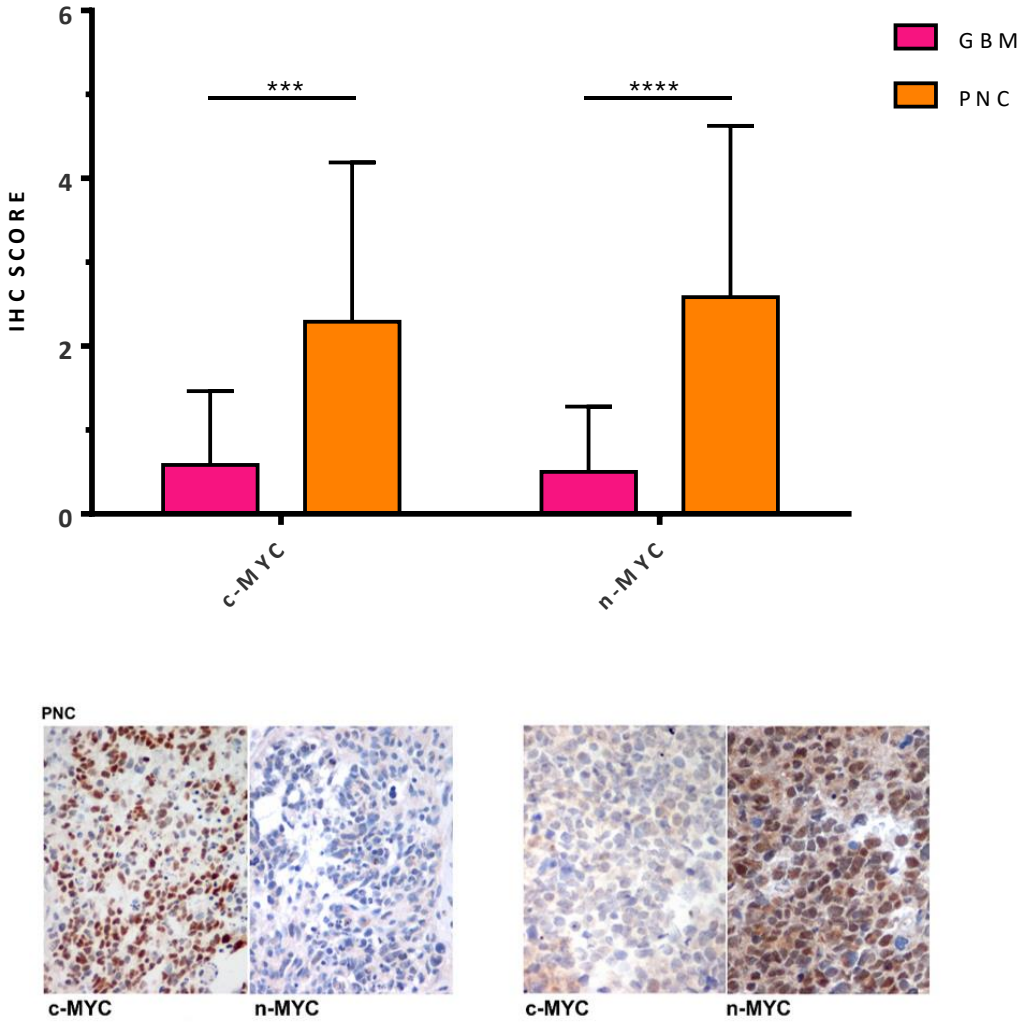


Figure 16. Expression of c-Myc and n-Myc in GBM-PNC
The histogram shows the mean value of c-Myc and n-Myc expression (quantitative analysis performed by applying a "score" from 1 to 6 as described in the "Materials and Methods" section) in the two components of GBM-PNC. *** p<0.001; **** p<0.0001, Student's t test. Representative images of c-Myc and n-Myc expression in two PNC components of GBM-PNC tumours, showing the correlation of mutual exclusivity between these factors (40X original magnification).

Expression of gene classifiers and attribution of transcriptional GBM subtypes

Recently, our group proposed a machine-learning algorithm for the classification of GBMs into transcriptional subtypes using an IHC approach (Orzan, Pagani et al. 2020). The analysis of the markers necessary for the classification was performed separately on both components of the GBM-PNC cohort and then loaded on the algorithm available at the following link: http://fisher.med.unibs.it:3838/GBMscore/.

The indicated biomarkers were selected among the others, because mostly inversely correlated and differentially expressed between different GBMs, allowing for discrimination and clustering. It is important to underline that it is not the level of expression of the single biomarkers, but the combination of the levels of expression of all the gene classifiers that permits to attribute the probability of the sample to belong to one of the three transcriptional subtypes (CL, PN or MES).

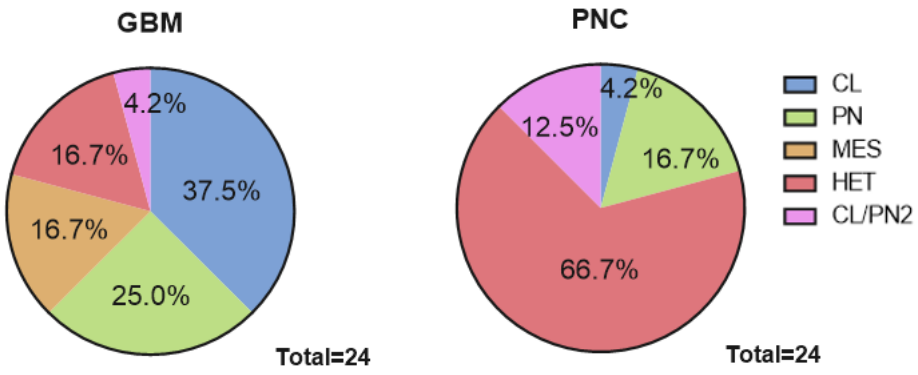


Figure 17. Distribution of the transcriptional subtypes in the GBM and PNC components
Each subtype is identified by a different color. The frequencies in percentage for each subtype are indicated.

For the GBM component, we found that the most represented subtype in our cohort is the Classical subtype (37.5%), followed by the Proneural (25.0%); 16.7% are attributed to the Mesenchymal subtype. 4.2% of our samples are classified as CL/PN2, a subtype characterized by the concomitant expression of EGFR at high levels and Proneural markers, while 16.7% of our samples have the same probability of belonging to each of the three subtypes (labelled as Heterogeneous; HET). In summary, the majority of our samples belong to CL or PN subtype (66.7%)(Figure 17).

We then analysed the PNC component with the same approach. As shown in the right plot, the vast majority of the samples could not be classified into one subtype and were tagged as HET (66.7%). 16.7% of the samples were classified as PN, while 4.2% were attributed to the CL subtype; 12.5% were classified as CL/PN2, while none as MES. From these data we can deduce that the PNC component of the GBM-PNC can not be classified using the algorithm from Orzan et al, a system that was designed for the classification of GBMs. This conclusion supports the evidence that the PNC component is constituted by cells with an undifferentiated phenotype.

GBM-PNC classification based on genome-wide DNA methylation profiles

Thereafter we further analysed our samples by a genome-wide DNA methylation profiling. In 2012, Sturm and colleagues identified six clusters for high grade gliomas analysing their genome-wide DNA

methylation profiles (Sturm, Witt et al. 2012). In 2016, they made a wider analysis and extended the methylation classification system to all the CNS tumours, comprising also the tumours of neuronal origin(Sturm, Orr et al. 2016). All these data were collected and converged into a platform publicly available where the users can upload their data and obtain the classification of their samples. (https://www.molecularneuropathology.org/mnp).

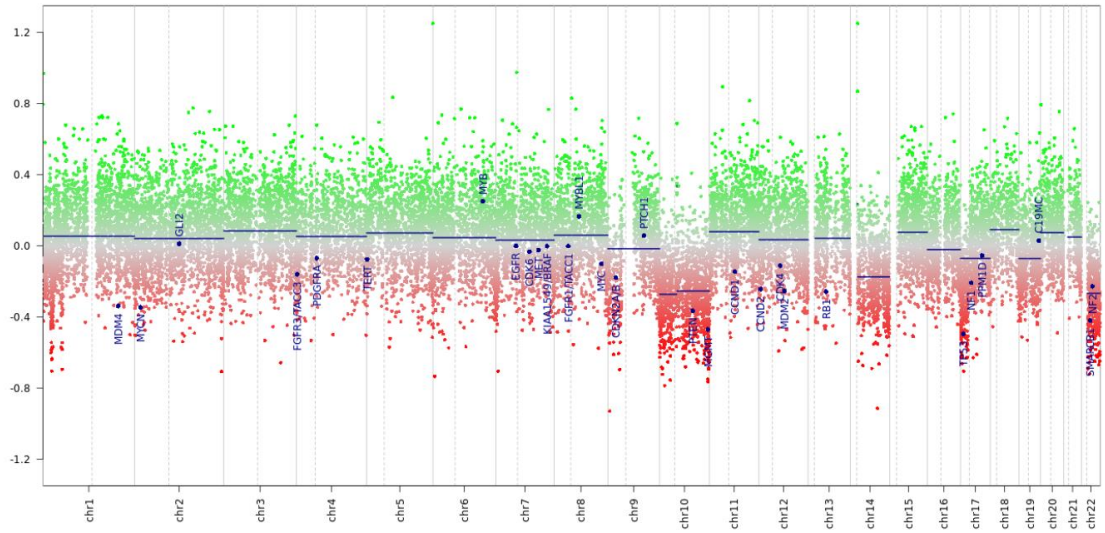
FFPE samples from the two separated components of 4 GBM-PNC of our cohort were sent to Diagenode s.a. for the analysis with the Infinium MethylationEPIC array BeadChip (850K) on Illumina platform. Raw data in .idat format were then uploaded on the platform in order to obtain the DNA methylation-based classification and the CNV profile. In order to establish whether there is a loss or gain of genetic material, we applied, for chromosomal regions, deviations ≥ 0.4 from the baseline, while, for single genes, values that deflect from the baseline ≥0.8.

In the following pages the results for the analysed samples are shown.

As shown in Figure 18, for Tumour#1 it was not possible to attribute a cluster of methylation for the GBM component. This sample was previously classified as Heterogeneous from the algorithm for the transcriptional subtype. The PNC component was assigned to the IDH wildtype class, even though with a low calibrated score; the attribution is conceivably to the RTK I or RTK II subclasses. The analysis of the CNV profile collimates with the classification given. For Tumour#2 (Figure 19), the attribution to the IDH mutant cluster is clear, with a high calibrated score, and even the CNV profile clearly matches; the mutation for IDH1 was previously assessed by IHC and pyrosequencing. For Tumour#3 (Figure 20), the methylation analysis resulted in very unlikely attribution to a class family and subclass, probably due to a low quality of the sample. As a matter of fact, Plexus Tumours, subclass B, are pediatric intraventricular neoplasms derived from the choroid plexus epithelium, with very different characteristics if compared to GBMs (Thomas, Sill et al. 2016). The CNV profile, indeed, revealed a typical IDH wildtype GBM profile, and the probable attribution is to RTK I o RTK II classes. This sample was previously attributed to the Classical transcriptional subtype for the GBM component, while the PNC was Heterogeneous. For Tumour#4, previously classified as Heterogeneous for both components, the attribution to the IDH wildtype, RTK I or II family, is clear, even though with a low calibrated score, and the CNV profile corresponds (Figure 21).

TUMOUR #1					
COMP.	**METHYLATION CLASS**		**CNV LOSS**	**CNV GAIN**	**CNV PROFILE**
GBM	**no matching methylation classes with calibrated score≥0.3**	Chr: Genes:	10		**IDH WILDTYPE glioblastoma** RTKI/RTKII
PNC	**methylation class family Glioblastoma, IDH wildtype 0.43** methylation class glioblastoma, IDH wildtype, subclass RTK I 0.2 methylation class glioblastoma, IDH wildtype, subclass RTK II 0.17	Chr: Genes:	10; 22 PTEN; MGMT; NF2		

GBM



PNC

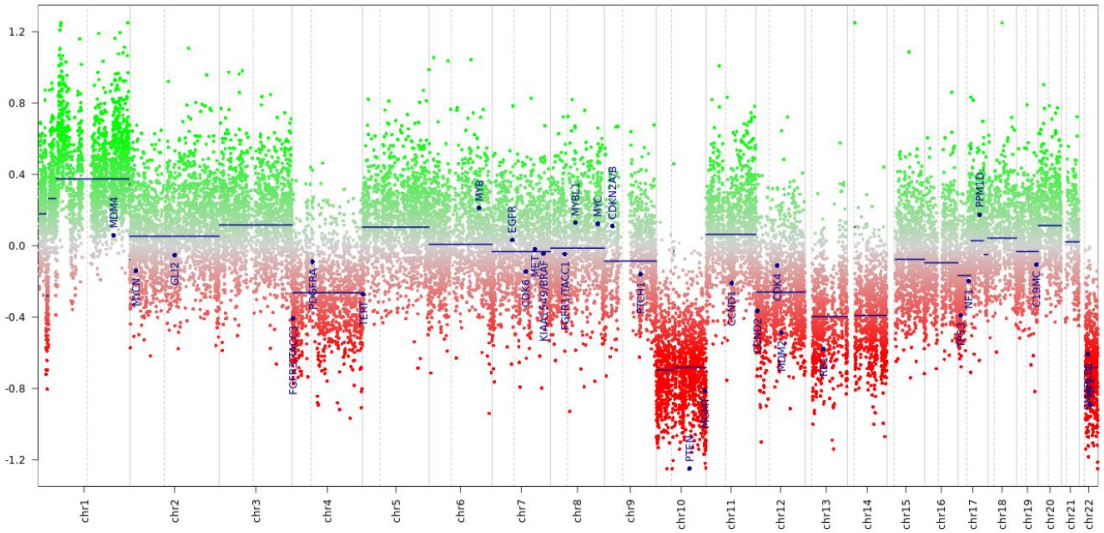
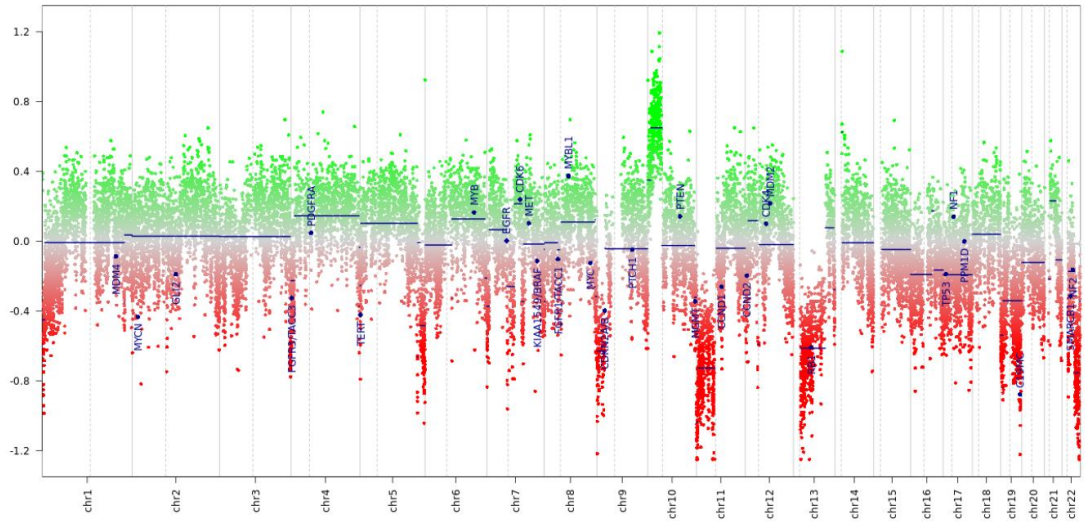


Figure 18. Results of the classification based on genome-wide DNA methylation profile for Tumour#1
At the top, the table with the methylation class attribution for the two separate components (left) and the summary of the CNV status (right). Central (GBM) and lower (PNC) panels are the depiction of chromosome 1 to 22 with the p-arm (left) and the q-arm (right) separated by a dotted line. Gains/amplifications represent positive, losses negative deviations from the baseline, with the 29 brain tumour relevant gene regions highlighted.

TUMOUR #2					
COMP.	**METHYLATION CLASS**		**CNV LOSS**	**CNV GAIN**	**CNV PROFILE**
GBM	**methylation class family Glioma; IDH mutant 0.99** IDH glioma, subclass high grade astrocytoma 0.91	Chr: Genes:	11; 13; 22 C19MC	10	**IDH MUTANT**
PNC	**methylation class family Glioma; IDH mutant 0.88** IDH glioma, subclass high grade astrocytoma 0.82	Chr: Genes:	7; 9; 11; 12; 16; 18; 19		**glioma**

GBM



PNC

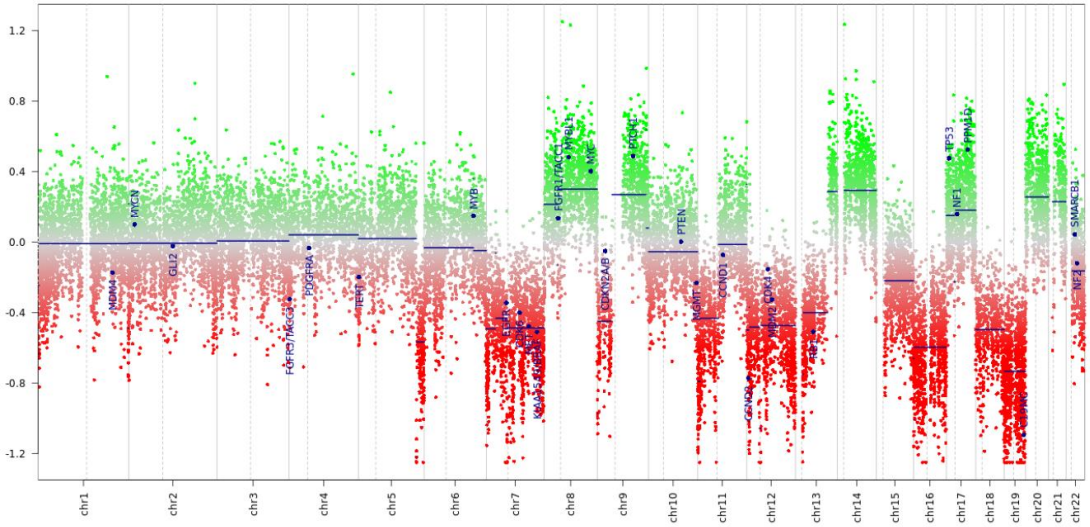


Figure 19. Results of the classification based on genome-wide DNA methylation profile for Tumour#2
At the top, the table with the methylation class attribution for the two separate components (left) and the summary of the CNV status (right). Central (GBM) and lower (PNC) panels are the depiction of chromosome 1 to 22 with the p-arm (left) and the q-arm (right) separated by a dotted line. Gains/amplifications represent positive, losses negative deviations from the baseline, with the 29 brain tumour relevant gene regions highlighted.

TUMOUR #3					
COMP.	**METHYLATION CLASS**		**CNV LOSS**	**CNV GAIN**	**CNV PROFILE**
GBM	**methylation class family Plexus Tumour 0.53** methylation class Plexus Tumour, subclass pediatric B 0.5	Chr: Genes:	10; 13; 22	MDM4; c-Myc, CDK4	**IDH WILDTYPE glioblastoma** RTKI/RTKII
PNC	**methylation class family Plexus Tumour 0.54** methylation class Plexus Tumour, subclass pediatric B 0.52	Chr: Genes:	10; 13; 22	MDM4; c-Myc, CDK4	

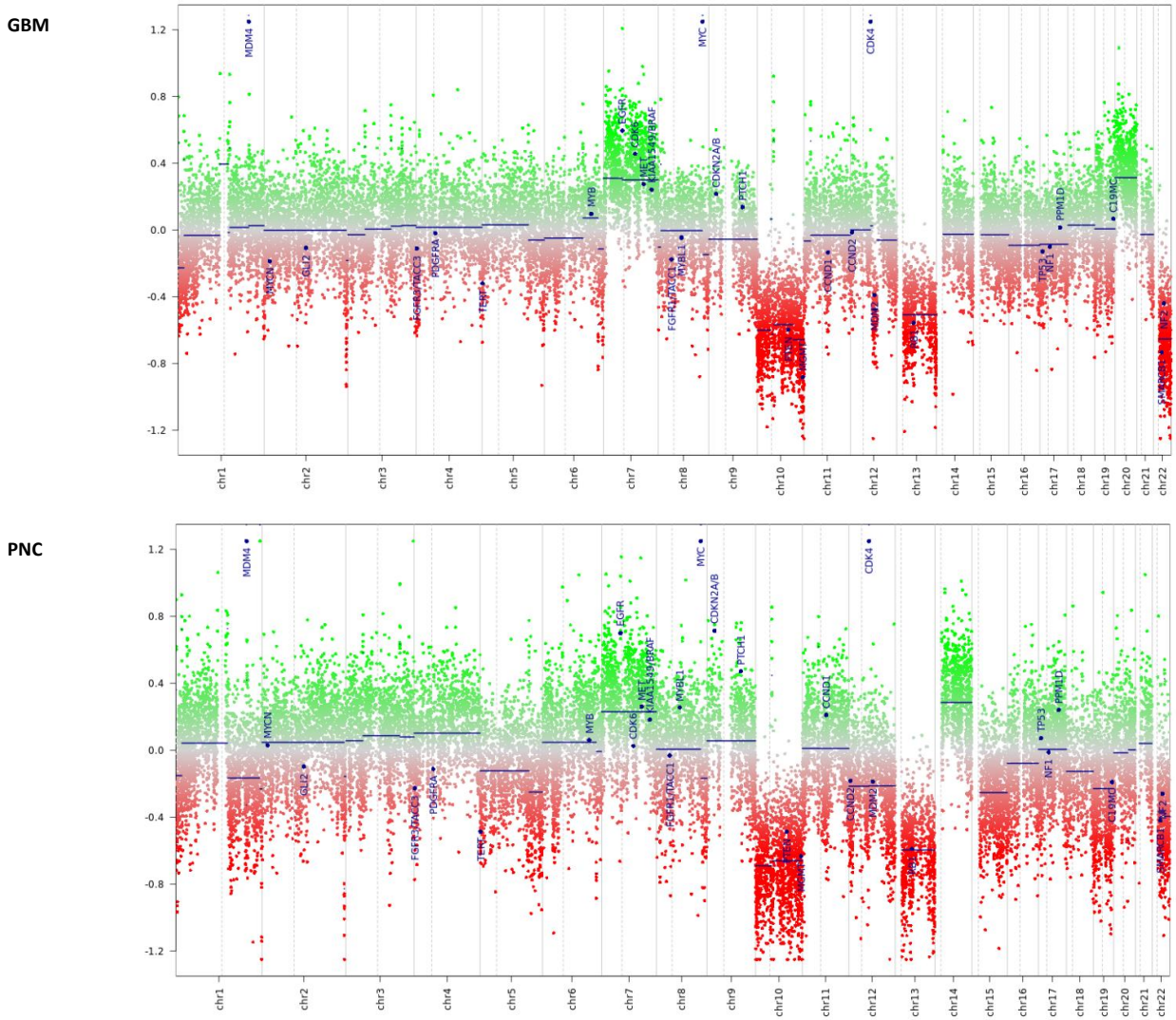
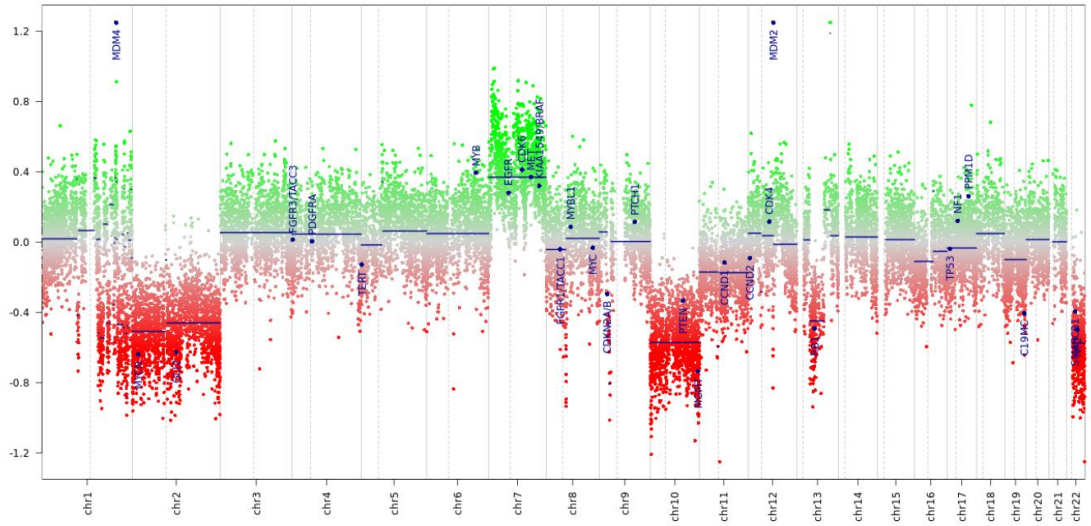


Figure 20. Results of the classification based on genome-wide DNA methylation profile for Tumour#3
At the top, the table with the methylation class attribution for the two separate components (left) and the summary of the CNV status (right). Central (GBM) and lower (PNC) panels are the depiction of chromosome 1 to 22 with the p-arm (left) and the q-arm (right) separated by a dotted line. Gains/amplifications represent positive, losses negative deviations from the baseline, with the 29 brain tumour relevant gene regions highlighted.

TUMOUR #4					
COMP.	**METHYLATION CLASS**		**CNV LOSS**	**CNV GAIN**	**CNV PROFILE**
GBM	**methylation class family Glioblastoma, IDH wildtype 0.38**	Chr:	1; 2; 10; 13; 22	7	
	methylation class glioblastoma, IDH wildtype, subclass RTK II 0.19	Genes:		MDM4; MDM2	
	methylation class glioblastoma, IDH wildtype, subclass RTK I 0.15				**IDH WILDTYPE**
PNC	**methylation class family Glioblastoma, IDH wildtype 0.53**	Chr:	1; 2; 10; 13; 22	7	**glioblastoma**
	methylation class glioblastoma, IDH wildtype, subclass RTK II 0.25	Genes:		MDM4; MDM2	RTKI/RTKII
	methylation class glioblastoma, IDH wildtype, subclass RTK I 0.23				

GBM



PNC

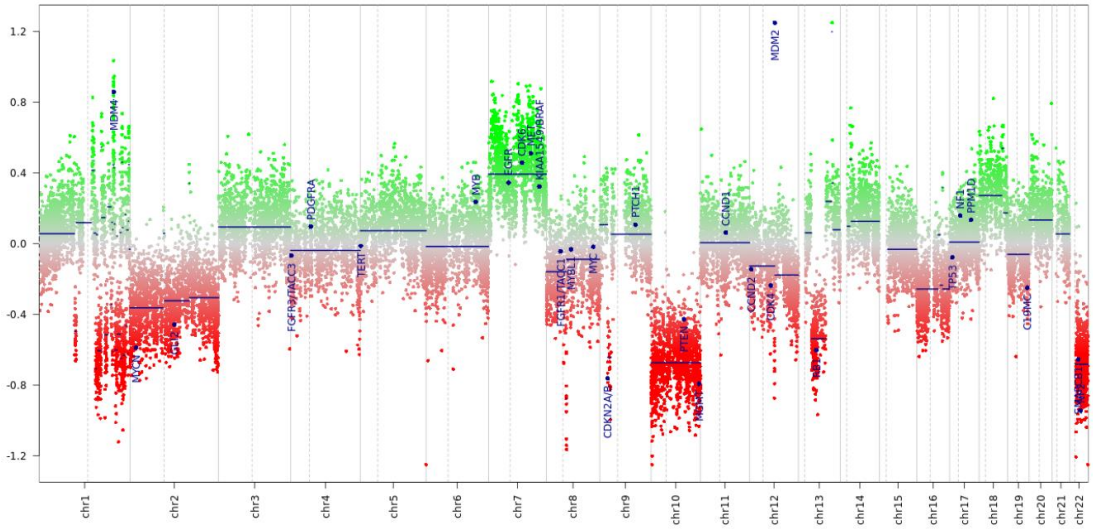


Figure 21. Results of the classification based on genome-wide DNA methylation profile for Tumour#4
At the top, the table with the methylation class attribution for the two separate components (left) and the summary of the CNV status (right). Central (GBM) and lower (PNC) panels are the depiction of chromosome 1 to 22 with the p-arm (left) and the q-arm (right) separated by a dotted line. Gains/amplifications represent positive, losses negative deviations from the baseline, with the 29 brain tumour relevant gene regions highlighted.

In summary, for all the samples analysed, it was possible to attribute a classification. Whereby the analysis of the methylation profile failed, the CNV analysis gave an indication.

It is noteworthy the fact that, despite their immunophenotypical diversity, GBM and PNC components were always classified in the same methylation cluster, even when the classification was unlike, suggesting the subsistence of the same methylation status. Also the CNV analyses of the two components showed a roughly comparable profile within the same patient, even though the PNC components generally showed a more unstable profile. This evidences suggest that the phenotypical differences between the two components have to be attributed to different genetics or to different transcriptional pathways, rather than epigenetic mechanisms.

The analyses of the genome-wide DNA methylation profiles and the CNV profiles assigned a classification even to samples that was not possible to attribute to a unique transcriptional profile (previously classified as HET). Interestingly, the 3 IDH-wildtype samples were assigned to the RTK I or RTK II subclasses, indistinguishable based on the CNV profile, but corresponding to the PN and the CL transcriptional subtypes respectively, that are the most represented subtypes in our sample cohort.

In conclusion, this integrated approach involving different methods of analysis allowed to classify these lesions known to be extremely heterogeneous.

Genetic Analysis by Next Generation Sequencing

In order to deeply investigate the nature of these lesions, we asked whether GBM and PNC components could harbour distinctive driver genetic alterations. To this aim, we performed a Next Generation Sequencing analysis in collaboration with the laboratory of Candiolo Cancer Institute and with the University of Torino. A custom NGS panel was designed to explore 75 genes that are relevant in GBM/PNC tumours, as described in Materials and Methods. We studied six sample pairs including the four pairs used for methylation analysis. Ten PBMCs from independent individuals were also sequenced using this custom NGS panel to build the reference sample for genetic analysis (see Materials and Methods). After DNA extraction and processing, the quality of final libraries was performed and equal amounts of libraries were pooled and sequenced on the Illumina MiSeq sequencer. After sequencing, the quality of library preparation and sequencing were evaluated using the following parameters: Number of READs sequenced; Percentage of enrichment on target; Coverage of depth; Median Depth and PCR deduplications (see Materials and Methods). Due to sequence quality, different limits of detection (LOD) were applied to each sample (Table 4).

Samples	READS	%_ENRICHMENT	DEPTH (X)	COVERAGE (%)	DUPLICATES (%)	LOD (VAF)
FFPE-Tumor#1-GBM	2827019	73,60%	1990	98,37%	29,98%	1,00%
FFPE-Tumor#1-PNC	2323603	71,55%	1359	98,35%	39,21%	1,00%
FFPE-Tumor#2-GBM	3494145	19,69%	375	98,38%	3,45%	1,30%
FFPE-Tumor#2-PNC	1861299	18,26%	171	98,43%	3,16%	2,90%
FFPE-Tumor#3-GBM	1745464	16,09%	131	98,20%	3,45%	3,80%
FFPE-Tumor#3-PNC	1598354	24,17%	110	98,29%	5,42%	4,50%
FFPE-Tumor#4-GBM	4074333	21,31%	334	98,39%	4,86%	1,50%
FFPE-Tumor#4-PNC	10066425	24,14%	1199	98,66%	7,87%	1,00%
FFPE-Tumor#5-GBM	1871502	72,89%	879	98,28%	45,16%	1,00%
FFPE-Tumor#5-PNC	1895464	72,67%	906	98,27%	43,40%	1,00%
FFPE-Tumor#6-GBM	3547456	67,52%	1745	98,35%	47,87%	1,00%
FFPE-Tumor#6-PNC	599275	68,34%	374	98,30%	30,70%	1,00%
PBMC-CT088	1472917	50,35%	966	98,47%	2,86%	n.a.
PBMC-CT102	1218991	51,52%	830	98,47%	2,67%	n.a.
PBMC-CT107	1462499	51,67%	968	98,62%	3,12%	n.a.
PBMC-CT112	1215084	51,24%	805	98,49%	2,65%	n.a.
PBMC-CT113	1226952	52,76%	813	98,47%	2,38%	n.a.
PBMC-CT122	1239679	50,76%	804	98,79%	2,90%	n.a.
PBMC-CT136	898745	51,08%	595	98,47%	2,02%	n.a.
PBMC-CT138	1292229	51,98%	860	98,74%	2,29%	n.a.
PBMC-CT139	1366623	51,18%	915	98,49%	2,47%	n.a.
PBMC-CTR004	1283396	50,55%	840	98,59%	2,62%	n.a.

Table 4. Post sequencing quality check
Quality check of raw data after sequencing. For each sample number of reads, percentage of target enrichment, average sequencing depth, target coverage, percentage of duplicate reads and limits of detection (LOD) are reported.

Identified likely clonal mutations i.e. displaying a variant allele frequency (VAF) higher than 10% are reported in Tables 5-10; copy number analysis is summarized in

Table 11. Moreover, likely subclonal mutations i.e displaying a VAF <10% are reported in Appendix 1-6.

As shown in Table 5, Tumour#1 harbours two independent *TP53* missense mutations and an *RB1* deletion leading to frameshift and ultimately to premature protein truncation. Regarding copy number alterations, Tumour#1 showed heterozygous loss of genes relevant for GBM diagnosis including *MGMT* and *TERT*. Interestingly, a copy of *EBF3* was lost in both GBM and PNC components (

Table 11). Private, low frequency variants identified in GBM and PNC components are listed in Appendix 1.

GENE	Mutation Coordinates (hg19)	Accession number	N Change	AA Change	Effect	COSMIC	GBM VAF (%)*	PNC VAF (%)*
RB1	chr13:48941637-48941642	NM_000321	c.948_951delTCTT	p. S318Nfs*13	frameshift	-	68,56	72,36
TP53	chr17:7578394	NM_000546	c.A536G	p.H179R	nonsynonymous	188	40,96	43,07
TP53	chr17:7578535	NM_000546	c.A395G	p.K132R	nonsynonymous	66	40,21	48,32

Table 5. Genetic alterations for Tumour#1
Summary of likely clonal mutations as identified for each of the two components of GBM-PNC in Tumour#1. Each mutation is described using: gene name, genomic coordinates (hg19), accession number, nucleotide change, amino acid change, number of occurrences in The Catalogue of Somatic Mutations in Cancer (COSMIC) database and VAF for each component.

Tumour#2 harbours 11 high frequency variants shared between the two components. As for Tumour#1, also in Tumour#2 *TP53* and *RB1* genes are altered in both components. Among the shared mutations, this tumour displays the already mentioned the *IDH1* p.R132H missense mutation and *ATRX* truncating alteration that may suggest a possible evolution from a lower grade glioma and should therefore be considered as a secondary GBM (Table 6). Indeed, no previous history of low grade glioma has been reported for this patient and the high genetic instability along with the alterations on the RB1 pathway indicate that this lesion has to be considered by default a Glioblastoma (Brat, Aldape et al. 2020). Private, low frequency variants identified in GBM and PNC components are listed in Appendix 2.

GENE	Mutation Coordinates (hg19)	Accession number	N Change	AA Change	Effect	COSMIC	GBM VAF (%)*	PNC VAF (%)*
ATRX	chrX:76888721-76888726	NM_000489	c.5104_5107delGAAA	p.E1702Yfs*22	frameshift	-	70,41	88,41
ERBB2	chr17:37855834	NM_001289936	c.C22A	p.P8T	nonsynonymous	-	51,20	42,22
IDH1	chr2:209113112	NM_001282387	c.G395A	p.R132H	nonsynonymous	3611	36,39	48,86
MLH1	chr3:37061893	NM_000249	c.T977C	p.V326A	nonsynonymous	-	52,23	65,34
MSH3	chr5:79950733	NM_002439	c.C187G	p.P63A	nonsynonymous	-	36,36	51,16
NTRK1	chr1:156851418	NM_002529	c.T2375A	p.L792Q	nonsynonymous	-	45,35	64,84
PIK3CA	chr3:178916937-178916941	NM_006218	c.325_327delGAA	p.E109del	in frame deletion	11	38,75	47,22
PIK3R4	chr3:130442401	NM_014602	c.C1838T	p.S613F	nonsynonymous	-	51,07	32,99
RB1	chr13:48951086-48951089	NM_000321	c.1249_1250delAA	p. R418Sfs*9	frameshift	-	13,68	45,35
TP53	chr17:7577526-7577527	NM_000546	c.775_776insGGAT	p.D259Gfs*6	frameshift	-	31,92	26,45
TP53	chr17:7577568	NM_000546	c.G713T	p.C238F	nonsynonymous	132	60,87	68,69

Table 6. Genetic alterations for Tumour#2
Summary of likely clonal mutations as identified for each of the two components of GBM-PNC in Tumour#2. Each mutation is described using: gene name, genomic coordinates (hg19), accession number, nucleotide change, amino acid change, number of occurrences in The Catalogue of Somatic Mutations in Cancer (COSMIC) database and VAF for each component.

