Synthetic Genetic Tracing of Molecular and Cellular Heterogeneity in Glioblastoma

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

Doctor rerum naturalium

(Dr. rer. nat.)

eingereicht an der

Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin

von

Matthias Jürgen Schmitt, M.Sc.

Präsidentin der Humboldt-Universität zu Berlin:

Prof. Dr. Julia von Blumenthal

Dekan der Lebenswissenschaftlichen Fakultät

der Humboldt-Universität zu Berlin:

Prof. Dr. Dr. Christian Ulrichs

Gutachter/innen

1. Prof. Dr. Markus Landthaler
2. Prof. Dr. Christoph Harms
3. Dr. Gaetano Gargiulo

Tag der mündlichen Prüfung: **16.12.2022**

Abstract

Glioblastoma (GBM) remains the most difficult primary solid tumor of the central nervous system despite the intensively growing body of research on its molecular and cellular characteristics. Whereas GBM treatment is aggressive and involves surgical resection, radiotherapy, and chemotherapy, tumor recurrence is unavoidable. GBM treatment resistance is associated with genetic and cellular heterogeneity, as well as phenotypic plasticity.

To improve understanding of Glioblastoma heterogeneity, we developed custom genetic tracing strategies for subtype-specific transcriptional states from Glioblastoma patient signatures. In GBM cells, our novel technology enabled us to identify intrinsic and non-cell autonomous determinants of cell fate commitment. *In vitro* and *in vivo*, we discovered that the mesenchymal GBM adapts in the presence of microenvironmental signaling and is regulated by inflammatory and differentiation programs. We demonstrated that cell fate commitment towards a mesenchymal state is adaptive and reversible and occurs through partially overlapping transcriptional responses, including external signaling and ionizing radiation. Importantly, using synthetic locus control regions (sLCRs), we were able to uncover crosstalk between innate immune cells and glioma-initiating cells, directing the tumor cells into a mesenchymal state linked to increased resistance to chemotherapy.

Here, we build on this innovative approach to trace cell fate transitions in complex biological settings, with a focus on the cellular crosstalk between malignant and non-tumor cells in the context of phenotypic plasticity and therapeutic resistance. Beyond that, this method offers the broad translational potential to be applied to other fields of research, including developmental biology or regenerative medicine.

Zusammenfassung

Das Glioblastom (GBM) repräsentiert den am schwierigsten zu behandelnden primären soliden Tumor des Zentralnervensystems dar, trotz der intensiv wachsenden Zahl von Studien zu seinen molekularen und zellulären Eigenschaften. Obwohl die GBM-Therapie aggressiv ist und chirurgische Resektion, Strahlentherapie und Chemotherapie umfasst, ist ein Wiederauftreten des Tumors unvermeidlich. Die GBM-Behandlungsresistenz ist mit genetischer und zellulärer Heterogenität sowie phänotypischer Plastizität verbunden.

Um das Verständnis der Heterogenität des Glioblastoms zu vertiefen, haben wir maßgeschneiderte genetische Tracing-Strategien für subtypspezifische Transkriptionszustände aus Glioblastom-Patientensignaturen entwickelt. In GBM-Zellen ermöglichte uns unsere neuartige Technologie, intrinsische und nicht-zellautonome Bestimmungsfaktoren von Zellzuständen zu identifizieren. *In vitro* und *in vivo* konnten wir zeigen, dass sich der mesenchymale GBM-Subtyp als adaptive Identität in Gegenwart von Mikroumgebungssignalen ausbildet und durch Entzündungs- und Differenzierungsprogramme reguliert wird. Wir haben gezeigt, dass die Ausbildung eines mesenchymalen Zellzustand adaptiv und reversibel ist und durch verschiedene Auslöser wie externer Signaltransduktion und ionisierende Strahlung mit teilweise überlappenden transkriptionellen Signaturen eingenommen werden kann. Insbesondere konnten wir mithilfe synthetischer Locus-Kontrollregionen (sLCRs) eine Interaktion zwischen Zellen des angeborenen Immunsystems und Glioma-Zellen aufdecken, wodurch die Tumorzellen in einen mesenchymalen Zustand versetzt wurden, der mit einer erhöhten Resistenz gegen Chemotherapie verbunden ist.

Hier bauen wir auf diesem innovativen Ansatz auf, um Übergänge von Zellzuständen in komplexen biologischen Umgebungen zu verfolgen, mit einem Schwerpunkt auf der zellulären Wechselwirkung zwischen gesunden und Tumorzellen im Zusammenhang mit phänotypischer Plastizität und therapeutischer Resistenz. Darüber hinaus bietet diese Methode ein breites translationales Potenzial für die Anwendung auf andere Forschungsgebiete, einschließlich der Entwicklungsbiologie oder der regenerativen Medizin.

I. Table of contents

II. Index

III. List of Figures

IV. List of Tables

1 Introduction

1.1 Glioblastoma

1.1.1 Epidemiology and histopathology of Glioblastoma

Gliomas are tumours of the brain that make up 80% of all malignant brain tumours. Gliomas were historically divided into four classes (grades I–IV) based on microscropic appearance and histological features, with grade IV, often known as Glioblastoma (GBM; Louis et al., 2007; Jovčevska et al., 2013; Weller et al., 2015). Adult gliomas are currently categorized into two types based on the mutational status of two important biomarkers, the isocitrate dehydrogenase genes IDH1 and IDH2. Lower histologic grades are associated with a better prognosis and a median survival of more than 12 years in IDH-mutant gliomas, but they frequently progress to higher grades and clinical behavior later in the disease's natural history.

GBMs fall into the category of Isocitrate dehydrogenase (IDH) wild-type (IDHwt) in the 2021 WHO classification of CNS malignancies, and they are the most aggressive kind of diffuse gliomas, with an overall mean survival rate of 12–17 months following diagnosis (Stupp et al., 2005; Louis et al., 2021). GBM has an incidence of 3.2 diagnoses per 100.000 inhabitants with a median age of 65 years and is slightly more prevalent in men than in women (Thakkar et al., 2014; Tamimi et al., 2017).

1.1.2 Therapeutic approaches for Glioblastoma

The current standard treatment regimen for newly diagnosed GBM patients was introduced in 2005, and it includes as a first-line intervention a maximum safe surgical resection. Since GBMs are extremely heterogeneous and prone to infiltrating neighboring tissues, total surgical resection is nearly impossible. Therefore, radiotherapy and concomitant administration of temozolomide (TMZ), an FDA-approved alkylating agent, are usually following surgery in the standard-of-care (Stupp et al., 2005; Weller et al., 2013). However, the outcome of this therapy is limited to a 14-month survival benefit and even less in elderly patients (Stupp et al., 2009; Scott et al., 2012). Unfortunately, virtually all patients show tumor regression after the first-line treatment with a 5-year recurrence rate as high as 90% (Weller et al., 2013). Once the tumor is recurring, therapeutic options are also limited and the success of second surgery or re-irradiation is modest and can only be applied to patients with favourable prognostic indications (Ringel et al., 2015; Suchorska et al., 2016; Kazmi et al., 2019; Birzu et al., 2021).

Since GBM is a highly vascularized tumor, anti-angiogenic therapies with Bevacizumab (a monoclonal antibody binding the VEGFR) have advanced as far as phase III clinical trials, but only shown an increase of progression-free-survival without improved overall-survival, which is why it is now used alone or in combination trials for recurrent GBM (Chinot et al., 2014; Gilbert et al., 2014; Birzu et al., 2021). Similarly, precision medicine specifically targeting aberrant signaling pathways discovered in GBM has gained momentum in an attempt to improve patient outcome. These include either monoclonal antibodies (e.g. Bevacizumab) or small molecule inhibitors targeting various receptor tyrosine kinases such as EGFR or PDGFR (Weller et al., 2013; Wilson et al., 2014). However, results from phase II clinical trials for multiple targeted therapies did not show clear results, although it is noteworthy that small sample sizes and very heterogenous pre-treatment histories of patients impede a systematic assessment (Wen et al., 2006; Raymone et al,. 2008; Raizer et al., 2009; van den Bent et al., 2009; Peereboom et al., 2010; Birzu et al., 2021).
Further experimental therapies have been developed and are currently in various phases of clinical trials. These approaches span immunotherapy including cancer-peptide vaccines, CAR-T cell therapies and check-point inhibition, oncolytic viral therapy, epigenetic drugs such as histone deacetylase inhibitors as well as targeting the tumor microenvironment, where the list is of course not exhaustive (Wilson et al., 2014; Wick et al., 2018; Birzu et al., 2021; Oronsky et al., 2021).
Glioblastoma remains one of the deadliest cancers even after decades of study. Improvement of patient outcomes beyond the standard-of-care has stalled, due to the addition of targeted therapies and novel approaches largely failing to significantly improve overall-survival and the lack of a standardized treatment regimen for inevitably recurrent tumors.
Taken together, this also condenses to the view that a one-fits-all therapy is and certainly will not be successful in the future. Instead, combinatorial approaches of multivariate therapies, as well as discovering biomarkers or prognostic parameters in order to stratify patients towards more favourable outcomes may prove beneficial.

1.1.3 Mechanisms of treatment resistance

Resistance to therapy has been historically largely revolving around the role of genetic mutations to overcome drug sensitivity in cancer research (Schmitt et al., 2016). This working model included the possibilities of mutations that either alter the structure of the target in order to alleviate binding of the inhibitor, the bypassing of the target through activation of downstream effectors or the upregulation and reliance on an alternative compensatory signaling pathway. Underlying to this form of genetic resistance mechanism is either the selection of pre-existing

resistant subclones with certain genetic alterations or the selective acquisition of new mutations that confer resistance (Hata et al., 2016; Boumahdi et al., 2020). The evolutionary trajectory of resistance is influenced by the degree of spatial tumor heterogeneity, as tumors with a high degree of subclonal variety would be more likely to harbour pre-resistant clones than tumors with a more homogenous clonal architecture, as well as the temporal dynamics underlying these processes (Andor et al., 2015; Maley et al., 2017).

More recently it has been increasingly recognised that genetic factors alone are likely not the only driver of therapy resistance and other possible non-genetic mechanisms are also to be taken into account (Sharma et al., 2010; Marine et al. 2020). These include adaptive responses of tumor cells towards more drug-tolerant cellular states through epigenetic and transcriptional reprogramming. Such a process is referred to as phenotypic plasticity and in normal cells represents a natural way to cope with perturbations in their environment such as developmental changes or injury. In the context of cancer, therapy-induced phenotypic plasticity offers tumor cells the possibility to shift towards a non-sensitive cellular state and further fuels the level of intratumoral heterogeneity (Gupta et al., 2019).

Adaptive responses can either be driven by tumor intrinsic response programmes or influenced by anatomical features and members of the microenvironment, which form together with cancer cells the tumor eco-system (Maley et al., 2017; Labrie et al., 2022). Furthermore, phenotypic plasticity in response to a certain therapeutic stress is not a one-way street and in principle allows tumor cells to partly revert to their initial state upon removal of treatment pressure. Both, genetic and non-genetic resistance mechanisms are not mutually exclusive and are likely to act in an interplay and influence each other with selection for those mechanisms which are ultimately having the highest success of conferring resistance (Maley et al., 2017; Boumahdi et al., 2019; Marine et al., 2020).

Which form or combination of resistance mechanisms are responsible for the poor response of Glioblastoma patients to virtually any tested treatment approach remains one of the biggest challenges in clinical management of the disease. What is known is that Glioblastoma exhibits a high degree of heterogeneity at the inter- and intra-tumoral as well as cellular level, as described in the following paragraphs, and that this heterogeneity is likely a major contribution to its therapeutic resistance.