Tumour#3 harbours high frequency shared mutations for *TP53*, *EXO1* and MSH3 (Table 7) and amplifications in *MDM4*, *MYC* and *CDK4* genes (Table 11). CDK4, in complex with cyclins, displaces RB1 from E2F leading to cell cycle progression. CDK4 amplification may mimic RB1 loss of function alterations (Knudsen, Nambiar et al. 2020). Interestingly GBM component harbours 2 private high frequency variants in *PTEN* and *EGFR* genes. On the contrary, the PNC component of Tumour#3 displays an exclusive huge *AKT3* gene amplification. Taken together these data suggest that

Tumour#3 arises from a common ancestor characterized by the presence of a set of point mutations and amplifications than other genomic lesions were acquired after the emergence of the GBM (*EGFR* and *PTEN* mutations) and PNC (*AKT3* amplification) components. Private, low frequency variants identified in GBM and PNC components are listed in Appendix 3.

GENE	Mutation Coordinates (hg19)	Accession number	N Change	AA Change	Effect	COSMIC	GBM VAF (%)*	PNC VAF (%)*
TP53	chr17:7578458	NM_000546	c.C472G	p.R158G	nonsynonymous	40	78,77	98,63
EXO1	chr1:242042364	NM_130398	c.A1828G	p.S610G	nonsynonymous	-	49,34	43,44
MSH3	chr5:79966024	NM_002439	c.A688G	p.T230A	nonsynonymous	-	47,24	56,38
EGFR	chr7:55229241	NM_005228	c.G1548T	p.W516C	nonsynonymous	-	18,99	-
PTEN	chr10:89692904-89692906	NM_000314	c.389_389delG	p. R130Qfs*4	frameshift	-	35,96	-

Table 7 Genetic alterations for Tumour#3
Summary of likely clonal mutations as identified for each of the two components of GBM-PNC in Tumour#2. Each mutation is described using: gene name, genomic coordinates (hg19), accession number, nucleotide change, amino acid change, number of occurrences in The Catalogue of Somatic Mutations in Cancer (COSMIC) database and VAF for each component.

Tumour#4 displayed mutations in *PTEN* and *MYC* and *RB1* genes in both components (Table 8). Copy number gains and/or amplification of *MDM2* and *MDM4* have been observed in both components, confirming previous observation obtained by CNV analysis of DNA methylation-assay. Of note, this patient was previously diagnosed with a well-differentiated liposarcoma showing *MDM2* amplification, suggesting the possible presence of a germline alteration predisposing to gene amplifications or a germline fragility of the *MDM2* locus.
Private, low frequency variants identified in GBM and PNC components are listed in Appendix 4.

GENE	Mutation Coordinates (hg19)	Accession number	N Change	AA Change	Effect	COSMIC	GBM VAF (%)*	PNC VAF (%)*
MYC	chr8:128750527	NM_002467	c.T64C	p.F22L	nonsynonymous	-	45,09	45,46
PTEN	chr10:89624305	NM_001304717	c.T598A	p.Y200N	nonsynonymous	5	75,83	57,95
RB1	chr13:49033949-49033950	NM_000321	c.2086_2087insTCAT	p.R696Ifs*26	frameshift	-	59,56	36,89

Table 8. Genetic alterations for Tumour#4
Summary of likely clonal mutations as identified for each of the two components of GBM-PNC in Tumour#4. Each mutation is described using: gene name, genomic coordinates (hg19), accession number, nucleotide change, amino acid change, number of occurrences in The Catalogue of Somatic Mutations in Cancer (COSMIC) database and VAF for each component.

Tumour#5 displayed high frequency shared mutations in *ERBB2*, *PDGFRA* and *EXO1*, a multi exonic deletion of RB1 (Table 9) and *MDM2* amplification. Interestingly, as reported in Appendix 5, Tumour#5 displays a huge number of low frequency variants (VAF<10%): we identified 58 variants shared by both components, 528 and 507 GBM and PNC-component specific alterations respectively (Appendix 5).

GENE	Mutation Coordinates (hg19)	Accession number	N Change	AA Change	Effect	COSMIC	GBM VAF (%)*	PNC VAF (%)*
ERBB2	chr17:37871547	NM_004448	c.C1157A	p.A386D	nonsynonymous	123	61,34	64,94
EXO1	chr1:242023871	NM_130398	c.A809T	p.D270V	nonsynonymous	-	38,89	39,76
PDGFRA	chr4:55146520	NM_006206	c.A2194G	p.M732V	nonsynonymous	-	42,09	38,70
RB1	na	NM_000321	c.?	p.0?	multiexonic deletion	-	na	na

Table 9. Genetic alterations for Tumour#5

Summary of likely clonal mutations as identified for each of the two components of GBM-PNC in Tumour#5. Each mutation is described using: gene name, genomic coordinates (hg19), accession number, nucleotide change, amino acid change, number of occurrences in The Catalogue of Somatic Mutations in Cancer (COSMIC) database and VAF for each component.

Tumour#6 display *NF1, TP53, PTEN, MET, MLH3, EXO5, RB1* and *PDGRFA* shared and likely clonal mutations. Tumour#6 GBM component harbours a private high frequency alteration in POLD1 gene (Table 10). As reported for Tumour#5, Tumour#6 displays a huge number of low frequency variants (VAF<10%): we identified 2 variants shared by both components, 83 and 597 GBM- and PNC-component specific alterations respectively (Appendix 6).

GENE	Mutation Coordinates (hg19)	Accession number	N Change	AA Change	Effect	COSMIC	GBM VAF (%)*	PNC VAF (%)*
EXO5	chr1:40980826	NM_022774	c.A610G	p.M204V	nonsynonymous	-	48,06	50,38
MET	chr7:116340214	NM_001127500	c.G1076A	p.R359Q	nonsynonymous	1	68,44	61,79
MLH3	chr14:75513222	NM_001040108	c.G3137A	p.R1046Q	nonsynonymous	-	59,54	97,50
NF1	chr17:29552261	NM_001042492	c.C1994T	p.S665F	nonsynonymous	-	89,07	96,49
PDGFRA	chr4:55127448	NM_006206	c.G236A	p.G79D	nonsynonymous	-	44,93	25,14
POLD1	chr19:50918229	NM_001256849	c.G2546A	p.R849H	nonsynonymous	-	41,15	-
PTEN	chr10:89692904	NM_001304717	c.C907T	p.R303*	stopgain	168	73,65	94,96
RB1	chr13:48916833-48916834	NM_000321	c.363_364insA	p. N123Kfs*8	frameshift	-	42,11	47,14
TP53	chr17:7577538	NM_000546	c.G743A	p.R248Q	nonsynonymous	842	78,23	95,19

Table 10. Genetic alterations for Tumour#6
Summary of likely clonal mutations as identified for each of the two components of GBM-PNC in Tumour#6. Each mutation is described using: gene name, genomic coordinates (hg19), accession number, nucleotide change, amino acid change, number of occurrences in The Catalogue of Somatic Mutations in Cancer (COSMIC) database and VAF for each component.

In GBMs, a hypermutator phenotype has been associated with TMZ-induced mutagenesis due to inactivation of the DNA mismatch repair pathway (MSH6, MSH2, MHS4, MSH5, PMS1, PMS2, MLH1, and MLH2). Notably, a rare subset of pre-treatment adult glioma patients with *de novo* hypermutator phenotype was identified by Sa et al. (Sa, Choi et al. 2019). Due to the high number of subclonal mutations found in Tumour#5 and Tumour#6, we can hypothesize a hypermutator phenotype. As reported by Sa and colleagues, Tumour#5 and Tumour#6 lack somatic mutations of IDH1 and MGMT promoter methylation and harbour different mutations in DNA repair encoding genes.

Furthermore,as shown in Figure 22, *RB1* gene is mutated or deleted in 5 Tumours out of 6 (83.3% of analysed samples) while in general GBM cohorts reported in literature only 12,8% of GBMs display *RB1* mutations (Jonsson, Lin et al. 2019). Interestingly the remaining Tumour (#3) harbours a *CDK4* gene amplification that, from a functional point of view, may mimic *RB1* loss (Knudsen, Nambiar et al. 2020). This mutation is present only in 12,8% of the tumours in a general GBM cohort (Jonsson, Lin et al. 2019). Noteworthy, 10-15% of children harbouring germline *RB1* mutations develop brain PNETs, commonly in the pineal gland (Saab, Rodriguez-Galindo et al. 2009, Miller, Rogers et al. 2011). Taken together these preliminary observations suggest that the *RB1* alterations

may play a key role for the origin of GBM-PNC, in contrast with GBMs that do not develop a PNC component.

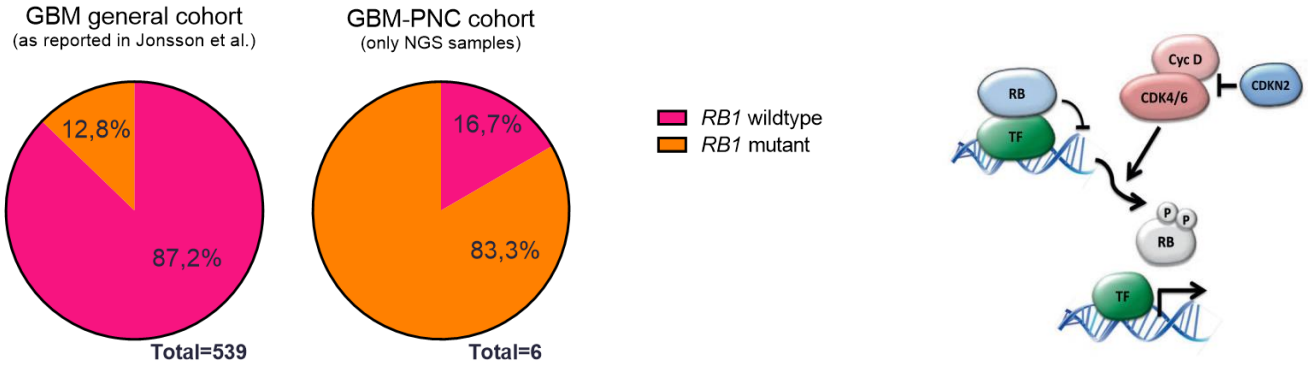


Figure 22. Distribution of *RB1* mutations in general GBM cohort and in our GBM-PNC cohort
Plots show the frequency of *RB1* mutations in a general GBM cohort (as reported by Jonsson et al.), on the left, and in the 6 samples analysed by NGS from our sample cohort, on the right. The frequencies in percentage for each subtype are indicated. Of note, the only *RB1* wildtype sample form our cohort harbours a *CDK4* gene amplification that, from a functional point of view, may mimic *RB1* loss. On the right, generic depiction of the core RB-pathway. CDK4/6 activity is stimulated by D-type cyclins, which can be inhibited by members of the CDKN2 gene family. CDK4/6 activity converges on RB-family members to mediate phosphorylation and functional inactivation (adapted from Knudesn et al.).

				Variation of allele Number											
Chr	Start	Stop	Gene	#1 GBM	#1 PNC	#2 GBM	#2 PNC	#3 GBM	#3 PNC	#4 GBM	#4 PNC	#5 GBM	#5 PNC	#6 GBM	#6 PNC
chr1	40980217	40981338	EXO5	2,63	2,87	1,73	1,34	1,48	1,29	2,14	1,86	1,86	1,78	1,85	1,81
chr1	76262671	76378572	MSH4	4,63	5,81	3,16	2,01	2,42	2,86	2,35	2,72	2,88	2,91	3,85	2,82
chr1	78414451	78444688	FUBP1	4,72	5,85	3,47	2,61	2,91	3,06	3,05	2,61	3,11	3,23	3,64	3,26
chr1	156785622	156851434	NTRK1	1,75	1,93	1,39	2,84	1,80	1,63	1,76	1,79	1,28	1,19	1,15	1,78
chr1	204494647	204518810	MDM4	3,81	4,45	2,69	2,04	14,56	41,17	4,98	10,11	3,04	3,08	3,01	2,71
chr1	226252053	226259180	H3F3A	4,07	4,34	3,01	2,13	2,33	2,31	1,57	1,95	3,10	3,27	3,36	3,68
chr1	242013728	242052902	EXO1	3,81	4,53	2,46	1,51	1,94	1,87	2,43	2,43	2,94	2,98	3,26	2,21
chr1	243663045	244006472	AKT3	4,75	5,71	3,79	2,57	3,70	133,47	3,04	2,87	3,73	3,79	4,62	3,56
chr2	16080821	16086219	MYCN	1,02	0,84	1,30	2,59	1,64	1,70	1,02	1,16	1,00	1,07	0,75	1,39
chr2	25457148	25536853	DNMT3A	1,67	1,43	1,98	3,53	2,09	2,26	1,48	1,47	1,72	1,75	1,18	1,78
chr2	47630331	47710088	MSH2	2,91	2,87	2,34	1,79	2,08	2,36	1,76	1,91	4,11	3,72	2,74	1,82
chr2	48010373	48033999	MSH6	2,28	2,22	1,85	1,20	1,39	1,39	1,41	1,49	3,14	2,86	2,05	1,41
chr2	158594043	158656005	ACVR1	2,76	2,80	2,20	1,44	1,66	1,56	2,06	1,72	3,40	3,63	2,92	1,77
chr2	190656536	190742162	PMS1	3,01	2,91	2,96	1,70	2,08	2,23	1,79	2,02	3,24	3,61	3,37	2,00
chr2	209101803	209116275	IDH1	2,82	3,02	2,25	1,49	1,81	1,80	1,82	1,85	3,52	3,77	2,99	1,81
chr2	212248340	213403254	ERBB4	2,92	2,88	2,42	1,49	1,93	1,86	2,06	1,97	3,38	3,55	2,99	1,91
chr2	227659726	227663454	IRS1	1,07	0,86	1,23	2,53	1,37	1,38	1,00	1,04	1,03	0,98	0,87	1,13
chr3	37035039	37092144	MLH1	3,66	3,58	2,42	1,75	2,31	2,29	2,78	2,72	3,74	3,92	3,31	1,57
chr3	130398159	130464062	PIK3R4	3,02	2,74	2,51	1,50	2,01	2,09	2,47	2,35	3,05	3,18	3,21	2,91
chr3	138374231	138478185	PIK3CB	3,28	3,25	2,64	1,60	2,15	2,36	2,44	2,56	3,16	3,43	4,01	3,34
chr3	178916614	178952152	PIK3CA	3,44	3,43	3,01	1,67	2,13	2,40	2,44	2,64	3,41	3,75	4,18	3,52
chr3	186507758	186522502	RFC4	3,37	3,25	2,55	1,71	2,02	2,09	2,35	2,50	3,70	3,82	3,82	3,49
chr4	1795662	1808989	FGFR3	1,07	0,75	0,90	2,68	1,62	2,36	1,32	1,40	0,84	0,78	0,67	1,34
chr4	39290381	39367861	RFC1	2,95	2,64	2,71	2,03	2,56	2,76	2,35	2,31	2,83	2,85	2,72	2,71
chr4	55124936	55161439	PDGFRA	2,30	2,05	1,92	1,27	1,50	1,34	1,99	2,10	2,11	2,05	2,19	2,01
chr5	1253843	1295104	TERT	1,07	1,08	0,82	2,75	1,49	1,69	1,51	1,58	1,04	1,02	0,75	0,98
chr5	67522504	67593429	PIK3R1	3,00	3,74	2,68	1,70	2,36	2,06	2,20	2,46	2,59	2,61	3,37	2,12
chr5	79950547	80171681	MSH3	3,14	4,01	2,70	1,78	2,27	2,13	2,22	2,55	2,70	2,64	3,45	2,08
chr5	176516604	176524677	FGFR4	1,26	1,34	1,04	2,19	1,67	1,78	1,73	1,74	0,90	0,83	0,92	1,48
chr6	26031878	26032288	HIST1H3B	1,86	1,56	2,98	4,01	2,07	3,25	1,96	1,92	1,44	1,33	1,97	4,14
chr6	31708244	31730308	MSH5	2,07	2,26	1,83	1,81	1,69	1,51	2,04	2,00	2,05	2,01	1,98	2,71
chr6	33286520	33290691	DAXX	1,83	1,72	2,28	2,66	1,90	1,68	1,82	1,76	1,60	1,59	1,80	2,87
chr7	6013030	6048650	PMS2	2,68	2,69	2,11	1,41	2,68	2,31	3,01	2,91	2,02	1,92	3,21	2,70
chr7	44154384	44163158	POLD2	1,42	1,27	1,59	2,23	2,44	2,62	2,68	2,35	1,62	1,48	1,72	3,14
chr7	55086971	55273310	EGFR	2,18	2,21	2,29	1,80	3,83	3,36	3,00	2,68	2,62	2,46	2,91	2,98
chr7	73646436	73668713	RFC2	2,54	2,30	2,10	1,68	2,75	2,48	3,34	3,18	2,83	2,69	2,76	3,05
chr7	116339139	116436178	MET	2,73	2,69	1,99	0,95	2,33	2,26	3,30	3,09	3,49	3,18	3,81	2,79
chr7	140434397	140624503	BRAF	3,28	3,44	2,74	1,40	3,31	3,21	3,55	3,43	4,01	4,00	4,85	3,30
chr8	38271146	38318624	FGFR1	1,72	1,66	1,72	2,27	1,38	1,31	1,87	1,80	1,30	1,20	1,50	3,18
chr8	128748840	128753204	MYC	1,19	1,14	2,14	3,07	19,75	79,23	1,56	1,56	0,74	0,73	1,08	2,91
chr9	21968228	21994330	CDKN2A	1,30	1,15	1,07	1,90	1,74	2,26	0,78	0,94	0,69	0,60	1,54	2,99
chr9	22005986	22008952	CDKN2B	1,44	1,18	1,56	2,54	2,55	2,28	1,01	1,07	0,72	0,65	1,81	3,42
chr9	87285664	87636352	NTRK2	2,24	2,43	2,78	2,39	1,91	1,93	2,24	2,10	1,53	1,34	3,09	2,74
chr9	139390523	139440238	NOTCH1	1,13	0,96	1,53	3,97	1,79	2,16	1,66	1,66	0,64	0,53	1,01	2,19
chr10	89623707	89725229	PTEN	2,21	2,56	3,03	1,87	1,49	1,34	1,52	1,89	2,16	2,27	2,80	2,59
chr10	123239095	123353331	FGFR2	1,51	1,59	2,04	1,67	1,06	0,89	1,32	1,43	1,34	1,37	1,37	1,70
chr10	131265480	131565261	MGMT	1,18	1,09	1,76	2,90	1,17	1,13	0,99	1,20	0,93	0,89	0,98	1,68
chr10	131636193	131762032	EBF3	1,02	1,09	1,47	2,58	1,24	1,03	0,98	1,19	1,09	1,06	0,93	1,78
chr11	67119438	67120870	POLD4	1,85	1,47	2,39	3,82	2,53	2,79	2,06	2,09	2,10	1,78	1,51	2,85
chr11	74303704	74351811	POLD3	2,83	2,81	2,26	1,93	2,34	2,47	2,62	2,18	3,69	3,97	2,52	2,83
chr12	56474085	56495839	ERBB3	1,95	1,86	1,66	1,08	1,45	1,02	1,93	2,02	2,07	2,00	1,62	1,11
chr12	58142308	58145500	CDK4	1,94	1,68	1,95	1,26	16,66	36,95	1,88	1,93	2,09	2,06	1,72	1,20
chr12	69202258	69233629	MDM2	3,00	2,87	3,01	1,79	1,74	2,34	94,35	57,72	63,11	61,27	2,97	1,62
chr12	118454633	118469083	RFC5	2,64	2,25	1,90	1,20	1,74	1,32	1,90	2,20	2,56	2,56	1,83	1,16
chr12	133201283	133263901	POLE	1,70	1,37	1,61	1,88	1,71	1,37	1,70	1,71	1,52	1,46	1,04	1,03
chr13	34392316	34540262	RFC3	2,31	2,16	1,65	1,34	1,36	1,41	2,65	2,52	3,64	3,77	2,18	2,06
chr13	48878049	49054207	RB1	2,53	2,60	2,07	1,66	1,79	1,65	1,39	1,79	1,16	1,35	2,94	2,54
chr14	75483785	75516358	MLH3	2,20	2,26	2,06	1,73	1,68	2,22	2,20	2,16	3,10	3,56	2,52	1,69
chr14	105236678	105258980	AKT1	1,15	1,07	1,19	3,43	1,47	2,18	1,83	1,70	1,34	1,36	0,97	1,39
chr15	88420166	88799384	NTRK3	2,05	2,15	2,11	1,74	1,53	1,05	2,11	2,09	2,20	2,04	1,88	2,29
chr15	89859982	89876985	POLG	1,57	1,49	2,03	2,27	1,57	1,21	1,72	1,61	1,46	1,38	1,31	1,93
chr15	90627498	90645622	IDH2	1,63	1,62	1,93	2,06	1,74	1,13	1,77	1,77	1,62	1,47	1,40	1,93

				Variation of allele Number											
Chr	Start	Stop	Gene	#1 GBM	#1 GBM	#2 GBM	#2 GBM	#3 GBM	#3 GBM	#4 GBM	#4 GBM	#5 GBM	#5 GBM	#6 GBM	#6 GBM
chr17	7572927	7579912	TP53	1,73	1,69	1,92	3,11	1,84	1,96	2,07	1,86	2,30	2,13	1,06	1,67
chr17	29422328	29701173	NF1	2,94	3,22	2,65	2,03	2,15	2,26	2,53	2,48	3,52	3,81	1,90	1,91
chr17	37855813	37884297	ERBB2	1,53	1,51	1,53	3,24	1,74	1,74	1,72	1,76	1,68	1,68	0,78	1,34
chr19	11094828	11172492	SMARCA4	1,57	1,48	1,87	2,61	3,10	2,25	1,96	1,89	2,03	2,02	1,08	1,23
chr19	40739779	40771174	AKT2	1,42	1,37	1,45	2,00	2,28	1,52	1,75	1,65	1,51	1,54	1,15	0,96
chr19	42776146	42799343	CIC	1,31	1,11	1,36	2,38	1,75	1,45	1,71	1,66	1,24	1,22	0,85	1,05
chr19	50902109	50921204	POLD1	1,23	1,11	1,05	2,16	1,77	1,42	1,69	1,71	1,17	1,20	0,86	0,81
chr20	5095929	5100444	PCNA	3,07	3,13	2,89	3,19	2,92	2,92	1,97	2,57	4,90	5,28	3,40	2,40
chr22	24129357	24176367	SMARCB1	1,13	1,17	2,08	2,83	1,13	1,25	1,10	1,32	1,68	1,60	1,38	1,25
chr22	29999988	30090791	NF2	1,40	1,64	1,85	1,82	1,12	1,29	1,18	1,45	2,22	2,14	1,91	1,26
chrX	76763829	77041487	ATRX	2,00	2,91	2,59	2,35	4,32	4,04	4,08	3,89	2,79	2,95	4,70	3,24

Table 11. Copy number analysis for Tumours#1-6

Copy Number Variation as evaluated for each of the two components of GBM-PNC in Tumours#1-6. For each gene, chromosome location is shown. Significant reduction of allele number (<1,2) and amplifications are highlighted in blue and red, respectively.

APPENDIX

GBM subclonal variants (VAF<10%)*							
DNMT3A c.G137A	FGFR1 c.G1651A	POLD1 c.C2825T	PTEN c.G544A	TERT c.G677A			

PNC subclonal variants (VAF<10%)*							
AKT2 c.C1292T	AKT2 c.G1195A	CDK4 c.G629A	CIC c.G7049A	EBF3 c.C101T	EBF3 c.G65A	ERBB3 c.G1469A	FGFR2 c.G443A
FGFR3 c.G1237A	FGFR3 c.G1343T	FGFR3 c.G2266A	FGFR3 c.G796A	FGFR3 c.G808A	FGFR3 c.G958A	IDH2 c.C1097T	IDH2 c.C16T
IDH2 c.G1039A	IRS1 c.G115A	IRS1 c.G974A	MGMT c.C43T	MGMT c.G710A	MYCN c.C653T	NOTCH1 c.G1511A	NOTCH1 c.G2605A
NOTCH1 c.G4972A	NOTCH1 c.G7591A	NTRK1 c.G220A	POLD1 c.C2987T	POLD1 c.G2724T	POLE c.G901A	PTEN c.C463A	TERT c.C1967T
TERT c.G2053A	TP53 c.C689T						

Appendix 1. Subclonal alterations identified in Tumour#1
List of private likely subclonal variations identified in Tumour#1 GBM and PNC components. *only variants showing VAF> of LOD are reported.

GBM subclonal variants (VAF<10%)*							
AKT1 c.C1328T	ATRX c.C4857A	ATRX c.C7156T	ATRX c.G1930A	ATRX c.G5104T	CIC c.C4580T	CIC c.C5446T	EGFR c.G1123A
FGFR4 c.G2315A	MSH4 c.G2699A	MSH6 c.G2539A	NF1 c.G4832A	NOTCH1 c.G1094A	PIK3R1 c.G1946A	POLD2 c.C1437A	POLD4 c.C242T
POLG c.A128G	RFC3 c.T245A						

PNC subclonal variants (VAF<10%)*							
ERBB4 c.G2110A	ERBB4 c.G956A	MLH3 c.G25A	PIK3CA c.G646A	PIK3R4 c.G2015A	POLG c.G1328A		

Appendix 2. Subclonal alterations identified in Tumour#2
List of private and likely subclonal variations identified in Tumour#2 GBM and PNC components. *only variants showing VAF> of LOD are reported.

GBM subclonal variants (VAF<10%)*							
FGFR3 c.G1543A	FGFR3 c.G524A	MLH1 c.G1435A	MSH6 c.C3254T	NTRK2 c.G2226A	PIK3CB c.C487T	PIK3R4 c.C2626T	POLE c.G6136A

Appendix 3. Subclonal alterations identified in Tumour#3
List of private and likely subclonal variations identified in Tumour#3 GBM component. *only variants showing VAF> of LOD are reported.

GBM subclonal variants (VAF<10%)*							
FGFR4 c.C214T	IDH2 c.C444A	NTRK2 c.A1397G					

PNC subclonal variants (VAF<10%)*							
IRS1 c.A812T	NTRK3 c.C380T						

Appendix 4. Subclonal alterations identified in Tumour#4
List of private and likely subclonal variations identified in Tumour#4 GBM and PNC components. *only variants showing VAF> of LOD are reported.

Common subclonal variants (VAF<10%)*							
AKT1 c.G1246A	AKT1 c.G451A	AKT2 c.G535A	AKT2 c.G1040A	CDKN2A c.C238T	CDKN2A c.C443T	CIC c.G6401A	DNMT3A c.C286T
EGFR c.C2408T	ERBB2 c.C2153T	ERBB3 c.G1331A	ERBB3 c.G3620A	FGFR2 c.G994A	FGFR2 c.G814A	FGFR2 c.G608A	FGFR3 c.C247T
FGFR3 c.G1930A	FGFR3 c.G196A	FGFR3 c.C188T	FGFR3 c.C256T	FGFR4 c.C1675T	FGFR4 c.G640A	FGFR4 c.G1660A	IDH2 c.G529A
IDH2 c.G266A	IRS1 c.C265T	IRS1 c.G1729A	IRS1 c.G3266A	MYC c.C31T	MYCN c.G538A	NF1 c.G37A	NOTCH1 c.G3271A
NOTCH1 c.G608A	NOTCH1 c.G524A	NOTCH1 c.C334T	NOTCH1 c.C64T	NOTCH1 c.G5725A	NOTCH1 c.G2813A	NOTCH1 c.C6749T	NOTCH1 c.G1801A
NOTCH1 c.G775A	NOTCH1 c.G701A	NOTCH1 c.G1348A	NOTCH1 c.C2338T	NTRK2 c.C1994T	NTRK2 c.G1849A	PIK3R3 c.G535A	PMS2 c.G2347A
PMS2 c.G1433A	POLD1 c.G3245A	POLD1 c.C211T	POLD1 c.G2624A	POLE c.G5609A	POLG c.G2420A	RB1 c.C58T	RFC3 c.C25T
SMARCA4 c.G2611A	SMARCA4 c.G2833A						

GBM subclonal variants (VAF<10%)*							
ACVR1 c.C919T	ACVR1 c.G563A	AKT1 c.G955A	AKT1 c.G143A	AKT1 c.G1195A	AKT1 c.G322A	AKT1 c.C520T	AKT1 c.G880A
AKT1 c.C1084T	AKT1 c.G808A	AKT1 c.G362A	AKT2 c.C226T	AKT2 c.G605A	ATRX c.G6103A	ATRX c.G4189A	ATRX c.G2342A
ATRX c.G1961A	ATRX c.G4639A	ATRX c.C6253T	ATRX c.C6532T	ATRX c.G7069A	ATRX c.G1451A	ATRX c.G7200A	ATRX c.G4202A
BRAF c.C1525T	BRAF c.C110T	CDK4 c.C515T	CDK4 c.C847T	CDK6 c.G700A	CDKN2A c.C56T	CDKN2A c.G445A	CDKN2A c.C319T
CDKN2A c.G226A	CDKN2A c.C104T	CDKN2A c.G250A	CDKN2A c.C257T	CDKN2A c.C292T	CDKN2A c.C385T	CDKN2A c.G373A	CDKN2A c.C218T
CDKN2A c.G167A	CDKN2B c.G38A	CDKN2B c.C305T	CDKN2B c.G302A	CIC c.G3433A	CIC c.G4493A	CIC c.C6920T	CIC c.G4732A
CIC c.C7457T	CIC c.G5824A	CIC c.C5314T	CIC c.C3179T	CIC c.G4594A	CIC c.G3152A	CIC c.C6898T	CIC c.C4916T
CIC c.C4853T	CIC c.C5057T	CIC c.G6509A	CIC c.G4318A	CIC c.G5983A	CIC c.G4001A	CIC c.C6475T	CIC c.C32T
CIC c.C4135T	CIC c.G3124A	CIC c.C6290T	DAXX c.G824A	DAXX c.G1480A	DAXX c.C1868T	DAXX c.C560T	DAXX c.G79A
DAXX c.G1078A	DNMT3A c.G841A	DNMT3A c.G379A	DNMT3A c.C1193T	DNMT3A c.C1046T	DNMT3A c.G1585A	DNMT3A c.C557T	DNMT3A c.G1373A
DNMT3A c.G2225A	EBF3 c.G32T	EBF3 c.G29A	EBF3 c.C94T	EBF3 c.G93A	EBF3 c.C935T	EBF3 c.G83A	EBF3 c.C647T
EBF3 c.G257A	EBF3 c.C1190T	EBF3 c.G1603A	EGFR c.C664T	EGFR c.G1592A	EGFR c.C2222T	EGFR c.C445T	EGFR c.C2645T
ERBB2 c.G40A	ERBB2 c.G61A	ERBB2 c.C649T	ERBB2 c.G3143A	ERBB2 c.C1955T	ERBB2 c.G3034A	ERBB2 c.G2524A	ERBB2 c.C917T
ERBB2 c.C2686T	ERBB3 c.G355A	ERBB3 c.C241T	ERBB3 c.C508T	ERBB3 c.C2567T	ERBB3 c.C2578T	ERBB3 c.G2921A	ERBB3 c.G587A
ERBB3 c.C1276T	ERBB3 c.C2510T	ERBB3 c.C637T	ERBB4 c.G58A	FGFR1 c.G2519A	FGFR1 c.G2558A	FGFR1 c.C1121T	FGFR1 c.G2365A
FGFR1 c.G64A	FGFR1 c.G421A	FGFR1 c.G403A	FGFR1 c.C301T	FGFR1 c.C2402T	FGFR1 c.G50A	FGFR1 c.C253A	FGFR1 c.G160A
FGFR1 c.C1784T	FGFR2 c.G1845A	FGFR2 c.C2224T	FGFR2 c.G202A	FGFR2 c.C241T	FGFR2 c.C290T	FGFR2 c.G1733A	FGFR2 c.G1721A
FGFR3 c.C2204T	FGFR3 c.G539A	FGFR3 c.G1879A	FGFR3 c.G940A	FGFR3 c.G670A	FGFR3 c.G1681A	FGFR3 c.G286A	FGFR3 c.G1925A
FGFR3 c.C667T	FGFR3 c.G1115A	FGFR3 c.G901A	FGFR3 c.G1451A	FGFR3 c.G1817A	FGFR3 c.G2134A	FGFR3 c.C1415T	FGFR3 c.C167T
FGFR3 c.C1252T	FGFR3 c.G131A	FGFR3 c.G473A	FGFR3 c.G811A	FGFR3 c.G1567A	FGFR3 c.G1345A	FGFR3 c.C1708T	FGFR3 c.G1894A
FGFR3 c.C2000A	FGFR3 c.G349A	FGFR3 c.C825A	FGFR3 c.G824A	FGFR3 c.G1966A	FGFR4 c.G883A	FGFR4 c.C1166T	FGFR4 c.G1144A
FGFR4 c.C1990T	FGFR4 c.G1106A	FGFR4 c.G1642A	FGFR4 c.C736T	FGFR4 c.G139A	FGFR4 c.C2119T	FGFR4 c.C88T	FGFR4 c.C1903T
FGFR4 c.G1909A	FGFR4 c.G1949A	FGFR4 c.C1121T	FGFR4 c.C2186T	FGFR4 c.G1426A	FGFR4 c.G862A	FGFR4 c.G562A	FUBP1 c.G857A
IDH1 c.G481A	IDH1 c.C932T	IDH1 c.G299A	IDH1 c.G59A	IDH2 c.C1097T	IDH2 c.G110A	IDH2 c.G1145A	IDH2 c.G448A
IDH2 c.G1069A	IDH2 c.G520A	IDH2 c.C1214T	IDH2 c.G1207A	IRS1 c.C1876T	IRS1 c.G3047A	IRS1 c.G3304A	IRS1 c.C1327T
IRS1 c.G148A	IRS1 c.G751A	IRS1 c.G401A	IRS1 c.G3211A	IRS1 c.G271A	IRS1 c.C2527T	IRS1 c.C2548T	IRS1 c.G773A
IRS1 c.G3425A	IRS1 c.C374T	IRS1 c.G1565A	IRS1 c.G404A	IRS1 c.G2740A	IRS1 c.C2992T	IRS1 c.C938T	IRS1 c.C3125T
IRS1 c.C2326T	IRS1 c.G22A	IRS1 c.C2389T	IRS1 c.G3242A	IRS1 c.G2528A	MET c.G3563A	MET c.G1030A	MGMT c.C118T
MGMT c.C667T	MGMT c.C349T	MGMT c.G710A	MGMT c.C64T	MGMT c.C616T	MLH1 c.C793T	MLH1 c.G677A	MLH1 c.C1561T
MLH1 c.G955A	MLH1 c.C1975T	MSH2 c.G187A	MSH3 c.G133A	MSH3 c.C56T	MSH4 c.G100A	MSH4 c.G172A	MSH5 c.C62T
MSH5 c.G29A	MSH5 c.G1901A	MSH6 c.G3992A	MYC c.G247A	MYC c.G331A	MYC c.G1055A	MYC c.C631T	MYC c.C602A
MYCN c.C455T	MYCN c.C545T	NF2 c.G1753A	NF2 c.C1762T	NF2 c.C1256T	NF2 c.C1270T	NF2 c.C1252T	NOTCH1 c.G6685A
NOTCH1 c.C6188T	NOTCH1 c.C2439G	NOTCH1 c.C388T	NOTCH1 c.G751A	NOTCH1 c.G7349A	NOTCH1 c.G1483A	NOTCH1 c.G5785A	NOTCH1 c.G5734A
NOTCH1 c.C3761T	NOTCH1 c.G665A	NOTCH1 c.G4715A	NOTCH1 c.G5561A	NOTCH1 c.C526T	NOTCH1 c.G5132A	NOTCH1 c.C2660T	NOTCH1 c.G1094A
NOTCH1 c.G5015A	NOTCH1 c.G1058A	NOTCH1 c.C2387T	NOTCH1 c.C6208T	NOTCH1 c.G3784A	NOTCH1 c.G214A	NOTCH1 c.G4057A	NOTCH1 c.G3122A
NOTCH1 c.C3394T	NOTCH1 c.G6509A	NOTCH1 c.G3758A	NOTCH1 c.G578A	NOTCH1 c.C6409T	NOTCH1 c.C6139T	NOTCH1 c.G515T	NOTCH1 c.C3679T
NOTCH1 c.G3664A	NOTCH1 c.G3334A	NOTCH1 c.G7034A	NOTCH1 c.C5900T	NOTCH1 c.G5092A	NOTCH1 c.C3824T	NOTCH1 c.G3880A	NOTCH1 c.C3482T
NOTCH1 c.G6359A	NOTCH1 c.G1027A	NOTCH1 c.G5348A	NOTCH1 c.G1337A	NOTCH1 c.G5446A	NOTCH1 c.G4422A	NOTCH1 c.C6665T	NOTCH1 c.C1373A
NOTCH1 c.G635A	NOTCH1 c.C2903T	NOTCH1 c.G658A	NOTCH1 c.G2303A	NOTCH1 c.G2249A	NOTCH1 c.C4921T	NOTCH1 c.G5336A	NOTCH1 c.C2863T
NOTCH1 c.G1628A	NOTCH1 c.G1847A	NOTCH1 c.C1774T	NOTCH1 c.G226A	NOTCH1 c.G4573A	NOTCH1 c.G4376A	NOTCH1 c.C4984T	NOTCH1 c.C6767T
NOTCH1 c.G2270A	NOTCH1 c.G1393A	NOTCH1 c.C5272T	NOTCH1 c.G3497A	NOTCH1 c.G1060A	NOTCH1 c.C6593T	NOTCH1 c.G211A	NOTCH1 c.G1978A
NOTCH1 c.C4639A	NOTCH1 c.G395A	NOTCH1 c.C2474T	NOTCH1 c.C7397T	NOTCH1 c.C5281T	NOTCH1 c.C5287T	NOTCH1 c.G2960A	NOTCH1 c.G161A
NOTCH1 c.G6070A	NOTCH1 c.G3295A	NOTCH1 c.G6637A	NOTCH1 c.C3886T	NOTCH1 c.G4723C	NOTCH1 c.G2807A	NOTCH1 c.G710A	NOTCH1 c.G6707A
NOTCH1 c.G6301A	NOTCH1 c.G6715A	NOTCH1 c.G715A	NOTCH1 c.C4157T	NOTCH1 c.G4711A	NOTCH1 c.G5227A	NOTCH1 c.C7645A	NOTCH1 c.C5459T
NOTCH1 c.G4313A	NOTCH1 c.C4446A	NOTCH1 c.G4714A	NOTCH1 c.G1343A	NOTCH1 c.C5944T	NTRK1 c.G880A	NTRK1 c.C1708T	NTRK1 c.G2231A
NTRK1 c.G40A	NTRK1 c.G1520A	NTRK1 c.G2128A	NTRK1 c.G281A	NTRK1 c.G659A	NTRK1 c.C1817T	NTRK1 c.G1114A	NTRK2 c.C406T
NTRK2 c.G961A	NTRK2 c.C1807T	NTRK2 c.C370T	NTRK2 c.G139A	NTRK2 c.G2245A	NTRK2 c.G1652A	NTRK2 c.G1423A	NTRK2 c.G107A
NTRK2 c.C1420T	NTRK3 c.G665A	NTRK3 c.C848T	NTRK3 c.G602A	NTRK3 c.C388T	NTRK3 c.G1397A	PDGFRA c.G3211A	PDGFRA c.C1672T
PDGFRA c.C2368T	PDGFRA c.C137T	PIK3R1 c.C374T	PIK3R1 c.G1786A	PIK3R3 c.C865T	PMS2 c.C2108T	PMS2 c.C1169T	PMS2 c.G625A
POLD1 c.C2563T	POLD1 c.G2794A	POLD1 c.G883A	POLD1 c.C2173T	POLD1 c.C2728T	POLD1 c.G433A	POLD1 c.G536A	POLD1 c.C40T
POLD1 c.C1556T	POLD1 c.G671A	POLD1 c.C1016T	POLD1 c.G2629A	POLD1 c.G613A	POLD1 c.G1786A	POLD1 c.G952A	POLD1 c.C232T
POLD1 c.G1295A	POLD1 c.C1054T	POLD1 c.C1072T	POLD1 c.C2534T	POLD1 c.C2983T	POLD1 c.C1954T	POLD1 c.G607A	POLD2 c.G340A
POLD2 c.G31A	POLD2 c.G939A	POLD2 c.G1414A	POLD2 c.C430T	POLD2 c.C569T	POLE c.G1015A	POLE c.C6803T	POLE c.C3890T
POLE c.G3971A	POLE c.G1837A	POLE c.C3817T	POLE c.C5467T	POLE c.C3109T	POLE c.C28T	POLE c.C4228T	POLE c.G40T
POLE c.C6605T	POLE c.G4825A	POLE c.G2540A	POLE c.G4339A	POLE c.G4117A	POLE c.C5632T	POLE c.C3229T	POLE c.G4559A
POLE c.G1332A	POLE c.G4112A	POLE c.G1337A	POLE c.G4663A	POLE c.C1280T	POLG c.G265A	POLG c.G419A	POLG c.G2597A
POLG c.G2543A	POLG c.G2753A	POLG c.G275A	POLG c.C2569T	POLG c.C3328T	POLG c.G2020A	POLG c.C635T	POLG c.G331A
POLG c.C2137T	POLG c.G3262A	POLG c.G2659A	POLG c.G1721A	POLG c.G289A	POLG c.C202T	POLG c.G617A	POLG c.C1385T
POLG c.G1132A	POLG c.C409T	POLG c.C3344T	PTEN c.G1669A	PTEN c.G461A	RB1 c.C1223T	RB1 c.C634A	RB1 c.G1095T
RB1 c.C130T	RB1 c.G2707A	RFC1 c.G3A	RFC1 c.G1163A	RFC2 c.C983T	RFC2 c.G401A	RFC4 c.C733T	SMARCA4 c.G1084A

GBM subclonal variants (VAF<10%)*							
SMARCA4 c.C1189T	SMARCA4 c.C1615T	SMARCA4 c.C2896T	SMARCA4 c.G4981A	SMARCA4 c.G4873A	SMARCA4 c.G5012A	SMARCA4 c.G2581A	SMARCA4 c.C3019T
SMARCA4 c.G4988A	SMARCA4 c.G4024A	SMARCA4 c.G3073A	SMARCA4 c.C4912T	SMARCA4 c.C944T	SMARCA4 c.C662T	SMARCA4 c.C4364T	SMARCA4 c.C731T
SMARCB1 c.G569A	SMARCB1 c.G110A	TERT c.C1336T	TERT c.G751A	TERT c.C968T	TERT c.G676A	TERT c.G652A	TERT c.C649T
TERT c.C358T	TERT c.C2312T	TERT c.C1307T	TERT c.G568A	TERT c.C2014T	TERT c.G2002A	TERT c.G1100A	TERT c.G916A
TERT c.C580T	TERT c.G398A	TERT c.C146T	TERT c.C1138T	TERT c.C1232T	TERT c.C142T	TERT c.C3164T	TERT c.G677A
TERT c.C1853T	TERT c.G1843A	TERT c.G835A	TERT c.G2351A	TERT c.G838A	TERT c.C1873T	TERT c.C823T	TERT c.G2702A
TP53 c.G1010A							

PNC subclonal variants (VAF<10%)*							
ACVR1 c.G484A	AKT1 c.G688A	AKT1 c.G1252A	AKT1 c.G722A	AKT1 c.C173T	AKT1 c.G818A	AKT1 c.G484A	AKT1 c.G1390A
AKT1 c.G1100A	AKT1 c.G665A	AKT1 c.C1399T	AKT2 c.G667A	AKT2 c.C782T	AKT2 c.G44A	AKT2 c.G338A	AKT2 c.C395T
AKT2 c.G784A	AKT2 c.G598A	ATRX c.G20A	ATRX c.G2053A	ATRX c.C6331T	ATRX c.G749A	ATRX c.G4511A	ATRX c.G1169A
ATRX c.C4684T	ATRX c.G4517A	BRAF c.G1085A	CDK4 c.G805A	CDK4 c.G737A	CDK4 c.G671A	CDK4 c.G626A	CDK6 c.G162A
CDK6 c.C640T	CDK6 c.G808A	CDK6 c.G260A	CDKN2A c.C25A	CDKN2A c.G466A	CDKN2A c.G244A	CDKN2A c.G220A	CDKN2A c.C227T
CDKN2A c.G200A	CDKN2A c.C14T	CDKN2A c.C395T	CDKN2A c.G331A	CDKN2A c.C332T	CDKN2B c.C340T	CDKN2B c.G349A	CDKN2B c.C359T
CDKN2B c.G186T	CDKN2B c.G182A	CIC c.C3920T	CIC c.C5792T	CIC c.G6284A	CIC c.G3679A	CIC c.G7441A	CIC c.C6044T
CIC c.C6877T	CIC c.G4093A	CIC c.G7072A	CIC c.G4988A	CIC c.C5378T	CIC c.G6619A	CIC c.G3764A	CIC c.G6728A
CIC c.C7469T	CIC c.G5315A	CIC c.G3620A	CIC c.C3719T	CIC c.C5764T	CIC c.C7048T	CIC c.C6421T	CIC c.G4660A
CIC c.G4621A	CIC c.G4127A	CIC c.G2926A	CIC c.C6556T	DAXX c.C1769T	DAXX c.G1970A	DAXX c.G32A	DAXX c.G925A
DAXX c.G1130A	DAXX c.C2030T	DAXX c.C904T	DAXX c.C1147T	DNMT3A c.C1516T	DNMT3A c.G1769A	DNMT3A c.G2683A	DNMT3A c.G629A
DNMT3A c.C976T	DNMT3A c.C484T	DNMT3A c.G1228A	DNMT3A c.G2246A	EBF3 c.G1216A	EBF3 c.G100A	EBF3 c.G1339A	EBF3 c.G1372A
EBF3 c.C1583T	EBF3 c.C1258T	EBF3 c.G35A	EBF3 c.G1345A	EBF3 c.C155T	EBF3 c.G1204A	EGFR c.C925T	EGFR c.G2930A
EGFR c.C2410T	EGFR c.G2030A	EGFR c.G665A	EGFR c.G1001A	ERBB2 c.G3757A	ERBB2 c.C833T	ERBB2 c.C590T	ERBB2 c.G2551A
ERBB2 c.C3436T	ERBB2 c.G2693A	ERBB2 c.C2057T	ERBB2 c.G1720A	ERBB2 c.C365T	ERBB2 c.G1789A	ERBB2 c.G1507A	ERBB2 c.C2404T
ERBB2 c.C1606T	ERBB3 c.C1831T	ERBB3 c.C3373T	ERBB3 c.C3925T	ERBB3 c.C1352T	ERBB3 c.G2446A	ERBB3 c.C3604T	ERBB4 c.C916T
FGFR1 c.C259T	FGFR1 c.G2250A	FGFR1 c.G338A	FGFR1 c.G1903A	FGFR1 c.G388A	FGFR1 c.G1364A	FGFR1 c.G663A	FGFR1 c.C320T
FGFR2 c.C607T	FGFR2 c.G307A	FGFR2 c.G65A	FGFR2 c.G1750A	FGFR2 c.C16T	FGFR2 c.G1933A	FGFR2 c.G151A	FGFR2 c.G2194A
FGFR2 c.G1778A	FGFR2 c.G2179A	FGFR2 c.G790A	FGFR2 c.C2249T	FGFR2 c.C2342T	FGFR2 c.C1022T	FGFR3 c.G707A	FGFR3 c.G436A
FGFR3 c.G2308A	FGFR3 c.G904A	FGFR3 c.G2255A	FGFR3 c.G244A	FGFR3 c.G841A	FGFR3 c.G2368A	FGFR3 c.C586T	FGFR3 c.G1978A
FGFR3 c.G1484A	FGFR3 c.G1063A	FGFR3 c.G1765A	FGFR3 c.C523T	FGFR3 c.G2141A	FGFR3 c.G1885A	FGFR3 c.G652A	FGFR3 c.C1969T
FGFR3 c.G1456A	FGFR3 c.C271T	FGFR3 c.G971A	FGFR4 c.C184T	FGFR4 c.C601T	FGFR4 c.G2335A	FGFR4 c.G2266A	FGFR4 c.G1276A
FGFR4 c.C1556T	FGFR4 c.G875A	FGFR4 c.G1190A	FGFR4 c.G1586A	FGFR4 c.C2081T	FGFR4 c.G1073A	FGFR4 c.C1100T	FGFR4 c.G2353A
FGFR4 c.C1544T	FGFR4 c.G2318A	FUBP1 c.G1094A	FUBP1 c.G548A	IDH2 c.C265T	IDH2 c.C140T	IDH2 c.G166A	IDH2 c.G331A
IDH2 c.G1067T	IRS1 c.C3050T	IRS1 c.C3418T	IRS1 c.G454A	IRS1 c.G3287A	IRS1 c.C1022T	IRS1 c.G103A	IRS1 c.C1738T
IRS1 c.C1618A	IRS1 c.G314A	IRS1 c.C1619T	IRS1 c.G974A	IRS1 c.G1799A	IRS1 c.G3419A	IRS1 c.C1103T	IRS1 c.C2018T
IRS1 c.G1366A	IRS1 c.G3223A	IRS1 c.G794A	IRS1 c.G2542A	IRS1 c.G1460A	IRS1 c.G1319A	IRS1 c.G1316A	IRS1 c.G2683A
IRS1 c.G1309A	IRS1 c.C3038T	IRS1 c.C904T	IRS1 c.C602T	IRS1 c.C3107T	MGMT c.C47T	MGMT c.G247A	MGMT c.G556A
MLH1 c.G1460A	MLH1 c.C925T	MSH2 c.G7A	MSH2 c.G1862A	MSH3 c.C1360T	MSH4 c.C38T	MSH5 c.C826T	MSH5 c.C1669T
MSH5 c.G868A	MSH5 c.C964T	MSH6 c.G230A	MYC c.G527A	MYCN c.C683T	MYCN c.G601A	MYCN c.C788T	MYCN c.C418T
MYCN c.G463A	MYCN c.G493A	MYCN c.G758A	MYCN c.G280A	NF1 c.G5609A	NF1 c.G7A	NF1 c.G5509A	NF2 c.G1300A
NF2 c.G1232A	NF2 c.G674A	NF2 c.G1303A	NOTCH1 c.G3506A	NOTCH1 c.G5422A	NOTCH1 c.G6991A	NOTCH1 c.C4987T	NOTCH1 c.G3496A
NOTCH1 c.C6691T	NOTCH1 c.G4252A	NOTCH1 c.G2116A	NOTCH1 c.G6782A	NOTCH1 c.G5017A	NOTCH1 c.G1957A	NOTCH1 c.G4672A	NOTCH1 c.G592A
NOTCH1 c.G7043A	NOTCH1 c.C812T	NOTCH1 c.G2176A	NOTCH1 c.G3001A	NOTCH1 c.G6536A	NOTCH1 c.C368T	NOTCH1 c.G1354A	NOTCH1 c.C116T
NOTCH1 c.C7355T	NOTCH1 c.G6733A	NOTCH1 c.G1078A	NOTCH1 c.G1039A	NOTCH1 c.C4981T	NOTCH1 c.G6343A	NOTCH1 c.C4786T	NOTCH1 c.C983T
NOTCH1 c.C185T	NOTCH1 c.G3905A	NOTCH1 c.C6979T	NOTCH1 c.G1816A	NOTCH1 c.C6890T	NOTCH1 c.C5099T	NOTCH1 c.G4078A	NOTCH1 c.C4864T
NOTCH1 c.C1423T	NOTCH1 c.G3598A	NOTCH1 c.G5951A	NOTCH1 c.G4898A	NOTCH1 c.C4744T	NOTCH1 c.G2713A	NOTCH1 c.G3646A	NOTCH1 c.G5551A
NOTCH1 c.C3770T	NOTCH1 c.G3442A	NOTCH1 c.C3399A	NOTCH1 c.G5608A	NOTCH1 c.G5885A	NOTCH1 c.C7451T	NOTCH1 c.G2086A	NOTCH1 c.C2407A
NOTCH1 c.C3859T	NOTCH1 c.G646A	NOTCH1 c.G1280A	NOTCH1 c.C4136T	NOTCH1 c.C6010T	NOTCH1 c.G7153A	NOTCH1 c.G1733A	NOTCH1 c.G1838A
NOTCH1 c.C7156A	NOTCH1 c.C907T	NOTCH1 c.C4571T	NOTCH1 c.C2951T	NOTCH1 c.G970A	NOTCH1 c.C6817T	NOTCH1 c.C6407T	NOTCH1 c.C6680T
NOTCH1 c.G6067A	NOTCH1 c.G5378A	NOTCH1 c.G4972A	NOTCH1 c.G6376A	NOTCH1 c.G5777A	NOTCH1 c.T4331C	NOTCH1 c.G1679T	NOTCH1 c.G5374A
NOTCH1 c.G2693A	NOTCH1 c.C6104T	NOTCH1 c.C4483T	NOTCH1 c.C4343T	NOTCH1 c.G7280A	NOTCH1 c.C5780T	NOTCH1 c.C1676T	NOTCH1 c.G6315A
NOTCH1 c.C7457T	NOTCH1 c.C7400T	NOTCH1 c.G2719A	NOTCH1 c.G3773A	NTRK1 c.C250T	NTRK1 c.C253T	NTRK1 c.C2074T	NTRK1 c.G1006A
NTRK1 c.G1162A	NTRK1 c.G1150A	NTRK1 c.G2143A	NTRK1 c.C710T	NTRK1 c.G62A	NTRK1 c.C1925T	NTRK1 c.C2084T	NTRK1 c.G2282A
NTRK1 c.G356A	NTRK1 c.C964T	NTRK1 c.G2035A	NTRK1 c.G1748A	NTRK1 c.C1301T	NTRK2 c.G2125A	NTRK2 c.C1979T	NTRK2 c.G1283A
NTRK2 c.C1319T	NTRK2 c.G1879A	NTRK2 c.C383T	NTRK2 c.G1816A	NTRK2 c.G653A	NTRK2 c.G1870A	NTRK2 c.G250A	NTRK2 c.G208A
NTRK2 c.C2071T	NTRK2 c.G85A	NTRK2 c.C2194T	NTRK3 c.G2089A	NTRK3 c.G1078A	PCNA c.C191T	PDGFRA c.C2810T	PDGFRA c.G1565A
PDGFRA c.G2741A	PDGFRA c.C2540T	PIK3CA c.C2554T	PIK3R3 c.C190T	PIK3R3 c.C451T	PIK3R3 c.G28A	PIK3R3 c.G22A	PMS2 c.G1438A
PMS2 c.G2378A	PMS2 c.G2128A	PMS2 c.G994A	PMS2 c.G1262A	PMS2 c.C1687T	POLD1 c.C1919T	POLD1 c.G2803A	POLD1 c.G2599A
POLD1 c.G103A	POLD1 c.G2903A	POLD1 c.C2162T	POLD1 c.G2546A	POLD1 c.G3047A	POLD1 c.C2825T	POLD1 c.G2668A	POLD1 c.C3244T
POLD1 c.G1486A	POLD1 c.C2684T	POLD1 c.C2861T	POLD1 c.C644T	POLD1 c.C1418T	POLD1 c.G497A	POLD1 c.C1814T	POLD1 c.G89A
POLD2 c.C1274T	POLD2 c.C41T	POLD2 c.G1441A	POLE c.C4555T	POLE c.G5596A	POLE c.G2225A	POLE c.C3475T	POLE c.C3958T
POLE c.G2395A	POLE c.C3155T	POLE c.G4594A	POLE c.G6610A	POLE c.C6539T	POLE c.C3932T	POLE c.G4603A	POLE c.C6391T
POLE c.G1663A	POLG c.G305A	POLG c.C418T	POLG c.G824A	POLG c.G2713A	POLG c.C784T	POLG c.C343T	POLG c.C647T
POLG c.C1903T	POLG c.G234A	POLG c.C2993T	POLG c.G2029A	POLG c.G3560A	POLG c.G88A	POLG c.G1246A	POLG c.C1831T
POLG c.G3409A	POLG c.G1910A	POLG c.C653T	PTEN c.C808T	PTEN c.G1220A	RB1 c.C746T	RB1 c.C1544A	RB1 c.G1638A
RB1 c.G781A	RB1 c.C5T	RB1 c.G43A	RB1 c.C2392T	RFC1 c.G1268A	RFC2 c.C62T	RFC3 c.C526T	SMARCA4 c.G592A
SMARCA4 c.G3928A	SMARCA4 c.G4297A	SMARCA4 c.C935T	SMARCA4 c.C4760T	SMARCA4 c.C3991T	SMARCA4 c.G4141A	SMARCA4 c.G4855A	SMARCA4 c.C2681T
SMARCA4 c.C4933T	SMARCA4 c.C1030T	SMARCA4 c.C2545T	SMARCB1 c.G64A	SMARCB1 c.C1037T	TERT c.G184A	TERT c.G172A	TERT c.C1967T
TERT c.C377T	TERT c.G587A	TERT c.G215A	TERT c.G1540A	TERT c.G1340A	TERT c.G2332A	TERT c.C2479T	TERT c.G2221A
TERT c.G1264A	TERT c.G3314A	TERT c.C697A	TERT c.G436A	TERT c.C601T	TERT c.C2911T	TERT c.C2564T	TERT c.C2092T
TERT c.G2152A	TERT c.G598A	TERT c.C383T	TERT c.G1885A	TERT c.C2005T	TERT c.C134T	TERT c.C770T	TERT c.C347T
TERT c.G2734A	TERT c.G268A	TERT c.C271T	TERT c.C1246T	TERT c.G298A	TP53 c.C140T	TP53 c.G145A	TP53 c.G1136A
TP53 c.C466T	TP53 c.C275T	TP53 c.G556A					

Appendix 5. Subclonal alterations identified in Tumour#5

List of shared and private, likely subclonal variations identified in Tumour#5 GBM and PNC components. *only variants showing VAF> of LOD are reported.

Common subclonal variants (VAF<10%)*							
IRS1 c.G304A	NTRK1 c.G1331A						

GBM subclonal variants (VAF<10%)*							
ACVR1 c.C869T	AKT1 c.G1228A	AKT1 c.G253A	AKT2 c.C1145T	AKT2 c.C226T	CDK6 c.G190A	CIC c.C3959T	CIC c.C3970T
CIC c.C4498T	CIC c.C4756T	CIC c.C5081T	CIC c.G4675A	CIC c.G5188A	DNMT3A c.G1165A	DNMT3A c.G2597A	DNMT3A c.G542A
EBF3 c.G1223A	ERBB2 c.C22A	ERBB2 c.C2908T	ERBB2 c.C694T	ERBB2 c.G3280A	ERBB2 c.G671A	FGFR2 c.C16T	FGFR2 c.G443A
FGFR3 c.C2297T	FGFR3 c.C2303T	FGFR3 c.C392T	FGFR3 c.C662T	FGFR3 c.C742T	FGFR3 c.G1226A	FGFR3 c.G1519A	FGFR3 c.G1681A
FGFR3 c.G169A	FGFR3 c.G1744A	FGFR3 c.G1885A	FGFR3 c.G222A	FGFR3 c.G2368A	FGFR3 c.G2420A	FGFR4 c.C1193T	IDH2 c.C59T
IDH2 c.G1051A	IDH2 c.G1058A	IRS1 c.C2617T	IRS1 c.G475A	MGMT c.C43T	MGMT c.C698T	MGMT c.G229A	MGMT c.G571A
MSH3 c.C172A	MYC c.G144C	MYCN c.C626T	NOTCH1 c.C4136T	PDGFRA c.C1126T	POLD1 c.C1823T	POLD1 c.C923T	POLD1 c.G1363A
POLD1 c.G2938A	POLD4 c.G71A	POLE c.C4708T	POLE c.G5162A	POLG c.147_158del12	POLG c.C1879T	POLG c.G2420A	POLG c.G3242A
PTEN c.C434T	RB1 c.C2221T	SMARCA4 c.C4066T	SMARCA4 c.C937T	SMARCA4 c.G1250A	SMARCA4 c.G2581A	SMARCA4 c.G2936A	SMARCA4 c.G4568A
TERT c.C2312T	TERT c.C394T	TERT c.G1283A	TERT c.G1562A	TERT c.G2368A	TERT c.G2419A	TERT c.G32A	TERT c.G442A
TERT c.G674A	TP53 c.G1085A	TP53 c.G217A					

PNC subclonal variants (VAF<10%)*							
ACVR1 c.G700A	ACVR1 c.C1123T	ACVR1 c.G193A	ACVR1 c.G1153A	ACVR1 c.G572A	AKT1 c.C628T	AKT1 c.C430T	AKT1 c.G431A
AKT1 c.C320A	AKT1 c.G665A	AKT1 c.G1195A	AKT1 c.C365T	AKT2 c.G480T	AKT2 c.G509A	AKT2 c.C590T	AKT2 c.C953T
AKT2 c.G1036A	AKT2 c.C1063T	AKT2 c.G29A	AKT3 c.C196T	ATRX c.G2678A	ATRX c.C6892T	ATRX c.A562G	ATRX c.C6532T
ATRX c.G6940A	ATRX c.G4189A	ATRX c.G4843A	ATRX c.C4889T	ATRX c.G6533A	BRAF c.1799_1803delTGAAA	BRAF c.C2246T	BRAF c.G175A
BRAF c.C1007T	CDK4 c.G719A	CDK4 c.G130A	CDK4 c.C412T	CDK6 c.G67A	CDK6 c.G244A	CDK6 c.C130T	CDKN2A c.C142T
CDKN2A c.G322A	CDKN2A c.G223A	CDKN2A c.G448A	CIC c.G5059A	CIC c.G4523A	CIC c.C6289T	CIC c.C5756T	CIC c.C5768T
CIC c.C6322T	CIC c.C4444T	CIC c.C4928T	CIC c.G6800A	CIC c.G5935A	CIC c.C3677A	CIC c.G4988A	CIC c.C6475T
CIC c.C4994T	CIC c.C6707T	CIC c.G5509A	CIC c.G4247A	CIC c.C6806T	CIC c.C5411T	CIC c.G3536A	CIC c.C6605A
CIC c.C5081G	CIC c.G59A	CIC c.C6875T	CIC c.C3221T	CIC c.G3073A	CIC c.C5096A	CIC c.C4009T	CIC c.C4000T
CIC c.G4046A	CIC c.C5042T	CIC c.C7120T	CIC c.C5599T	CIC c.G5189T	CIC c.G3124A	CIC c.C4759T	CIC c.C3179T
CIC c.C5635T	CIC c.G5998A	CIC c.C3268T	CIC c.G3820A	CIC c.G3292A	CIC c.C4583T	CIC c.C4558T	DAXX c.G1480A
DAXX c.C521T	DNMT3A c.G1021A	DNMT3A c.G2473A	DNMT3A c.C1721T	DNMT3A c.G1142A	DNMT3A c.G841A	DNMT3A c.G626A	DNMT3A c.G2123A
DNMT3A c.G2207A	DNMT3A c.G137A	DNMT3A c.G2300A	EBF3 c.G127A	EBF3 c.G103A	EBF3 c.G737A	EBF3 c.G29A	EBF3 c.C94T
EBF3 c.G523A	EBF3 c.G93A	EGFR c.G3226A	EGFR c.C2408T	EGFR c.G2006A	ERBB2 c.G2933A	ERBB2 c.C418A	ERBB2 c.C2440T
ERBB2 c.C100T	ERBB2 c.G23A	ERBB2 c.C3593T	ERBB2 c.C581A	ERBB2 c.G2546A	ERBB2 c.C3466A	ERBB2 c.G3166A	ERBB2 c.G1889A
ERBB2 c.C1281A	ERBB2 c.G1489A	ERBB2 c.G764A	ERBB2 c.C1366T	ERBB3 c.G1891A	ERBB3 c.C1181T	ERBB3 c.C508T	ERBB3 c.G992A
ERBB3 c.C826A	ERBB3 c.G3574A	ERBB3 c.G3809A	ERBB3 c.G3272A	ERBB3 c.G2587A	ERBB3 c.G388T	ERBB3 c.C2767A	ERBB3 c.G310A
ERBB3 c.G1469A	ERBB3 c.C4018T	ERBB3 c.G3987T	ERBB3 c.G3941A	ERBB3 c.C2895A	ERBB3 c.C3406T	ERBB3 c.C695T	ERBB3 c.G2460T
ERBB3 c.G1490A	ERBB3 c.G970T	ERBB3 c.G1529A	ERBB3 c.G337T	ERBB3 c.C1462T	ERBB3 c.G118A	ERBB3 c.G3845A	ERBB3 c.C1433T
ERBB3 c.G3841A	ERBB3 c.G3330T	ERBB4 c.C916T	ERBB4 c.G1585A	ERBB4 c.C779T	ERBB4 c.C851T	ERBB4 c.C1177T	ERBB4 c.G2105A
ERBB4 c.G643T	ERBB4 c.C50T	ERBB4 c.G1463A	ERBB4 c.G421A	ERBB4 c.G1574A	ERBB4 c.G499A	ERBB4 c.C3638T	EXO1 c.C1009T
EXO1 c.C632T	EXO1 c.C1862T	EXO1 c.C1996T	FGFR1 c.G1802A	FGFR1 c.G905A	FGFR1 c.C2537T	FGFR1 c.G1135A	FGFR1 c.C658T
FGFR1 c.G1240A	FGFR1 c.G1856A	FGFR2 c.C1732T	FGFR2 c.C2434T	FGFR2 c.G892A	FGFR2 c.C1935A	FGFR2 c.G2038A	FGFR2 c.C2408T
FGFR2 c.G2104A	FGFR2 c.G182A	FGFR2 c.G505A	FGFR2 c.G151A	FGFR2 c.C512A	FGFR2 c.C2360T	FGFR2 c.G391A	FGFR3 c.C503T
FGFR3 c.G952A	FGFR3 c.G824A	FGFR3 c.C1708T	FGFR3 c.G1718A	FGFR3 c.C1340A	FGFR3 c.C1388T	FGFR3 c.G2081A	FGFR3 c.G1817A
FGFR3 c.C115A	FGFR4 c.G731A	FGFR4 c.C1405T	FGFR4 c.C2186T	FGFR4 c.C1300T	FGFR4 c.C449T	FGFR4 c.C175T	FGFR4 c.G314A
FGFR4 c.G1120A	FGFR4 c.C451T	FGFR4 c.C1046T	FGFR4 c.C2119T	FGFR4 c.C1513T	FGFR4 c.G352T	FGFR4 c.C1309T	FGFR4 c.G185A
FGFR4 c.C1190T	FGFR4 c.G1745A	FGFR4 c.G659A	FGFR4 c.G1583A	FUBP1 c.C89T	IDH1 c.C664T	IDH1 c.G59A	IDH1 c.G211A
IDH2 c.C140T	IDH2 c.C445T	IDH2 c.G653A	IDH2 c.G663A	IDH2 c.G622T	IRS1 c.C3269T	IRS1 c.C1244T	IRS1 c.G3370A
IRS1 c.G2327A	IRS1 c.C2560T	IRS1 c.G3151A	IRS1 c.C3075A	IRS1 c.G3398A	IRS1 c.C602T	IRS1 c.C818T	IRS1 c.G3650A
IRS1 c.C1337T	IRS1 c.C2278T	IRS1 c.C58T	IRS1 c.C1558T	IRS1 c.C2368T	IRS1 c.G119A	IRS1 c.G2107A	IRS1 c.C1706T
IRS1 c.G688A	IRS1 c.G607A	MDM2 c.G808A	MDM2 c.G781A	MDM2 c.G58A	MDM4 c.G611A	MDM4 c.C661T	MET c.C4151T
MET c.G367A	MET c.G2770A	MET c.C3091T	MET c.G1406A	MGMT c.C616T	MGMT c.C206T	MLH1 c.G884T	MLH1 c.G1595A
MLH1 c.G566A	MLH1 c.A1556G	MLH1 c.G1816T	MLH3 c.C2806T	MSH2 c.G759A	MSH2 c.G998A	MSH2 c.G82A	MSH2 c.G1145A
MSH2 c.G2150A	MSH2 c.G632A	MSH2 c.G631A	MSH2 c.G1495A	MSH3 c.C199G	MSH3 c.G151A	MSH3 c.G1151A	MSH3 c.G2686A
MSH3 c.G895A	MSH3 c.G377A	MSH3 c.G2521A	MSH3 c.G3136A	MSH3 c.G2239A	MSH3 c.C1720T	MSH3 c.C2200T	MSH3 c.G799A
MSH4 c.G1540A	MSH4 c.C716T	MSH4 c.G2051A	MSH4 c.C1661T	MSH4 c.G64A	MSH4 c.G2699A	MSH5 c.C1580T	MSH5 c.G1774A
MSH6 c.G2736A	MSH6 c.C3202T	MSH6 c.C3991T	MSH6 c.C2764T	MSH6 c.C2962T	MSH6 c.G3725A	MSH6 c.G2153A	MSH6 c.C3299T
MSH6 c.C3254T	MSH6 c.G2635T	MYC c.C272T	MYC c.G1246A	MYCN c.G832A	MYCN c.C668A	MYCN c.C406A	MYCN c.C1117T
MYCN c.G274A	MYCN c.C1003T	MYCN c.C1073T	NF1 c.G1366A	NF1 c.C4084T	NF1 c.G7046A	NF1 c.G5299A	NF1 c.G5603A
NF1 c.G5572A	NF1 c.C1958T	NF1 c.G2936A	NF1 c.C355T	NF1 c.G1567A	NF1 c.G3649A	NF1 c.G4660A	NF1 c.C4721T
NF1 c.C6521T	NF1 c.G3262A	NF1 c.C6557T	NF1 c.C8440T	NF1 c.C5309T	NF1 c.C4787A	NF2 c.C1396T	NF2 c.G172A
NF2 c.C1126T	NF2 c.T1493A	NF2 c.G1114A	NF2 c.G1712A	NOTCH1 c.G1661A	NOTCH1 c.G4898A	NOTCH1 c.G1543A	NOTCH1 c.G3839A
NOTCH1 c.G3887A	NOTCH1 c.G52A	NOTCH1 c.G2297A	NOTCH1 c.G6745A	NOTCH1 c.C6749T	NOTCH1 c.C362T	NOTCH1 c.G2890A	NOTCH1 c.G6260A
NOTCH1 c.G6379A	NOTCH1 c.G5260A	NOTCH1 c.C2903T	NOTCH1 c.G1915A	NOTCH1 c.G3685A	NOTCH1 c.C4685T	NOTCH1 c.C7420T	NOTCH1 c.G4583A
NOTCH1 c.C388T	NOTCH1 c.G2735A	NOTCH1 c.G3121A	NOTCH1 c.G448A	NOTCH1 c.G7432A	NOTCH1 c.G2864A	NOTCH1 c.C3835T	NOTCH1 c.G7340A
NOTCH1 c.C4822T	NOTCH1 c.G4828A	NOTCH1 c.G2693A	NOTCH1 c.G6685A	NOTCH1 c.G4682A	NOTCH1 c.C5183T	NTRK1 c.G119A	NTRK1 c.C865T
NTRK1 c.C1301T	NTRK1 c.G1276A	NTRK1 c.G100A	NTRK1 c.C1105T	NTRK2 c.G1279A	NTRK2 c.G1939A	NTRK2 c.G1623T	NTRK2 c.C1726T
NTRK2 c.C1019T	NTRK2 c.G1030A	NTRK3 c.G1826A	NTRK3 c.G1397A	NTRK3 c.G1759A	NTRK3 c.C821T	NTRK3 c.C100T	NTRK3 c.G1402A
PCNA c.G778A	PDGFRA c.G2494A	PDGFRA c.G250A	PDGFRA c.G577A	PDGFRA c.C1031T	PDGFRA c.C832A	PIK3CA c.G1437A	PIK3CA c.G2453A
PIK3CB c.G1363A	PIK3CB c.C1520T	PIK3CB c.G446A	PIK3CB c.C615A	PIK3R1 c.C506T	PIK3R1 c.G1541A	PIK3R1 c.G514A	PIK3R1 c.G961T
PIK3R1 c.C1924T	PIK3R1 c.C2149T	PIK3R1 c.G1796A	PIK3R1 c.G1096A	PIK3R1 c.C686T	PIK3R3 c.G935A	PIK3R3 c.C59T	PIK3R3 c.G909T
PIK3R4 c.G1025A	PIK3R4 c.G3286A	PIK3R4 c.C2836T	PIK3R4 c.C886T	PIK3R4 c.C3716T	PIK3R4 c.C3395T	PMS1 c.C2444T	PMS1 c.G271A