1.1.4 The genomic landscape of GBM

Historically, characterization of brain tumors has relied on histologic tissue assessment and grading based on the degree of anaplasia to provide a diagnosis for patients (Louis et al., 2007). However, exclusive histophatological glioma classification does not provide sufficient information on the biological properties of a tumor (e.g. treatment response) and consequently is insufficient for accurate patient stratification (van den Bent et al., 2010). As advances in sequencing technologies have emerged, vast amounts of genetic information from hundreds of brain tumor samples have been generated and are now used in conjunction with histopathology in a multi-layered approach to define brain tumor subtypes.

The identification of IDH mutations in gliomas was a major turning point in our understanding of this disease. Mutations of IDH result in accumulation of the oncometabolite d-2-hydroxyglutarate (d-2-HG), which mechanistically leads to abnormal DNA and histone methylation, eventually resulting in widespread CpG island hypermethylation (Parsons et al., 2008; Yan et al., 2008; Balss et al., 2008; Noushmehr et al., 2010; Weller et al., 2015). IDH mutations are most prevalent in low-grade gliomas, associated with younger patient ages and longer overall survival, and are characteristic of secondary GBMs (Yan et al., 2009; Cohen et al., 2013; Weller et al., 2015). They are considered to be one of the early-occurring genetic alterations during tumor intiation, although IDH mutant alone was not sufficient to develop gliomas in mouse models (Sasaki et al., 2012; Suzuki et al., 2015).

GBMs are IDH-wildtype, WHO grade IV gliomas and reside among the most extensively genetically characterized tumor types due to large-scale sequencing efforts by The Cancer Genome Atlas consortium (TCGA 2008). GBMs usually display diverse and complex genomic patterns with several genetic mutations and chromosomal aberrations, such as the frequently observed gain of chromosome 7 and monosomy of chomosome 10. However, the bandwidth of mutations tends to revolve around three distinct biological pathways connected to known oncogenic pathways. Firstly, activating mutations of members of the receptor tyrosine kinase MAPK/PI3K signaling pathways, including deletion of phosphatase and tensin homolog (PTEN), amplifications of epidermal growth factor receptor (EGFR) and platelet-derived growth factor receptor alpha (PDGFRa) or mutations of Neurofibromin 1 (NF1) and PI3 kinases. Secondly, alterations in proto-oncogenes associated with the p53 tumor suppressor pathway, including TP53 and murine double minute 2 and 4 (MDM2, MDM4), leading to an inhibition of apoptosis. The third major pathway affected by frequent genomic alterations is the tumor suppressive retinoblastoma (RB1) pathway, with cyclin-dependent kinases (CDK4, CDK6, CCND2, CDKN2A/B) as regulators of cell cycle regulation displaying disruptive abberations. Typically, mutations within a single pathway

occur at variable frequencies and are mutually exclusive, with the exception of receptor tyrosine kinases in the PI3K pathway. Additionally, mutations of TERT and ATRX are common, which are key regulators of chromatin remodelling and telomere length (TCGA 2008; Verhaak et al., 2010; Brennan et al., 2013; Aldape et al., 2015; Ceccarelli et al., 2016). Around 40% of IDH-wild-type Glioblastomas have EGFR amplification, and a majority of these tumors also have a genomic rearrangement resulting in exon 2–7 deletion. This mutation results in the production of EGFR variant III, which lacks the extracellular ligand-binding region of the deleted exons but is constitutively active (Lee et al., 2006; Weller et al., 2015; Aldape et al., 2015).

1.1.5 Molecular subtype classification of GBM

Although much information about the genetics of the disease has been generated, the high degree of genomic and histopathological heterogeneity has prevented major breakthroughs in the development of effective treatments. Purely genotype-targeted therapies have not significantly improved patient outcome, since they either display poor brain penetrance associated with severe patient toxicity or are overcome by evading mechanisms by the tumor (Jain 2018; Le Rhun et al., 2019). For instance, targeted inhibition of EGFR, which is an established target and oncogenic driver in GBM, has not yet shown therapeutic effects, since there is evidence of the existence of extrachromosomal DNA (ecDNA) which increases its expression and offers a straightforward escape route (Nathanson et al., 2014).

In an effort to resolve inter-patient variability and to integrate further molecular profiles for GBM classification, TCGA combined data from three distinct platforms into a single coherent dataset. Based on bulk tumor gene expression data from several hundred GBMs, they discovered gene expression profiles that defined four transcriptional subtypes of GBM, which they termed proneural, mesenchymal, classical, and neural (Verhaak et al., 2010).

The proneural subtype (PN) has been linked to specific point mutations in IDH1 and TP53, and amplification of PDGFRA expression. Characteristic markers of the oligodendrocyte lineage and neuronal development (PDGFRA, OLIG2, NKX2-2, DCX, DLL3, ASCL1, TCF4, SOX2/4/10/11) are signature genes of the proneural subtype, whereas EGFR and PTEN are usually wildtype. A majority of secondary GBMs (i.e. those arising from lower grade gliomas) as well as patients diagnosed below the age of 40 years are marked by a proneural subtype signature (Phillips et al., 2006; Noushmehr et al., 2010; Verhaak et al., 2010).

The classical (CL) subtype is especially marked by chromosome 10 loss and chromosome 7 amplification, which encodes the EGFR gene, resulting in high levels of EGFR expression with

frequent point mutations or the presence of EGFRvIII. Further characteristics include deletion of the CDKN2A gene and absence of mutations, which are common in the other two subtypes like IDH1, PDGFRA, TP53 and NF1 (Verhaak et al., 2010).

The mesenchymal (MES) subtype is associated with a hemizygous deletion of the NF1 gene and deletion of PTEN and featured the expression of markers genes of mesenchymal differentiation and astrocytic genes (S100A1, CHI3L1, CD44, MET), as well as being associated with inflammation, the NFkB signaling pathway, high degrees of necrosis and infiltration of immune cells (Phillips et al., 2006; Verhaak et al., 2010; Wang et al., 2017). A later study focusing on glioma-intrinsic gene signatures suggested the neural subtype to be an artifact of healthy brain tissue, likely stemming from the tumor borders. Through filtering steps for environmental contaminants, the number of transcriptional subtypes was reduced to mesenchymal, classical and proneural. Although the approximately 50 marker genes per subtype were considerably smaller than the 840 genes of the previous TCGA subtype clusters, about 50% of signature genes were identical (Wang et al., 2017).

Besides gene expression, also DNA methylation profiling has been used to build a classification for gliomas (Galbraith & Snuderl, 2022). The first comprehensive study comparing the methylomes of glioma patients classified them into three groups, one of which had a significant degree of hypermethylation at a wide number of genomic loci and was consequently designated the glioma-CpG island methylator phenotype (G-CIMP). Tumors that fell into the G-CIMP category showed a tendency for improved patient survival and were highly associated with mutations in IDH1 and the proneural gene expression subtype (Noushmehr et al., 2010). A later study compiled DNA methylation patterns of adult and pediatric glioma patients and similarly identified six distinct clusters that predominantly separated younger from adult patients, which further subdivided based on the presence of mutations in IDH (Sturm et al., 2012). The role of IDH mutations as a clinical biomarker and disease classifier was further corroborated in a follow-up study by Ceccarelli et al., who performed comprehensive molecular profiling of 1,122 adult gliomas and a pan-glioma methylation profiling spanning multiple adult and pediatric brain tumors (Ceccarelli et al., 2016; Capper et al., 2018). As a result, the WHO adopted in its updated classification system of central nervous tumors the use of DNA methylation and specifically delineated IDH mutated gliomas from Glioblastomas, which are defined as IDH wildtype (Louis et al., 2021).

Despite the identification of molecular classifiers based on transcriptional signatures, their clinical impact remained obscure. The initially reported survival advantadge of proneural patients in the study of Verhaak et al., 2010 was later corrected by stratifying patients according to IDH

mutational status, with IDH wildtype patients of neither transcriptional subtype displaying a marked survival difference in newly diagnosed Glioblastoma (Sturm et al., 2012; Ceccarelli et al., 2016). Ultimately, molecular classification did not prompt the development of new therapies, despite the fact that these years of extensive studies shed light on the complex relationship between genetic mutations, modifications in DNA methylation, chromatin remodelling, and gene expression, and provided understanding of how dysregulation at various levels may contribute to the pathobiology of GBM.

1.2 Cell-intrinsic and external factors of GBM intra-tumoral heterogeneity

1.2.1 Transcriptional signatures define a continuum of cellular states and plasticity

Although the derivation of a molecular classification system for GBM by the TCGA has proven valuable to increase the understanding of the molecular variability of GBMs between patients (inter-tumoral heterogeneity), the diversity within individual tumors (intra-tumoral heterogeneity) could not be assessed with bulk tumor sequencing methods. This was addressed by moving from bulk tumor analyses to spatially-resolved tumor multi-region sampling and later also to the single-cell level. Sottoriva et al. proved that various transcriptome subtypes exist inside the same tumor by evaluating multiple diverse GBM tumor regions (Sottoriva et al., 2013). In the first single-cell study on GBM, Patel et al. demonstrated that single tumor cells within the same tumor can adopt all three TCGA-defined molecular subtype signatures (Patel et al., 2014). Taken together, these studies show that numerous transcriptome profiles corresponding to distinct cellular states coexist within a GBM tumor and that the expression profile of bulk tumors is a composite of GBM cells with very diverse transcriptional states.

Much attention has been drawn to the establishment of a concept of self-renewing glioma stem cells (GSCs) at the apex of tumor cell hierarchy and as a source for tumor initiation and maintenance (Singh et al., 2003; Lathia et al., 2015; Lan et al., 2017; Mitchell et al., 2021). As a broad functional definition, GSCs have been ascribed the capacity to propagate a tumor when transplanted into a recipient animal and recapitulate its characteristic histology. Further properties of GSCs have been inferred from this functional definition, such as their ability of self-renewal and being the drivers of tumor relapse as a result of their resistance to chemo- and radiotherapy (Bao et al., 2006; Liu et al., 2006; Chen et al., 2012; Gimple et al., 2019; Prager et al., 2020). Traditionally, putative markers enriching for GSCs were used for isolation from tumor samples

such as CD133, CD44, CD15, SOX2, L1CAM, PDGFRA and EGFR (Singh et al., 2004; Galli et al., 2006; Lee et al., 2006; Piccirillo et al., 2006; Son et al., 2009; Anido et al., 2010). However, a clear definition of what a GSC is, what functional properties are included, and what surface markers are presented is interpreted differently by multiple groups in the field and therefore sparks some controversy (Parada et al., 2017; Prager et al., 2019; Prager et al., 2020; Suvà & Tirosh, 2020).

With the advent of single-cell RNA-sequencing of GBMs, the granularity of cellular states among malignant cells could be greatly enhanced, which seemed to increasingly challenge the theory of a strict unidirectional hierarchy in GBM. Among IDH-wildtype glioma, it has been shown that single tumor cells occupy multiple cellular states, which all contain proliferating cells and are able to individually give rise to tumors (Patel et al., 2014; Neftel et al., 2019; Bhaduri et al., 2020). Each cellular state displayed a characteristic transcriptional signature either associated with neurodevelopmental cells or a mesenchymal programme (Astrocyte (AC)-like, Neural progenitor cell (NPC)-like, Oligodendrocyte progenitor cell (OPC)-like and Mesenchymal (MES)-like) and displayed variable expression patterns of all previously established GSC-isolation markers. Furthermore, cellular state plasticity was described, with tumor cells of each isolated cellular state being able to switch their fate and reconstruct the initial heterogeneity of the tumors they were extracted from (Neftel et al., 2019).

These discoveries lay the ground for the concept of "GSC-multiplicity," in which GSC-like features are understood as transient biological properties of a cellular state, which can be acquired or lost rather than as a discrete entity atop a hierarchical stem cell hierarchy (Suvà & Tirosh, 2020). Indeed, among putative GSC populations from other studies, a certain bifurcation into a more proliferative/stemness and developmental-like programme (overlapping with OPC/NPC or PN) and a second trajectory driven by differentiation and inflammatory injury-response (overlapping with AC/MES or CL) appears to be a feature that is common among multiple reports (Fig. 1; Neftel et al., 2019; Bhaduri et al., 2020; Couturier et al., 2020; Richards et al., 2021; Yabo et al., 2021).

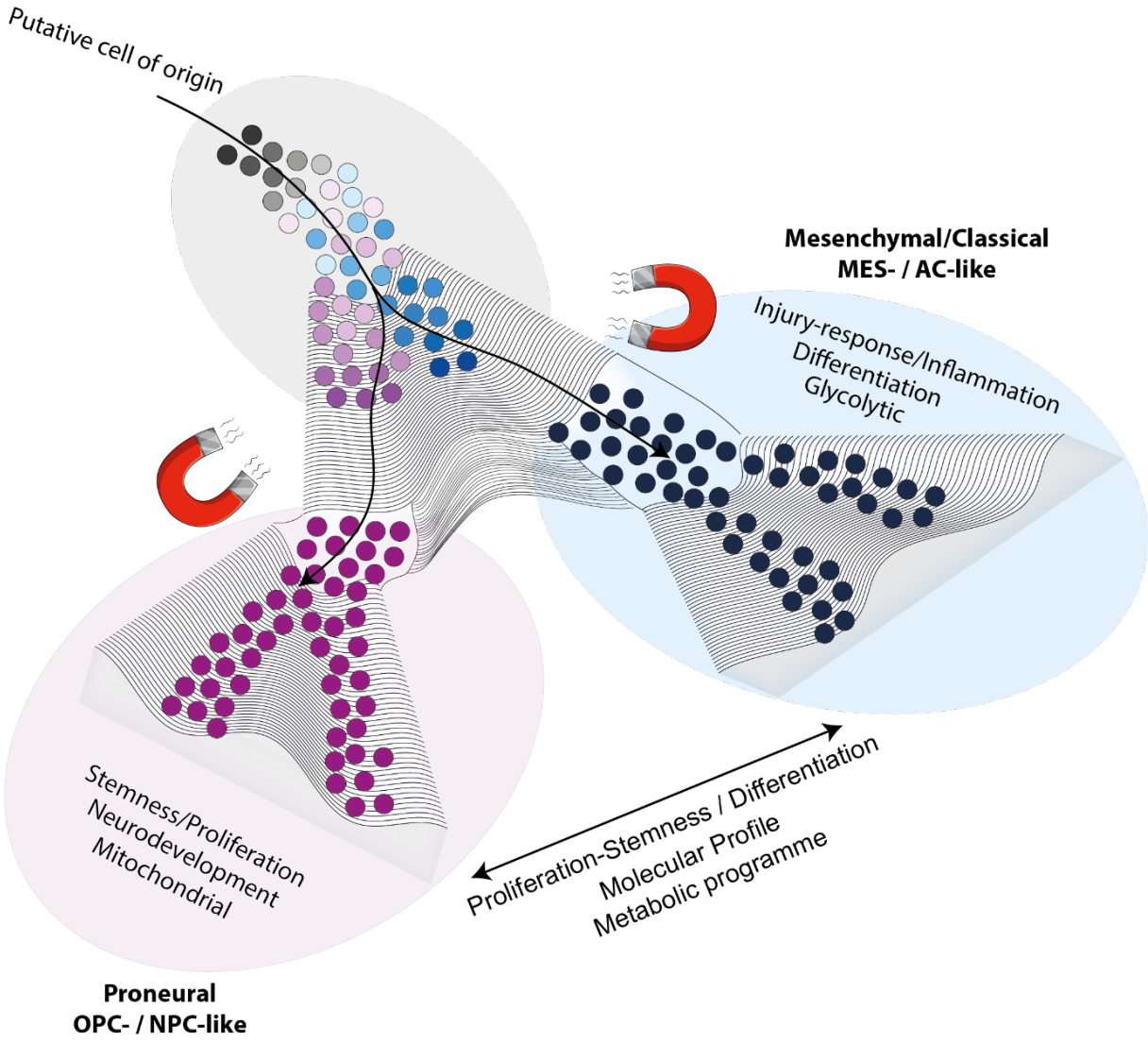


Figure 1 - Dynamic cell state transitions within a continuum of cellular states as drivers of phenotypic heterogeneity.

GBMs contain tumor cells with a variety of cellular states, with many cells displaying intermediate states, indicating phenotypic fate transitions. The transcriptional signatures aligned to the two principal axes of variation contrast the stemness and differentiation status, metabolic programmes, and molecular profiles, which are related with TCGA-subtypes or neurodevelopmental-like profiles from single tumor cells. Cell fate decisions can be influenced by various state-attractors, are not terminal and characterised by a dynamic and reversible nature.

However, fate commitment to one state is not an ultimate decision, and GBM cells can interconvert flexibly in response to various stimuli in the tumor microenvironment or external signaling, as well as re-establish stem-cell-like characteristics from more differentiated states (Dirkse et al., 2019).

1.2.2 *Mechanisms and clinical implications of phenotypic plasticity*

Apart from intratumoral heterogeneity on the genetic level, phenotypic plasticity of tumor cells is a key challenge in the development of effective treatment strategies, as indicated in the previous paragraphs. (Fig. 2). This concept is defined as the ability of tumor cells to switch between cellular states dynamically and reversibly, each of which is marked by a very characteristic profile of biochemical and biophysical properties (Cabrera et al., 2015; Gupta et al., 2018). A flexible transition along a continuum of more differentiated and slowly proliferating cell fates towards tumor-promoting stem-like states allows tumor cells to adapt and respond to various perturbations during tumor formation, progression and eventually develop therapeutic resistance.

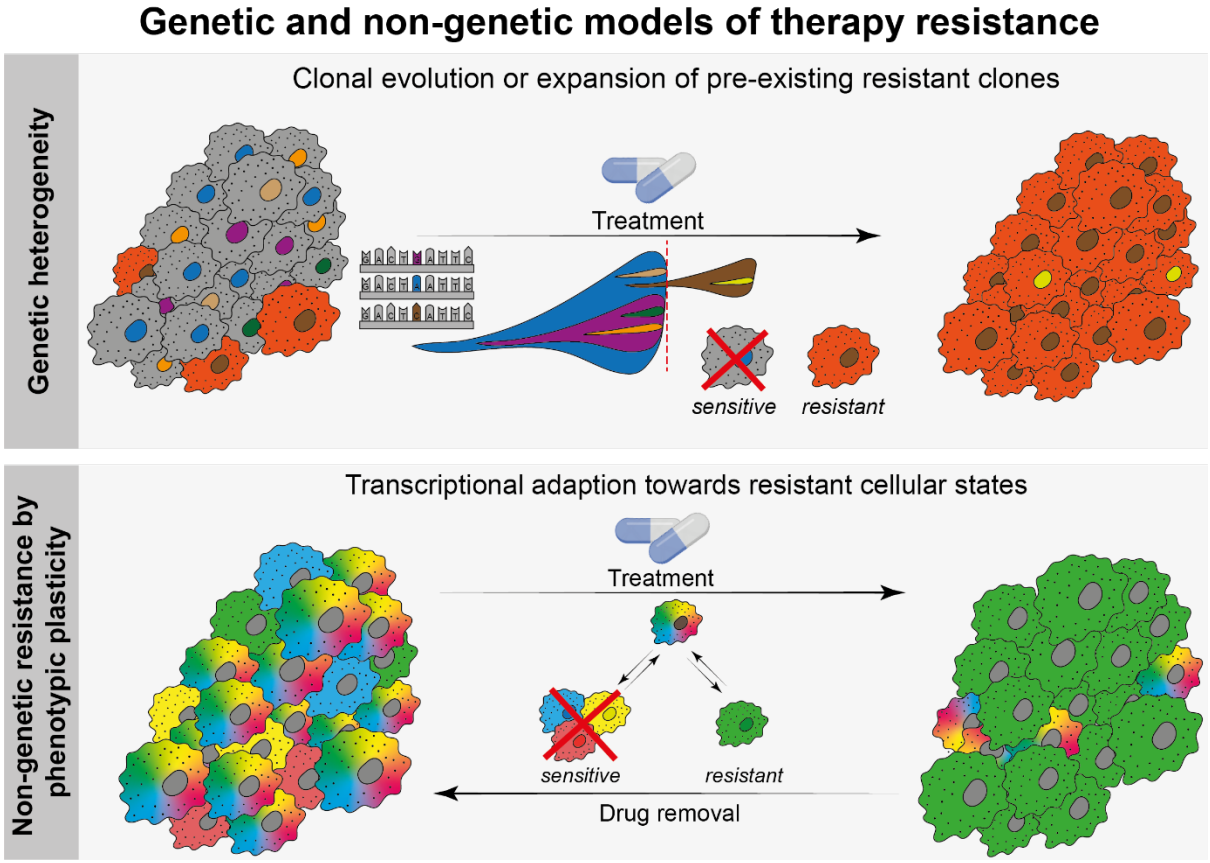


Figure 2 - Mechanisms of genetic and non-genetic therapy resistance.

Tumors are composed of cells with varying sensitivity to systemic or targeted therapies. While sensitive tumor cells are eradicated, treatment resistance in a subset of tumor cells can be mediated by multiple mechanisms. Genetic drug resistance includes either the selection of pre-existing resistant clones with distinct mutations (illustrated by differentially colored nuclei; top panel) and their subsequent clonal expansion or the evolution of novel resistance-confering mutations under selective pressure. Non-genetic mechanisms include an adaptive response of drug-tolerant cancer cells and their transcriptional remodelling towards resistant cellular states (bottom panel). This process of phenotypic plasticity likewise allows the re-establishment of initial cellular states after removal of the therapeutic intervention.

The mechanisms underlying cell state transitions are very complex and involve exogenous stimuli as well as cell-autonomous factors, and can likewise be triggered or amplified by therapeutic interventions (Fig. 3; Das et al., 2020; Thankamony et al., 2020; Marine et al., 2020). Glioma cell adaptability to distinct tumour microenvironments involves a non-restrictive transcriptional programme that enables them to transition between cell states. In this regard, epigenetic modifications are crucial to enable swift and, unlike genetic alterations, reversible cell state transitions depending on the activation or repression of particular genes. Chaligne et al. observed that targets of the histone modifying Polycomb complexes, which leave repressive chromatin marks, are characteristically hypomethylated in stem-like cells of IDH wildtype GBM (Chaligne et al., 2021). Similarly, Johnson et al. discovered significant levels of methylation in regulatory regions of genes involved in cell differentiation and stress response and identified marked changes in methylation patterns of glioma cells that were challenged with stressors such as hypoxia or ionizing radiation (Johnson et al., 2021).

One very well described example of a cell fate transition is the process of epithelial-mesenchymal-transition (EMT), where cancer cells lose epithelial properties, such as apico-basal polarization and cell-cell contacts, and acquire mesenchymal traits, which are linked to an increased migratory and invasive phenotype with enhanced resistance to therapy and apoptosis (Kalluri & Weinberg, 2010; Shibue & Weinberg, 2017; Dongre & Weinberg, 2019). An analogous case to EMT in GBM is the process of proneural-to-mesenchymal transition (PMT), which has been described upon treatment and tumor recurrence. This expression has been termed due to the fact that this sort of cell fate transition is typically observed in the proneural subtype of GBM, but more commonly is also referred to as mesenchymal transition or transdifferentiation. PMT and EMT share similarities with regards to their cellular properties, molecular marker expression and pathological features.