PNC subclonal variants (VAF<10%)*							
PMS1 c.C341T	PMS1 c.C2782T	PMS2 c.G590A	PMS2 c.G2418A	PMS2 c.G652A	PMS2 c.G1228A	PMS2 c.G1675A	PMS2 c.A299G
POLD1 c.G3001A	POLD1 c.G3016A	POLD1 c.G2293A	POLD1 c.G2743A	POLD1 c.G2000A	POLD1 c.G3188A	POLD1 c.C554T	POLD1 c.C1366A
POLD1 c.G2713A	POLD1 c.C2404A	POLD1 c.C715T	POLD1 c.C2461A	POLD1 c.G2599A	POLD1 c.C455T	POLD1 c.G1836T	POLD1 c.G1157A
POLD1 c.G1594A	POLD1 c.C945A	POLD1 c.C355T	POLD1 c.G758A	POLD1 c.C376A	POLD1 c.G1621A	POLD1 c.G536A	POLD1 c.C606A
POLD1 c.G1328A	POLD1 c.C251A	POLD1 c.G613A	POLD1 c.G2275A	POLD1 c.G163T	POLD1 c.C623A	POLD1 c.G695A	POLD2 c.C472T
POLD2 c.C910T	POLD2 c.G724A	POLD4 c.C113A	POLE c.C3994T	POLE c.G6727A	POLE c.G1009T	POLE c.C4306T	POLE c.G4660A
POLE c.G5206A	POLE c.C3958T	POLE c.C941A	POLE c.C691T	POLE c.G556A	POLE c.C5066T	POLE c.C1823T	POLE c.G2488T
POLE c.C3121T	POLE c.G6233A	POLE c.G2930A	POLE c.G5863A	POLE c.T6100C	POLE c.G3971A	POLE c.G3652A	POLE c.C3769T
POLE c.A4240T	POLE c.G1059T	POLE c.G592T	POLE c.T3701C	POLE c.G1215T	POLE c.G1630A	POLG c.152_153insTGT	POLG c.G1592A
POLG c.G3631A	POLG c.G2798A	POLG c.G2458A	POLG c.C892T	POLG c.G1894A	POLG c.G917A	POLG c.G1309A	POLG c.C2977T
POLG c.C1553T	POLG c.G88A	PTEN c.G1217A	PTEN c.G1523A	RB1 c.C2728T	RB1 c.G1667A	RB1 c.C53T	RB1 c.G1334A
RB1 c.T2048C	RB1 c.C1735T	RB1 c.G764A	RB1 c.G298A	RB1 c.G1352A	RB1 c.G2752A	RB1 c.G113A	RFC1 c.G1909A
RFC1 c.G2329A	RFC1 c.G2042A	RFC1 c.C1438T	RFC1 c.G1451A	RFC1 c.G2683A	RFC3 c.G449A	RFC3 c.C823T	RFC3 c.G343A
RFC4 c.G181A	RFC4 c.G88A	RFC4 c.C908T	RFC4 c.G359A	RFC5 c.G287A	RFC5 c.C724T	SMARCA4 c.C2210T	SMARCA4 c.G1274A
SMARCA4 c.C2048A	SMARCA4 c.G4334A	SMARCA4 c.C941T	SMARCA4 c.G3703A	SMARCA4 c.G4501A	SMARCA4 c.G2546A	SMARCA4 c.C1787T	SMARCA4 c.G832A
SMARCA4 c.G4427A	SMARCA4 c.G5029A	SMARCA4 c.G3464A	SMARCA4 c.G3436A	SMARCA4 c.G373T	SMARCA4 c.G3599A	SMARCA4 c.G4051A	SMARCA4 c.C931T
SMARCA4 c.C910T	SMARCA4 c.C1010T	SMARCA4 c.G2654A	SMARCA4 c.G3912T	SMARCA4 c.C1537T	SMARCA4 c.G710A	SMARCA4 c.G4305T	SMARCA4 c.G3575A
SMARCB1 c.G305A	SMARCB1 c.C1037T	SMARCB1 c.G550A	SMARCB1 c.G310T	SMARCB1 c.G386A	SMARCB1 c.G43A	SMARCB1 c.C493T	TERT c.G275A
TERT c.C2495T	TERT c.G1370A	TERT c.G428A	TERT c.G397A	TERT c.G2456A	TERT c.C1168T	TERT c.C1072T	TERT c.C3226T
TERT c.C1372A	TERT c.G268A	TERT c.C2633T	TERT c.G274A	TERT c.G538A	TERT c.C560T	TERT c.C407A	TERT c.G430A
TERT c.G1238A	TERT c.A3023G	TERT c.C2650A	TERT c.C377T	TERT c.G895A	TERT c.G1169A	TERT c.A1195T	TP53 c.C637T
TP53 c.G880A	TP53 c.C541T	TP53 c.C112A	TP53 c.G734A	TP53 c.C749A			

Appendix 6. Subclonal alterations identified in Tumour#6

List of shared and private, likely subclonal variations identified in Tumour#6 GBM and PNC components. *only variants showing VAF> of LOD are reported.

PART II

CONTRIBUTION OF EBF3 TO THE NEURONAL CELL COMMITMENT

INTRODUCTION

In the first part of this work, we characterized a cohort of Glioblastomas with Primitive Neuronal Component. IHC analysis showed a glial profile for the GBM component, as expected, while the PNC component, not surprisingly, was usually negative or barely positive for GBM typical markers, as GFAP and EGFR. Notably, we found a novel biomarker selectively expressed in the PNC and constantly negative in the GBM component: EBF3 (Figure 23). *EBF3* is a transcriptional factor belonging to a highly conserved four gene family, with an important role in neurogenesis and neuronal migration. Specific mutations of *EBF1* and *EBF3* have been observed in different types of tumours. Particularly is frequently deleted or methylated in both primary and secondary glioblastomas. Moreover, our group demonstrated that in medulloblastoma, a malignant embryonal brain tumour, *EBF3* is highly expressed and acts as a major master regulator of neuronal differentiation. *EBF3* expression in early undifferentiated progenitors promotes neuronal differentiation, but a sustained *EBF3* expression in medulloblastoma correlates with an immature phenotype, and contributes to promote neoplastic progression. Since EBF3 is highly expressed in the PNC component of the GBM-PNC, and is persistently negative in the GBM component, we hypothesize that it could have a major role in driving the divergent differentiation of the two components and we decided to further investigate the potential role of this factor *in vitro* and *in vivo*.

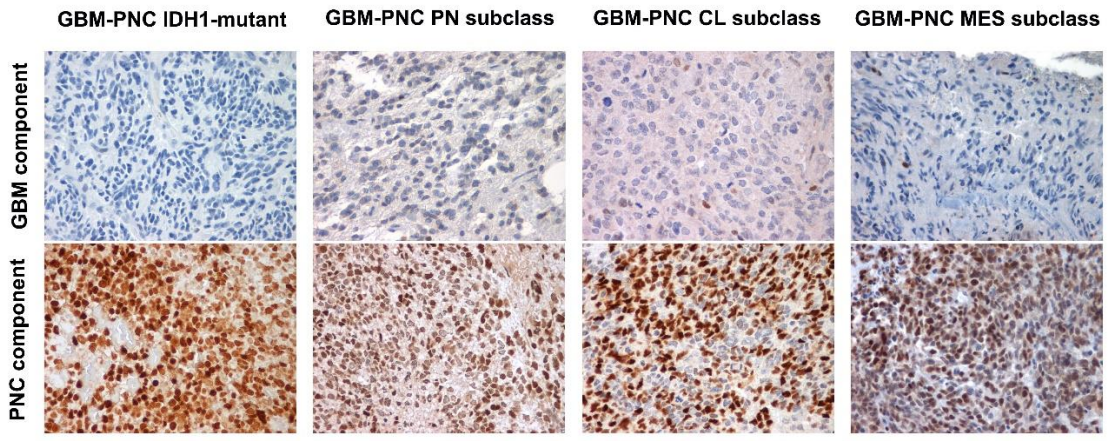


Figure 23. EBF3 expression in GBM-PNC
EBF3 is selectively and constantly expressed in the PNC component of the GBM-PNC tumours previously analysed (lower images), regardless of the transcriptional subtypes. GBM component is constantly negative (upper images) (40X original magnification).

Early-B cell Factors (EBFs)

Early B-cell factors (*EBF*s), a four gene family (*EBF1, EBF2, EBF3* and *EBF4*), are DNA-binding transcriptional factors with an atypical zinc-finger and helix-loop-helix motif. They are highly preserved among different species; Ebf orthologs were found in mouse, rat, *C.elegans*, *Drosophila, Xenopus* and *D.rerio*. These factors have an important role in cellular differentiation during the development in different organs and tissues including B-cell differentiation (Medina, Pongubala et al. 2004, Hagman and Lukin 2005), neurogenesis (Garel, Marin et al. 1997, Dubois, Bally-Cuif et al. 1998, Garcia-Dominguez, Poquet et al. 2003), and bone development (Kieslinger, Folberth et al. 2005). In humans, *EBF* genes are located in the 5q34, 8p21.2, 10q26.3, and 20p13 chromosomal regions (Liao 2009).

Ebfs are characterized by a modular structure and well-defined functional domains (Figure 24) (Liao 2009). This protein family is linked directly to the DNA consensus sequence (5'-CCCNNGGG-3') in form of homo- or hetero-dimer (Wang, Betz et al. 2002). The DNA binding domain (DBD) is close to the NH2-terminal sequence and is composed of about 200 residues in which the sequence is preserved during the evolution. This domain contains an atypical zinc finger motif (H-X3-C-X2-C-X5-C), characteristic of the EBFs family and necessary for the DNA binding; because this motif is shorter than other zinc-binding structures, it is also known as “zinc knuckle”(Fields, Ternyak et al. 2008). Next to the DBD there is a sequence that resembles the IPT/TIG fold domain, whose function is not yet well-determined, even though it could be involved in homo- or etero-dimerisation with other proteins (Liao 2009). At the COOH-terminal domain of the IPT/TIG domain is located a helix-loop-helix (HLH) atypical domain, probably involved in the dimerization. The second and the third helixes, differently from the typical HLH domain, have a very similar aminoacidic sequence, probably due to an exonic duplication (Liao 2009). The COOH-terminal domain, less preserved and with less residues of Serin, Threonine and Proline, plays an important role in the transcriptional activation (Hagman, Gutch et al. 1995).

In the CNS, EBFs have a key role in neuronal differentiation and regional specification (Garel, Marin et al. 1997). During mouse embryogenesis, *ebf1*, *ebf2* and *ebf3* are expressed in early post-mitotic neurons along the entire rostro-caudal axis of the developing CNS, with overlapping patterns, except in the forebrain where each of them is restricted to specific regions (Garel, Marin et al. 1997, Wang, Tsai et al. 1997). Interestingly, *ebfs* expression is restricted to early post-mitotic neuronal progenitors and their function is required for coupling neuronal differentiation and cell cycle exit (Garcia-Dominguez, Poquet et al. 2003); in particular, data are consistent with an early role of *ebf1*

and *ebf3*, coincidental and redundant with that of *ebf2*. Also in human organoids, *EBF3* is found among the genes upregulated at the earlier time point. This is consistent with the evidence that *EBF3* is relevant in the process of neurogenesis and neuronal migration and in the differentiation towards the neuronal cell fate (Smits, Magni et al. 2020).

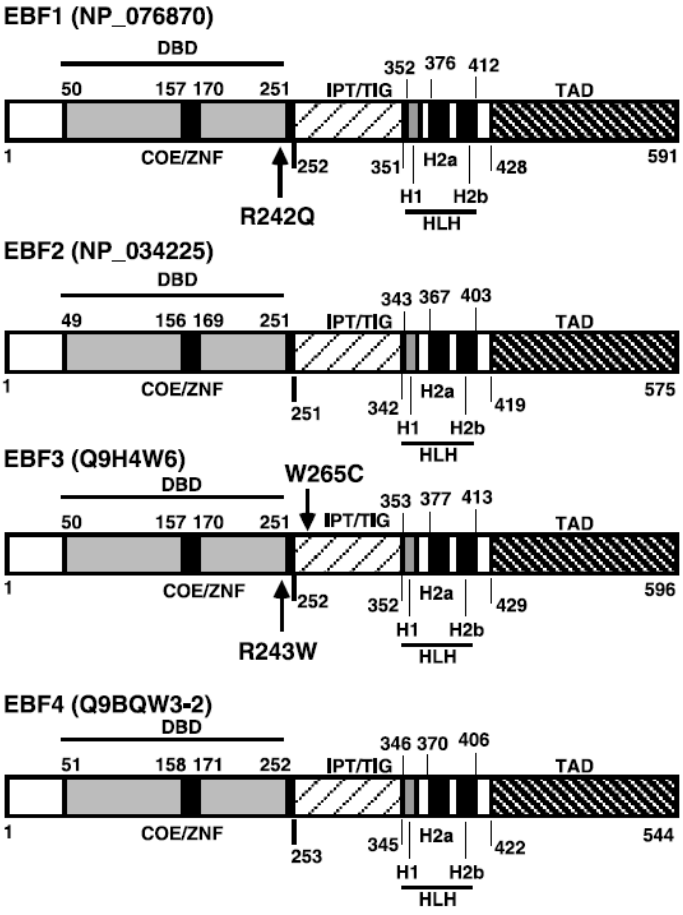


Figure 24. Structural features of EBF family of transcription factors
Schematic representation of the structure of the four paralogues of *EBF* family and their corresponding accession numbers (GenBank or SwissProt database). Specific domains are shown in shaded or striped boxes. Numbers refer to the position of amino acid residue in each protein. Cancer specific somatic mutations in *EBF1* and *EBF3* are indicated by arrows (Liao 2009).

It has been reported that EBF transcriptional factors could exert a role in tumourigenesis. In fact, low-frequency specific mutations in the *EBFs* gene family, and in particular in *EBF1* and *EBF3*, have been observed in different types of tumours, such as pancreatic cancer (Jones, Zhang et al. 2008) and GBM (Parsons, Jones et al. 2008), suggesting that *EBFs* may contribute to the development of neoplasms by acting as oncosuppressors (Zardo, Tiirikainen et al. 2002, Liao 2009). In fact, the inactivation of *EBF* genes could imply an expansion of undifferentiated progenitor cells, leading to tumour development (Chen, Cheung et al. 2002, Zardo, Tiirikainen et al. 2002, Mullighan, Goorha et al. 2007). Particularly, *EBF3* gene, mapping on human chromosome 10q, is frequently deleted or methylated in both primary and secondary glioblastomas, as well as in other tumour types, where

epigenetic alterations are linked to tumour progression and metastatic ability (Chatterjee, Stockwell et al. 2017, Rodger, Chatterjee et al. 2019). In particular, in glioblastomas, *EBF3* inactivation frequently occurs along with the inactivation of the *PTEN* locus in 10q22 and representing a well-established molecular alteration during gliomagenesis (Xiao, Wu et al. 2002, Xiao, Yin et al. 2005). A genome-wide analyses using integrated genomic and epigenetic screenings revealed that *EBF3* is inactivated by methylation or deletion in about 50% of grade II, 83% of grade III and 90% of grade IV brain tumours (Zardo, Tiirikainen et al. 2002). Consistent with these observations, *EBF3* is expressed in normal CNS cells but is silenced in brain tumour cells, suggesting a potential role as tumour suppressor gene in the brain.

In literature, it is reported that *EBF3* expression in tumour cells *in vitro* regulates a gene expression program in which genes involved in cell cycle arrest, such as the Cip/Kip p21 cyclin-dependent kinase inhibitor (CDKI), are selectively up-regulated, and, in the same time, genes involved in cell proliferation (cyclins and CDKs) and survival (Daxx and Mcl-1) are repressed (Figure 25). EBF3 may directly mediate their transcriptional activation or repression through interacting with specific binding sites in the promoters of these genes. Indeed, p21 and p27 promoters contain a putative EBF-binding site for the binding of EBF3. Moreover, EBF3 may induce apoptosis as caspase-3 activation and cleavage of PARP were observed (Zhao, Niu et al. 2006).

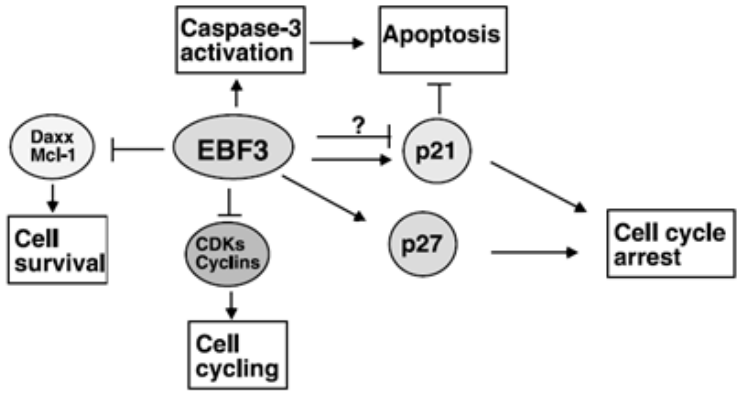


Figure 25. A potential pathway involving EBF3
EBF3 regulates the expression of genes involved in cell growth and proliferation. Specifically, expression of EBF3 activates the expression of p21 and p27 at early time points, which induce cell cycle arrest. At late time points, p21 and p27 are downregulated and apoptosis of tumour cells occur through caspase-3 activation and PARP-1 (poly [ADP-ribose] polymerase 1) cleavage. EBF3 expression also leads to marked downregulation of cyclins and CDKs as well as proteins involved in cell survival such as Daxx and Mcl-1 (Liao 2009).

Gene expression profile studies revealed a significant down-regulation of *EBF3* expression or epigenetic silencing also in other cancers, as liver, prostate and head and neck squamous cell

carcinoma (Chen, Cheung et al. 2002) in line with the hypothesis that *EBF3* may act as tumour suppressor gene. On the contrary, our group demonstrated that in medulloblastoma, a malignant embryonal brain tumour, *EBF3* is highly expressed and acts as a major master regulator of neuronal differentiation. In fact, *EBF3* expression in early undifferentiated progenitors promotes neuronal differentiation. However, a sustained *EBF3* expression in medulloblastoma correlates with an immature phenotype, confirmed by lack of expression of mature markers of neuronal differentiation, and contributes to promote neoplastic progression (Corno, Pala et al. 2012). These data indicate that *EBF3*, according to specific cell contest, may act as an oncogene.

Since EBF3 is highly expressed in the PNC component of the GBM-PNC of our cohort, and is persistently negative in the GBM component, we hypothesize that it could have a major role in driving the divergent differentiation of the two components. In order to elucidate its function, we created a CRISPR/Cas9 EBF3 knockout model using a GBM stem cell line expressing EBF3. Concurrently we forced an EBF3 negative GSC line to overexpress *EBF3* by lentiviral transduction.

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas9 technology

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) is a microbial nuclease system involved in defense against invading phages and plasmids (Figure 26). CRISPR loci in microbial hosts contain a combination of CRISPR-associated (Cas) genes as well as non-coding RNA elements capable of programming the specificity of the CRISPR-mediated nucleic acid cleavage.

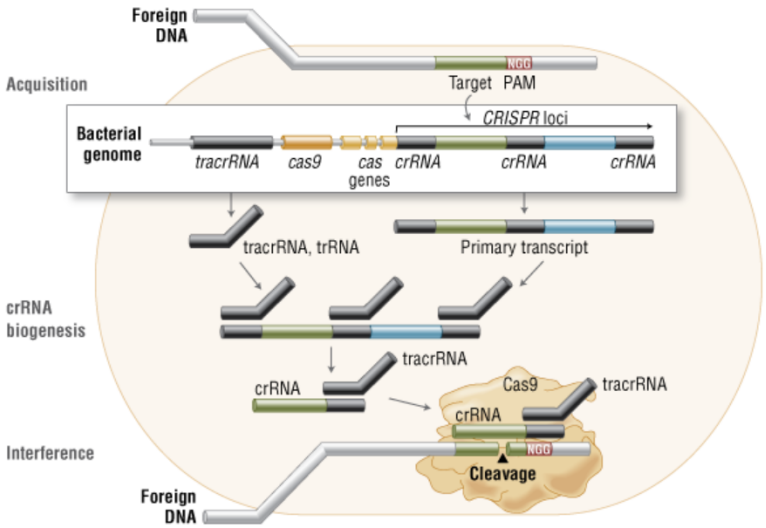


Figure 26. CRISPR/Cas9 system *in vivo*: bacterial adaptive immunity
In the acquisition phase, foreign DNA is incorporated into the bacterial genome at the CRISPR loci. CRISPR loci are then transcribed and processed into crRNA during crRNA biogenesis. During interference, Cas9 endonuclease complexed with a crRNA and separate tracrRNA cleaves foreign DNA containing a 20-nucleotide crRNA complementary sequence adjacent to the PAM sequence (Reis, Hornblower et al. 2014).

Three types of CRISPR mechanisms have been identified: Type I, Type II and Type III (Shabbir, Hao et al. 2016); Type II from Streptococcus pyogenes is the most studied. Invading DNA from viruses or plasmids is cut into small fragments (around 20 bps) and incorporated as “spacers” into a CRISPR locus amidst a series of short repetitive sequences (repeats). The loci are transcribed and transcripts are processed to generate small RNAs (crRNA), each harbouring a variable sequence transcribed from the invading DNA, known as the “protospacer” sequence, and part of the CRISPR repeat. Each crRNA hybridizes with a second “transactivating” RNA (tracrRNA) and they complex with the CRISPR-associated protein 9 (Cas9) nuclease. The protospacer-encoded portion of the crRNA directs Cas9 to cleave the complementary target-DNA sequences, generating double-strand breaks (DSB) (Sander and Joung 2014). The double-stranded endonuclease activity of Cas9 also requires the presence of a short conserved motif (2-5 bp), known as protospacer-associated motif (PAM). The PAM sequence is about 2-6 nucleotides downstream of the DNA sequence targeted by the guide RNA and the Cas nuclease cuts 3-4 nucleotides upstream of it. Each nuclease recognizes a specific PAM sequence: the most commonly-used Cas9 from Streptococcus pyogenes recognizes the PAM sequence 5’-NGG-3’(Swarts, Mosterd et al. 2012).

The simplicity of this system makes it a powerful platform for genome editing, transcriptional perturbation and epigenetic modulation. In 2012 Charpentier, Doudna and collaborators developed a two-components system, combining the tracrRNA and the crRNA into a single chimeric guide RNA (sgRNA) to induce sequence-specific DSBs and targeted genome editing in the presence of the Cas9 protein (Jinek, Chylinski et al. 2012). Briefly, 20 nucleotides at the 5’ end of the sgRNA direct Cas9 to a specific target DNA site, upstream of a 5’-NGG PAM sequence, using standard RNA-DNA complementarity base-pairing rules. With this system, Cas9 can be directed to any DNA sequence of the form N_{20}-NGG in order to introduce DSBs in the genomic DNA site complementary to the sgRNA sequence. For this outstanding discovery, in 2020 Charpentier and Doudna were awarded with the Nobel prize for Chemistry.The DSB induced by the Cas9 activates the double-strand break repair machinery in the cell. DSBs can be repaired by the Non-Homologous End Joining (NHEJ) pathway, resulting in insertions and/or deletions, disrupting the targeted locus. Alternatively, if a donor template with homology with the target locus is supplied, the DSB may be repaired by the homology-directed repair (HDR) pathway, allowing a precise replacement of the mutation (Figure 27).

Among the aforementioned alternatives, in the present work, Type II system with Cas9 from Streptococcus pyogenes was used; this system activates NHEJ for DNA repair of DSBs and the generation of indels in the target gene.

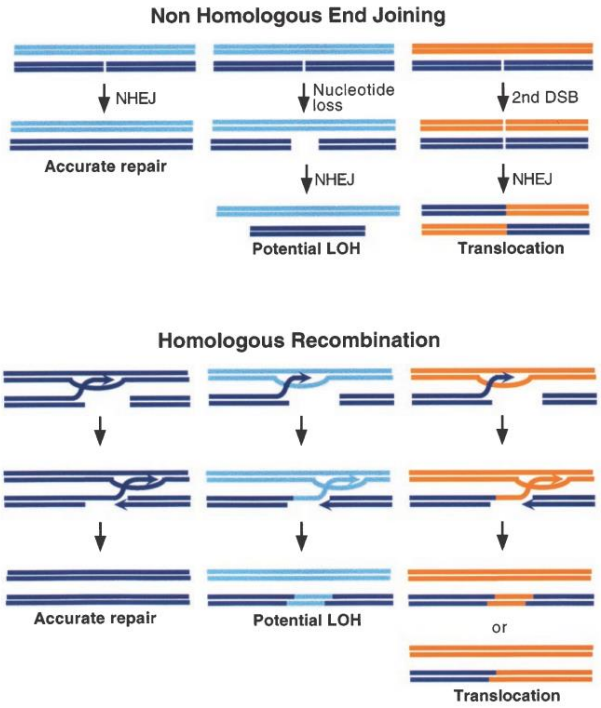


Figure 27. Mechanism of Double Strand Break repair
Upper figure shows Non-Homologous End Joining (NHEJ) mechanism of DSB repair with potential benefits and risks. Lower panel shows Homologous Recombination mechanism of DSB repair with potential benefits and risks (Ferguson and Alt 2001).

Use of lentiviral vectors for the transduction of Glioblastoma cells

Transfection with DNA constructs, such as plasmids, has been shown to be inefficient in Glioblastoma (Hagemann, Meyer et al. 2006). In contrast, viral gene delivery is highly efficient. In particular, lentiviral vectors, whose genome is reverse transcribed and stably integrated into the genome of host cells, have been widely used for the stable transduction glioblastoma stem cells. Lentiviral vectors can be engineered for CRISPR/Cas9 technique by co-expressing a mammalian codon-optimized Cas9 nuclease along with a single guide RNA (sgRNA) to facilitate genome editing (Shalem, Sanjana et al. 2014). Unlike other viruses (e.g., retroviruses), lentiviral infection does not depend on cell cycle status. Lentiviral particles can transduce fast cycling tumour cells, as well as quiescent or slowly dividing glioblastoma stem cells, and differentiated post-mitotic cells (Freed and Martin 1994).

As summarized in Figure 28, lentiviral systems are based on HIV-1, which encodes 3 large (polycistronic) open reading frames: *gag*, encoding proteins of the lentiviral core; *pol*, encoding

proteins required for replication (reverse transcriptase, integrase, protease); and *env*, encoding the glycoproteins that cover the viral surface (Sakuma, Barry et al. 2012). The HIV-1 envelope protein Env naturally targets CD4-positive cells. However, modern lentiviral transduction systems work with "pseudotyped" lentiviruses, in which the env gene has been replaced with the vesicular stomatitis virus envelope glycoprotein G (VSV-G) (Akkina, Walton et al. 1996). This envelope protein brings along a much broader tropism, since VSV-G targets phosphatidylserine on the surface of host cells.

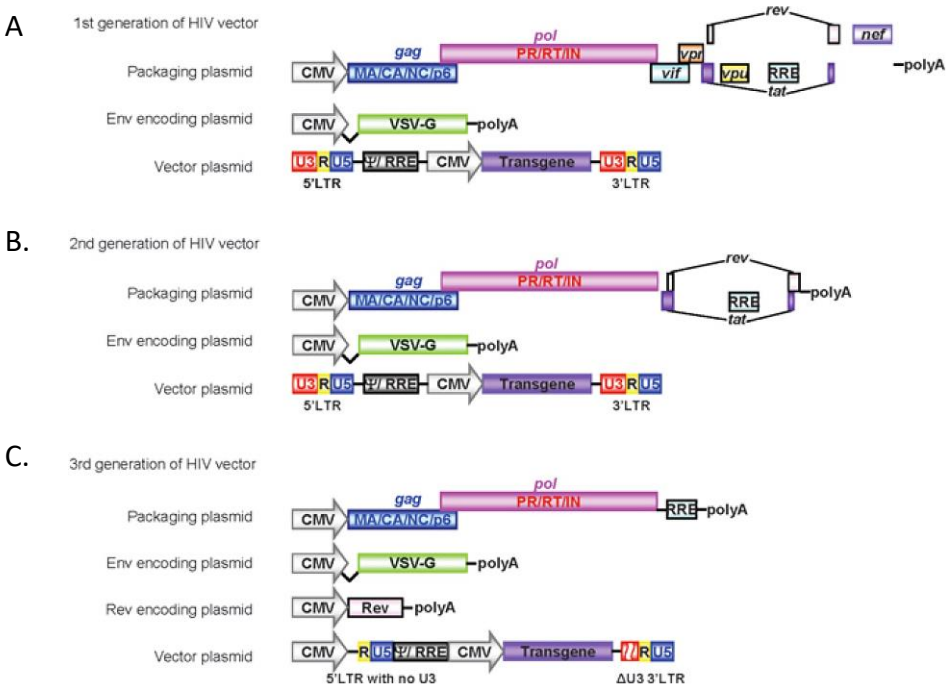


Figure 28. Schematic representation of HIV vectors
A. The first generation of HIV vectors includes all of the viral proteins, except Env protein, in a packaging plasmid. VSV-G is provided by a different plasmid. HIV vector plasmid contains LTRs and the transgene is expressed under a strong viral promoter such as the CMV promoter. B. For the second generation of HIV vectors, all of the accessory proteins are excluded from the packaging plasmid. Similar to the first generation of HIV vectors, expression of glycoprotein and transgene are provided by different plasmids. C. The third generation of HIV vectors requires four different plasmids. In addition to the three plasmids (i.e. a packaging plasmid, an Env-encoding plasmid and a vector plasmid), Rev protein is provided by a different plasmid. The vector plasmid is also modified by deleting the U3 region from 5'-LTR and partially deleting 3'-LTR to reduce the possible production of replication-competent viruses, and a strong viral promoter such as RSV or CMV is inserted for expression of the vector (Sakuma, Barry et al. 2012).

This pseudotyping allows lentiviral particles to infect essentially all cell types, a property known as pantropism. In addition, the VSV-G envelope confers higher stability of viral particles, allowing easier handling and harsher methods of concentration. Today's third generation lentiviral packaging systems have drastically ameliorated safety issues. Lentiviral vectors are replication-incompetent in host cells. Only in producer cells can viral particles be assembled. This is accomplished by splitting up the viral genome into separate plasmids that only in combination can form functional viral particles.

These systems consist of three packaging plasmids in combination with one transfer plasmid. Usually, one packaging plasmid carries the gag and pol operons of the viral genome, one plasmid carries the vsv-g envelope gene, and one plasmid encodes for the required regulatory gene rev. The transfer plasmid carries the transgene of interest, flanked by long terminal repeats (LTRs), as well as additional elements required for proper lentiviral packaging and genomic integration of the transgenic cassette. Only when all four plasmids are taken up by a cell will this system produce viral particles. These viruses can integrate into a host genome, but not further replicate (reviewed in great detail in (Sakuma, Barry et al. 2012). In addition, third-generation lentiviral vectors are self-inactivating (SIN). This safety modification refers to elimination of promoter function in viral LTRs during reverse transcription of the viral genome prior to integration into the host DNA. The SIN feature ensures that the viral LTR will not ectopically promote transcription of putative cellular proto-oncogenes after integration (Zufferey, Dull et al. 1998). In this system, viral transgene expression is dependent on promoters built into the modified viral genome.

MATERIALS AND METHODS

Neurosphere derivation from patients' specimens and culture

Neurospheres were isolated from GBM post-surgery samples as described in Galli et al. (Galli, Binda et al. 2004). GBM-PNC BT483 GSCs were derived in collaboration with the Laboratory of Cancer Stem Cell Research (Dr. C.Boccaccio) at Candiolo Cancer Institute, FPO IRCCS (Turin). GBM L0512 GSCs were derived in collaboration with the Neural Stem Cell Biology Unit (Dr. R.Galli) of San Raffaele Scientific Institute, Milan. Tissue samples were provided by Surgical Departments according to the ethical requirements of the institutional committee on human experimentation. Briefly, tissues are mechanical dissociated and digested with collagenase type I (Thermo Fisher Scientific). Single-cell suspensions were plated at clonal density (1-10 cells/µl) in standard medium containing: DMEM/F-12 (Euroclone), 2 mM glutamine (Euroclone), penicillin-streptomycin (1:100, EuroClone), B-27 (1:50, Thermo Fisher Scientific), human recombinant fibroblast growth factor 2 (bFGF, 20 ng/ml; Peprotech) and epidermal growth factor (EGF, 20 ng/ml; Peprotech). During the first passages, medium was replaced or supplemented with fresh growth factors until cells started to grow forming floating aggregates. To expand cultures, spheres were mechanically dissociated, counted with trypan blue to evaluate the number of live cells, and then re-plated as single cells at clonal density in complete fresh medium. Cells were incubated at 37°C, 5% CO_2, H_2O saturated atmosphere.

CRISPR/CAS9 mediated genome editing using a lentiviral vector

CRISPR/CAS9 mediated genome editing was used to generate clonal *EBF3* knockout GBM-PNC GSCs lines with a lentiviral vector.

Target selection

The gRNA chosen to target EBF3 gene was designed using the online resource: https://chopchop.cbu.uib.no/. The gRNA was designed to uniquely target the EBF3 gene within the genome and it was chosen for the absence of self-dimerization and off-target activity.

The sequence of the gRNA used was: **TGT ATC GAC AAC AGT CAA CG TGG** (see Figure 29).

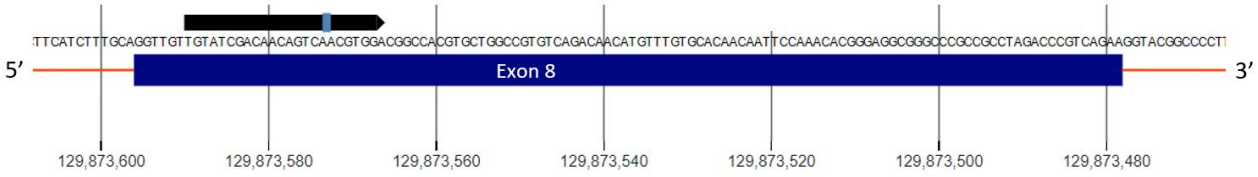


Figure 29. EBF3 gRNA sequence in the genome
The chosen gRNA targets a specific sequence in the 8th exon of EBF3 gene. For *S.pyogenes* Cas9, 20-bp targets (black box) must be followed at their 3' ends by 5'-NGG PAM site. Cas9 nuclease cuts 3-nt upstream of the PAM sequence. The cleavage site is indicated by the light blue box. The blue box represents the exon, while red lines represent introns.

Molecular cloning: oligos annealing and cloning

We used LentiCRISPRv2GFP (plasmid#82416 Addgene) (Walter, Venancio et al. 2017) an all-in-one 3rd generation lentiviral vector expressing GFP alongside Cas9 and a sgRNA cloning site (Figure 30). In this vector SpCas9 is driven by the human Elongation Factor-1α (EF1-α) promoter and the gRNA by the human U6 polymerase III promoter.

This expression vector contains a type IIS restriction site flanking the promoter-gRNA construct. Type IIS restriction enzymes, such as Esp3I, are different from traditional restriction enzymes because they cleave outside their recognition sequence, creating four base flanking overhangs. Since these overhangs are not part of the recognition sequence, they can be customized for the direct ligation of the gRNA into the plasmid.