In IDH-wildtype GBM, the mesenchymal subtype has shown a tendency to worse overall patient survival and displayed more aggressive traits compared to other signatures (Phillips et al., 2006; Verhaak et al., 2010; Segerman et al., 2016; Wang et al., 2017). Patients with initial non-mesenchymal tumors tend to shift their dominant subtype towards mesenchymal states in recurrent tumors after first-line therapy (Wang et al., 2017; Varn et al., 2021). On the subject of how mesenchymal transition may be coordinated mechanistically, a number of studies have previously been published, indicating a complex interaction of cell-intrinsic and microenvironmental variables, which will be explored in the following section.

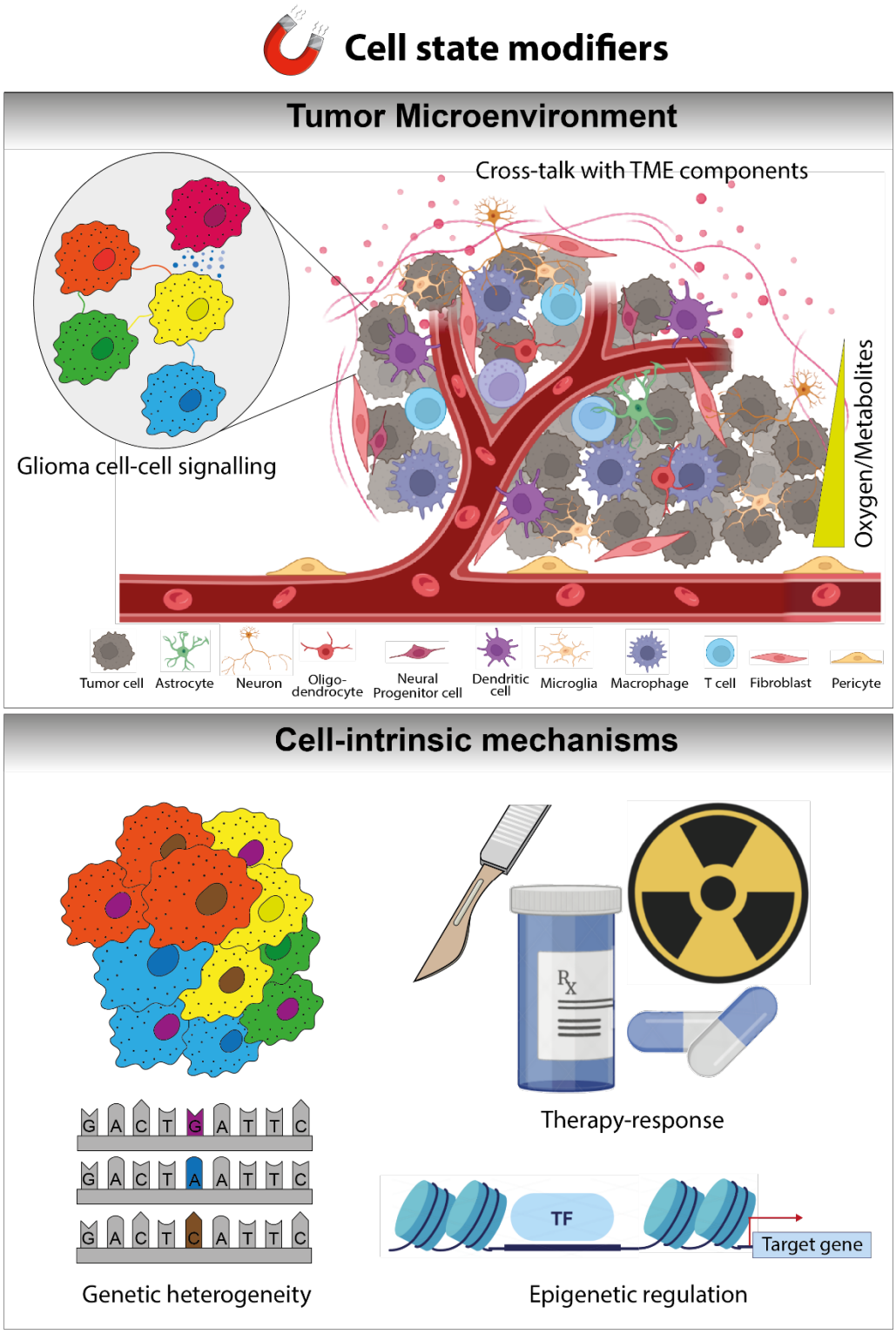


Figure 3 - Microenvironmental and intrinsic modifiers of cellular state plasticity.

Cellular plasticity in GBM is governed by a combination of external, as well as cell-intrinsic mechanisms. Extrinsic cues originate from the tumor microenvironment which consists of diverse tumor niches with varying availability of metabolites and oxygen, as well as neighbouring cells of the immune system or the healthy brain parenchyma that can engage in paracrine signaling with tumor cells. Cell-intrinsic factors affecting phenotypic heterogeneity and plasticity include the underlying genetic background of a cell, inherent core transcriptional networks that can be activated or amplified in response to stress, such as therapy, as well as the permissiveness of the epigenome as a driver of transcriptional remodelling.

1.2.3 Treatment-related cellular plasticity

With the current standard of care for newly diagnosed GBM patients, comprising maximal tumor resection followed by radiochemotherapy, the median survival only marginally improves (Stupp et al., 2005). Together with a high rate of recurring tumors and a 2-year survival of less than 20%, the treatment failures have been partly attributed to therapy-induced phenotypic plasticity and, in particular, a PMT. Following surgical excision and different therapies, a typical trajectory of tumor regrowth towards a mesenchymal phenotype has been described in multiple studies (Phillips et al., 2006; Wood et al., 2016; Wang et al., 2017; Varn et al., 2021). As part of the standard of care, radiotherapy was hypothesized to potentially promote mesenchymal transition. Upon exposure to radiation, upregulation of the mesenchymal marker genes CD44, CHI3L1, VIM and COL1A1 together with NFkB pathway activation as well as the induction of transcriptional master regulators of mesenchymal GBM, such as CEBPb and STAT3, have been described (Bhat et al., 2013; Halliday et al., 2014; Lau et al., 2015; Minata et al., 2019; Behnan et al., 2019). Importantly, Bhat et al. have showed that glioma-initiating cells underwent PMT in response to TNFa treatment and NFkB activation. In this study, as well as in others, mesenchymally committed tumor cells have proven more resistant to radiotherapy, which offers an explanation for the poor response of recurrent tumors (Bhat et al., 2013; Segerman et al., 2016).

Due to the fact that GBMs display a high degree of vascularization, tumors were suspected to show an improved response to antiangiogenic therapy. However, the use of Bevacizumab did not have an effect on overall-survival and has only marginally improved the progression-free survival of patients from 1.5 to 4.2 months (Friedman et al., 2009; Lai et al., 2011; Chinot et al., 2014; Wick et al., 2016; Li et al., 2017). In a study by Piao et al., anti-angiogenic treatment triggered the activation of genes implicated in a mesenchymal programme, invasion and migration, as well as stimulating the infiltration of immune cells, leaving room for speculation of PMT as the cause for treatment failure (Piao et al., 2012; Piao et al., 2013). Additionally, anti-angiogenic therapy has been found to increase glucose metabolism in tumor cells, resulting in increased lactic acid production, which is also a common feature of hypoxia-response (Fack et al., 2015; Harris, 2002).

Whether TMZ treatment, as the second arm of standard therapy, also plays a role in mesenchymal transition is not well investigated yet, although a TMZ-induced hypermutation phenotype has been described, which might trigger some of the cascades leading to mesenchymal differentiation (Johnson et al., 2014; Barthel et al., 2019). Taken together, these reports highlight that phenotypic plasticity and, in particular, mesenchymal transition, represents a way for tumor cells to overcome therapeutic pressure and a major hurdle in the development of new therapies.

1.2.4 The role of hypoxia on shaping tumor cell plasticity

Oxygen is one of the most essential components of cellular metabolism for most eukaryotes, and a lack of oxygen leads to a state that is called hypoxia. Due to the fast proliferating nature of solid tumors, cancer cells are often located far away from nutrient and oxygen-supplying blood vessels, which makes hypoxia a hallmark of many cancers (Harris, 2002). Especially in the brain, oxygenation can vary significantly depending on the region, and in brain tumors oxygen levels as low as 1.5% O2 have been described (Beppu et al., 2002).

GBM tumors are characterised by distinct anatomic components, such as tumor bulk, the infiltrative front, microvascular proliferation and pseudopalisading cells around necrosis (Hambardzumyan & Bergers, 2015; Puchalski et al., 2018). The latter two create hypoxic microenvironments, as necrotic foci are characterised by abscence of vascularization and hyper-proliferative blood vessels are usually functionally aberrant, which limits the efficiency of oxygen supply. Mechanistically, the cellular response to hypoxia is mediated by the stabilisation and nuclear translocation of the hypoxia inducible factor (HIF) transcription factors, which induce transcription of a large variety of downstream targets involved in angiogenesis, glycolosis, autophagy, and invasion (Jain et al., 2007; Monteiro et al., 2017; Chédeville & Madureira, 2021).

In GBM, HIF1a has been proposed to directly control mesenchymal transition through the regulation of ZEB1, consistent with its role in promoting EMT in a variety of other cancers (Jacobs et al., 2015; Tam et al., 2020). In a number of recent studies, transcriptional states of mesenchymal Glioblastoma have been repeatedly associated with expression signatures of hypoxia response (Neftel et al., 2019; Wang L et al., 2019; Couturier et al., 2020; Pine et al., 2020; Richards et al., 2021). signalingHowever, whether a lack of oxygen is ultimately the driver of mesenchymal transition or the independent activation of the HIF transcriptional network and simultaneous promotion of a mesenchymal signature may explain the connection remains to be experimentally tested.

1.2.5 Metabolic alterations related to phenotypic plasticity

Given the diverse set of tumor niches with varying metabolite availability and tumor cells' energy-demanding high proliferative index, a modulated metabolism may also be able to influence phenotypic plasticity and cellular state. A well-described phenomenon observed in many cancer cells is an increase in glycolytic activity at the expense of the tricarboxylic acid cycle in order to sustain the high energy consumption for tumor maintenance and growth, which is known as the Warburg effect. This process is characterized by increased glucose absorption and elevated aerobic

glycolysis, which converts glucose to pyruvate, ultimately resulting in increased lactate generation (Vander Heiden et al., 2009; Liberti et al., 2016).

In mesenchymal GBM, glycolytic activity has been reported to be greatly increased as compared to the proneural subtype, and mesenchymal signatures have been repetitively associated with necrotic tumor regions, which are typically hypoxic and prone to anaerobic glycolytic processes (Verhaak et al., 2010; Mao et al., 2013; Agnihotri & Zadeh, 2016; Yin et al., 2017; Varn et al., 2021). Increased levels of lactate in the extracellular space of tumors have been demonstrated to induce acute inflammation through NFkB pathway activation, trigger the infiltration of various immune cells and promote the polarization of macrophages into a pro-inflammatory M2 state, all of which are closely connected to mesenchymal GBM signatures (Shime et al., 2008; Végran et al., 2011; Rivera & Bergers, 2013; Colegio et al., 2014; Noy & Pollard, 2014).

Furthermore, a recent study attempted to characterise GBM using biological pathways instead of genetic or transcriptional subtyping. They found that tumor cells were aligning along two major axes: proliferation/neurodevelopment and cellular metabolism. The metabolic category was subdivided into a mitochondrial (MTC) and a glycolytic/plurometabolic (GPM) component, the latter of which demonstrated substantial changes in glycolysis, carbohydrate and lipid metabolism and was strongly related to a mesenchymal signature (Garofano et al., 2021).

Although these findings suggest that certain metabolic characteristics may influence cellular plasticity and, in particular, the mesenchymal signature, the specific molecular mechanisms and the question of whether altered metabolism is the cause or an effect of mesenchymal GBM remains to be elucidated.