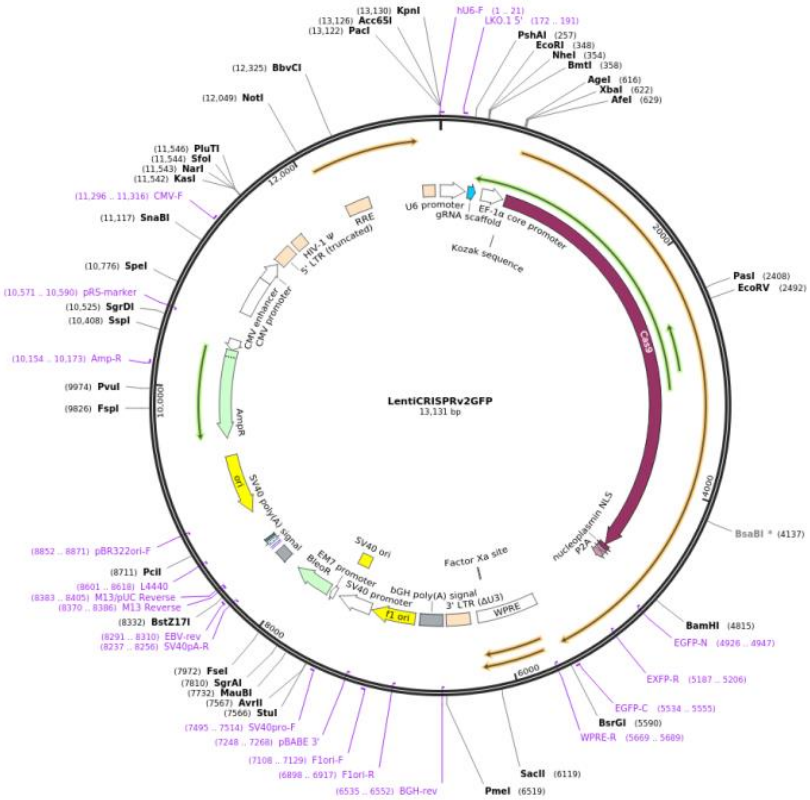


Figure 30. Map of LentiCRISPRv2-based vector

LentiCRISPRvsGFP lentiviral vector contains GFP alongside a *Streptococcus pyogenes* Cas9 driven by the human Elongation Factor-1α (EF1-α) promoter; the gRNA expression is driven by the human U6 polymerase III promoter.

In order to clone the target sequence into the LentiCRISPRv2 backbone (Figure 31), we followed the protocol provided by the Zhang's laboratory, available on https://www.addgene.org/crispr/reference/#protocols (Sanjana, Shalem et al. 2014, Shalem, Sanjana et al. 2014).

Briefly, 10 µg of plasmid was digested using 6U of Esp3I (BsmBI,Thermo Scientific), following the manufacturer's protocol. After the inactivation of the enzyme, 2U of shrimp alkaline phosphatase (SAP, Roche) were added to catalyze the phosphorylation of 5' phosphates. The digested product was run on an agarose gel and purified using the Wizard SV gel and PCR Clean-up System (Promega), according to the manufacturer's instructions.

In order to clone the target sequence into the digested LentiCRISPRV2 backbone, the oligo pairs encoding the 20-nt guide sequences were synthesized of the following form:

5'-CACCGNNNNNNNNNNNNNNNNNNN-3'
3'CNNNNNNNNNNNNNNNNNNNCAAA-5'

The overhang sequences 5'-CACC-3' (sense oligo) and 3'-CAAA-5' (antisense oligo) are complementary to the overhangs generated by the type IIS restriction enzyme Esp3I. The G in the sense oligo corresponds to the first nucleotide of the gRNA and it is necessary for efficient U6-driven expression. The sense oligo sequence corresponds to the exact sequence of the genomic target. The sequences of the oligos for the gRNA are:

- **Sense primer**: CACCGTGTATCGACAACAGTCAACG
- **Antisense primer**: AAACCGTTGACTGTTGTCGATACAC

100 picomol of each of the two oligos were annealed using T4 polynucleotide kinase (T4PNK, New England Biolabs) in a thermal cycler, using the following parameters: 37°C, 30 minutes; 95°C, 5 minutes; ramp down to 25°C at 5°C/minute. The ligation reaction was performed using 2U of T4 DNA Ligase (Promega) 1 hour at room temperature with 50 ng of Esp3I digested LentiCRISPRv2GFP and 1 µl of annealed oligos.

The correct insertion of the target sequence was analysed by Sanger sequencing using the following primers:

- **For pU6 universal:** 5'-GAGGGCCTATTTCCCATGATTCC–3'
- **Rev EBF3 gRNA:** 5'-TAGCTCTAAAACCGTTGACTGTT-3'

Cycle sequencing was carried out using BigDye Terminator v3.1 Cycle Sequencing kit (Applied Biosystems).and analysed on a 3130 DNA Analyzer ABI capillary electrophoresis system (Applied Biosystems). Sequences were then analysed using Snapgene Viewer software, v.5.2.2 (https://www.snapgene.com/snapgene-viewer/)

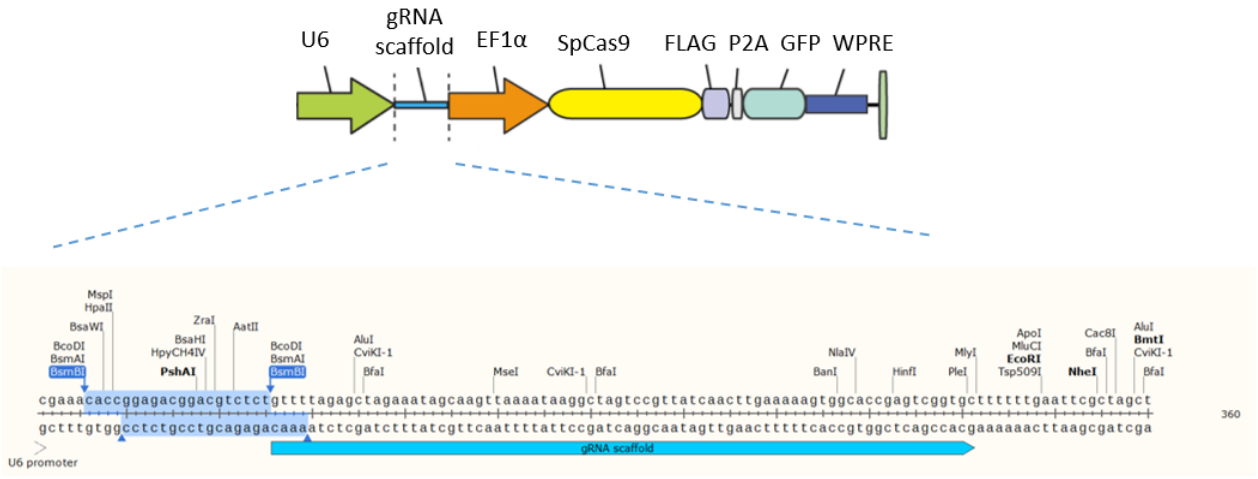


Figure 31. Detail of LentiCRISPRv2GFP cloning site (one vector system)
This plasmid contains two expression cassettes, hSpCas9 and the chimeric guide RNA. The vector can be digested using BsmBI, and a pair of annealed oligos can be cloned into the single guide RNA scaffold. The oligos are designed based on the target site sequence (20bp) and need to be flanked on the 3' end by a 3bp NGG PAM sequence.

At the end of the ligation, the plasmid was used to transform Stbl3™ Chemically Competent *E.*Coli (Thermo Fisher Scientific) following manufacturer's instructions. This strain reduces the frequency of homologous recombination of unstable regions that could occur in lentiviral plasmids due to the presence of long terminal repeats. Bacteria were subsequently plated on ampicillin selection plates. Positive colonies were screened for the presence of the target sequence by Sanger sequencing, as previously reported.

Lentiviral vectors production

In order to produce the lentiviral particles, we used HEK 293T cells, a cell line derived from the embryonic kidney cell line HEK 293 cell line. HEK 293T express a mutant version of the SV40 large T antigen, that is essential for replicating plasmids containing a SV40 origin of replication (Figure 32).

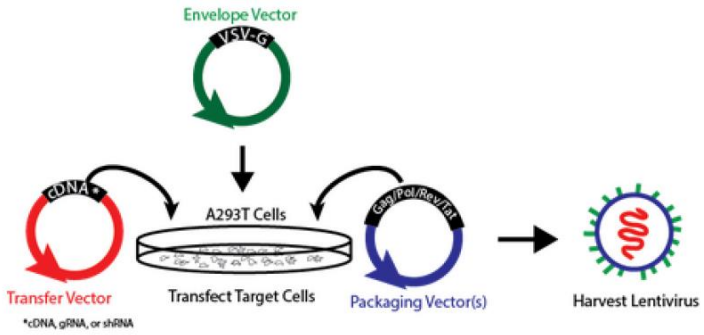


Figure 32. Viral production schematic
The 3rd generation transfer vector LentiCRISPRv2GFP, coding for EBF3 sgRNA, is transfected in HEK293T cells with a 2nd generation packaging system (plasmid psPAX2) and one envelope plasmid (p-CMV-VSV-G). Supernatant containing viral particles is collected at 24 and 48 hours (from https://www.addgene.org/protocols/lentivirus-production/).

For lentiviral production, we used a 2nd generation packaging system composed of the packaging plasmid psPAX2 (plasmid#12260 Addgene; gift from Didier Trono) and the envelope plasmid pCMV-VSV-G (plasmid#8454 Addgene) (Stewart, Dykxhoorn et al. 2003) (Figure 33).

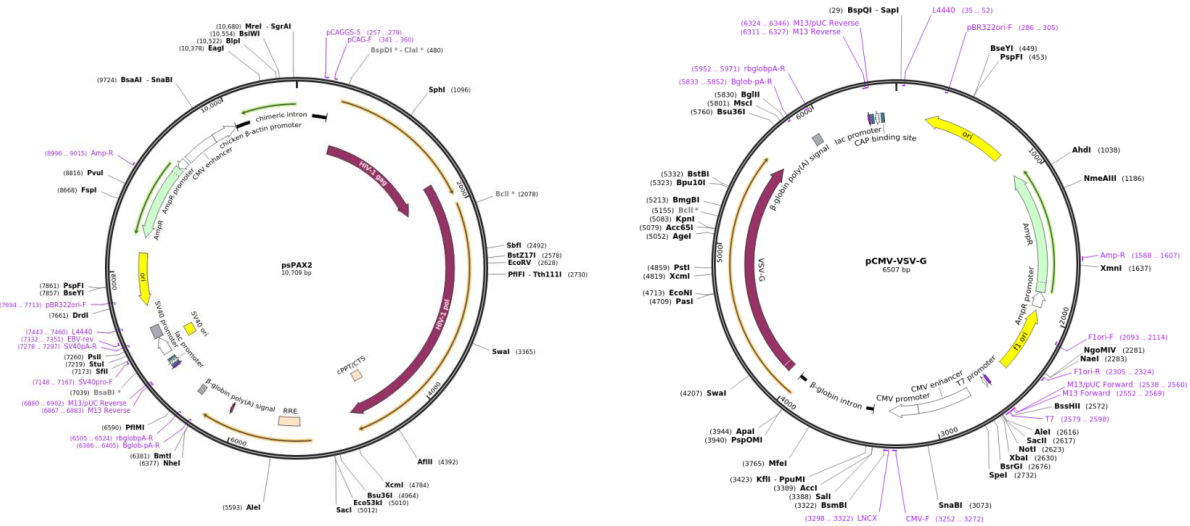


Figure 33. Maps of psPAX2 and pCMV-VSV-G plasmids
A 2nd generation packaging system composed of the packaging plasmid psPAX2 and pCMV-VSV-G plasmids are used for the packaging of the lentiviral vectors.

Cells were plated in a 15 cm tissue culture plate at the density of 65.000 cells/cm^2 in 10% FBS DMEM high glucose+L-glutamine and transfected with calcium phosphate. A mixture of DNA and calcium phosphate was prepared by mixing 20 µg of LentiCRISPRv2GFP, 15 µg of psPAX2 and 10 µg of pCMV-VSV-G with $CaCl_2$ at a final concentration of 250 mM (Merck Life Science). The precipitate was obtained by dropwise addition of the DNA-$CaCl_2$ mixture to the same volume of a 2X solution of Hepes Buffered Saline while bubbling air through it. The $CaPO_4$ precipitate was added to the cells and then, after incubation at 37°C overnight, the medium was replenished. Supernatant containing viral particles was collected after 24 and 48 hours. After filtration with a 0.45 µm pore size filter, the supernatant was centrifuged at 23.000 rpm for 2 hours at 4°C. The pellet was then resuspended in PBS, aliquoted and stored at -80°C.

Lentiviral Transduction of Primary Glioblastoma Stem Cells

BT483 neurospheres were dissociated as single cell suspension and plated at the density of 150.000 cells/cm^2 in a 24-well format plate in GBM stem cell medium. 2 µl of lentivirus were added dropwise and the plate was incubated at 37°C. After 16 hours, medium was replaced. After 48 hours, cells started to express GFP and the efficiency of the transduction was checked by fluorescence microscopy. Cells were then seeded at a density of 0.5 cells/well in a 96-well format plate.

Clonal colonies were screened by Sanger sequencing. Briefly, total DNA was isolated using the PureLink® Genomic DNA Mini Kit (Thermo Fisher Scientific), according to manufacturer's instruction.The genomic region spanning the EBF3 sgRNA site was amplified using PCR. Primers used for PCR were:

- **EBF3 Seq For** 5'-GGCTTGTATTTTGCCTAACAGG-3'
- **EBF3 Seq Rev** 5'-GGATGGTGCAAAGAAAAAGAAG-3'

PCR was performed using 150 ng of DNA and Xtra Taq Polymerase (Genespin). The product of PCR was processed for enzymatic removal of primers and nucleotides using EuroSAP PCR Enzymatic Clean-up kit (Euroclone), according to manufacturer's instructions. Cycle sequencing was carried out using BigDye Terminator v3.1 Cycle Sequencing kit (Applied Biosystems), as previously described. Sequencing products were purified using Performa® Spin Columns (Edge ByoSystems) and analysed on a 3130 DNA Analyzer ABI capillary electrophoresis system (Applied Biosystems). Sequences were then analysed using Chromas Lite 2.66 software (http://www.technelysium.com.au/chromas_lite.html).

GSCs culture and viral transduction

For viral transduction of GSCs L0512 with *EBF3*, a mouse *ebf3* cDNA clone was provided by Dr. Rossella Galli (San Raffaele Institute, Milan). Since EBF3 is a highly conserved protein, and the sequence of the human and mouse EBF3 proteins have a high identity and similarity (Figure 34), we decide to use a mouse construct already available in our laboratory. *ebf3* cDNA was cloned into the monocistronic transfer lentiviral vector (LV) pCCL.sin.cPPT.PGK.GFP.WPRE11. GFP was excided and substituted with FLAG-EBF3 cassette. In order to excide GFP, BamHI and SalI cloning sites were used. After cutting with BamHI the plasmid was blunted (T4 polimerase; New England Biolabs, Ipswich, MA), cut with SalI and de-phosphorylated using CIP enzyme (New

England Biolabs). From pCDNA-FLAG-EBF3 plasmid, full length FLAG-EBF3 was excided with XhoI and Pme. FLAG-EBF*3* was inserted into the lentiviral vector by using T4 ligase (New England Biolabs). Sister cultures were infected with the same vector coding for GFP, pCCL.sin.cPPT.PGK.GFP.WPRE11, as mock control. GSCs were transduced with 1x107 TU/ml of each LV for 16 hours.

A

>sp|Q9H4W6.2|COE3_HUMAN RecName: Full=Transcription factor COE3; AltName: Full=Early B-cell factor 3; Short=EBF-3; AltName: Full=Olf-1/EBF-like 2; Short=O/E-2; Short=OE-2
MFGIQENIPRGGTTMKEEPLGSGMNPVRSWMHTAGVVDANTAAQSGVGLARAHFEKQPPSNLRKSNFFHF
VLALYDRQGQPVEIERTAFVDFVEKEKEPNNEKTNNGIHYKLQLLYSNGVRTEQDLYVRLIDSMTKQAIV
YEGQDKNPEMCRVLLTHEIMCSRCCDKKSCGNRNETPSDPVIIDRFFLKFFLKCNQNCLKNAGNPRDMRR
FQVVVSTTVNVDGHVLAVSDNMFVHNNSKHGRRARRLDPSEGTAPSYLENATPCIKAISPSEGWTTGGAT
VIIIGDNFFDGLQVVFGTMLVWSELITPHAIRVQTPPRHIPGVVEVTLSYKSKQFCKGAPGRFVYTALNE
PTIDYGFQRLQKVIPRHPGDPERLPKEVLLKRAADLVEALYGMPHNNQEIILKRAADIAEALYSVPRNHN
QIPTLGNNPAHTGMMGVNSFSSQLAVNVSETSQANDQVGYSRNTSSVSPRGYVPSSTPQQSNYNTVSTSM
NGYGSGAMASLGVPGSPGFLNGSSANSPYGIVPSSPTMAASSVTLPSNCSSTHGIFSFSPANVISAVKQK
SAFAPVVRPQASPPPSCTSANGNGLQAMSGLVVPPM

>sp|O08791.1|COE3_MOUSE RecName: Full=Transcription factor COE3; AltName: Full=Early B-cell factor 3; Short=EBF-3; AltName: Full=Olf-1/EBF-like 2; Short=O/E-2; Short=OE-2
MFGIQENIPRGGTTMKEEPLGSGMNPVRSWMHTAGVVDANTAAQSGVGLARAHFEKQPPSNLRKSNFFHF
VLALYDRQGQPVEIERTAFVDFVEKEKEPNNEKTNNGIHYKLQLLYSNGVRTEQDLYVRLIDSMTKQAIV
YEGQDKNPEMCRVLLTHEIMCSRCCDKKSCGNRNETPSDPVIIDRFFLKFFLKCNQNCLKNAGNPRDMRR
FQVVVSTTVNVDGHVLAVSDNMFVHNNSKHGRRARRLDPSEGTAPSYLENATPCIKAISPSEGWTTGGAT
VIIIGDNFFDGLQVVFGTMLVWSELITPHAIRVQTPPRHIPGVVEVTLSYKSKQFCKGAPGRFVYTALNE
PTIDYGFQRLQKVIPRHPGDPERLPKEVLLKRAADLVEALYGMPHNNQEIILKRAADIAEALYSVPRNHN
QIPTLGNTPAHTGMMGVNSFSSQLAVNVSETSQANDQVGYSRNTSSVSPRGYVPSSTPQQSNYNTVSTSM
NGYGSGAMANLGVPGSPGFLNGSSANSPYGIVPSSPTMAASSVTLPSNCSSTHGIFSFSPANVISAVKQK
SAFAPVVRPQASPPPSCTSANGNGLQAMSGLVVPPM

B

```
# Aligned_sequences: 2
# Length: 596
# Identity:     594/596 (99.7%)
# Similarity:   595/596 (99.8%)
# Gaps:           0/596 (0.0%)
# Score: 3124.0
#=======================================

Q9H4W6.2         1 MFGIQENIPRGGTTMKEEPLGSGMNPVRSWMHTAGVVDANTAAQSGVGLA     50
                   |||||||||||||||||||||||||||||||||||||||||||||||||||
O08791.1         1 MFGIQENIPRGGTTMKEEPLGSGMNPVRSWMHTAGVVDANTAAQSGVGLA     50

Q9H4W6.2        51 RAHFEKQPPSNLRKSNFFHFVLALYDRQGQPVEIERTAFVDFVEKEKEPN    100
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||
O08791.1        51 RAHFEKQPPSNLRKSNFFHFVLALYDRQGQPVEIERTAFVDFVEKEKEPN    100

Q9H4W6.2       101 NEKTNNGIHYKLQLLYSNGVRTEQDLYVRLIDSMTKQAIVYEGQDKNPEM    150
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||
O08791.1       101 NEKTNNGIHYKLQLLYSNGVRTEQDLYVRLIDSMTKQAIVYEGQDKNPEM    150

Q9H4W6.2       151 CRVLLTHEIMCSRCCDKKSCGNRNETPSDPVIIDRFFLKFFLKCNQNCLK    200
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||
O08791.1       151 CRVLLTHEIMCSRCCDKKSCGNRNETPSDPVIIDRFFLKFFLKCNQNCLK    200

Q9H4W6.2       201 NAGNPRDMRRFQVVVSTTVNVDGHVLAVSDNMFVHNNSKHGRRARRLDPS    250
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||
O08791.1       201 NAGNPRDMRRFQVVVSTTVNVDGHVLAVSDNMFVHNNSKHGRRARRLDPS    250

Q9H4W6.2       251 EGTAPSYLENATPCIKAISPSEGWTTGGATVIIIGDNFFDGLQVVFGTML    300
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||
O08791.1       251 EGTAPSYLENATPCIKAISPSEGWTTGGATVIIIGDNFFDGLQVVFGTML    300

Q9H4W6.2       301 VWSELITPHAIRVQTPPRHIPGVVEVTLSYKSKQFCKGAPGRFVYTALNE    350
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||
O08791.1       301 VWSELITPHAIRVQTPPRHIPGVVEVTLSYKSKQFCKGAPGRFVYTALNE    350

Q9H4W6.2       351 PTIDYGFQRLQKVIPRHPGDPERLPKEVLLKRAADLVEALYGMPHNNQEI    400
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||
O08791.1       351 PTIDYGFQRLQKVIPRHPGDPERLPKEVLLKRAADLVEALYGMPHNNQEI    400

Q9H4W6.2       401 ILKRAADIAEALYSVPRNHNQIPTLGNNPAHTGMMGVNSFSSQLAVNVSE    450
                   ||||||||||||||||||||||||||.||||||||||||||||||||||||
O08791.1       401 ILKRAADIAEALYSVPRNHNQIPTLGNTPAHTGMMGVNSFSSQLAVNVSE    450

Q9H4W6.2       451 TSQANDQVGYSRNTSSVSPRGYVPSSTPQQSNYNTVSTSMNGYGSGAMAS    500
                   ||||||||||||||||||||||||||||||||||||||||||||||||||:
O08791.1       451 TSQANDQVGYSRNTSSVSPRGYVPSSTPQQSNYNTVSTSMNGYGSGAMAN    500

Q9H4W6.2       501 LGVPGSPGFLNGSSANSPYGIVPSSPTMAASSVTLPSNCSSTHGIFSFSP    550
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||
O08791.1       501 LGVPGSPGFLNGSSANSPYGIVPSSPTMAASSVTLPSNCSSTHGIFSFSP    550

Q9H4W6.2       551 ANVISAVKQKSAFAPVVRPQASPPPSCTSANGNGLQAMSGLVVPPM    596
                   |||||||||||||||||||||||||||||||||||||||||||||
O08791.1       551 ANVISAVKQKSAFAPVVRPQASPPPSCTSANGNGLQAMSGLVVPPM    596
```

Figure 34. Alignment of human EBF3 protein and mouse EBF3 protein
Panel A shows the sequence of EBF3 human protein (UniprotKB-Q9H4W6.2 COE3_HUMAN) and the sequence of the mouse ortholog (UniprotKB-O08791 COE3_MOUSE). Panel B shows the sequence alignment obtained with EMBOSS Needle using the Needleman-Wunsch algorithm (https://www.ebi.ac.uk/Tools/psa/).

Nucleic acid extraction

Genomic DNA (gDNA) was extracted from established neurospheres using the PureLink Genomic kit (Euroclone) following manufacturer's instructions. Extracted purified nucleic acids were quantified with Nanodrop mySPEC (VWR) and used as described.

Total RNA was extracted by Trizol method (Ambion, Thermo Fisher Scientific) following manufacturer's instructions. 1 µg of total RNA was submitted to a reverse-transcription reaction using iSCRIPT cDNA synthesis kit (Bio-Rad) in a final 20µL volume.

Reverse Transcription-Quantitative Polymerase Chain Reaction (RT-qPCR)

RT-qPCR was performed on a ViiA7 Thermal cycler (Thermo Fisher Scientific). Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) mRNA was used as endogenous reference for the relative quantification.

The following assays was performed using Gene Expression Taqman® with the Fast Advanced Master Mix (Thermo Fisher Scientific) system:

- Hs GADPH Gene Expression Taqman® Hs99999905_m1;
- Hs EBF1 Gene Expression Taqman® Hs00395518_m1;
- Hs EBF2 Gene Expression Taqman® Hs00970594_m1;
- Hs EBF3 Gene Expression Taqman® Hs01008793_m1;
- Hs EBF4 Gene Expression Taqman® Hs00985603_g1;
- Hs TUBB3 Gene Expression Taqman® Hs00964963_g1;
- Hs GFAP Gene Expression Taqman® Hs00909236_m1;
- Hs NESTIN Gene Expression Taqman® Hs04187831_g1;
- Hs SOX2 Gene Expression Taqman® Hs01053049_s1.

The following assay was performed using iTaq™ Universal SYBR® Green Supermix (Bio-Rad):

- Hs GAPDH For: 5' ACAGTTGCCATGTAGACC 3';
- Hs GAPDH Rev: 5' TTGAGCACAGGGTACTTTA 3';
- Mm Ebf3 For: 5´ AGAGCCGAACAACGAGAAAA 3´;
- Mm Ebf3 Rev: 5´ GCACATCTCCGGATTCTTGT 3´;

Western blot

Cells were scraped from dishes in ice-cold Lysis Buffer (Thermo Fisher Scientific) containing Phosphatase and Protease Inhibitor Cocktail (ThermoFisher). Lysates were centrifuged at 14000 rpm for 15 min at 4°C and proteins were quantified using the Bradford protein assay (Euroclone). 30-60 µg of protein extracts were resolved by SDS-PAGE using Bolt™ 8%-10% Bis-Tris Plus pre-cast gels (Thermo Fisher Scientific), then transferred to PVDF Transfer Membrane (Thermo Fisher Scientific). Membranes were processed and incubated overnight at 4°C with the following primary antibodies: mouse monoclonal anti-actin (clone AC-40, 1:500, Sigma Aldrich); mouse monoclonal anti-EBF3 (clone 8D6, 1:3000, Abnova); mouse monoclonal anti-

GFAP (clone 2E1, 1:500, Santa Cruz Biotechnologies) rabbit polyclonal anti-Nestin (clone 10C2, 1:1000, Millipore). After incubation with horseradish peroxydase-conjugated secondary antibodies (Euroclone), membranes were developed with ECL Star Enhanced Chemiluminescent Substrate (Euroclone) and visualized with LI-COR Imaging System (Odissey). Band intensity was quantified by densitometric analysis using ImageJ software.

Short-Term Self-Renewal Assay (Clonogenic Assay)

For the clonogenic assay, cells derived from the dissociation of clonal single neurospheres were seeded into uncoated 24-well plates in proliferation medium at a density of 50 cells/well. Cells were counted in each well 2 h after plating, in order to obtain a "baseline" value to be used for establishing the efficiency of secondary sphere generation. Neurospheres generated in each well were counted 20 days after plating. The number of secondary neurospheres was normalized to the number of cells originally seeded in each well.

Long-Term Self-Renewal Assay (Growth Curve/Population Assay)

For growth curve/population analyses, cells were seeded at a density of 8000 cells/cm^2 in proliferation medium. After 7 days, spheres were collected and cells counted by trypan blue exclusion after spheres disgregation. These steps were repeated for 3 subculturing passages. At any passage, the relative amplification rate was calculated by dividing the amount of counted cells by the number of cells originally seeded. The amplification rate was multiplied for the number of cells originally seeded and this number plotted in correspondence of every subculturing passage. In this way, a growth curve was generated, whose slope provides information as to the frequency of putative Cancer Stem Cells (CSCs).

Differentiation assay

To assess for multipotency, cells were plated at a density of 25000 cells/cm^2 in a 12 wells plate previously coated with Geltrex™ LDEV-Free Reduced Growth Factor Basement Membrane Matrix (Gibco) in standard medium containing human recombinant bFGF (20 ng/ml) and EGF (20 ng/ml). After 4 hours, medium was changed and fresh medium deprived of mitogens and supplemented with 2% Fetal Bovine Serum (Gibco) was added. After 7 days, cells were collected for RNA extraction and subsequent RT-qPCR.

Cell blocks preparation

2 to 3 million cells were resuspended in physiological solution (NaCl 0,9%); plasma and tromboplastin were added dropwise until the formation of a clotted sphere. The sphere was immediately formalin fixed and paraffin embedded. Cell blocks were cut into 2 µm sections and processed as previously described (see PART I, Materials and Methods-Immunostaining on paraffin-embedded sections)

Animal procedures

Mice were handled in agreement with the guidelines conforming to current Italian regulations for the protection of animals used for scientific purposes (D.lgs 26/2014) that implements EU Directive 2010/63/EU

for animal experimentations. Procedures were approved by the Ethical Committee for Animal Experimentation of the Ospedale Policlinico San Martino and by the Italian Ministry of Health. The experiments were performed with the BALB/c mouse strain NOD.CB17-Prkcd^{scid}/NCrHsd (Envigo). Mice were anesthetized with a cocktail containing Fentanyl/Midazolam/Metadominine (0.05/5/0.5 mg/kg, respectively), a 5 mm scalp incision was performed to expose the bregma area. 1 µl of suspension, containing transduced cells, were injected by a Hamilton syringe guided by a stereotaxic apparatus at bregma coordinates: anterior-posterior, 1.0 mm; lateral, 1.5 mm left and 2.5 mm below the skull surface. After surgery skin was sutured with adsorbable wire (Johnson & Johnson Ethicon Sutures Vicryl 4/0) and mice were awakened by a specific antidote cocktail containing Atipamezol, Flumacenil and Naloxone (2.5/0.5/1.2 mg/kg, respectively). Mice were monitored daily after transplant and killed at first signs of neurological distress or at experiment endpoint. Brains were removed and fixed in 10% formalin. 2–4 µm thick paraffin-embedded tissue sections were used for hematoxylin and eosin staining and immunohistochemistry as previously described (see PART I, Materials and Methods-Immunostaining on paraffin-embedded sections).

Statistical analysis

Values were expressed as mean ± standard deviation (SD). In vitro experiments were repeated at least three times. Statistical analyses were performed using GraphPad Prism Software (GraphPad Software Inc.). Unpaired two-sided Student's t –test was used. P -value < 0.05 was considered to be significant.

* $p<0.05$; ** $p<0.01$; *** $p<0.001$; **** $p<0.0001$.

RESULTS

Generation and characterization of GSCs knockout and overexpressing EBF3 clones

To understand the putative role of EBF3 in the neuronal cell commitment and in the development of the dual phenotype of GBM-PNC, we generated *EBF3* knockout and EBF3 overexpressing cancer stem cells.

For *EBF3* knockout, we selected a GSC line derived from a GBM-PNC (BT483) that constitutively expressed high levels of EBF3. Clones were generated using CRISPR/Cas9 technology. BT483 were transduced with a lentiviral vector containing LentiCRISPRv2GFP (plasmid#82416 Addgene) (Walter, Venancio et al. 2017), an all-in-one 3rd generation lentiviral plasmid expressing GFP alongside Cas9. To the point of the sgRNA cloning site, we cloned a sgRNA targeting exon 8 of *EBF3* gene. Wildtype controls were obtained transducing BT483 with a lentivirus containing the LentiCRISPRv2GFP without *EBF3* sgRNA.

GFP expressing clones were screened by Sanger sequencing. Finally, we chose 2 wildtype clones, denominated as “Beatrice” (BT483-WT-Bea) and “Christopher” (“BT483-WT-Chr”), and 4 knockout clones, denominated as “Dave” (BT483-KO-Dav), “Elizabeth” (BT483-KO-Eli), “Frances” (BT483-KO-Fra) and Harry (BT483-KO-Har), for further analysis. In Figure 35 mutations obtained are shown.

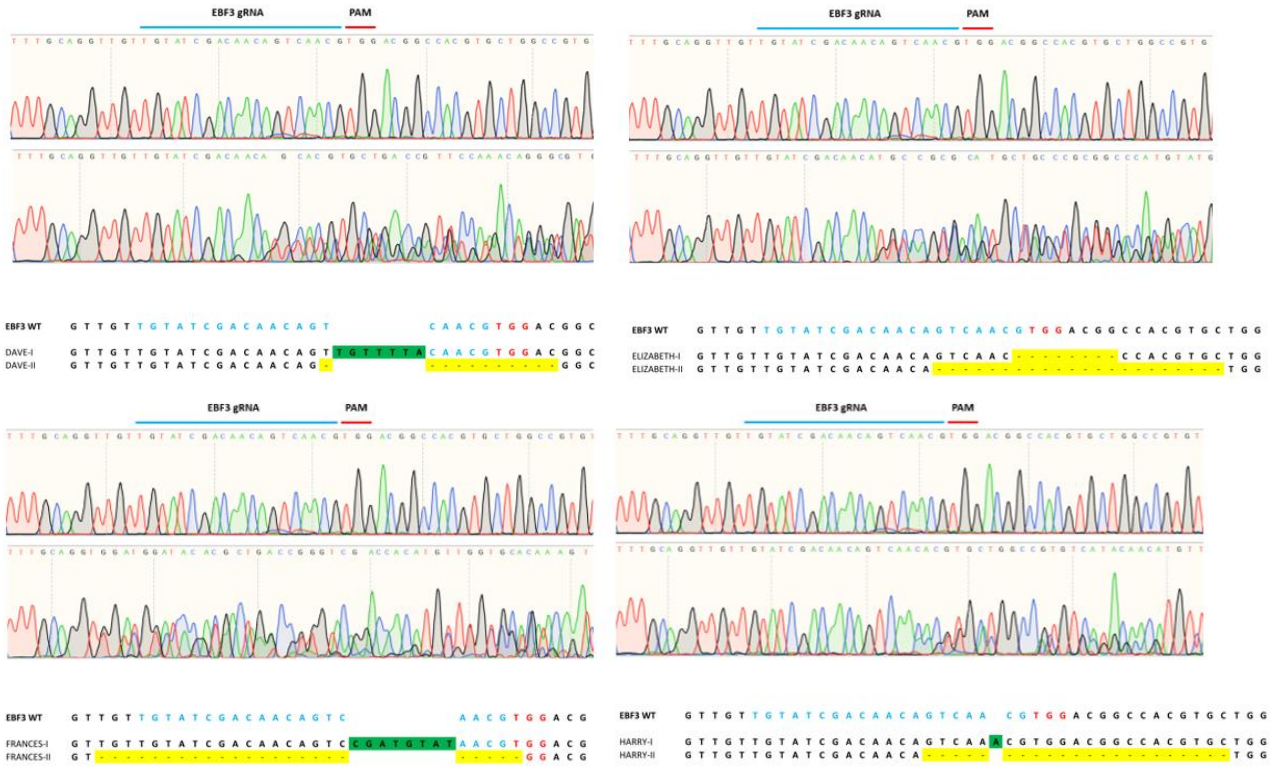


Figure 35. Sequence analysis of EBF3 knockout clones

Sequence chromatograms of the genomic region containing the EBF3 exon 8 gRNA for each of the 4 selected knockout clones, as compared to the wildtype sequence (upper line). The sgRNA region is highlighted in blue, the PAM region in red. Insertions are depicted as green boxes, deletions are depicted as yellow boxes.

EBF3 overexpressing cells were obtained by transducing EBF3-negative L0512 GSCs with a lentiviral vector coding for mouse *ebf3* full-length cDNA and featuring the FLAG epitope as a tag (L0512-EBF3). The same line was transduced with a lentiviral vector coding for GFP, as a mock control (L0512-GFP). BT483 cell clones under proliferative conditions did not show any change in morphology as compared to their mock control, showing the typical aspect of floating neurospheres, as shown in Figure 36. Contrariwise, L0512-EBF3, do not form sphere-like aggregates, growing as adherent cells.

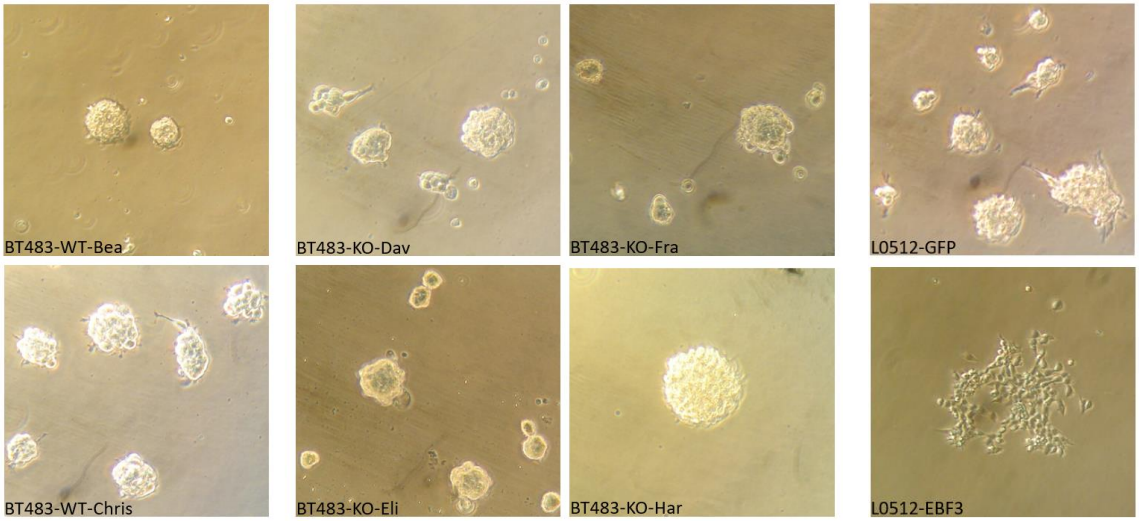


Figure 36. Morphology of BT483 and L0512 cell clones under proliferative conditions
Pictures in bright-field microscopy of the cell lines used in this work (400X).