1.2.6 Impact of the tumor microenvironment on GBM plasticity

The GBM tumor microenvironment (TME) is composed of diverse niches with distinct cell types, histological features, and biological processes that undergo dynamic changes throughout tumor initiation and progression (Hambardzumyan & Bergers, 2015). Tumor cells have evolved the ability to adapt to the particular conditions found in each TME niche, and direct interactions between tumor cells and non-malignant cells are believed to be critical in sustaining cellular plasticity (Prager et al., 2020; Mitchell et al., 2021; Venkatesh et al., 2019).

The TME is inhabited by a large diversity of cell types in the healthy human brain, including neurons, astrocytes, fibroblasts, pericytes, endothelial cells and immune cells (Charles et al., 2012; Quail & Joyce, 2013; Tomaszewski et al., 2019). Tumor associated macrophages and microglia

(TAM) make up the highest fraction of immune cells and can reach up to 50% of the total tumor mass (Hambardzumyan et al., 2016; Quail & Joyce, 2017; Darmanis et al., 2017).

Numerous studies have drawn a relationship between an enrichment of pro-inflammatory and immunological signatures and the mesenchymal subtype. In comparison to proneural and classical subtypes, mesenchymal tumors had the lowest purity score, suggesting the presence of infiltrating non-neoplastic cells in this subtype (Wang et al., 2017). This drop in purity was shown to be related to NF1 deletion, a critical regulator of the mesenchymal subtype of GBM (Marques et al., 2021). Higher resolution single-cell studies have confirmed these results and shown that the mesenchymal signature was in fact amplified in tumor cells, as opposed to being an artifact of the large number of infiltrating immune cells in bulk-level profiling (Neftel et al., 2019; Wang LB et al., 2021).

TAMs are known to be capable of releasing a range of pro- and anti-inflammatory factors, as well as pro-angiogenic and extracellular matrix remodelling cytokines (Bhat et al., 2013; Engler et al., 2013; Szulzewsky et al., 2016; Poon et al., 2017; Sa et al., 2020). In the tumor microenvironment, TAMs align along a variety of cellular states associated with decreased expression of homeostatic microglia core signature genes and enhanced signatures related to inflammation, metabolism and hypoxia. In turn, GBM cells release a broad spectrum of cytokines and chemokines, which have the ability to attract and activate TAMs (Broekman et al., 2018; Sankowski et al., 2019). Besides TAMs, infiltration of all types of T cells (CD8+ cytotoxic, regulatory, CD4+ helper) is also strongly associated with mesenchymal tumors, when compared against proneural and classical (Rutledge et al., 2013; Kaffes et al., 2019; Wang LB, et al., 2021).

Besides immune cells, other non-neoplastic cell types also engage in interactions with GBM cells (Wang X et al., 2018; Pavlyukov et al., 2018). Recent evidence revealed that physical interactions between tumor cells and neurons form a synapse-like connection, that plays an important role in their interplay through molecular and electrochemical communication that is able to promote tumor growth and invasion (Venkatesh et al., 2019; Venkataramani et al., 2019). Apart from paracrine signaling, direct interactions of GBM cells with one another via exosomes or long tumor microtubes have been shown to increase stemness and resistance to radiochemotherapy (Osswald et al., 2015; Weil et al., 2017; Xie et al., 2021).

In summary, this raises the possibility that bidirectional communication exists between the tumor microenvironment and GBM cells, and that this communication may establish an environment that promotes a mesenchymal signature in GBM cells. However, so far this represents correlative evidence and a direct causal relationship remined to be experimentally proven, alongside the mechanistical details and whether these interactions entail both direct cell-to-cell contact and paracrine pathways.

1.3 Lineage tracing and fate mapping

Lineage tracing and fate mapping are two similar but distinct techniques for determining the identity or behavior of cells and their progeny. Fate mapping tries to reconstruct and describe the developmental origin of adult tissue during the maturation process of embryonal tissue at a certain developmental stage. In contrast, lineage tracing focuses on identifying the relationship between a given cell and its progeny at each division in order to construct cellular hierarchies. Although derived from fate mapping, the objective of lineage tracing is to reconstruct lineage hierarchy down to the level of a single cell, therefore compromising spatial resolution (VanHorn & Morris, 2021).

Lineage tracing has proven useful to advance insights into Glioblastoma biology in the past, such as the search for the cell-of-origin and the dynamics of tumor maintenance and relapse after therapy. For instance, Liu et al. labelled neural stem cells (NSC) with a mosaic fluorescent reporter and traced the derived cell lineages (neurons, astrocytes, OPCs, and oligodendrocytes) throughout tumorigenesis in glioma mouse models. Using this lineage tracing method, they were able to demonstrate that the resulting GFP-positive tumors were predominantly composed of cells of the OPC lineage, with no proliferation of GFP-positive NSC, astrocytes, or neurons (Liu et al., 2011). In a similar mouse model, the lab of Luis Parada employed lineage tracing of GFP-tagged NSCs to discover glioma cells with a quiescent phenotype that are less proliferative than the bulk of tumur cells and capable of rebuilding the tumor following chemotherapy-induced growth arrest (Chen et al., 2012).

Instead of utilising fluorescent reporters, cellular barcoding was employed to monitor the cell fate of subpopulations of human Glioblastoma cells in orthotopic transplantation-based mice models. This led to the discovery of a hierarchical model of slow-cycling stem cells which give rise to a fast-proliferating progenitor population that is in turn generating non-replicating cells. Furthermore, the authors could identify a rare subset of glioma cells that appeared to be pre-resistant to chemotherapy and able to re-expand the tumor mass after therapy (Lan et al., 2017).

Despite the enormous progress made by lineage tracing approaches in understanding Glioblastoma tumor formation, maintenance, and response to therapy, it remains difficult to interpret observations and distinguish causality from correlation in order to construct a reliable model for clinical improvement (Gimple et al., 2022).

1.4 Specific Aims

In most solid tumors, intratumoral genetic heterogeneity is widely believed to contribute to therapy resistance and tumor recurrence through clonal selection of already resistant clones, although non-genetic resistance mechanisms are also starting to gain increasing attention (Marine et al., 2020; Bell et al., 2020). The development of targeted therapies together with advanced longitudinal profiling has further underlined this fact (McGranahan et al., 2017). Although for GBM after the standard of care, there is undoubtedly an emergence of new mutations, and in extreme cases, a hypermutated phenotype, evidence of clonal selection is rarely observed. In fact, initial tumor-driving mutations of early tumor formation seem to remain, and later occurring mutations add on top of these largely in a spatially enclosed and stochastic manner (Johnson et al., 2013; Kim et al., 2015; Wang J et al., 2016; Barthel et al., 2019; Touat et al., 2020; Schaettler et al., 2022). This suggests that intratumoral heterogeneity on the genetic level alone and clonal selection of therapy-resistant tumor cells do not represent the primary factors of therapy failure in the majority of patients.

Rather than that, phenotypic plasticity with multi-factorial variables operating individually on a specific tumor cell and in conjunction with the whole tumor and its microenvironment offers several pathways to overcome resistance to all present treatment approaches. Some of these processes are of paracrine nature, with individual glioma cells secreting soluble factors to neighbouring tumor cells or the TME and in turn receiving feedback in the form of cytokines and chemokines. In interplay with other cell-intrinsic mechanisms, they offer glioma cells the possibility of dynamic cell state transitions during tumor initiation, progression, or under therapeutic pressure. Each cellular state manifests on a phenotypic level with different properties of differentiation/stemness status, proliferative capacity and activated metabolic pathways and on the molecular level with the acquisition of distinct transcriptional signatures (Fig. 4; Azam et al., 2020; Nicholson & Fine, 2021; Yabo et al., 2021; Kim et al., 2021; Wang Z et al., 2021).

However, the detailed molecular mechanisms and causal relationships revolving around phenotypic plasticity and their implications for clinical management of GBM are still the subject of ongoing research.

Layers of heterogeneity in Glioblastoma

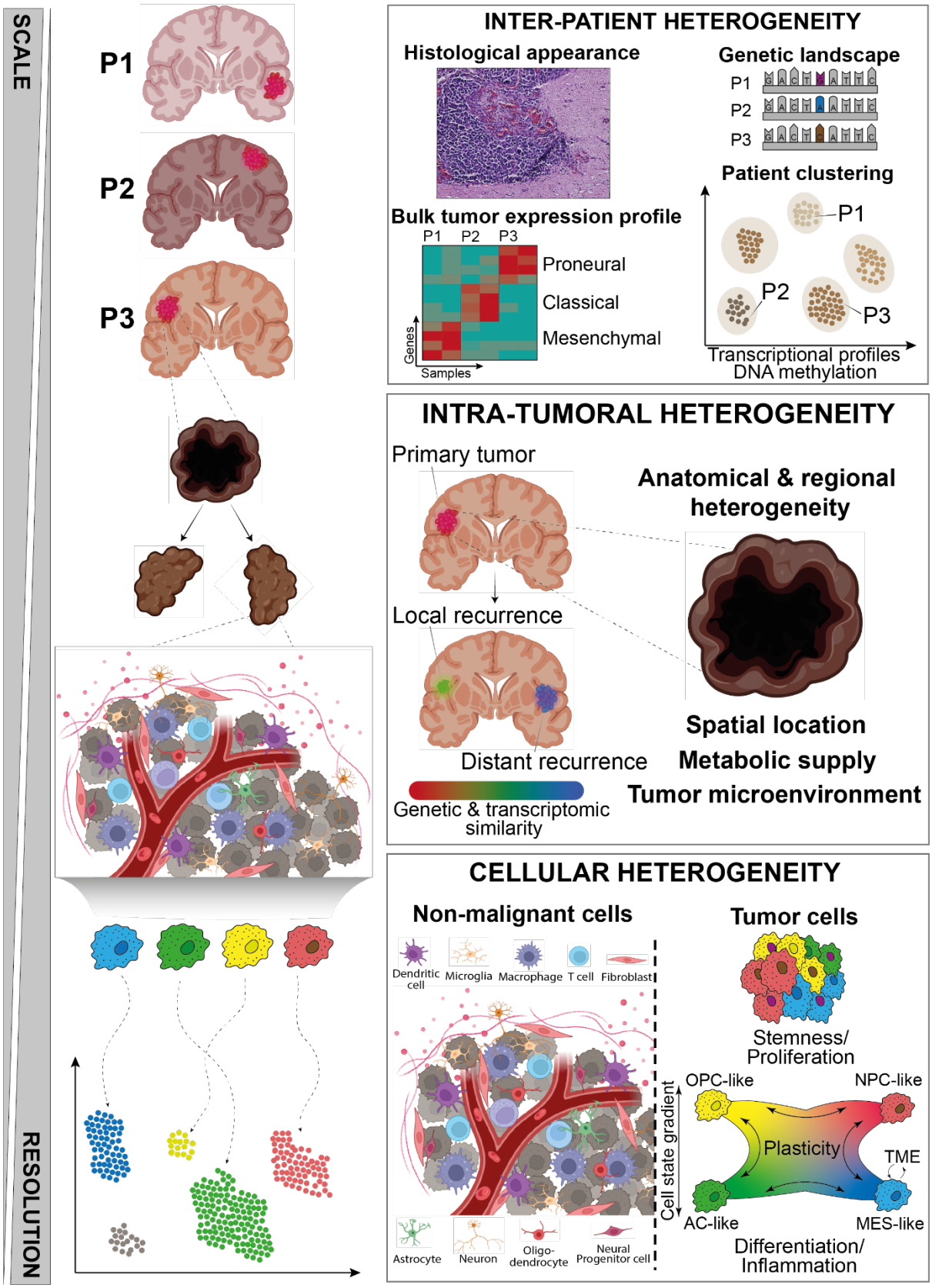


Figure 4 - The layers of heterogeneity in GBM.