All the knockout or overexpressing clones were analysed for EBF3 expression at both transcriptional and protein levels.

RT-qPCR with specific probes revealed a net decrease of EBF3 expression in knockout clones and increase in EBF3 expression in L0512EBF3 (Figure 37).

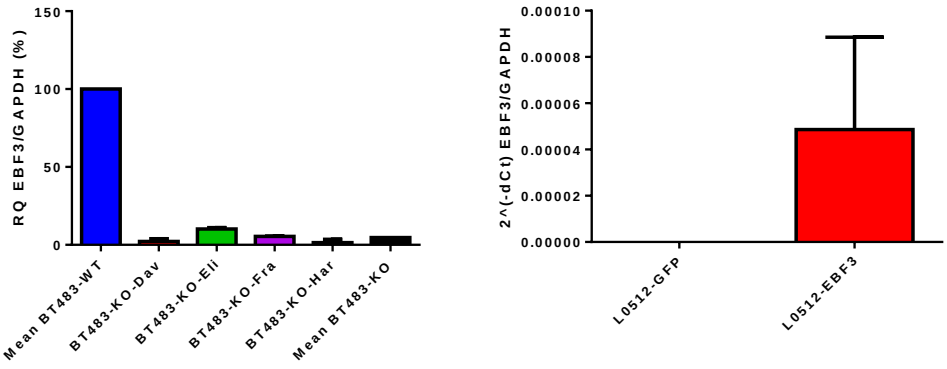


Figure 37. *EBF3* mRNA expression in BT483 and L0512 cell clones
EBF3 mRNA expression was normalized on *GAPDH* housekeeping gene; fold increase of *EBF3* mRNA in knockout clones was obtained with respect to the mean dCt value of the wildtype clones.

Of note, the analysis in Western blot showed a positive signal for all the clones, as shown in Figure 38, even though densitometric quantification revealed a lower amount of protein for the knockout BT483 clones and for the L0512-GFP mock cell line. Also IHC on cell block revealed a positive signal in BT483 knockout clones and L0512-GFP.

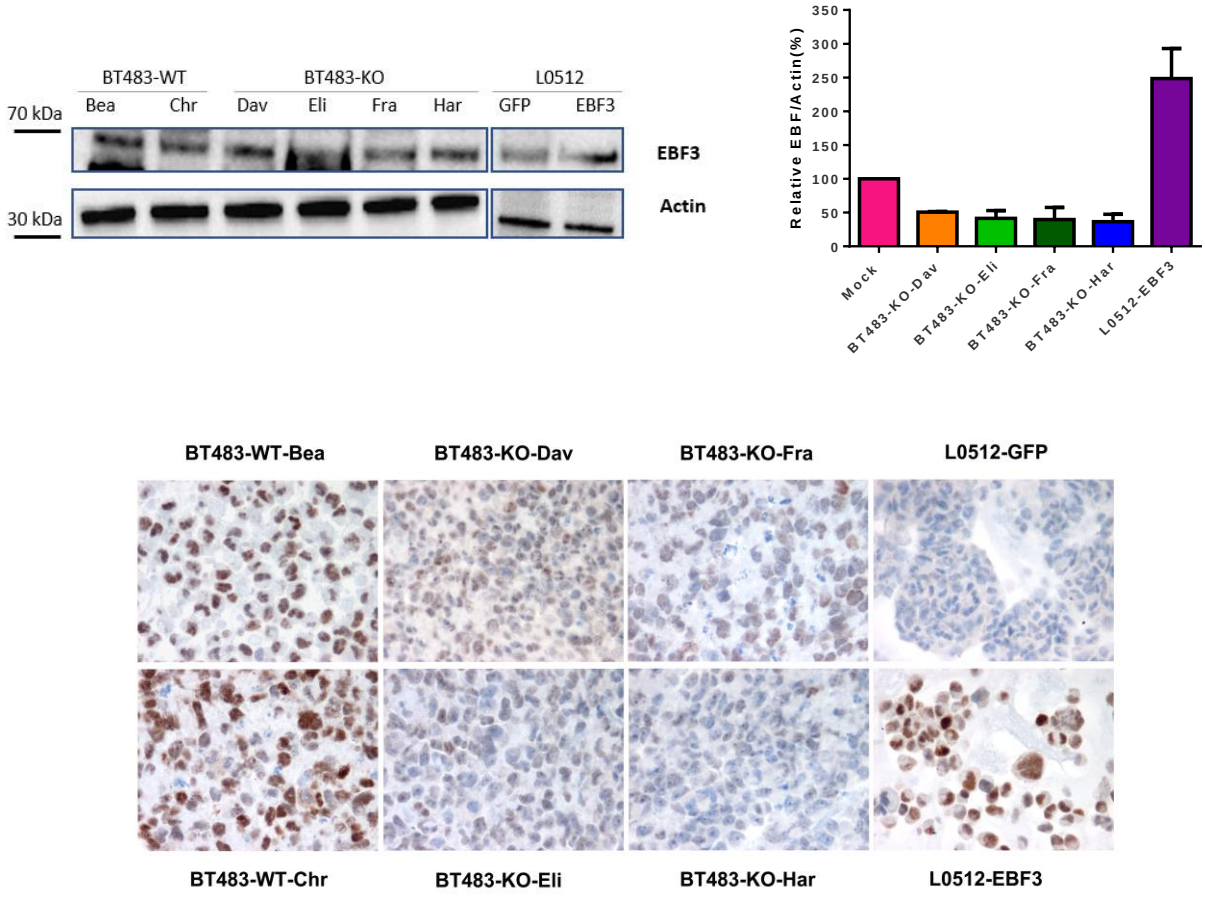


Figure 38. EBF3 protein expression in BT483 and L0512 clones
EBF3 expression was analysed by Western blot (upper panel). On the left, the picture of a representative experiment. Protein levels were analysed by densitometric quantification of 3 independent experiments, as shown in the right plot. Protein expression was normalized on actin housekeeping protein; fold increase was obtained with respect to their matched mock-transduced samples. In the lower panel, IHC for EBF3 on cell blocks of the different clones.

In a previous paper from our group (*EBF1 is expressed in pericytes during angiogenesis and contributes to the pericyte cell commitment*; Pagani et al. SUBMITTED) we demonstrated that EBF3 antibody used in this work is not specific for EBF3, but is able to recognise also the other EBF family members. Since no EBF3 specific antibody is available, we decided to investigate by RT-qPCR whether the antibody positivity could be attributable to other EBFs. Not surprisingly we found that all of the cell lines expressed both EBF1 and EBF4, being negative for EBF2 expression (Figure 39). As a matter of fact, it is known from the literature that, in the CNS, all EBF family members have a key role in neuronal differentiation and regional specification (Garel, Marin et al. 1997) and that

these factors are expressed at different time points and in different regions. However we did not observe any significant change in *EBF1* and *EBF4* mRNAs amount among the BT483 cell clones that could be attributable to the knockout of *EBF3* and could be interpreted as a mechanism of compensation. Of note, mock-transduced L0512-GFP cells do not express *EBF1*, while EBF3 overexpressing L0512-EBF3 do.

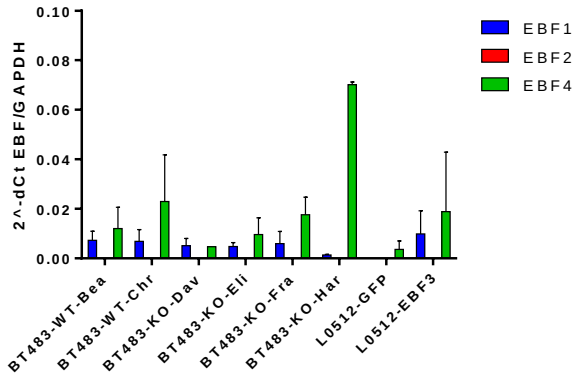


Figure 39. EBFs mRNA expression in BT483 and L0512 clones
*EBF*s mRNA expression was normalized on *GAPDH* housekeeping gene.

Effect of EBF3 on cell proliferation and self-renewal ability

It is reported in the literature that EBF3 is expressed in normal brain cells but is silenced in brain tumour cells (Zardo, Tiirikainen et al. 2002) and there is also evidence suggesting that EBF3 has a tumour suppressive role in other types of cancer (Chen, Cheung et al. 2002) Functional studies revealed that EBF3-mediated tumour suppression could be explained by EBF3 repression of genes required for cell proliferation and survival and the activation of those involved in cell cycle arrest (Zhao, Niu et al. 2006). We then investigated whether EBF3 suppression or overexpression could influence cell proliferation and stem cell self-renewal ability.

For growth curve assay, cells were plated and counted after trypan blue exclusion for 3 subculturing passages. Surprisingly, BT483 EBF3-knockout clones showed a lower kinetic of expansion, as compared to the wildtype BT483 clones. No differences were observed between L0512 mock and overexpressing clones (Figure 40). The same difference in self-renewal ability was observed when serial clonogenic assays were performed (Figure 41).

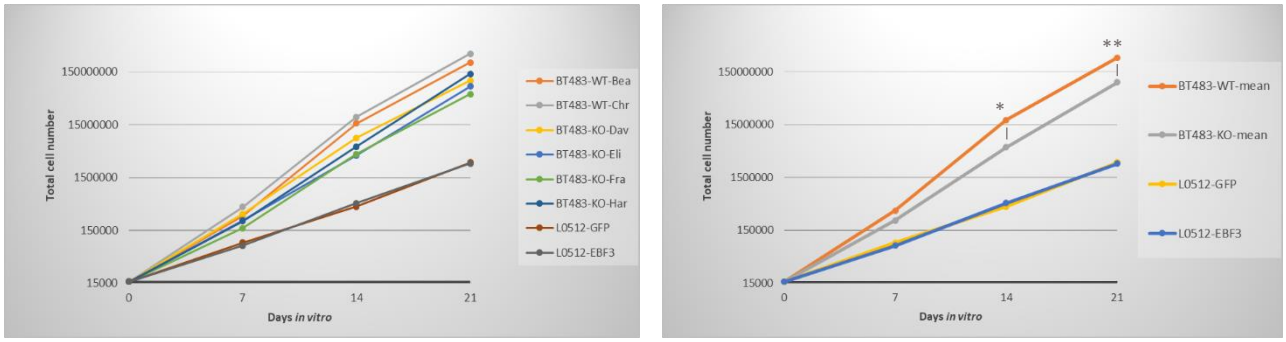


Figure 40. Self-renewal ability of BT483 and L0512 cell clones
Upper plot showing long-term growth curves of BT483 and L0512 cell lines. Three independent experiments. Mean BT483-WT Vs. Mean BT483-KO, 14 days: **p<0,01; Mean BT483-WT Vs. Mean BT483-KO, 21 days: *p<0,05, Student's t test.

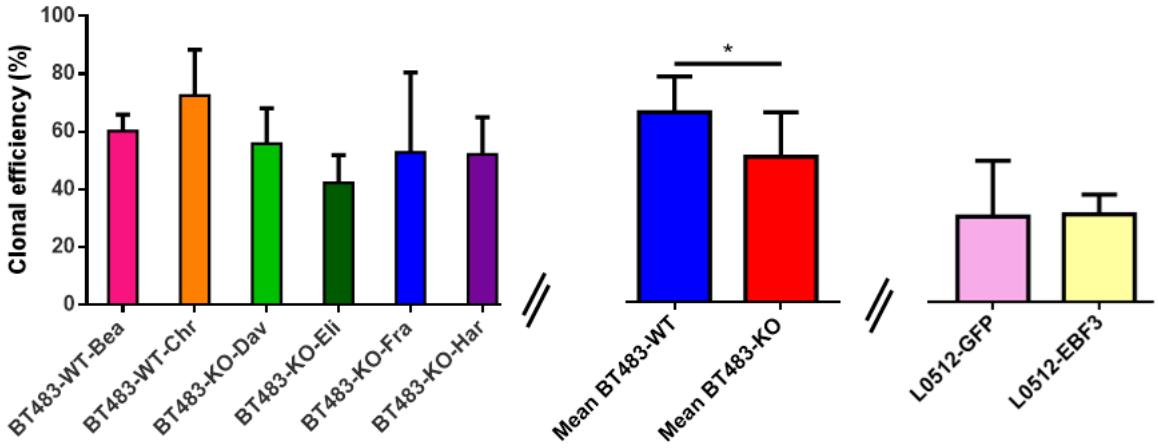


Figure 41. Clonogenic assay of BT483 and L0512 cell clones
Left plot shows clonogenic assay of BT483 cell lines. Three independent experiments. Central plot: Mean BT483-WT Vs. Mean BT483-KO,*p<0.05, Student's t test. Right plot: clonogenic assay of L0512 cell lines, showing no differences in self-renewal properties.

Effect of EBF3 on cell fate differentiation

It is reported in the literature that *EBF3* is relevant in the process of neurogenesis and in the differentiation towards the neuronal cell fate (Smits, Magni et al. 2020). In fact, *EBF3* expression in early undifferentiated progenitors promotes neuronal differentiation. However, a sustained *EBF3* expression in medulloblastoma correlates with an immature phenotype, confirmed by lack of expression of mature markers of neuronal differentiation, and contributes to promote neoplastic progression (Corno, Pala et al. 2012).

We then investigated the expression of differentiation markers in transduced GSCs both under proliferative and differentiative conditions.

Glial fibrillary acidic protein (GFAP) is a biomarker of astrocytes and, in GBM-PNC, we showed that it is expressed by the glial component, while the PNC component is negative. We analysed the

expression levels of this biomarker in GSC clones under proliferative conditions. In EBF3-knockout BT483 clones we found a certain variability in the levels of GFAP expression at both protein and transcriptional level; even though the global levels appear to be higher as compared to the wildtype control, the mean difference is not statistically significant (Figure 42). Contrariwise, transduced L0512-EBF3 express a lower level of GFAP protein as compared to their mock control and mRNA is barely detectable.

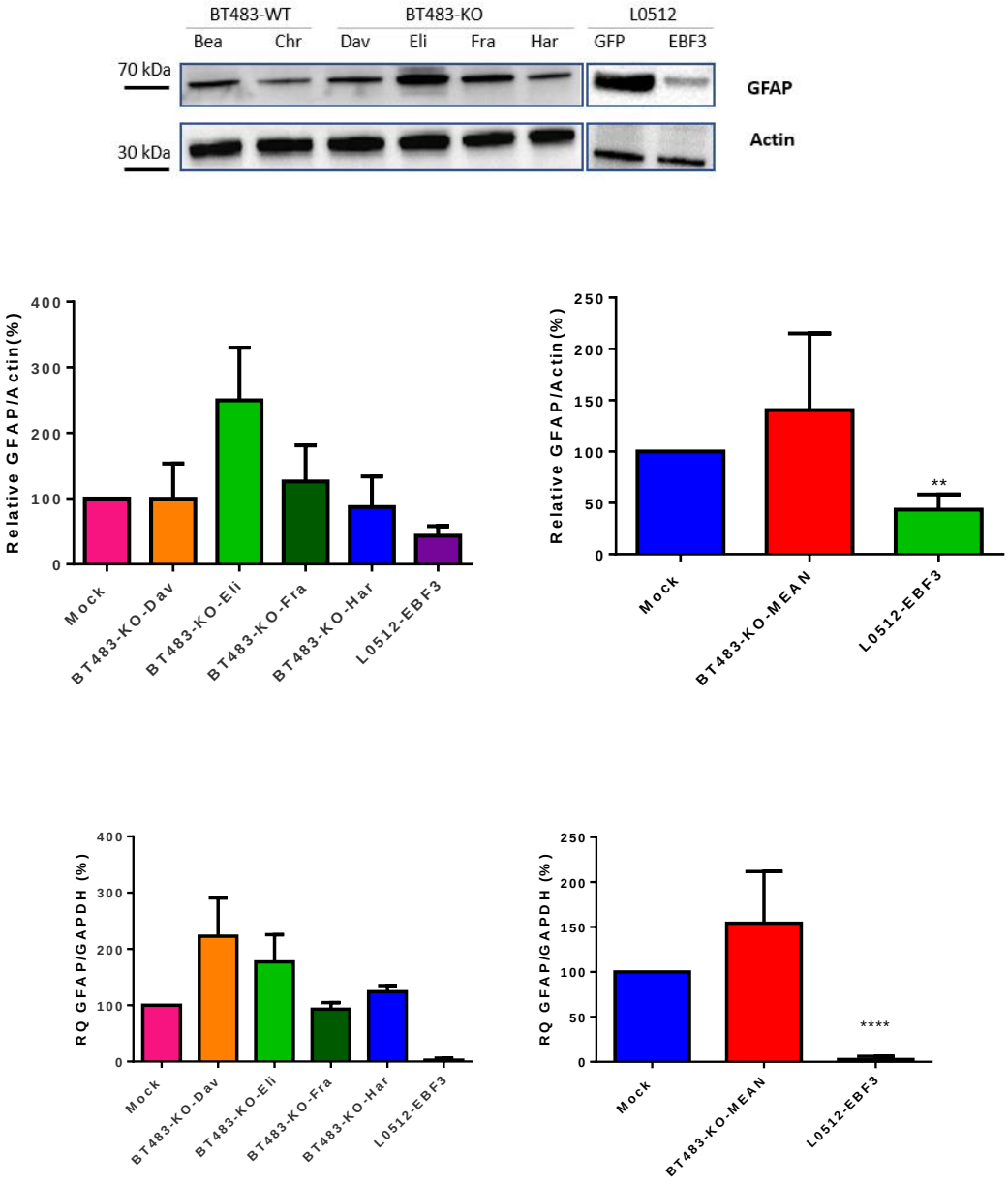
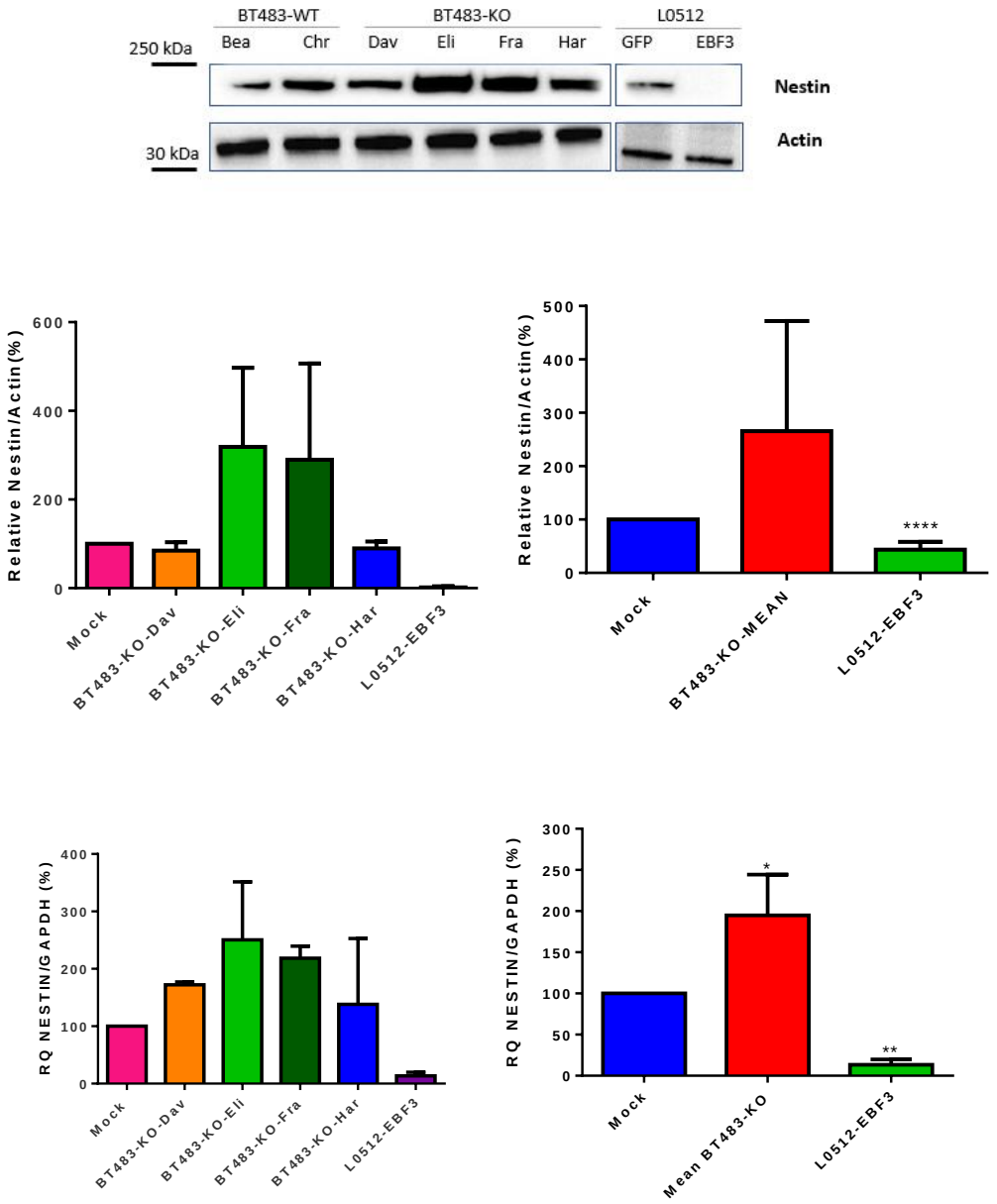


Figure 42. GFAP expression in BT483 and L0512 cell clones at protein and transcriptional level
GFAP expression was analysed by Western blot, as shown in the upper panel: on the left, the picture of one representative experiment. Protein levels were analysed by densitometric quantification of 3 independent experiments, as shown in the plot. Lower plot: *GFAP* mRNA expression was normalized on *GAPDH* housekeeping gene; fold increase of *GFAP* mRNA in transduced clones was obtained with respect to their matched mock-transduced samples. ** $p<0.01$; **** $p<0.0001$, Student's t test.

Nestin is generally a recognized marker of undifferentiated CNS cells; it is more expressed in the GBM component of GBM-PNC, with a much lower expression in the PNC. We thus assessed Nestin expression in transduced GSC cell lines. In line with our previous observation, Nestin was expressed at higher level in EBF3-negative cell lines (BT483-KO and L0512-GFP), and at lower level in EBF3 expressing cell lines, i.e. BT483-WT and L0512-EBF3 (Figure 43).



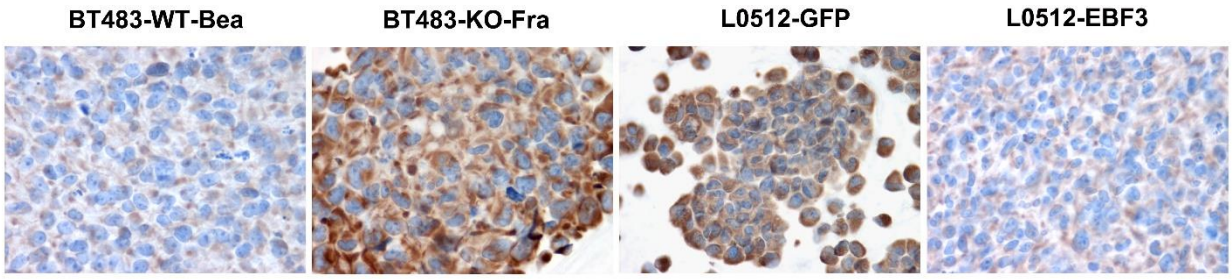


Figure 43. Nestin expression in BT483 and L0512 cell clones at protein and transcriptional level
Nestin expression was analysed by Western blot. Upper panel: on the left, the picture of a representative experiment. Protein levels were analysed by densitometric quantification of 3 independent experiments, as shown in the plot. Lower plot: *Nestin* mRNA expression was normalized on *GAPDH* housekeeping gene; fold increase of mRNA in transduced clones was obtained with respect to their matched mock-transduced samples. * p<0.05; ** p<0.01; **** p<0.0001, Student's t test. Lower image: representative pictures of Nestin IHC on positive (BT483-KO-Fra and L0512-GFP) and negative (BT483-WT-Bea and L0512-EBF3) cell clones.

Whe then induced to terminally differentiate GSC cell lines through removal of mitogens and exposure to fetal bovine serum, as previously described (Galli, Binda et al. 2004)

All of the GSC cell lines acquired the appearance of cells undergoing differentiation, as shown by bright-field microscopy. On the contrary, L0512-EBF3 showed a more immature phenotype, consistent in poorly ramified cells with polygonal appearance (Figure 44).

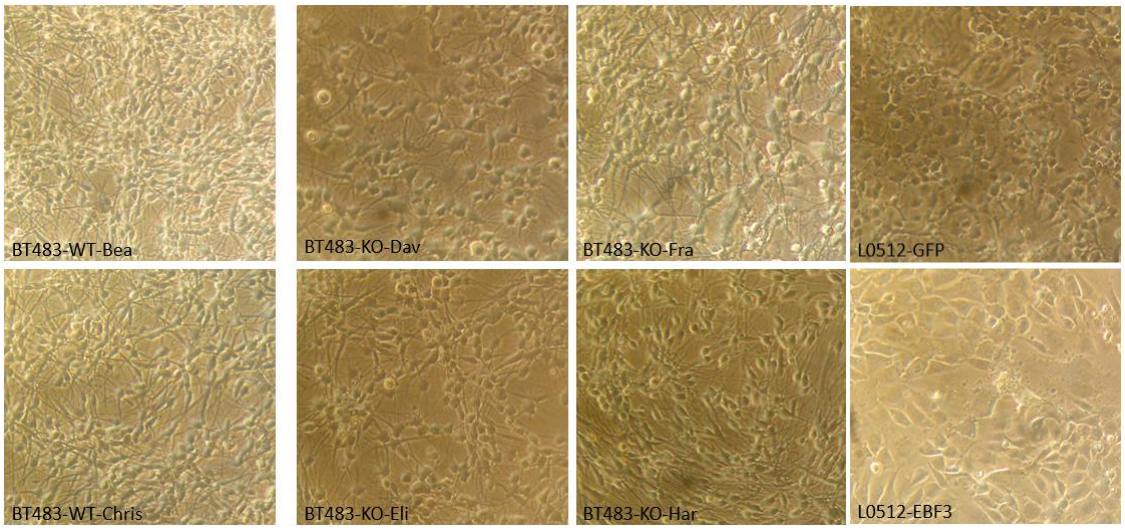


Figure 44. Morphology of BT483 and L0512 cell clones under differentiative conditions
Pictures in bright-field microscopy of the GSC transduced cell lines used in this work. BT483 clones (left and central images) show a mature phenotype. L0512-EBF3 clone shows a more immature phenotype, as compared to the mock control (right panel) (400X).

SOX2 is a biomarker indicating an immature profile and it is usually expressed by GSCs. As expected, mRNA levels of SOX2 in each of the clones cultured under differentiative conditions resulted in a net decrease as compared to their stem cell counterpart (Figure 45), indicating an efficient differentiation process.

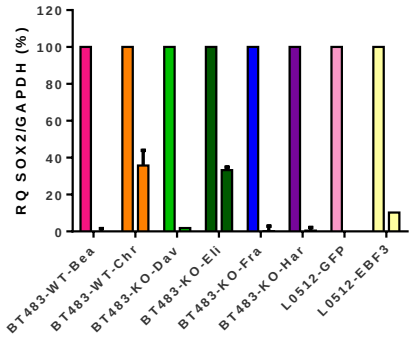


Figure 45. mRNA SOX2 expression in BT483 and L0512 cell clones
SOX2 mRNA expression was normalized on *GAPDH* housekeeping gene; fold increase of *SOX2* mRNA in each differentiated cell line was obtained with respect to their matched GSCs.

It usually accepted that stem cells should be capable to generate cells belonging to the three lineages of the CNS, namely neurons (expressing βIII-tubulin), astrocytes (expressing GFAP) and oligodendrocytes (expressing Gal-C), demonstrating full multipotency. Since the two components of GBM-PNC tumours have a dual glial and primitive neuronal phenotype, in order to evaluate the *EBF3* possible effect on the differentiation process, we only assessed the expression of βIII-tubulin and GFAP in differentiated transduced clones and compared the levels of expression to their stem cell counterpart. As shown in Figure 46, levels of *GFAP*, marker of glial lineage, reaches overall higher levels in all the clones, as compared to *TUBB3*, the gene coding for βIII-tubulin protein. Of note, BT483 wildtype clones, constitutively expressing EBF3, and L0512-EBF3, whose expression is ectopically induced by viral transduction, showed lower fold increase of both GFAP and TUBB3, as compared to their matched GSC line. This could indicate that EBF3 expression, both constitutive or induced, could lead to an impaired ability of differentiation.

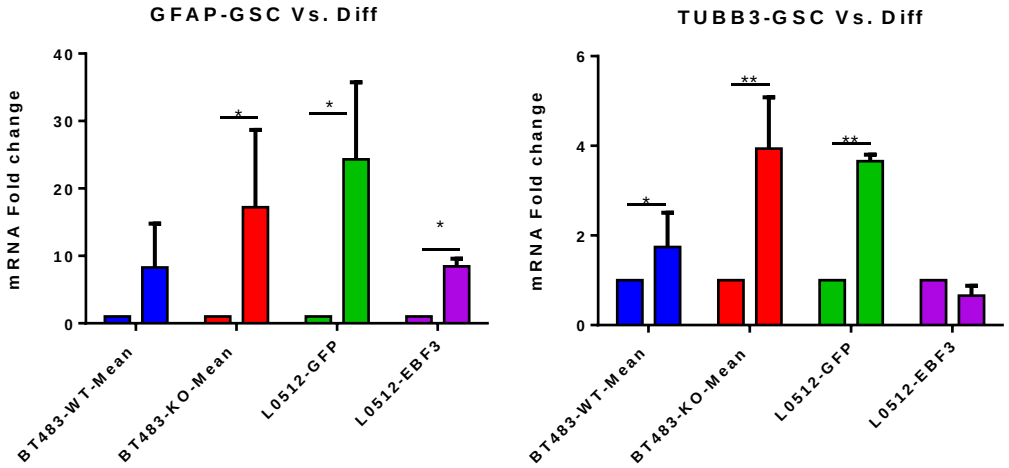


Figure 46. Expression of differentiation markers in GSC Vs. differentiated cell lines
GFAP (left plot) and *TUBB3* (right plot) mRNA expression was normalized on GAPDH housekeeping gene; fold increase of mRNA in differentiated cell lines was obtained with respect to their matched GSCs. * $p<0.05$; ** $p<0.01$, Student's t test.

In vivo transplantation of BT483 transduced clones

Mice were transplanted with transduced BT483 paired clones and divided in three experimental groups (4 animals for each group):

1. "Wildtype": BT483-WT-Bea and BT483-WT-Chr;
2. "HE": BT483-KO-Har and BT483-KO-Eli
3. "FD": BT483-KO-Fra and BT483-KO-Dave

We planned to sacrifice the animals at two different time-points. The first, after 105 days post-transplantation, represents an usual average time for early tumour development; the second will be at first signs of neurological distress. We thus euthanized two mice for each group at 105 days after the transplant; at this time point animals did not show any neurological symptoms. The remaining animals for each group are still alive and will be monitored daily and will be sacrificed later for the second time-point at first signs of neurological distress.

As shown in Figure 47, tumours at the first time-point were still at an early stage of development. Lesions did not develop as a solid mass, but mostly as small clusters of neoplastic cells with some cells infiltrating the surrounding parenchyma. Albeit these features are related to an early tumour development, immunophenotypical analysis allowed to enlight differences between the experimental groups. As shown in Figure 47, the wildtype transplanted mice showed clusters of neoplastic cells exhibiting a biphasic immunophenotype, some of them being positive for EBF3 and others for GFAP, as highlighted by both single and double immunostainings. Of note, the knockout transplanted animals showed tumour cells more homogeneously distributed with part of the cells being GFAP positive and others GFAP negative, but with no EBF3 expression. Interestingly, in line with the *in vitro* results, Nestin was intensely and diffusely expressed in KO-transplanted tumours, while in WT-transplanted mice the expression was less intense and focal. We did not further investigate the phenotype of the xenografts due to the poorly representative material and the small size of tumours. The complete characterization of the tumours will be done on the second time-point.

L0512 cell lines were not available at the time of the transplant and will be transplanted in future experiments.

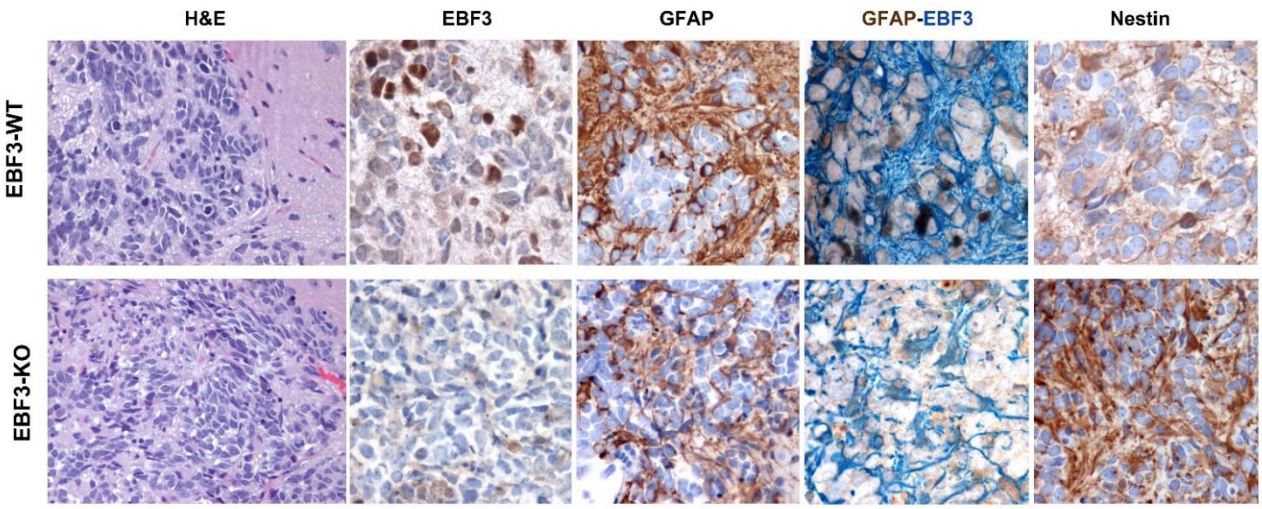


Figure 47. Immunophenotypical analysis of BT483-derived xenografts
Upper images show the wildtype tumours, lower images show knockout tumours. Immunostainings are performed as indicated. Panels are from 40X (H&E) and 60X (other panels) original magnification.

DISCUSSION

“Glioblastoma with Primitive Neuronal Component” is a rare histological variant of Glioblastoma, consistent of a mixed glial component with nodules of immature cells that display an early neuronal differentiation. Considering their rarity, we collected a significant cohort of patients (n=24) in order to deeply characterize the immunophenotypical and molecular features of these tumours. In contrast with other Glioblastomas, patients with GBM-PNC have a lower median age at diagnosis (59.5 y/o) and a peculiar topographical distribution, with the temporal and frontal lobes being the most affected brain regions (75% of the cases). Indeed, the large majority of Glioblastomas with a proneural/classical transcriptional signature and/or with *IDH1* mutation share the same topographical distribution (Sturm, Witt et al. 2012). Our cohort is actually enriched in *IDH1* mutant tumours (16.7%). Moreover, the most represented transcriptional subtypes, according to our recently developed machine learning algorithm (Orzan, Pagani et al. 2020), are the Classical and the Proneural subtypes, globally reaching the 66.7%. This peculiar distribution may be related to a tumour origin from a common neurogenetic brain region. Immunohistochemical analysis confirms the biphasic phenotype of these tumours. The glial component consistently shows positivity for the most common glial associated markers, such as GFAP, YAP1, CD44, Vimentin and EGFR, while the PNC component shows positivity for neuronal/embryonal markers, such as Synaptophysin and NeuN. These immunophenotypes are mutually exclusive, suggesting a divergent differentiation of the two components. Noteworthy, c-Myc and n-Myc are basically expressed only in the PNC component. In fact, it has been reported that expression of these markers is a typical feature of the primary CNS-PNETs with which the PNC components share a similar immunophenotype. Interestingly, both components show positivity for early stem cell markers, such as SOX1 and SOX2, and for markers linked to neurogenesis, such as DCX and βIIITubulin. This observation points out to a prevalent immature phenotype of both components, albeit specifically prone to a glial or neuronal differentiation, respectively. It has to be underlined that Nestin, a widely recognized neural stem cell marker, is preferentially, and almost exclusively, expressed within the glial component. Indeed, our data show that Nestin is expressed at a significantly higher level in the *EBF3*-knockout cell clones and is downregulated in cells forced to express EBF3. The same trend was found for GFAP. As previously reported, EBF3 plays a key role in neuronal cell fate differentiation. Indeed, EBF3 is selectively expressed within the PNC component. All together, these data suggest that overexpression of EBF3 may prevent the expression of glial markers. As a matter of fact, when GSC clones are induced to terminally differentiate, EBF3-expressing clones, both constitutively and whose expression is ectopically induced, showed lower expression fold increase of both *GFAP* and

TUBB3, as compared to their matched EBF3-negative counterpart. Altogether, data suggest that sustained expression of EBF3 impairs the differentiation ability of GSCs and correlates with an immature phenotype. On the other hand, when EBF3 is not constitutively expressed or is inactivated by knockout generation, GSCs upregulate GFAP and Nestin, along with an improved ability to differentiate along the glial lineage. This evidence indicates that Nestin, albeit considered a neural stem cell marker, in a neoplastic setting may be related to a maturation process otherwise blocked by sustained EBF3 expression. Accordingly, constitutively EBF3-expressing GBM-PNC GSCs (BT483-WT) show a higher kinetic of expansion and an improved self-renewal ability, as compared to the *EBF3*-KO clones. Surprisingly, the EBF3 negative GSCs (L0512-GFP) do not show any differences in cell proliferation and self-renewal ability as compared to their EBF3-overexpressing clone (L0512-EBF3). This different behavior may be due to the fact that GBM-PNC derived cells have an overall different genetic background in which sustained EBF3 expression contributes in maintaining a more immature phenotype.

In order to elucidate the possible distinctive traits of GBM-PNC, we investigated the major GBM molecular alterations within the two different components. As previously shown, transcriptional subclassification indicates that the glial component is enriched of proneural and classical subtypes. Of note, the majority of the PNC components could not be classified (66.7%), supporting the evidence that this component is mainly constituted by cells with an undifferentiated phenotype. Notably, the presence or absence of molecular alterations, such as IDH1-R123H mutation, loss of ATRX and overexpression of p53, are common features in both components. As a matter of fact, methylation-based classification classified the two components from the same tumour in the same group: one as IDH-mutant GBM, the others as RTK I/II GBMs. Remarkably, RTK I/II and IDH-mutant tumours correspond to the CL and PN transcriptional subgroups, in line with our previous observation. Indeed, the CL and PN GBM subgroups are tightly related, being the major difference the expression of EGFR that characterizes the CL subgroup. We have previously shown that EGFR expression is crucial for gliomagenesis and that protein overexpression strictly correlates with gene amplification (Mazzoleni, Politi et al. 2010). Accordingly, in EGFR positive cases, protein expression was found selectively in the GBM components and FISH analysis revealed gene amplification in the overexpressing samples. Nevertheless, it is not clear whether the gene amplification occurs selectively only in one of the two components, due to the difficulty to clearly distinguish the two components during FISH analysis in fluorescence microscopy. Moreover, the CNV analyses revealed the same amount of genetic material in both components. This issue is worth to be deeply clarified,

however present data suggest that EGFR expression may be modulated by a specific transcriptional mechanism, independently from the present or absent of gene amplification. Altogether, these observations suggest that the two components may arise from a common progenitor cell with subsequent divergent differentiation, with one component assuming a glial phenotype (GBM), the other with an immature morphological aspect (PNC). In addition, CNV analysis showed for the GBM component the typical GBM profile with gain on chromosome 7 and loss on chromosome 10. Interestingly, the PNC component was endowed with a greater genetic instability, showing a higher degree of loss/gain of genetic material on different chromosomes, according to its immature phenotype. Data are in line with the observation that the PNC component does not activate TERT and/or ALT as telomere maintenance mechanism, a feature that may contribute to tumour instability.

In order to test whether GBM and PNC components harbour distinctive genetic alterations, we further investigated the molecular profile of GBM-PNC with a custom NGS gene panel designed to explore 75 genes that are relevant in GBM and PNC tumours. We performed the analysis on the separated components from six representative GBM-PNC, including the four pairs used for methylation analysis. Data show that all the samples harbour high frequency mutations shared between the two components on frequently GBM altered genes, such as *TP53*, *PTEN*, *IDH1*, *TERT*, *PDGFRA*, *MDM2/4*. Interestingly, all samples harbour private molecular alterations specific for each component, suggesting that GBM-PNC tumours may arise from a common ancestor characterized by a set of mutations/amplifications, while other genomic lesions are acquired after the emergence of either GBM and PNC components. As previously described, in GBMs a hypermutator phenotype is usually associated with TMZ-induced mutagenesis due to inactivation of the DNA mismatch repair pathway (*MSH6*, *MSH2*, *MHS4*, *MSH5*, *PMS1*, *PMS2*, *MLH1*, and *MLH2*). Notably, a rare subset of pre-treated adult GBMs with a *de novo* hypermutator phenotype was identified by Sa et al. (Sa, Choi et al. 2019). Interestingly, two out of six tumours analysed by NGS show a high number of subclonal mutations, such that we can hypothesize a hypermutator phenotype. In line with what reported by Sa and colleagues, these two tumours lack somatic mutations of *IDH1* and *MGMT* promoter methylation and harbour different mutations in DNA repair encoding genes. This high mutational rate was found in both components with a significantly higher load in the PNC component, in line with the previous observation that the PNC component is genetically unstable. We noticed that in 5 samples out of 6 a copy of the *EBF3* gene is lost. Specifically, 3 samples show allelic loss in both components, 1 sample only in the GBM and 1 in the PNC component. It is well known that the *EBF3*

gene is frequently deleted or methylated in glioblastoma; however, albeit with an allelic loss, EBF3 is expressed in the PNC component, suggesting a specific transcriptional mechanism enabling EBF3 sustained expression. As previously discussed, the expression of EBF3 is strictly related to the PNC component and contributes to the maintenance of the immature phenotype.

Preliminary data from *in vivo* xenografts obtained with EBF3 expressing GBM-PNC derived GSCs show the growth of a biphasic tumour resembling the GBM-PNC, with neoplastic aggregates respectively positive or negative for GFAP and EBF3. Data suggest that GBM-PNC GSCs are endowed with the capacity to modulate EBF3 expression. On the contrary, xenografts from the corresponding EBF3-KO clones show an EBF3 negative tumour with a prominent glial aspect and a more widely distributed GFAP positivity. We can assume that GBM-PNC tumours, differently from the "conventional" GBMs, may have a different genetic background that regulates the differential EBF3 expression, allowing for the divergent emergence of the two components. A possible genetic hallmark of the GBM-PNC tumours may be represented by mutations in the *RB1* gene. In fact, *RB1* gene is mutated or deleted in 5 tumours out of the 6 analysed tumours (83.3%), while in the GBM cohorts reported in literature only 12.8% of GBMs display *RB1* mutations (Jonsson, Lin et al. 2019). Interestingly, the only *RB1* wildtype tumour harbours a *CDK4* gene amplification that, from a functional point of view, may mimic the *RB1* loss. In addition, it is reported that up to 15% of children harbouring germline *RB1* mutations develop CNS-PNETs, commonly in the pineal gland (Saab, Rodriguez-Galindo et al. 2009, Miller, Rogers et al. 2011), tumours that display similar features of the PNC component. Taken together these observations suggest that the RB1 pathway alterations may play a role in the origin of GBM-PNC. Interestingly, it is reported that in an *in vitro* model EBF3 directly mediate the transcription, through the interaction with specific binding sites in the promoter, of Cip/Kip p21 cyclin-dependent kinase inhibitor (CDKI), inducing cell cycle arrest through the inhibition of the RB1 pathway (Zhao, Niu et al. 2006). It has to be clarified how *RB1* mutations and regulation of EBF3 expression may interact with each other within the GBM-PNC context.

Altogether, data indicate that GBM-PNCs take their origin from common immature progenitor cells from which the two components diverge. These cells are prone to an intrinsic differentiation ability, as happens for neuroectodermic progenitors in normal brain development, and may undergo a subsequent differentiation by the accumulation of additional molecular alterations. However, both components, albeit committed toward a specific cell lineage, remain immature due to a block of the maturation processes. In this contest, both the genetic background and the modulation of EBF3 expression may play a crucial role.

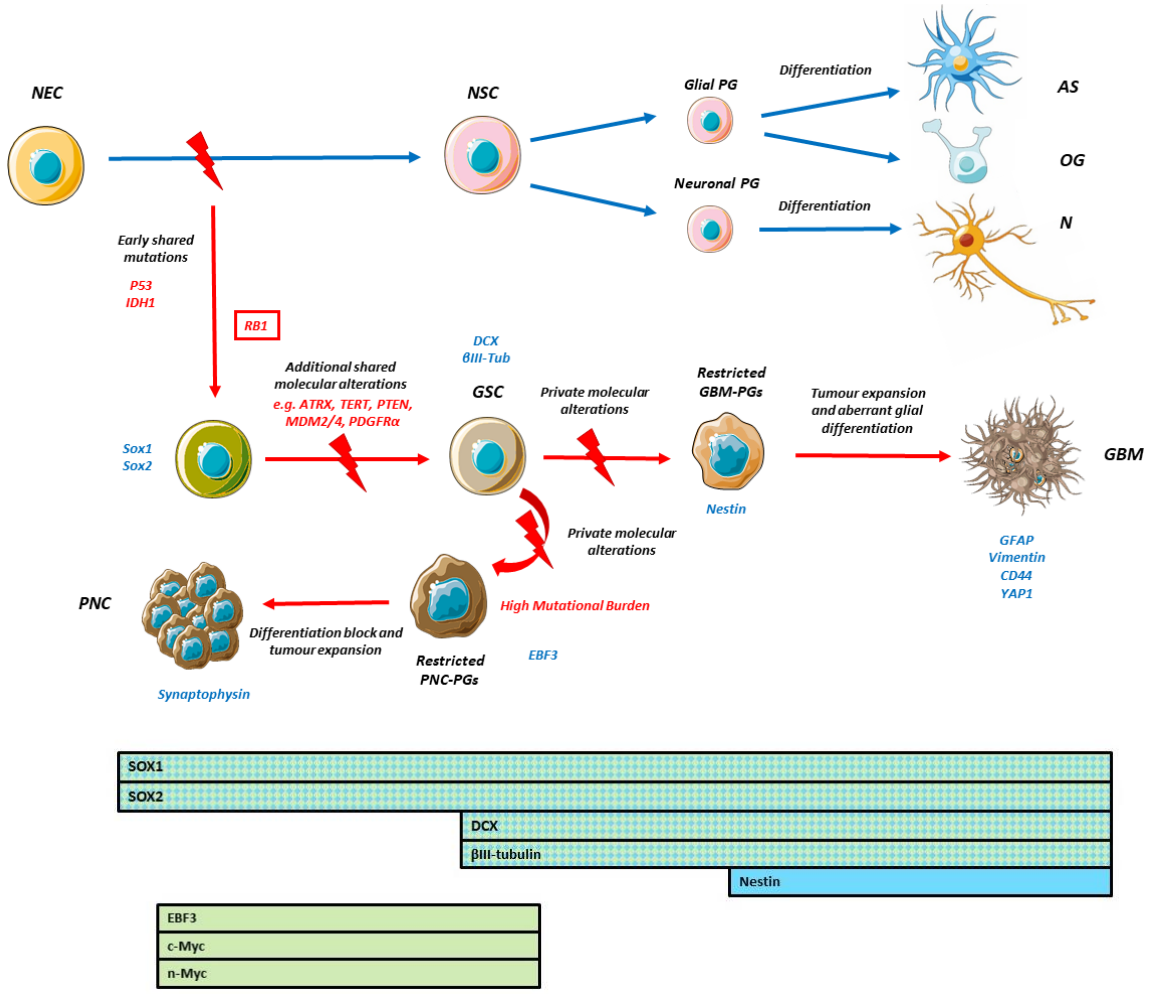


Figure 48. Schematic representation of the proposed model of GBM-PNC origin
The image summarizes our proposed model of stem cell hierarchy and division of cancer stem cells giving origin to the GBM-PNC, as compared to normal CNS cell origin. We propose that early driver mutations of the original Sox1 and Sox2 expressing neuroepithelial stem cells are at the origin of GBM-PNC. Additional shared molecular alteration will give rise to a putative common GSC from which the tumour will diverge into two different restricted progenitors (PGs) by the accumulation of private molecular alterations. The expansion of these restricted progenitors will lead to the final phenotype of the GBM-PNC. Molecular alterations are depicted in red, biomarkers are depicted in blue. Strips at the bottom show the time window of expression of the indicated markers in the GBM-PNC model. The figure was adapted using a template on the Servier medical art website (available online: www.servier.com) licensed under the creative commons attribution 3.0 unported license (available online:http://creativecommons.org/license/by/3.0/) and on "Shutterstock, Designua" (https://www.shutterstock.com/search/designua). Editing: Vanessa Delbarba-Servizi per la comunicazione.

The study shed a light on this rare GBM variant. In addition, we are confident that these biphasic tumours represent a model of tumour development and the comprehension of the underlying mechanism driving divergent differentiation could help understanding the origin of malignant brain tumours, as glioblastomas and primitive neuronal malignant tumours. In addition, the discovery of novel biomarkers and/or genetic alterations specifically involved in the generation of the different GBM-PNC components, reflecting the biology of both glial and embryonal brain tumours, may help to individuate novel possible druggable molecules for the treatment of malignant brain tumours.

LIST OF ABBREVIATIONS

(s)gRNA	(single chimeric) guide RNA
A	Amplified
ADC	Apparent diffusion coefficient
ALT	Alternative lengthening of telomeres
Cas	CRISPR-associated
CBS	Circular binary segmentation
CDKI	Cyclin-dependent kinase inhibitor
CEP	Chromosome Enumeration probe
Chr	Chromosome
CL	Classical
CMV	Human cytomegalovirus
CNS	Central Nervous System
CNV	Copy Number Variation
COSMIC	Catalogue of Somatic Mutations in Cancer
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CSC	Cancer Stem Cells
CT	Computed tomography
DBD	DNA binding domain
DSB	Double-strand breaks
FFPE	Formalin-Fixed Paraffin-Embedded samples
FISH	Fluorescence In Situ Hybridization (FISH)
GBM	Glioblastoma
GBM-PNC	Glioblastoma with Primitive Neuronal Component
GFP	Green Fluorescent Protein
GSC	Glioblastoma Stem Cells
H&E	Hematoxylin and eosin
HDR	Homology-directed repair
HET	Heterogeneous
HLH	Helix-loop-helix
IHC	Immunohistochemistry
IPT/TIG	Immunoglobulin-like fold, Plexins, Transcription factors
KO	Knockout
KPS	Karnofsky Performance Score
LOD	Limits of detection
LTR	Long Terminal Repeats
LV	Lentivirus

MES	Mesenchymal
MRI	Magnetic Resonance Imaging
NA	Not amplified
NEC	Neuroectoderm
NGS	Next Generation Sequencing
NHEJ	Non-Homologous End Joining
NOS	Not otherwise specified
OS	Overall survival
PAM	Protospacer-associated motif
PBMCs	Peripheral blood mononuclear cells
PG	Progenitors
PML	ALT-associated promyelocytic leukemia nuclear bodies
PN	Proneural
PNET	Primitive Neuroectodermal Tumours
POL	Polysomic
RPA	Recursive partitioning analysis
RSV	Rous sarcoma virus
RTK	Receptor Tyrosine Kinase
SD	Standard deviation
short crRNAs	Repeat/spacer-derived
SIN	Self-inactivating
TCGA	The Cancer Genome Atlas Research Network
TMM	Telomere maintenance mechanism
TMZ	Temozolomide
tracrRNA	Transactivating RNA
TU	Transducing Units
VAF	Variant allele frequency
VSV-G	Vesicular Stomatitis virus envelope glycoprotein G
WHO	World Health Organization
WT	Wildtype

…….

REFERENCES

- Akkina, R. K., R. M. Walton, M. L. Chen, Q. X. Li, V. Planelles and I. S. Chen (1996). "High-efficiency gene transfer into CD34+ cells with a human immunodeficiency virus type 1-based retroviral vector pseudotyped with vesicular stomatitis virus envelope glycoprotein G." J Virol **70**(4): 2581-2585.
- Ali, S., N. M. Joseph, A. Perry, R. F. Barajas, Jr. and S. Cha (2014). "Apparent diffusion coefficient in glioblastoma with PNET-like components, a GBM variant." J Neurooncol **119**(2): 353-360.
- Baretti, M. and D. T. Le (2018). "DNA mismatch repair in cancer." Pharmacol Ther **189**: 45-62.
- Beier, D., P. Hau, M. Proescholdt, A. Lohmeier, J. Wischhusen, P. J. Oefner, L. Aigner, A. Brawanski, U. Bogdahn and C. P. Beier (2007). "CD133(+) and CD133(-) glioblastoma-derived cancer stem cells show differential growth characteristics and molecular profiles." Cancer Res **67**(9): 4010-4015.
- Brat, D. J., K. Aldape, H. Colman, D. Figrarella-Branger, G. N. Fuller, C. Giannini, E. C. Holland, R. B. Jenkins, B. Kleinschmidt-DeMasters, T. Komori, J. M. Kros, D. N. Louis, C. McLean, A. Perry, G. Reifenberger, C. Sarkar, R. Stupp, M. J. van den Bent, A. von Deimling and M. Weller (2020). "cIMPACT-NOW update 5: recommended grading criteria and terminologies for IDH-mutant astrocytomas." Acta Neuropathol **139**(3): 603-608.
- Brennan, C. W., R. G. Verhaak, A. McKenna, B. Campos, H. Noushmehr, S. R. Salama, S. Zheng, D. Chakravarty, J. Z. Sanborn, S. H. Berman, R. Beroukhim, B. Bernard, C. J. Wu, G. Genovese, I. Shmulevich, J. Barnholtz-Sloan, L. Zou, R. Vegesna, S. A. Shukla, G. Ciriello, W. K. Yung, W. Zhang, C. Sougnez, T. Mikkelsen, K. Aldape, D. D. Bigner, E. G. Van Meir, M. Prados, A. Sloan, K. L. Black, J. Eschbacher, G. Finocchiaro, W. Friedman, D. W. Andrews, A. Guha, M. Iacocca, B. P. O'Neill, G. Foltz, J. Myers, D. J. Weisenberger, R. Penny, R. Kucherlapati, C. M. Perou, D. N. Hayes, R. Gibbs, M. Marra, G. B. Mills, E. Lander, P. Spellman, R. Wilson, C. Sander, J. Weinstein, M. Meyerson, S. Gabriel, P. W. Laird, D. Haussler, G. Getz and L. Chin (2013). "The somatic genomic landscape of glioblastoma." Cell **155**(2): 462-477.
- Campbell, B. B., N. Light, D. Fabrizio, M. Zatzman, F. Fuligni, R. de Borja, S. Davidson, M. Edwards, J. A. Elvin, K. P. Hodel, W. J. Zahurancik, Z. Suo, T. Lipman, K. Wimmer, C. P. Kratz, D. C. Bowers, T. W. Laetsch, G. P. Dunn, T. M. Johanns, M. R. Grimmer, I. V. Smirnov, V. Larouche, D. Samuel, A. Bronsema, M. Osborn, D. Stearns, P. Raman, K. A. Cole, P. B. Storm, M. Yalon, E. Opocher, G. Mason, G. A. Thomas, M. Sabel, B. George, D. S. Ziegler, S. Lindhorst, V. M. Issai, S. Constantini, H. Toledano, R. Elhasid, R. Farah, R. Dvir, P. Dirks, A. Huang, M. A. Galati, J. Chung, V. Ramaswamy, M. S. Irwin, M. Aronson, C. Durno, M. D. Taylor, G. Rechavi, J. M. Maris, E. Bouffet, C. Hawkins, J. F. Costello, M. S. Meyn, Z. F. Pursell, D. Malkin, U. Tabori and A. Shlien (2017). "Comprehensive Analysis of Hypermutation in Human Cancer." Cell **171**(5): 1042-1056.e1010.
- Capper, D., D. T. W. Jones, M. Sill, V. Hovestadt, D. Schrimpf, D. Sturm, C. Koelsche, F. Sahm, L. Chavez, D. E. Reuss, A. Kratz, A. K. Wefers, K. Huang, K. W. Pajtler, L. Schweizer, D. Stichel, A. Olar, N. W. Engel, K. Lindenberg, P. N. Harter, A. K. Braczynski, K. H. Plate, H. Dohmen, B. K. Garvalov, R. Coras, A. Hölsken, E. Hewer, M. Bewerunge-Hudler, M. Schick, R. Fischer, R. Beschorner, J. Schittenhelm, O. Staszewski, K. Wani, P. Varlet, M. Pages, P. Temming, D. Lohmann, F. Selt, H. Witt, T. Milde, O. Witt, E. Aronica, F. Giangaspero, E. Rushing, W. Scheurlen, C. Geisenberger, F. J. Rodriguez, A. Becker, M. Preusser, C. Haberler, R. Bjerkvig, J. Cryan, M. Farrell, M. Deckert, J. Hench, S. Frank, J. Serrano, K. Kannan, A. Tsirigos, W. Brück, S. Hofer, S. Brehmer, M. Seiz-Rosenhagen, D. Hänggi, V. Hans, S. Rozsnoki, J. R. Hansford, P. Kohlhof, B. W. Kristensen, M. Lechner, B. Lopes, C. Mawrin, R. Ketter, A. Kulozik, Z. Khatib, F. Heppner, A. Koch, A. Jouvet, C. Keohane, H. Mühleisen, W. Mueller, U. Pohl, M. Prinz, A. Benner, M. Zapatka, N. G. Gottardo, P. H. Driever, C. M. Kramm, H. L. Müller, S. Rutkowski, K. von Hoff, M. C. Frühwald, A. Gnekow, G. Fleischhack, S. Tippelt, G. Calaminus, C. M. Monoranu, A. Perry, C. Jones, T. S. Jacques, B. Radlwimmer, M. Gessi, T. Pietsch, J. Schramm, G. Schackert, M. Westphal, G. Reifenberger, P. Wesseling, M. Weller, V. P. Collins, I. Blümcke, M. Bendszus, J. Debus, A. Huang, N. Jabado, P. A. Northcott, W. Paulus, A. Gajjar, G. W. Robinson, M. D. Taylor, Z. Jaunmuktane, M. Ryzhova, M. Platten, A. Unterberg, W. Wick, M. A. Karajannis, M. Mittelbronn, T. Acker, C. Hartmann, K. Aldape, U. Schüller, R. Buslei, P. Lichter, M. Kool, C. Herold-Mende, D. W. Ellison, M. Hasselblatt, M. Snuderl, S. Brandner, A. Korshunov, A. von Deimling and S. M. Pfister (2018). "DNA methylation-based classification of central nervous system tumours." Nature **555**(7697): 469-474.
- Capper, D., D. Stichel, F. Sahm, D. T. W. Jones, D. Schrimpf, M. Sill, S. Schmid, V. Hovestadt, D. E. Reuss, C. Koelsche, A. Reinhardt, A. K. Wefers, K. Huang, P. Sievers, A. Ebrahimi, A. Schöler, D. Teichmann, A. Koch, D. Hänggi, A. Unterberg, M. Platten, W. Wick, O. Witt, T. Milde, A. Korshunov, S. M. Pfister and A. von Deimling (2018). "Practical implementation of DNA methylation and copy-number-based CNS tumour diagnostics: the Heidelberg experience." Acta Neuropathol **136**(2): 181-210.
- Ceccarelli, M., F. P. Barthel, T. M. Malta, T. S. Sabedot, S. R. Salama, B. A. Murray, O. Morozova, Y. Newton, A. Radenbaugh, S. M. Pagnotta, S. Anjum, J. Wang, G. Manyam, P. Zoppoli, S. Ling, A. A. Rao, M. Grifford, A. D. Cherniack, H. Zhang, L. Poisson, C. G. Carlotti, D. P. Tirapelli, A. Rao, T. Mikkelsen, C. C. Lau, W. K. Yung, R. Rabadan,

J. Huse, D. J. Brat, N. L. Lehman, J. S. Barnholtz-Sloan, S. Zheng, K. Hess, G. Rao, M. Meyerson, R. Beroukhim, L. Cooper, R. Akbani, M. Wrensch, D. Haussler, K. D. Aldape, P. W. Laird, D. H. Gutmann, H. Noushmehr, A. Iavarone, R. G. Verhaak and T. R. Network (2016). "Molecular Profiling Reveals Biologically Discrete Subsets and Pathways of Progression in Diffuse Glioma." Cell **164**(3): 550-563.

- Chatterjee, A., P. A. Stockwell, A. Ahn, E. J. Rodger, A. L. Leichter and M. R. Eccles (2017). "Genome-wide methylation sequencing of paired primary and metastatic cell lines identifies common DNA methylation changes and a role for EBF3 as a candidate epigenetic driver of melanoma metastasis." Oncotarget **8**(4): 6085-6101.
- Chen, J., Y. Li, T. S. Yu, R. M. McKay, D. K. Burns, S. G. Kernie and L. F. Parada (2012). "A restricted cell population propagates glioblastoma growth after chemotherapy." Nature **488**(7412): 522-526.
- Chen, R., M. C. Nishimura, S. M. Bumbaca, S. Kharbanda, W. F. Forrest, I. M. Kasman, J. M. Greve, R. H. Soriano, L. L. Gilmour, C. S. Rivers, Z. Modrusan, S. Nacu, S. Guerrero, K. A. Edgar, J. J. Wallin, K. Lamszus, M. Westphal, S. Heim, C. D. James, S. R. VandenBerg, J. F. Costello, S. Moorefield, C. J. Cowdrey, M. Prados and H. S. Phillips (2010). "A hierarchy of self-renewing tumor-initiating cell types in glioblastoma." Cancer Cell **17**(4): 362-375.
- Chen, X., S. T. Cheung, S. So, S. T. Fan, C. Barry, J. Higgins, K. M. Lai, J. Ji, S. Dudoit, I. O. Ng, M. Van De Rijn, D. Botstein and P. O. Brown (2002). "Gene expression patterns in human liver cancers." Mol Biol Cell **13**(6): 1929-1939.
- Cominelli, M., S. Grisanti, S. Mazzoleni, C. Branca, L. Buttolo, D. Furlan, B. Liserre, M. F. Bonetti, D. Medicina, V. Pellegrini, M. Buglione, R. Liserre, S. Pellegatta, G. Finocchiaro, P. Dalerba, F. Facchetti, M. Pizzi, R. Galli and P. L. Poliani (2015). "EGFR amplified and overexpressing glioblastomas and association with better response to adjuvant metronomic temozolomide." J Natl Cancer Inst **107**(5).
- Corno, D., M. Pala, M. Cominelli, B. Cipelletti, K. Leto, L. Croci, V. Barili, F. Brandalise, R. Melzi, A. Di Gregorio, L. S. Sergi, L. S. Politi, L. Piemonti, A. Bulfone, P. Rossi, F. Rossi, G. G. Consalez, P. L. Poliani and R. Galli (2012). "Gene signatures associated with mouse postnatal hindbrain neural stem cells and medulloblastoma cancer stem cells identify novel molecular mediators and predict human medulloblastoma molecular classification." Cancer Discov **2**(6): 554-568.
- Corti, G., A. Bartolini, G. Crisafulli, L. Novara, G. Rospo, M. Montone, C. Negrino, B. Mussolin, M. Buscarino, C. Isella, L. Barault, G. Siravegna, S. Siena, S. Marsoni, F. Di Nicolantonio, E. Medico and A. Bardelli (2019). "A Genomic Analysis Workflow for Colorectal Cancer Precision Oncology." Clin Colorectal Cancer **18**(2): 91-101.e103.
- Crisafulli, G., B. Mussolin, A. Cassingena, M. Montone, A. Bartolini, L. Barault, A. Martinetti, F. Morano, F. Pietrantonio, A. Sartore-Bianchi, S. Siena, F. Di Nicolantonio, S. Marsoni, A. Bardelli and G. Siravegna (2019). "Whole exome sequencing analysis of urine trans-renal tumour DNA in metastatic colorectal cancer patients." ESMO Open **4**(6).
- De Bacco, F., E. Casanova, E. Medico, S. Pellegatta, F. Orzan, R. Albano, P. Luraghi, G. Reato, A. D'Ambrosio, P. Porrati, M. Patane, E. Maderna, B. Pollo, P. M. Comoglio, G. Finocchiaro and C. Boccaccio (2012). "The MET oncogene is a functional marker of a glioblastoma stem cell subtype." Stem Cells **72**(17): 4537-4550.
- De Bacco, F., A. D'Ambrosio, E. Casanova, F. Orzan, R. Neggia, R. Albano, F. Verginelli, M. Cominelli, P. L. Poliani, P. Luraghi, G. Reato, S. Pellegatta, G. Finocchiaro, T. Perera, E. Garibaldi, P. Gabriele, P. M. Comoglio and C. Boccaccio (2016). "MET inhibition overcomes radiation resistance of glioblastoma stem-like cells." EMBO Mol Med **8**(5): 550-568.
- Doetsch, F., I. Caillé, D. A. Lim, J. M. García-Verdugo and A. Alvarez-Buylla (1999). "Subventricular zone astrocytes are neural stem cells in the adult mammalian brain." Cell **97**(6): 703-716.
- Dolecek, T. A., J. M. Propp, N. E. Stroup and C. Kruchko (2012). "CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2005-2009." Neuro Oncol **14 Suppl 5**: v1-49.
- Dubois, L., L. Bally-Cuif, M. Crozatier, J. Moreau, L. Paquereau and A. Vincent (1998). "XCoe2, a transcription factor of the Col/Olf-1/EBF family involved in the specification of primary neurons in Xenopus." Curr Biol **8**(4): 199-209.
- Dulai, M. S., C. M. Bosanko, A. M. Wang, D. S. Horoupian, S. Boodin, P. Y. Chen and J. D. Wilson (2004). "Mixed cystic gliosarcoma and primitive neuroectodermal tumor: a case report." Clin Neuropathol **23**(5): 218-222.
- Ferguson, D. O. and F. W. Alt (2001). "DNA double strand break repair and chromosomal translocation: lessons from animal models." Oncogene **20**(40): 5572-5579.
- Fields, S., K. Ternyak, H. Gao, R. Ostraat, J. Akerlund and J. Hagman (2008). "The 'zinc knuckle' motif of Early B cell Factor is required for transcriptional activation of B cell-specific genes." Mol Immunol **45**(14): 3786-3796.
- Freed, E. O. and M. A. Martin (1994). "HIV-1 infection of non-dividing cells." Nature **369**(6476): 107-108.
- Galli, R., E. Binda, U. Orfanelli, B. Cipelletti, A. Gritti, S. De Vitis, R. Fiocco, C. Foroni, F. Dimeco and A. Vescovi (2004). "Isolation and characterization of tumorigenic, stem-like neural precursors from human glioblastoma." Cancer Res. **64**(19): 7011-7021. doi: 7010.1158/0008-5472.CAN-7004-1364.
- Garcia-Dominguez, M., C. Poquet, S. Garel and P. Charnay (2003). "Ebf gene function is required for coupling neuronal differentiation and cell cycle exit." Development **130**(24): 6013-6025.

- Garel, S., F. Marin, M. G. Mattei, C. Vesque, A. Vincent and P. Charnay (1997). "Family of Ebf/Olf-1-related genes potentially involved in neuronal differentiation and regional specification in the central nervous system." Dev Dyn **210**(3): 191-205.
- Genschel, J., S. J. Littman, J. T. Drummond and P. Modrich "Isolation of MutSβ from human cells and comparison of the mismatch repair specificities of MutSβ and MutSα."
- Genschel, J., S. J. Littman, J. T. Drummond and P. Modrich (1998). "Isolation of MutSβ from human cells and comparison of the mismatch repair specificities of MutSβ and MutSα."
- Gimple, R. C., S. Bhargava, D. Dixit and J. N. Rich (2019). "Glioblastoma stem cells: lessons from the tumor hierarchy in a lethal cancer." Genes Dev **33**(11-12): 591-609.
- Haase, S., M. B. Garcia-Fabiani, S. Carney, D. Altshuler, F. J. Núñez, F. M. Méndez, F. Núñez, P. R. Lowenstein and M. G. Castro (2018). "Mutant ATRX: uncovering a new therapeutic target for glioma." Expert Opin Ther Targets **22**(7): 599-613.
- Hagemann, C., C. Meyer, J. Stojic, S. Eicker, S. Gerngras, S. Kühnel, K. Roosen and G. H. Vince (2006). "High efficiency transfection of glioma cell lines and primary cells for overexpression and RNAi experiments." J Neurosci Methods **156**(1-2): 194-202.
- Hagman, J., M. J. Gutch, H. Lin and R. Grosschedl (1995). "EBF contains a novel zinc coordination motif and multiple dimerization and transcriptional activation domains." Embo j **14**(12): 2907-2916.
- Hagman, J. and K. Lukin (2005). "Early B-cell factor 'pioneers' the way for B-cell development." Trends Immunol **26**(9): 455-461.
- Heaphy, C. M., A. P. Subhawong, S. M. Hong, M. G. Goggins, E. A. Montgomery, E. Gabrielson, G. J. Netto, J. I. Epstein, T. L. Lotan, W. H. Westra, I. M. Shih, C. A. Iacobuzio-Donahue, A. Maitra, Q. K. Li, C. G. Eberhart, J. M. Taube, D. Rakheja, R. J. Kurman, T. C. Wu, R. B. Roden, P. Argani, A. M. De Marzo, L. Terracciano, M. Torbenson and A. K. Meeker (2011). "Prevalence of the alternative lengthening of telomeres telomere maintenance mechanism in human cancer subtypes." Am J Pathol **179**(4): 1608-1615.
- Hegi, M. E., A. C. Diserens, T. Gorlia, M. F. Hamou, N. de Tribolet, M. Weller, J. M. Kros, J. A. Hainfellner, W. Mason, L. Mariani, J. E. Bromberg, P. Hau, R. O. Mirimanoff, J. G. Cairncross, R. C. Janzer and R. Stupp (2005). "MGMT gene silencing and benefit from temozolomide in glioblastoma." N Engl J Med **352**(10): 997-1003.
- Hovestadt, V., M. Remke, M. Kool, T. Pietsch, P. A. Northcott, R. Fischer, F. M. Cavalli, V. Ramaswamy, M. Zapatka, G. Reifenberger, S. Rutkowski, M. Schick, M. Bewerunge-Hudler, A. Korshunov, P. Lichter, M. D. Taylor, S. M. Pfister and D. T. Jones (2013). "Robust molecular subgrouping and copy-number profiling of medulloblastoma from small amounts of archival tumour material using high-density DNA methylation arrays." Acta Neuropathol **125**(6): 913-916.
- Huse, J. T. and H. E.C. (2010). "Targeting brain cancer: advances in the molecular pathology of malignant glioma and medulloblastoma." Nature Reviews Cancer **10**(5): 319-331.
- Huse, J. T., H. S. Phillips and C. W. Brennan (2011). "Molecular subclassification of diffuse gliomas: seeing order in the chaos." Glia **59**(8): 1190-1199.
- Idilli, A. I., F. Pagani, E. Kerschbamer, F. Berardinelli, M. Bernabé, M. L. Cayuela, S. Piazza, P. L. Poliani, E. Cusanelli and M. C. Mione (2020). "Changes in the Expression of Pre-Replicative Complex Genes in hTERT and ALT Pediatric Brain Tumors." Cancers (Basel) **12**(4).
- Ignatova, T. N., V. G. Kukekov, E. D. Laywell, O. N. Suslov, F. D. Vrionis and D. A. Steindler (2002). "Human cortical glial tumors contain neural stem-like cells expressing astroglial and neuronal markers in vitro." Glia **39**(3): 193-206.
- Ishizawa, K., S. Kan-nuki, H. Kumagai, T. Komori and T. Hirose (2002). "Lipomatous primitive neuroectodermal tumor with a glioblastoma component: a case report." Acta Neuropathol **103**(2): 193-198.
- Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna and E. Charpentier (2012). "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." Science **337**(6096): 816-821.
- Johnson, B. E., T. Mazor, C. Hong, M. Barnes, K. Aihara, C. Y. McLean, S. D. Fouse, S. Yamamoto, H. Ueda, K. Tatsuno, S. Asthana, L. E. Jalbert, S. J. Nelson, A. W. Bollen, W. C. Gustafson, E. Charron, W. A. Weiss, I. V. Smirnov, J. S. Song, A. B. Olshen, S. Cha, Y. Zhao, R. A. Moore, A. J. Mungall, S. J. M. Jones, M. Hirst, M. A. Marra, N. Saito, H. Aburatani, A. Mukasa, M. S. Berger, S. M. Chang, B. S. Taylor and J. F. Costello (2014). "Mutational Analysis Reveals the Origin and Therapy-Driven Evolution of Recurrent Glioma."
- Jones, S., X. Zhang, D. W. Parsons, J. C. Lin, R. J. Leary, P. Angenendt, P. Mankoo, H. Carter, H. Kamiyama, A. Jimeno, S. M. Hong, B. Fu, M. T. Lin, E. S. Calhoun, M. Kamiyama, K. Walter, T. Nikolskaya, Y. Nikolsky, J. Hartigan, D. R. Smith, M. Hidalgo, S. D. Leach, A. P. Klein, E. M. Jaffee, M. Goggins, A. Maitra, C. Iacobuzio-Donahue, J. R. Eshleman, S. E. Kern, R. H. Hruban, R. Karchin, N. Papadopoulos, G. Parmigiani, B. Vogelstein, V. E. Velculescu and K. W. Kinzler (2008). "Core signaling pathways in human pancreatic cancers revealed by global genomic analyses." Science **321**(5897): 1801-1806.