GBM tumors display a high degree of heterogeneity, which manifests at multiple levels of biological resolution. On inter-patient scale, tumors present with different histological features and bulk-tissue molecular profiles, allowing clustering of patients into subtypes. Within an invidiual patient, the spatio-temporal distribution of primary and recurrent tumors and particular tumor niches are highly heterogeneous and influential on tumor biology. At the highest resolution on single-cell level, cellular heterogeneity is characterised by the cellular composition of the tumor microenvironment, as well as cellular states of individual tumor cells.

Directly connecting to this loose end, this thesis presents a novel approach to generating synthetic genetic tracing reporters of GBM cellular states to further broaden the understanding and consequences of tumor heterogeneity in GBM and has the following specific aims:

<u>Aim 1:</u> The main objective of this work is to develop and validate genetic tracing vectors (sLCR) for Glioblastoma subtypes that will be used to study cell fate transitions.

<u>Aim 2:</u> Exploration of the molecular and cellular mechanisms underlying Glioblastoma heterogeneity through the use of sLCRs.

<u>Aim 3:</u> Investigate the mechanisms underlying treatment response and resistance using sLCRs.

2 Results

2.1 Generation of Glioblastoma-subtype synthetic Locus Control Regions (sLCRs)

Although branded as a genetic disease, epigenetic and transcriptional dysregulation are important hallmarks of cancer, which are partly influenced and super-imposed by genomic mutations and external factors (Hanahan et al., 2011; Bradner et al., 2017; Sengupta et al., 2017). This complexity is especially pronounced in Glioblastoma, where overarching discrete gene expression clusters, which might delineate malignant cells of different origin or clonal evolution and could be considered "cell types", are interfused with further diversity of tumor cells adopting a continuum of cellular states. Taken together, these mechanisms fuel critical disease characteristics, such as tumor progression, metastasis and therapy resistance (Dagogo-Jack et al., 2017; Suvà & Tirosh, 2019; Marusyk et al., 2020).

To trace Glioblastoma subtype identities and get a better understanding of the underlying molecular mechanisms, we developed synthetic genetic tracing vectors, which allow identification of discrete cellular states and monitoring of cell-fate transitions (Fig. 5A). The generation of synthetic locus control regions (sLCRs) underlies the assumption that cell fate determination is regulated by transcription factors that bind to cis-regulatory elements (CREs) to activate tissue- or cell state-specific gene expression signatures within a target-phenotype. Consequently, these vectors aim to transform the entire regulatory network of CREs within a target-phenotype into a synthetic stretch of DNA, which is driving the expression of a fluorescent reporter protein.

From patient-derived TCGA datasets, we extracted for each Glioblastoma subtype (Mesenchymal [MES], Proneural [PN] and Classical [CL]; Verhaak et al., 2010; Fig. 5B and Table 2) a list of differentially regulated genes (DRGs) for genes that are highly associated with each subtype (signature genes) and their regulators (genes coding for transcription factors). To compile a list with genomic coordinates of the CREs that represent the regulatory network for each Glioblastoma subtype, all known transcription factor binding sites (TFBS) around the genomic loci of all signature genes were identified. As exemplified for the case of the MES-GBM targeting MGT#1-sLCR, the output of this analysis is a list of genomic coordinates with CREs and their associated binding probabilities for the target phenotype-specific TFs (Ext. Fig. 5A). In each case, certain CREs contained clusters of varying numbers of TFBS that appeared together (Fig. 5C + Ext. Fig. 5B). Due to experimental constraints for the length of an sLCR and consequently the number of CREs to be included, a selection needs to be made according to the following criteria: covering a high transcription factor binding sites number (i) and diversity (ii), as well as a known

distance from the nearest endogenous transcriptional start site (iii). With some TFBS often clustering together at certain CREs (Ext. Fig. 5B), such regulatory elements need to be selected in order to maximize the output for criteria (i) and (ii) (Fig. 5D). Additionally, to meet criteria (iii), the last CRE to be selected should be close to a natural transcriptional start site in order to facilitate transcription of the downstream reporter.

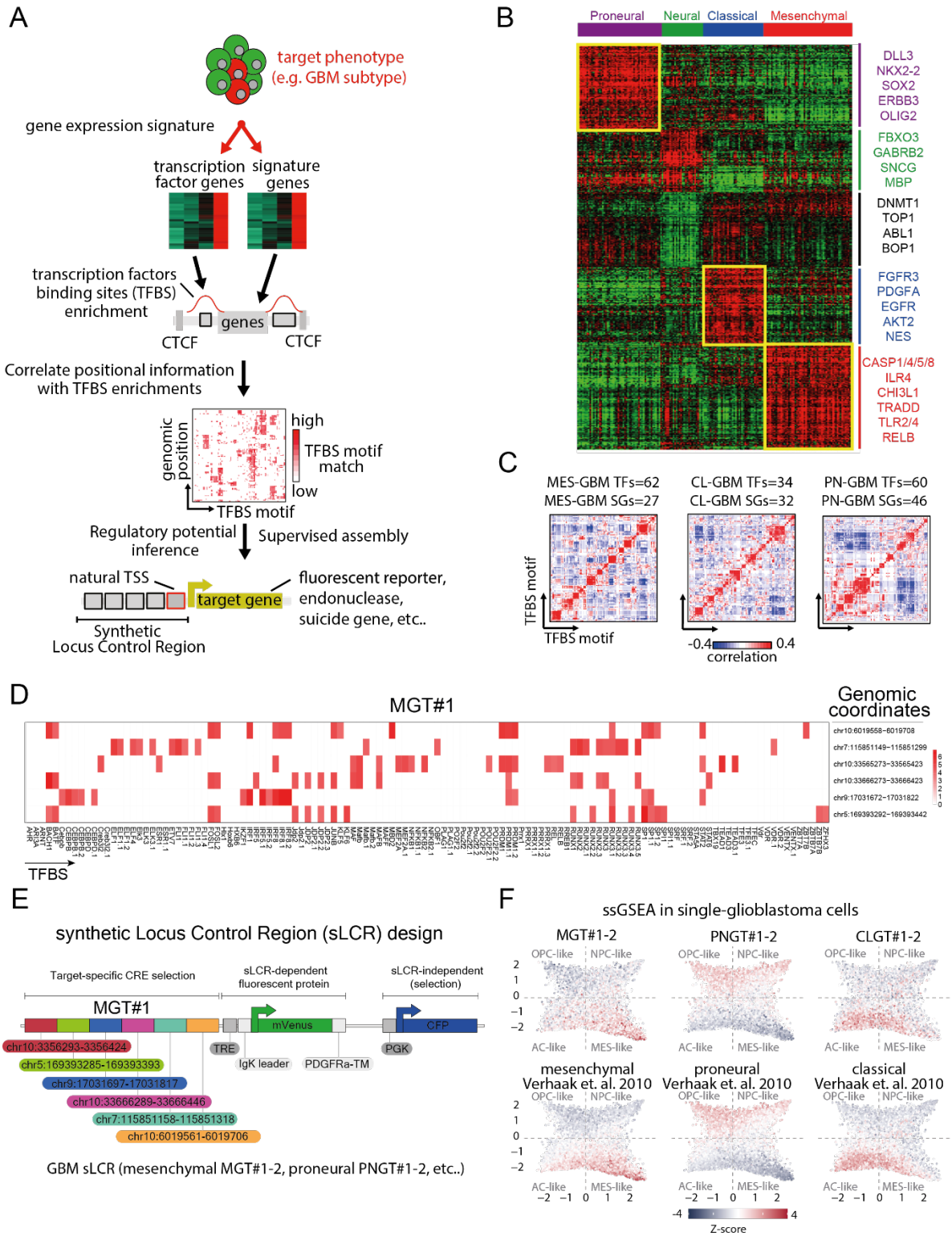


Figure 5 - Generation of Glioblastoma-subtype synthetic Locus Control Regions (sLCRs).

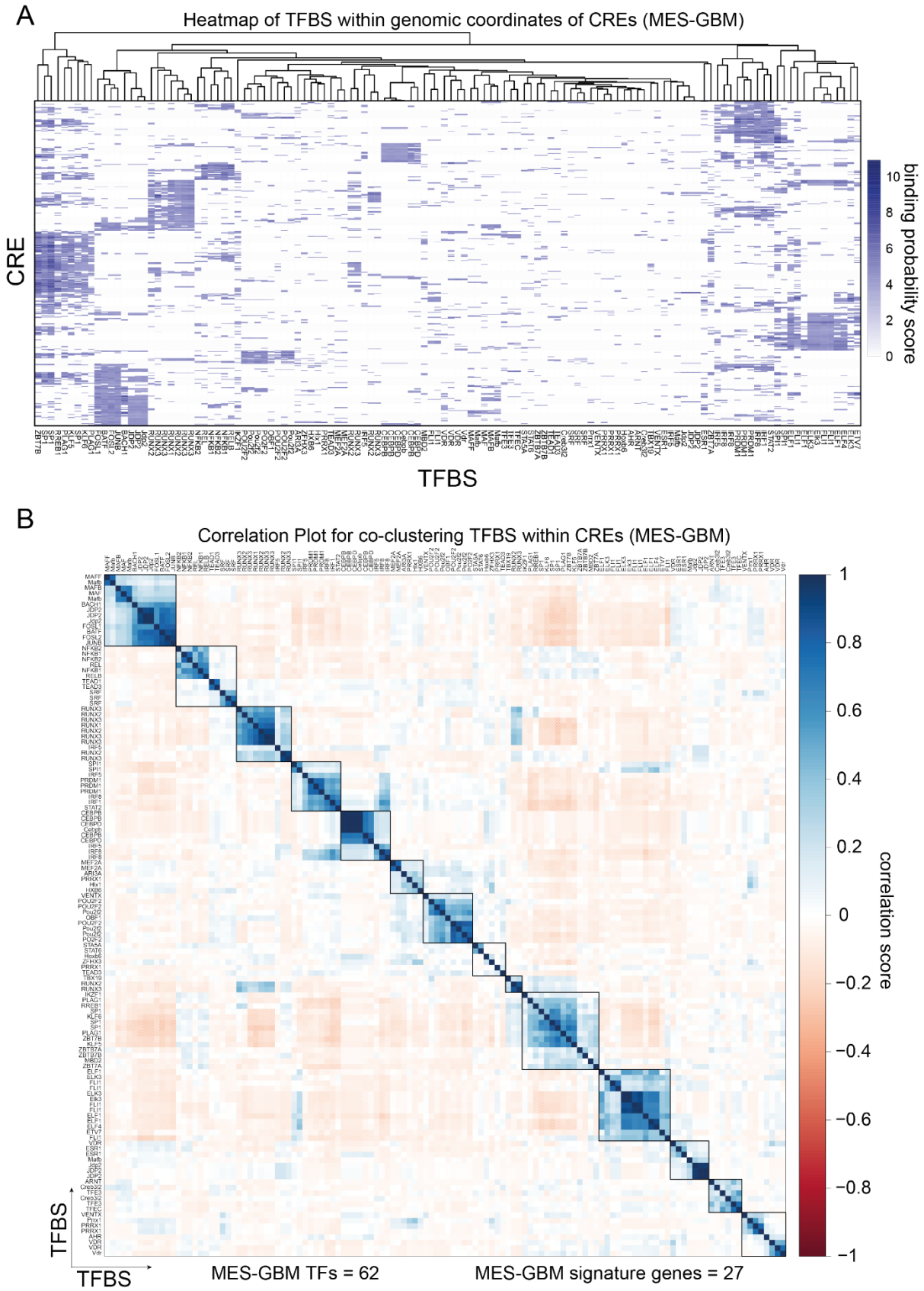
(**A**) Schematic illustration of the development of Glioblastoma sLCRs using gene expression data. (**B**) TCGA Glioblastoma molecular subtype clusters derived from patients (adapted from Verhaak et al.., 2010). (**C**) Heatmaps of pairwise correlations between strong TFBS motifs at Glioblastoma subtype-specific loci. Above each panel, the number of transcription factors and signature genes employed in the analysis is shown. MES=Mesenchymal; PN=Proneural; CL=Classical. (**D**) Heatmap of chosen CREs used to create the MGT#1-sLCR and associated TF binding probability. (**E**) Diagram of the MGT#1-sLCR with CREs, a sLCR-dependent fluorescent reporter, and a sLCR-independent selection fluorophore. (**F**) ssGSEA normalized scores for input genes for the corresponding sLCRs are shown above (Methods). In each quadrant, the cell states identified by (Neftel et al., 2019) are indicated, and the original single-cell position is kept in the two-dimensional representation (Methods). For a head-to-head comparison, TCGA subtypes (Verhaak et al., 2010) are shown below.