- Jonsson, P., A. L. Lin, R. J. Young, N. M. DiStefano, D. M. Hyman, B. T. Li, M. F. Berger, A. Zehir, M. Ladanyi, D. B. Solit, A. G. Arnold, Z. K. Stadler, D. Mandelker, M. E. Goldberg, J. Chmielecki, M. Pourmaleki, S. Q. Ogilvie, S. S. Chavan, A. T. McKeown, M. Manne, A. Hyde, K. Beal, T. J. Yang, C. P. Nolan, E. Pentsova, A. Omuro, I. T. Gavrilovic, T. J. Kaley, E. L. Diamond, J. B. Stone, C. Grommes, A. Boire, M. Daras, A. F. Piotrowski, A. M. Miller, P. H. Gutin, T. A. Chan, V. S. Tabar, C. W. Brennan, M. Rosenblum, L. M. DeAngelis, I. K. Mellinghoff and B. S. Taylor (2019). "Genomic Correlates of Disease Progression and Treatment Response in Prospectively Characterized Gliomas." Clin Cancer Res **25**(18): 5537-5547.
- Joseph, N. M., J. Phillips, S. Dahiya, M. F. M, T. Tihan, D. J. Brat and A. Perry (2013). "Diagnostic implications of IDH1-R132H and OLIG2 expression patterns in rare and challenging glioblastoma variants." Mod Pathol **26**(3): 315-326.
- Kandoth, C., N. Schultz, A. D. Cherniack, R. Akbani, Y. Liu, H. Shen, A. G. Robertson, I. Pashtan, R. Shen, C. C. Benz, C. Yau, P. W. Laird, L. Ding, W. Zhang, G. B. Mills, R. Kucherlapati, E. R. Mardis, D. A. Levine and C. G. A. R. Network (2013). "Integrated genomic characterization of endometrial carcinoma." Nature **497**(7447): 67-73.
- Kaplan, K. J. and A. Perry (2007). "Gliosarcoma with primitive neuroectodermal differentiation: case report and review of the literature." J Neurooncol **83**(3): 313-318.
- Kieslinger, M., S. Folberth, G. Dobreva, T. Dorn, L. Croci, R. Erben, G. G. Consalez and R. Grosschedl (2005). "EBF2 regulates osteoblast-dependent differentiation of osteoclasts." Dev Cell **9**(6): 757-767.
- Knudsen, E. S., R. Nambiar, S. R. Rosario, D. J. Smiraglia, D. W. Goodrich and A. K. Witkiewicz (2020). "Pan-cancer molecular analysis of the RB tumor suppressor pathway."
- Kulis, M. and M. Esteller (2010). "DNA methylation and cancer." Adv Genet **70**: 27-56.
- Liao, D. (2009). "Emerging roles of the EBF family of transcription factors in tumor suppression." Mol Cancer Res **7**(12): 1893-1901.
- Lie, D. C., H. Song, S. A. Colamarino, G. L. Ming and F. H. Gage (2004). "Neurogenesis in the adult brain: new strategies for central nervous system diseases." Annu Rev Pharmacol Toxicol **44**: 399-421.
- Liu, G., X. Yuan, Z. Zeng, P. Tunici, H. Ng, I. R. Abdulkadir, L. Lu, D. Irvin, K. L. Black and J. S. Yu (2006). "Analysis of gene expression and chemoresistance of CD133+ cancer stem cells in glioblastoma." Mol Cancer **5**: 67.
- Louis, D. N., A. Perry, G. Reifenberger, A. von Deimling, D. Figarella-Branger, W. K. Cavenee, H. Ohgaki, O. D. Wiestler, P. Kleihues and D. W. Ellison (2016). "The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary." Acta Neuropathol **131**(6): 803-820.
- Mangerel, J., A. Price, P. Castelo-Branco, J. Brzezinski, P. Buczkowicz, P. Rakopoulos, D. Merino, B. Baskin, J. Wasserman, M. Mistry, M. Barszczyk, D. Picard, S. Mack, M. Remke, H. Starkman, C. Elizabeth, C. Zhang, N. Alon, J. Lees, I. L. Andrulis, J. S. Wunder, N. Jabado, D. L. Johnston, J. T. Rutka, P. B. Dirks, E. Bouffet, M. D. Taylor, A. Huang, D. Malkin, C. Hawkins and U. Tabori (2014). "Alternative lengthening of telomeres is enriched in, and impacts survival of TP53 mutant pediatric malignant brain tumors." Acta Neuropathol **128**(6): 853-862.
- Mazzoleni, S., L. S. Politi, M. Pala, M. Cominelli, A. Franzin, L. Sergi Sergi, A. Falini, M. De Palma, A. Bulfone, P. L. Poliani and R. Galli (2010). "Epidermal growth factor receptor expression identifies functionally and molecularly distinct tumor-initiating cells in human glioblastoma multiforme and is required for gliomagenesis." Cancer Res **70**(19): 7500-7513.
- McDonald, K. L., J. McDonnell, A. Muntoni, J. D. Henson, M. E. Hegi, A. von Deimling, H. R. Wheeler, R. J. Cook, M. T. Biggs, N. S. Little, B. G. Robinson, R. R. Reddel and J. A. Royds (2010). "Presence of alternative lengthening of telomeres mechanism in patients with glioblastoma identifies a less aggressive tumor type with longer survival." J Neuropathol Exp Neurol **69**(7): 729-736.
- Medina, K. L., J. M. Pongubala, K. L. Reddy, D. W. Lancki, R. Dekoter, M. Kieslinger, R. Grosschedl and H. Singh (2004). "Assembling a gene regulatory network for specification of the B cell fate." Dev Cell **7**(4): 607-617.
- Miller, S., H. A. Rogers, P. Lyon, V. Rand, M. Adamowicz-Brice, S. C. Clifford, J. T. Hayden, S. Dyer, S. Pfister, A. Korshunov, M. A. Brundler, J. Lowe, B. Coyle and R. G. Grundy (2011). "Genome-wide molecular characterization of central nervous system primitive neuroectodermal tumor and pineoblastoma." Neuro Oncol **13**(8): 866-879.
- Mullighan, C. G., S. Goorha, I. Radtke, C. B. Miller, E. Coustan-Smith, J. D. Dalton, K. Girtman, S. Mathew, J. Ma, S. B. Pounds, X. Su, C. H. Pui, M. V. Relling, W. E. Evans, S. A. Shurtleff and J. R. Downing (2007). "Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia." Nature **446**(7137): 758-764.
- Muzny, D. M., Bainbridge, M. N., Chang, K., Dinh, H. H., Drummond, J. A., Fowler, G., et al. , and C. G. A. Network (2012). "Comprehensive molecular characterization of human colon and rectal cancer." Nature **487**(7407): 330-337.
- Narayanan, A., F. Gagliardi, A. L. Gallotti, S. Mazzoleni, M. Cominelli, L. Fagnocchi, M. Pala, I. S. Piras, P. Zordan, N. Moretta, E. Tratta, G. Brugnara, L. Altabella, G. Bozzuto, P. Gorombei, A. Molinari, R. A. Padua, A. Bulfone, L. S. Politi, A. Falini, A. Castellano, P. Mortini, A. Zippo, P. L. Poliani and R. Galli (2018). "The proneural gene ASCL1 governs the transcriptional subgroup affiliation in glioblastoma stem cells by directly repressing the mesenchymal gene NDRG1." Cell Death Differ.

- Neftel, C., J. Laffy, M. G. Filbin, T. Hara, M. E. Shore, G. J. Rahme, A. R. Richman, D. Silverbush, M. L. Shaw, C. M. Hebert, J. Dewitt, S. Gritsch, E. M. Perez, L. N. Gonzalez Castro, X. Lan, N. Druck, C. Rodman, D. Dionne, A. Kaplan, M. S. Bertalan, J. Small, K. Pelton, S. Becker, D. Bonal, Q. D. Nguyen, R. L. Servis, J. M. Fung, R. Mylvaganam, L. Mayr, J. Gojo, C. Haberler, R. Geyeregger, T. Czech, I. Slavc, B. V. Nahed, W. T. Curry, B. S. Carter, H. Wakimoto, P. K. Brastianos, T. T. Batchelor, A. Stemmer-Rachamimov, M. Martinez-Lage, M. P. Frosch, I. Stamenkovic, N. Riggi, E. Rheinbay, M. Monje, O. Rozenblatt-Rosen, D. P. Cahill, A. P. Patel, T. Hunter, I. M. Verma, K. L. Ligon, D. N. Louis, A. Regev, B. E. Bernstein, I. Tirosh and M. L. Suvà (2019). "An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma." Cell. **178**(4): 835-849.e821. doi: 810.1016/j.cell.2019.1006.1024. Epub 2019 Jul 1018.
- Network, C. G. A. (2012). "Comprehensive molecular characterization of human colon and rectal cancer." Nature **487**(7407): 330-337.
- Network, C. G. A. R. (2008). "Comprehensive genomic characterization defines human glioblastoma genes and core pathways." Nature **455**(7216): 1061-1068.
- Noushmehr, H., D. J. Weisenberger, K. Diefes, H. S. Phillips, K. Pujara, B. P. Berman, F. Pan, C. E. Pelloski, E. P. Sulman, K. P. Bhat, R. G. Verhaak, K. A. Hoadley, D. N. Hayes, C. M. Perou, H. K. Schmidt, L. Ding, R. K. Wilson, D. Van Den Berg, H. Shen, H. Bengtsson, P. Neuvial, L. M. Cope, J. Buckley, J. G. Herman, S. B. Baylin, P. W. Laird and K. Aldape (2010). "Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma." Cancer Cell **17**(5): 510-522.
- Ohgaki, H. and P. Kleihues (2013). "The definition of primary and secondary glioblastoma." Clin Cancer Res **19**(4): 764-772.
- Orzan, F., F. De Bacco, G. Crisafulli, S. Pellegatta, B. Mussolin, G. Siravegna, A. D'Ambrosio, P. M. Comoglio, G. Finocchiaro and C. Boccaccio (2017). "Genetic Evolution of Glioblastoma Stem-Like Cells From Primary to Recurrent Tumor." **35**(11): 2218-2228.
- Orzan, F., F. Pagani, M. Cominelli, L. Triggiani, S. Calza, F. De Bacco, D. Medicina, P. Balzarini, P. P. Panciani, R. Liserre, M. Buglione, M. M. Fontanella, E. Medico, R. Galli, C. Isella, C. Boccaccio and P. L. Poliani (2020). "A simplified integrated molecular and immunohistochemistry-based algorithm allows high accuracy prediction of glioblastoma transcriptional subtypes." Lab Invest **13**(10): 020-0437.
- Parsons, D. W., S. Jones, X. Zhang, J. C. Lin, R. J. Leary, P. Angenendt, P. Mankoo, H. Carter, I. M. Siu, G. L. Gallia, A. Olivi, R. McLendon, B. A. Rasheed, S. Keir, T. Nikolskaya, Y. Nikolsky, D. A. Busam, H. Tekleab, L. A. Diaz, Jr., J. Hartigan, D. R. Smith, R. L. Strausberg, S. K. Marie, S. M. Shinjo, H. Yan, G. J. Riggins, D. D. Bigner, R. Karchin, N. Papadopoulos, G. Parmigiani, B. Vogelstein, V. E. Velculescu and K. W. Kinzler (2008). "An integrated genomic analysis of human glioblastoma multiforme." Science **321**(5897): 1807-1812.
- Perry, A., C. R. Miller, M. Gujrati, B. W. Scheithauer, S. C. Zambrano, S. C. Jost, R. Raghavan, J. Qian, E. J. Cochran, J. T. Huse, E. C. Holland, P. C. Burger and M. K. Rosenblum (2009). "Malignant gliomas with primitive neuroectodermal tumor-like components: a clinicopathologic and genetic study of 53 cases." Brain Pathol **19**(1): 81-90.
- Pfeifer, G. P., Y. H. You and A. Besaratinia (2005). "Mutations induced by ultraviolet light." Mutat Res **571**(1-2): 19-31.
- Phillips, H. S., S. Kharbanda, R. Chen, W. F. Forrest, R. H. Soriano, T. D. Wu, A. Misra, J. M. Nigro, H. Colman, L. Soroceanu, P. M. Williams, Z. Modrusan, B. G. Feuerstein and K. Aldape (2006). "Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis." Cancer Cell **9**(3): 157-173.
- Piccirillo, S. G., R. Combi, L. Cajola, A. Patrizi, S. Redaelli, A. Bentivegna, S. Baronchelli, G. Maira, B. Pollo, A. Mangiola, F. DiMeco, L. Dalprà and A. L. Vescovi (2009). "Distinct pools of cancer stem-like cells coexist within human glioblastomas and display different tumorigenicity and independent genomic evolution." Oncogene **28**(15): 1807-1811.
- Pleasance, E. D., P. J. Stephens, S. O'Meara, D. J. McBride, A. Meynert, D. Jones, M. L. Lin, D. Beare, K. W. Lau, C. Greenman, I. Varela, S. Nik-Zainal, H. R. Davies, G. R. Ordoñez, L. J. Mudie, C. Latimer, S. Edkins, L. Stebbings, L. Chen, M. Jia, C. Leroy, J. Marshall, A. Menzies, A. Butler, J. W. Teague, J. Mangion, Y. A. Sun, S. F. McLaughlin, H. E. Peckham, E. F. Tsung, G. L. Costa, C. C. Lee, J. D. Minna, A. Gazdar, E. Birney, M. D. Rhodes, K. J. McKernan, M. R. Stratton, P. A. Futreal and P. J. Campbell (2010). "A small-cell lung cancer genome with complex signatures of tobacco exposure." Nature **463**(7278): 184-190.
- Quant, E. C. and P. Y. Wen (2010). "Novel medical therapeutics in glioblastomas, including targeted molecular therapies, current and future clinical trials." Neuroimaging Clin N Am **20**(3): 425-448.
- Reis, Hornblower, Robb and Tzertzinis (2014). CRISPR/Cas9 & Targeted Genome Editing: New Era in Molecular Biology, *NEB expressions Issue*. **I**.
- Roberts, S. A. and D. A. Gordenin (2014). "Hypermutation in human cancer genomes: footprints and mechanisms." Nat Rev Cancer **14**(12): 786-800.

- Roberts, S. A., M. S. Lawrence, L. J. Klimczak, S. A. Grimm, D. Fargo, P. Stojanov, A. Kiezun, G. V. Kryukov, S. L. Carter, G. Saksena, S. Harris, R. R. Shah, M. A. Resnick, G. Getz and D. A. Gordenin (2013). "An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers." Nat Genet **45**(9): 970-976.
- Rodger, E. J., A. Chatterjee, P. A. Stockwell and M. R. Eccles (2019). "Characterisation of DNA methylation changes in EBF3 and TBC1D16 associated with tumour progression and metastasis in multiple cancer types." Clin Epigenetics **11**(1): 114.
- Sa, J. K., S. W. Choi, J. Zhao, Y. Lee, J. Zhang, D. S. Kong, J. W. Choi, H. J. Seol, J. I. Lee, A. Iavarone, R. Rabadan and D. H. Nam (2019). "Hypermutagenesis in untreated adult gliomas due to inherited mismatch mutations." Int J Cancer **144**(12): 3023-3030.
- Saab, R., C. Rodriguez-Galindo, K. Matmati, J. E. Rehg, S. H. Baumer, J. D. Khoury, C. Billups, G. Neale, K. J. Helton and S. X. Skapek (2009). "p18Ink4c and p53 Act as tumor suppressors in cyclin D1-driven primitive neuroectodermal tumor." Cancer Res **69**(2): 440-448.
- Sakuma, T., M. A. Barry and Y. Ikeda (2012). "Lentiviral vectors: basic to translational." Biochem J **443**(3): 603-618.
- Sanai, N., A. Alvarez-Buylla and M. S. Berger (2005). "Neural stem cells and the origin of gliomas." N Engl J Med **353**(8): 811-822.
- Sander, J. D. and J. K. Joung (2014). "CRISPR-Cas systems for editing, regulating and targeting genomes." Nat Biotechnol **32**(4): 347-355.
- Sanjana, N. E., O. Shalem and F. Zhang (2014). "Improved vectors and genome-wide libraries for CRISPR screening." Nat Methods **11**(8): 783-784.
- Shabbir, M. A., H. Hao, M. Z. Shabbir, H. I. Hussain, Z. Iqbal, S. Ahmed, A. Sattar, M. Iqbal, J. Li and Z. Yuan (2016). "Survival and Evolution of CRISPR-Cas System in Prokaryotes and Its Applications." Front Immunol **7**: 375.
- Shalem, O., N. E. Sanjana, E. Hartenian, X. Shi, D. A. Scott, T. Mikkelson, D. Heckl, B. L. Ebert, D. E. Root, J. G. Doench and F. Zhang (2014). "Genome-scale CRISPR-Cas9 knockout screening in human cells." Science **343**(6166): 84-87.
- Singh, S. K., C. Hawkins, I. D. Clarke, J. A. Squire, J. Bayani, T. Hide, R. M. Henkelman, M. D. Cusimano and P. B. Dirks (2004). "Identification of human brain tumour initiating cells." Nature **432**(7015): 396-401.
- Smits, L. M., S. Magni, K. Kinugawa, K. Grzyb, J. Luginbühl, S. Sabate-Soler, S. Bolognin, J. W. Shin, E. Mori, A. Skupin and J. C. Schwamborn (2020). "Single-cell transcriptomics reveals multiple neuronal cell types in human midbrain-specific organoids." Cell Tissue Res **382**(3): 463-476.
- Song, X., R. Andrew Allen, S. Terence Dunn, K. M. Fung, P. Farmer, S. Gandhi, T. Ranjan, A. Demopoulos, M. Symons, M. Schulder and J. Y. Li (2011). "Glioblastoma with PNET-like components has a higher frequency of isocitrate dehydrogenase 1 (IDH1) mutation and likely a better prognosis than primary glioblastoma." Int J Clin Exp Pathol **4**(7): 651-660.
- Stewart, S. A., D. M. Dykxhoorn, D. Palliser, H. Mizuno, E. Y. Yu, D. S. An, D. M. Sabatini, I. S. Chen, W. C. Hahn, P. A. Sharp, R. A. Weinberg and C. D. Novina (2003). "Lentivirus-delivered stable gene silencing by RNAi in primary cells." RNA **9**(4): 493-501.
- Stieber, D., A. Golebiewska, L. Evers, E. Lenkiewicz, N. H. Brons, N. Nicot, A. Oudin, S. Bougnaud, F. Hertel, R. Bjerkvig, L. Vallar, M. T. Barrett and S. P. Niclou (2014). "Glioblastomas are composed of genetically divergent clones with distinct tumourigenic potential and variable stem cell-associated phenotypes." Acta Neuropathol **127**(2): 203-219.
- Stiver, S. I., X. Tan, L. F. Brown, E. T. Hedley-Whyte and H. F. Dvorak (2004). "VEGF-A angiogenesis induces a stable neovasculature in adult murine brain." J Neuropathol Exp Neurol **63**(8): 841-855.
- Stupp, R., M. E. Hegi, W. P. Mason, M. J. van den Bent, M. J. Taphoorn, R. C. Janzer, S. K. Ludwin, A. Allgeier, B. Fisher, K. Belanger, P. Hau, A. A. Brandes, J. Gijtenbeek, C. Marosi, C. J. Vecht, K. Mokhtari, P. Wesseling, S. Villa, E. Eisenhauer, T. Gorlia, M. Weller, D. Lacombe, J. G. Cairncross and R. O. Mirimanoff (2009). "Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial." Lancet Oncol **10**(5): 459-466.
- Sturm, D., S. Bender, D. T. Jones, P. Lichter, J. Grill, O. Becher, C. Hawkins, J. Majewski, C. Jones, J. F. Costello, A. Iavarone, K. Aldape, C. W. Brennan, N. Jabado and S. M. Pfister (2014). "Paediatric and adult glioblastoma: multiform (epi)genomic culprits emerge." Nat Rev Cancer **14**(2): 92-107.
- Sturm, D., B. A. Orr, U. H. Toprak, V. Hovestadt, D. T. W. Jones, D. Capper, M. Sill, I. Buchhalter, P. A. Northcott, I. Leis, M. Ryzhova, C. Koelsche, E. Pfaff, S. J. Allen, G. Balasubramanian, B. C. Worst, K. W. Pajtler, S. Brabetz, P. D. Johann, F. Sahm, J. Reimand, A. Mackay, D. M. Carvalho, M. Remke, J. J. Phillips, A. Perry, C. Cowdrey, R. Drissi, M. Fouladi, F. Giangaspero, M. Łastowska, W. Grajkowska, W. Scheurlen, T. Pietsch, C. Hagel, J. Gojo, D. Lötsch, W. Berger, I. Slavc, C. Haberler, A. Jouvet, S. Holm, S. Hofer, M. Prinz, C. Keohane, I. Fried, C. Mawrin, D. Scheie, B. C. Mobley, M. J. Schniederjan, M. Santi, A. M. Buccoliero, S. Dahiya, C. M. Kramm, A. O. von Bueren, K. von Hoff, S. Rutkowski, C. Herold-Mende, M. C. Frühwald, T. Milde, M. Hasselblatt, P. Wesseling, J. Rößler, U. Schüller, M. Ebinger, J. Schittenhelm, S. Frank, R. Grobholz, I. Vajtai, V. Hans, R. Schneppenheim, K. Zitterbart, V. P. Collins, E. Aronica, P. Varlet, S. Puget, C. Dufour, J. Grill, D. Figarella-Branger, M. Wolter, M. U. Schuhmann, T. Shalaby, M.

Grotzer, T. van Meter, C. M. Monoranu, J. Felsberg, G. Reifenberger, M. Snuderl, L. A. Forrester, J. Koster, R. Versteeg, R. Volckmann, P. van Sluis, S. Wolf, T. Mikkelsen, A. Gajjar, K. Aldape, A. S. Moore, M. D. Taylor, C. Jones, N. Jabado, M. A. Karajannis, R. Eils, M. Schlesner, P. Lichter, A. von Deimling, S. M. Pfister, D. W. Ellison, A. Korshunov and M. Kool (2016). "New Brain Tumor Entities Emerge from Molecular Classification of CNS-PNETs." Cell **164**(5): 1060-1072.

- Sturm, D., H. Witt, V. Hovestadt, D. A. Khuong-Quang, D. T. Jones, C. Konermann, E. Pfaff, M. Tönjes, M. Sill, S. Bender, M. Kool, M. Zapatka, N. Becker, M. Zucknick, T. Hielscher, X. Y. Liu, A. M. Fontebasso, M. Ryzhova, S. Albrecht, K. Jacob, M. Wolter, M. Ebinger, M. U. Schuhmann, T. van Meter, M. C. Frühwald, H. Hauch, A. Pekrun, B. Radlwimmer, T. Niehues, G. von Komorowski, M. Dürken, A. E. Kulozik, J. Madden, A. Donson, N. K. Foreman, R. Drissi, M. Fouladi, W. Scheurlen, A. von Deimling, C. Monoranu, W. Roggendorf, C. Herold-Mende, A. Unterberg, C. M. Kramm, J. Felsberg, C. Hartmann, B. Wiestler, W. Wick, T. Milde, O. Witt, A. M. Lindroth, J. Schwartzentruber, D. Faury, A. Fleming, M. Zakrzewska, P. P. Liberski, K. Zakrzewski, P. Hauser, M. Garami, A. Klekner, L. Bognar, S. Morrissy, F. Cavalli, M. D. Taylor, P. van Sluis, J. Koster, R. Versteeg, R. Volckmann, T. Mikkelsen, K. Aldape, G. Reifenberger, V. P. Collins, J. Majewski, A. Korshunov, P. Lichter, C. Plass, N. Jabado and S. M. Pfister (2012). "Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma." Cancer Cell **22**(4): 425-437.
- Swarts, D. C., C. Mosterd, M. W. van Passel and S. J. Brouns (2012). "CRISPR interference directs strand specific spacer acquisition." PLoS One **7**(4): e35888.
- Takami, H., A. Yoshida, S. Fukushima, H. Arita, Y. Matsushita, T. Nakamura, M. Ohno, Y. Miyakita, S. Shibui, Y. Narita and K. Ichimura (2015). "Revisiting TP53 Mutations and Immunohistochemistry--A Comparative Study in 157 Diffuse Gliomas." Brain Pathol **25**(3): 256-265.
- Tamai, S., M. Kinoshita, H. Sabit, T. Furuta, K. Miyashita, K. Yoshimura, T. Homma, K. Harada and M. Nakada (2019). "Case of metastatic glioblastoma with primitive neuronal component to the lung." Neuropathology **39**(3): 218-223.
- Tanaka, S., D. N. Louis, W. T. Curry, T. T. Batchelor and J. Dietrich (2013). "Diagnostic and therapeutic avenues for glioblastoma: no longer a dead end?" Nat Rev Clin Oncol **10**(1): 14-26.
- Thomas, A. A., C. W. Brennan, L. M. DeAngelis and A. M. Omuro (2014). "Emerging therapies for glioblastoma." JAMA Neurol **71**(11): 1437-1444.
- Thomas, C., M. Sill, V. Ruland, A. Witten, S. Hartung, U. Kordes, A. Jeibmann, R. Beschorner, K. Keyvani, M. Bergmann, M. Mittelbronn, T. Pietsch, J. Felsberg, C. M. Monoranu, P. Varlet, P. Hauser, A. Olar, R. G. Grundy, J. E. Wolff, A. Korshunov, D. T. Jones, M. Bewerunge-Hudler, V. Hovestadt, A. von Deimling, S. M. Pfister, W. Paulus, D. Capper and M. Hasselblatt (2016). "Methylation profiling of choroid plexus tumors reveals 3 clinically distinct subgroups." Neuro Oncol **18**(6): 790-796.
- Verhaak, R. G., K. A. Hoadley, E. Purdom, V. Wang, Y. Qi, M. D. Wilkerson, C. R. Miller, L. Ding, T. Golub, J. P. Mesirov, G. Alexe, M. Lawrence, M. O'Kelly, P. Tamayo, B. A. Weir, S. Gabriel, W. Winckler, S. Gupta, L. Jakkula, H. S. Feiler, J. G. Hodgson, C. D. James, J. N. Sarkaria, C. Brennan, A. Kahn, P. T. Spellman, R. K. Wilson, T. P. Speed, J. W. Gray, M. Meyerson, G. Getz, C. M. Perou, D. N. Hayes and N. Cancer Genome Atlas Research (2010). "Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1." Cancer Cell **17**(1): 98-110.
- Vitucci, M., D. N. Hayes and C. R. Miller (2011). "Gene expression profiling of gliomas: merging genomic and histopathological classification for personalised therapy." Br J Cancer **104**(4): 545-553.
- Walter, D. M., O. S. Venancio, E. L. Buza, J. W. Tobias, C. Deshpande, A. A. Gudiel, C. Kim-Kiselak, M. Cicchini, T. J. Yates and D. M. Feldser (2017). "Systematic in vivo inactivation of chromatin regulating enzymes identifies Setd2 as a potent tumor suppressor in lung adenocarcinoma." Cancer Res **77**(7): 1719-1729.
- Wang, J., E. Cazzato, E. Ladewig, V. Frattini, D. I. Rosenbloom, S. Zairis, F. Abate, Z. Liu, O. Elliott, Y. J. Shin, J. K. Lee, I. H. Lee, W. Y. Park, M. Eoli, A. J. Blumberg, A. Lasorella, D. H. Nam, G. Finocchiaro, A. Iavarone and R. Rabadan (2016). "Clonal evolution of glioblastoma under therapy." Nat Genet **48**(7): 768-776.
- Wang, Q., B. Hu, X. Hu, H. Kim, M. Squatrito, L. Scarpace, A. C. deCarvalho, S. Lyu, P. Li, Y. Li, F. Barthel, H. J. Cho, Y. H. Lin, N. Satani, E. Martinez-Ledesma, S. Zheng, E. Chang, C. E. Gabriel Sauve, A. Olar, Z. D. Lan, G. Finocchiaro, J. J. Phillips, M. S. Berger, K. R. Gabrusiewicz, G. Wang, E. Eskilsson, J. Hu, T. Mikkelsen, R. A. DePinho, F. Muller, A. B. Heimberger, E. P. Sulman, D. H. Nam and R. G. W. Verhaak (2018). "Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment." Cancer Cell **33**(1): 152.
- Wang, S. S., A. G. Betz and R. R. Reed (2002). "Cloning of a novel Olf-1/EBF-like gene, O/E-4, by degenerate oligo-based direct selection." Mol Cell Neurosci **20**(3): 404-414.
- Wang, S. S., R. Y. Tsai and R. R. Reed (1997). "The characterization of the Olf-1/EBF-like HLH transcription factor family: implications in olfactory gene regulation and neuronal development." J Neurosci **17**(11): 4149-4158.
- Westphal, M. and K. Lamszus (2011). "The neurobiology of gliomas: from cell biology to the development of therapeutic approaches." Nature Reviews Neuroscience **12**(9): 495-508.

- Wharton, S., I. Whittle, D. Collie, H. Bell and J. Ironside (2001). Gliosarcoma with areas of primitive neuroepithelial differentiation and extracranial metastasis . *Clinical neuropathology*. **20:** 212-218.
- Wood, M. D., G. F. Reis, D. E. Reuss and J. J. Phillips (2016). "Protein Analysis of Glioblastoma Primary and Posttreatment Pairs Suggests a Mesenchymal Shift at Recurrence." J Neuropathol Exp Neurol **75**(10): 925-935.
- Xiao, A., H. Wu, P. P. Pandolfi, D. N. Louis and T. Van Dyke (2002). "Astrocyte inactivation of the pRb pathway predisposes mice to malignant astrocytoma development that is accelerated by PTEN mutation." Cancer Cell **1**(2): 157-168.
- Xiao, A., C. Yin, C. Yang, A. Di Cristofano, P. P. Pandolfi and T. Van Dyke (2005). "Somatic induction of Pten loss in a preclinical astrocytoma model reveals major roles in disease progression and avenues for target discovery and validation." Cancer Res **65**(12): 5172-5180.
- Xu, G. and J. Y. Li (2018). "CDK4, CDK6, cyclin D1, p16(INK4a) and EGFR expression in glioblastoma with a primitive neuronal component." J Neurooncol **136**(3): 445-452.
- Xu, G., H. Zheng and J. Y. Li (2019). "Next-generation whole exome sequencing of glioblastoma with a primitive neuronal component." Brain Tumor Pathol **36**(3): 129-134.
- Yan, H., D. W. Parsons, G. Jin, R. McLendon, B. A. Rasheed, W. Yuan, I. Kos, I. Batinic-Haberle, S. Jones, G. J. Riggins, H. Friedman, A. Friedman, D. Reardon, J. Herndon, K. W. Kinzler, V. E. Velculescu, B. Vogelstein and D. D. Bigner (2009). "IDH1 and IDH2 mutations in gliomas." N Engl J Med **360**(8): 765-773.
- Yao, K., X. L. Qi, X. Mei and T. Jiang (2015). "Gliosarcoma with primitive neuroectodermal, osseous, cartilage and adipocyte differentiation: a case report." Int J Clin Exp Pathol **8**(2): 2079-2084.
- Zardo, G., M. I. Tiirikainen, C. Hong, A. Misra, B. G. Feuerstein, S. Volik, C. C. Collins, K. R. Lamborn, A. Bollen, D. Pinkel, D. G. Albertson and J. F. Costello (2002). "Integrated genomic and epigenomic analyses pinpoint biallelic gene inactivation in tumors." Nat Genet **32**(3): 453-458.
- Zhao, L. Y., Y. Niu, A. Santiago, J. Liu, S. H. Albert, K. D. Robertson and D. Liao (2006). "An EBF3-mediated transcriptional program that induces cell cycle arrest and apoptosis." Cancer Res **66**(19): 9445-9452.
- Zufferey, R., T. Dull, R. J. Mandel, A. Bukovsky, D. Quiroz, L. Naldini and D. Trono (1998). "Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery." J Virol **72**(12): 9873-9880.

RINGRAZIAMENTI

Il presente lavoro di tesi non sarebbe stato possibile senza la preziosa collaborazione di tanti colleghi e collaboratori.
Innanzitutto ringrazio il prof. Poliani per avermi dato fiducia ed avermi accolto nel suo laboratorio, per i suoi preziosi insegnamenti e per il tempo dedicato alla stesura del presente lavoro di tesi.
Ringrazio la collega d.ssa Cominelli per il prezioso supporto e per il lavoro di istologia ed immunoistochimica.
Ringrazio i colleghi del Laboratorio di Ricerca ed i colleghi del Laboratorio di Diagnostica della sezione di Anatomia Patologica, in particolare il Laboratorio di Biologia Molecolare, e tutto il personale del reparto.
Ringrazio tutti colleghi della Sezione di Biologia Molecolare del DMMT diretta dalla prof.ssa Poli, in particolare la d.ssa Gryzik e la d.ssa Asperti, per il design e la realizzazione del sistema CRISPR/Cas9.
Ringrazio il dott. La Via della Sezione di Biologia e Genetica del DMMT per l'aiuto nella realizzazione del vettore lentivirale.
Ringrazio la d.ssa Galli del laboratorio dell'Unità di Biologia delle Cellule Staminali Neurali, Istituto San Raffaele, per aver fornito le cellule staminali di Glioblastoma e per gli insostituibili consigli.
Ringrazio le colleghe del Laboratorio per la Ricerca sulle Staminali Tumorali diretto dalla prof. Boccaccio presso L'Istituto di Candiolo IRCCS, d.ssa De Bacco e d.ssa Orzan, per aver fornito le cellule staminali di GBM-PNC e per le analisi molecolari in NGS, nonché per la costante e proficua collaborazione.

GIRL POWER!

Ringrazio il dott. Crisafulli del Gruppo di Bio Informatica del Laboratorio di Oncologia Molecolare, Istituto di Candiolo IRCCS, per le analisi bioinformatiche dei dati di NGS.
Ringrazio il prof. Dalerba, Department of Pathology and Cell Biology, Columbia University, per la revisione del lavoro di tesi, per i consigli illuminanti e per la continua disponibilità.
Ringrazio il prof. Conti per la revisione del lavoro di tesi.
Ringrazio la d.ssa Delbarba per il supporto grafico (grazie Vane).
Ringrazio l'Associazione "dedicato A te" per il loro impegno profuso a favore della ricerca e per il loro generoso sostegno, sempre in ricordo di Andrea.

Infine ringrazio chiunque mi abbia insegnato qualcosa nel corso degli anni di studio e di professione.

Ad maiora semper