In accordance with these parameters, we aimed to generate sLCRs for genetic tracing of mesenchymal, classical and proneural Glioblastoma subtypes, which will hereafter be referred to as MGT, CLGT, and PNGT. This was accomplished by combining 5-6 of the identified cis-regulatory elements (representing 40-60% of the regulatory potential) into a lentiviral vector where the sLCR is driving a fluorescent protein alongside with a second sLCR-independent nuclear H2B-CFP fusion protein driven by a ubiquitously active PGK promoter to facilitate sorting of transduced cells (e.g. MGT#1; Fig. 5E).

To test whether the sLCRs we have generated are specifically labelling their targeted Glioblastoma subtype identity, we performed single-sample gene set enrichment analysis (ssGSEA) on a patient-derived Glioblastoma single-cell dataset. Consistently, ssGSEA showed a high enrichment score for the input genes of each individual reporter, corresponding to known cell states of freshly purified single Glioblastoma cells (Neftel et al., 2019), and their related TCGA subtype (Fig. 5F).

Extended Figure 5 - Generation of Glioblastoma-subtype synthetic Locus Control Regions (sLCRs).

(**A**) Heatmap of the top 500 selected CREs (rows) for Glioblastoma mesenchymal subtype with binding probability for TFBS (columns). (**B**) Heatmap of pairwise correlations between strong TFBS motifs at Glioblastoma mesenchymal subtype-specific loci. Number of MES-GBM transcription factors = 62. Number of MES-GBM signature genes = 27.


A
Heatmap of TFBS within genomic coordinates of CREs (MES-GBM)
CRE
TFBS
binding probability score
10
8
6
4
2
0
B
Correlation Plot for co-clustering TFBS within CREs (MES-GBM)
TFBS
TFBS
MES-GBM TFs = 62
MES-GBM signature genes = 27
correlation score
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1

2.2 sLCRs identify their target phenotype in vitro in human glioma-initiating cells

To test the functionality and relative expression *in vitro*, we next transduced proneural and mesenchymal sLCRs (synthetic reporters representing two opposite Glioblastoma subtypes) into a near-isogenic pair of human glioma-initiating cells (hGICs). These cells were derived from spontaneously immortalized human neural progenitor cells (Stock et al., 2012) and genetically manipulated with commonly identified mutations of patients in the TCGA-GBM cohort (Fig. 6A) and have a proneural-like expression signature, possibly inherited from the cell of origin. In addition to PTEN-knockdown and shared spontaneous aberrations, the near-isogenic pair background consists of TP53 and NF1 depletion (IDH-WT hGICs) or IDH1R132H and TP53R273H mutants' over-expression (IDH-mut hGICs; Fig. 6A). Both lines are capable of inducing high grade tumor formation upon intracranial orthotopic injection into the brains of immuno-compromised mice, with full penetrance and a median survival of 36 days (IDH-WT hGICs or 26 days (IDH-mut hGICs; Fig. 6B).

When lentivirally-transduced and selected with the mesenchymal reporter MGT#1 and the proneural reporter PNGT#2, live-cell imaging and FACS quantification showed that PNGT#2 is more expressed than MGT#1 in both hGIC lines, with MGT#1-expression being relatively lower in the IDH-mutant cells (Fig. 6C, top and Fig. 6D). We aimed to further test the specificity of our mesenchymal and proneural reporters by extending the panel of cell lines to patient-derived glioma stem cells (GSCs) and cancer cell lines of non-brain-tumor cell lines with epithelial (lung and breast cancer) and non-epithelial (leukemia) origin. After transduction and FACS selection of each cell line with MGT#1 and PNGT#2, we quantified the expression of each reporter using RT-qPCR and normalised through endogenous GAPDH and the number of reporter integrations into genomic DNA (see methods).

In accordance with the intended design, both sLCRs were well expressed in the patient-derived GSCs and displayed low expression in "off-target" cell lines, such as leukemia cell lines (Fig. 6C, bottom). Interestingly, the mesenchymal reporter MGT#1 showed high expression levels in cells with a well-established mesenchymal background (GBM166 and MDA-231) as opposed to cells of epithelial identity (H2122 and MCF-7), independently of the tissue of origin and was able to trace epithelial-mesenchymal-transition (EMT) through TGFB-signaling in A549 epithelial lung-cancer cells (A549 vs. A549+TGF-β; Fig. 6C, bottom and Ext. Fig. 6).

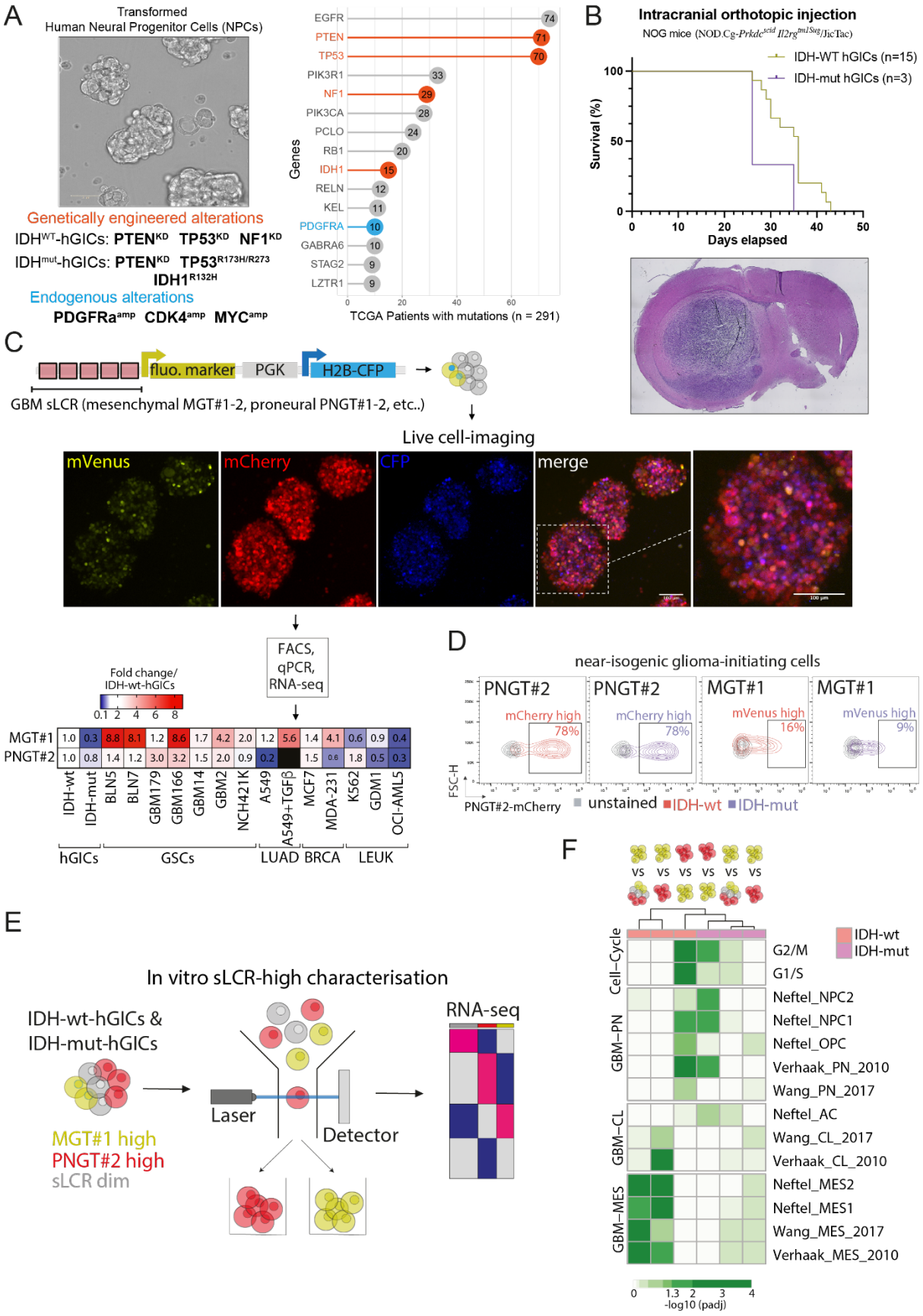

A
Transformed
Human Neural Progenitor Cells (NPCs)
Genetically engineered alterations
IDH^WT-hGICs: PTEN^KD TP53^KD NF1^KD
IDH^mut-hGICs: PTEN^KD TP53^R173H/R273 IDH1^R132H
Endogenous alterations
PDGFRa^amp CDK4^amp MYC^amp
EGFR 74
PTEN 71
TP53 70
PIK3R1 33
NF1 29
PIK3CA 28
PCLO 24
RB1 20
IDH1 15
RELN 12
KEL 11
PDGFRA 10
GABRA6 10
STAG2 9
LZTR1 9
Genes
TCGA Patients with mutations (n = 291)
B
Intracranial orthotopic injection
NOG mice (NOD.Cg-Prkdc^scid Il2rg^tm1Sug/JicTac)
IDH-WT hGICs (n=15)
IDH-mut hGICs (n=3)
Survival (%)
Days elapsed
C
fluo. marker
PGK
H2B-CFP
GBM sLCR (mesenchymal MGT#1-2, proneural PNGT#1-2, etc..)
Live cell-imaging
mVenus
mCherry
CFP
merge
FACS, qPCR, RNA-seq
Fold change/ IDH-wt-hGICs
0.1 2 4 6 8
MGT#1 1.0 0.3 8.8 8.1 1.2 8.6 1.7 4.2 2.0 1.2 5.6 1.4 4.1 0.6 0.9 0.4
PNGT#2 1.0 0.8 1.4 1.2 3.0 3.2 1.5 2.0 0.9 0.2 1.5 0.6 1.8 0.5 0.3
IDH-wt
IDH-mut
BLN5
BLN7
GBM179
GBM166
GBM14
GBM2
NCH421K
A549
A549+TGFβ
MCF7
MDA-231
K562
GDM1
OCI-AML5
hGICs
GSCs
LUAD
BRCA
LEUK
D
near-isogenic glioma-initiating cells
PNGT#2
mCherry high 78%
PNGT#2
mCherry high 78%
MGT#1
mVenus high 16%
MGT#1
mVenus high 9%
FSC-H
PNGT#2-mCherry
unstained
IDH-wt
IDH-mut
E
In vitro sLCR-high characterisation
IDH-wt-hGICs &
IDH-mut-hGICs
MGT#1 high
PNGT#2 high
sLCR dim
Laser
Detector
RNA-seq
F
VS VS VS VS VS VS
IDH-wt
IDH-mut
Cell–Cycle
G2/M
G1/S
GBM–PN
Neftel_NPC2
Neftel_NPC1
Neftel_OPC
Verhaak_PN_2010
Wang_PN_2017
GBM–CL
Neftel_AC
Wang_CL_2017
Verhaak_CL_2010
GBM–MES
Neftel_MES2
Neftel_MES1
Wang_MES_2017
Verhaak_MES_2010
0 1.3 2 3 4
-log10 (padj)

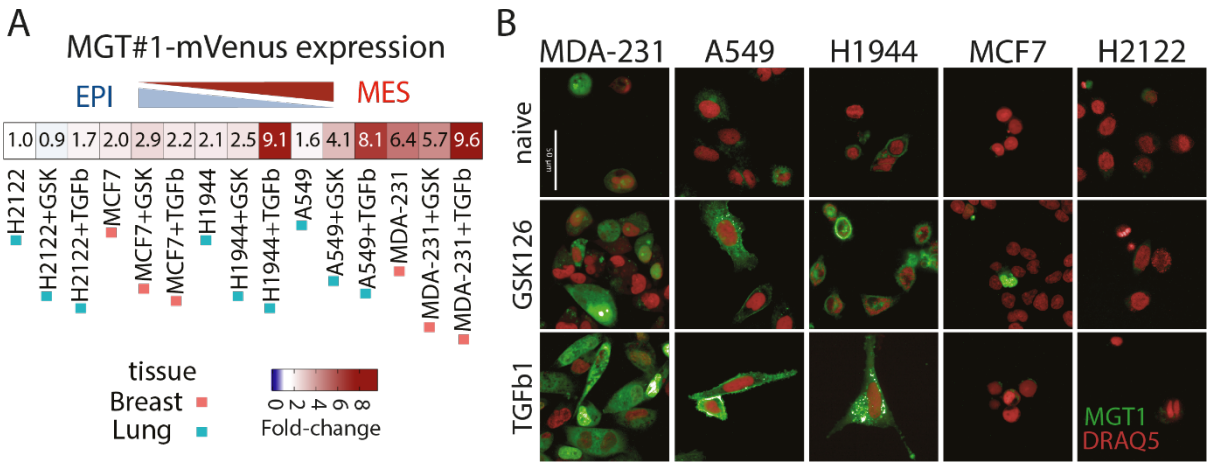
Figure 6 - sLCRs are capable of identifying their target phenotype *in vitro* in human glioma-initiating cells.

(**A**) Left, genetically engineered alterations and brightfield image of the human glioma-initiating cell (hGIC) model used in this study. Right, a summary of the most commonly registered aberrations discovered in the TCGA Glioblastoma patient cohort. (**B**) Survival analysis and H&E staining of tumors in NOG mice following orthotopic intracranial injection of hGICs. (**C**) Schematic illustration of an sLCR and fluorescence live-cell imaging of reporter expression is shown above. The heatmap below shows the expression of MGT#1 and PNGT#2 genes normalized by GAPDH and the number of integrations relative to hGICs. hGICs=human glioma initiating cells; GSCs=glioma stem cells. Additionally, selected non-brain tumor cell lines are included. (**D**) FACS profile of IDH-wt-hGICs and IDH-mut-hGICs transduced with the indicated reporters, followed by FACS sorting for the reporter-independent marker H2B-CFP. (**E**) Schematic illustration of the transcriptional profiling of bulk, MGT#1- and PNGT#2-expressing hGICs. (**F**) Heatmap representing GSEA-adjusted p-values (see Methods) for the specified Glioblastoma subtypes/state-signatures and comparisons in the corresponding hGIC lines.

As our hGICs are cultured as neurospheres under serum-free stem cell conditions, the ambivalence between differentiation and self-renewal, together with individual single-cell decisions and intrinsic stochastic fluctuations can drive some degree of heterogeneity among the population (Höving et al., 2021). Such heterogeneity may be reflected by differences in transcriptional regulation and cellular states of individual cells and therefore we intended to test whether subsets of hGICs with higher expression of either sLCR would be more homogeneous in terms of gene expression profiles than the population average. FACS-purified IDH-mut-hGICs and IDH-WT-hGICs based on high expression of either reporter could be distinguished by RNA-seq and principal component analysis (Fig. 6E). From the resulting profiles, we next performed gene-set enrichment analysis (GSEA) for mesenchymal and proneural gene-sets from patient-derived bulk and single-cell data. In the context of IDH-WT-hGICs, high expression of MGT#1 enriched for mesenchymal gene sets in contrast to PNGT#2-high as well as bulk unsorted cells, which were enriched for proneural and cell cycle-related gene sets (Fig. 6F).

Interestingly, mesenchymal Glioblastoma gene sets were more enriched in IDH-wild-type than IDH-mutant cells, in line with observations from Fig. 6C-D. Taken together, the data suggest that sLCR are reliably indicative of different cellular states, as high expression *in vitro* corresponds to an enrichment of expression profiles for their respective target phenotypes.

In summary, we have developed a method that can transform cis-regulatory elements of transcriptomic signatures from biologically- or clinically-relevant phenotypes into synthetic locus control regions to allow genetic tracing of cellular identities and cell fate transitions.



Extended Figure 6 - Mesenchymal-sLCR MGT#1 provides tissue-independent information on mesenchymal identity.

(**A**) Heatmap depicting the expression of the MGT#1 reporter in lung and breast cancer cell lines treated as indicated. GAPDH expression was used to normalize gene expression, and the number of integrations was calculated relative to the H2122 cell line (see Methods). (**B**) Live cell imaging of mVenus (EMT reporter) and DRAQ5 (nuclear dye) in the indicated cell lines and conditions.

2.3 Synthetic genetic tracing in vivo reveals Glioblastoma heterogeneity

It was shown using cellular barcoding and single-cell RNA-sequencing that in patient-derived xenograft (PDX) models, Glioblastoma cells with a certain subtype identity may create tumors with malignant cells recapitulating the whole panel of cellular states (Neftel et al., 2019). This study reinforces the view that tumor cells in the mouse brain microenvironment are capable of dynamically transitioning between cellular states and highlights the exceptionally plastic nature of GBM cellular states.

To test whether sLCRs would be suitable for genetic tracing of tumor cell fate transitions *in vivo*, we intracranially transplanted IDH-wt-hGICs with MGT#1 into immunodeficient mice (n=10). All animals developed high grade tumors, with large areas of the mouse brain being infiltrated (Fig. 7A). Counterstaining of the tissue for MGT#1 revealed that the reporter was primarily expressed in areas inside of the tumor, in many cases in particular at the invasive front to the healthy mouse brain (Fig. 7A-B). Although the majority of tumor cells did not show a signal for MGT#1, these cells still had visible H2B-CFP puncta, facilitated through mitotic chromatin condensation in dividing cells, ruling out the possibility that these cells did not harbour a functional integration cassette of the reporter (Fig. 7B-D). Consequently, this suggests that MGT#1 expression informs about functional intratumoral heterogeneity *in vivo* with a subset of tumor cells acquiring a pronounced mesenchymal identity.

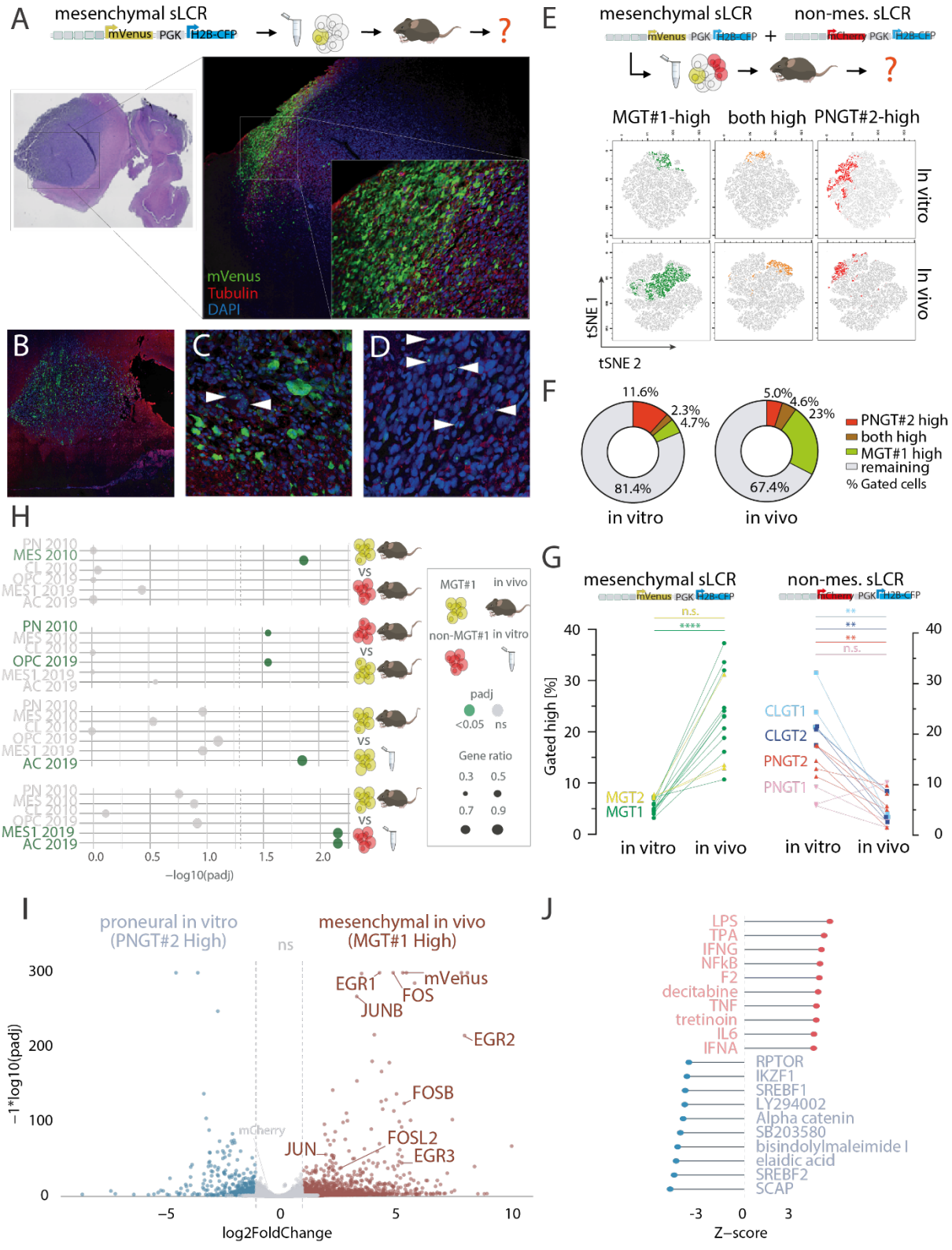

A
mesenchymal sLCR
mVenus PGK H2B-CFP
mVenus
Tubulin
DAPI
B
C
D
E
mesenchymal sLCR
non-mes. sLCR
mVenus PGK H2B-CFP
mCherry PGK H2B-CFP
MGT#1-high
both high
PNGT#2-high
In vitro
In vivo
tSNE 1
tSNE 2
F
11.6%
2.3%
4.7%
81.4%
5.0%
4.6%
23%
67.4%
in vitro
in vivo
PNGT#2 high
both high
MGT#1 high
remaining
% Gated cells
H
PN 2010
MES 2010
CL 2010
OPC 2019
MES1 2019
AC 2019
VS
MGT#1
in vivo
non-MGT#1
in vitro
padj
<0.05
ns
Gene ratio
0.3
0.5
0.7
0.9
0.0
0.5
1.0
1.5
2.0
−log10(padj)
G
mesenchymal sLCR
non-mes. sLCR
n.s.
****
**
Gated high [%]
MGT2
MGT1
CLGT1
CLGT2
PNGT2
PNGT1
in vitro
in vivo
I
proneural in vitro
(PNGT#2 High)
ns
mesenchymal in vivo
(MGT#1 High)
EGR1
mVenus
FOS
JUNB
EGR2
FOSB
FOSL2
EGR3
JUN
mCherry
−1*log10(padj)
log2FoldChange
J
LPS
TPA
IFNG
NFkB
F2
decitabine
TNF
tretinoin
IL6
IFNA
RPTOR
IKZF1
SREBF1
LY294002
Alpha catenin
SB203580
bisindolylmaleimide I
elaidic acid
SREBF2
SCAP
Z−score

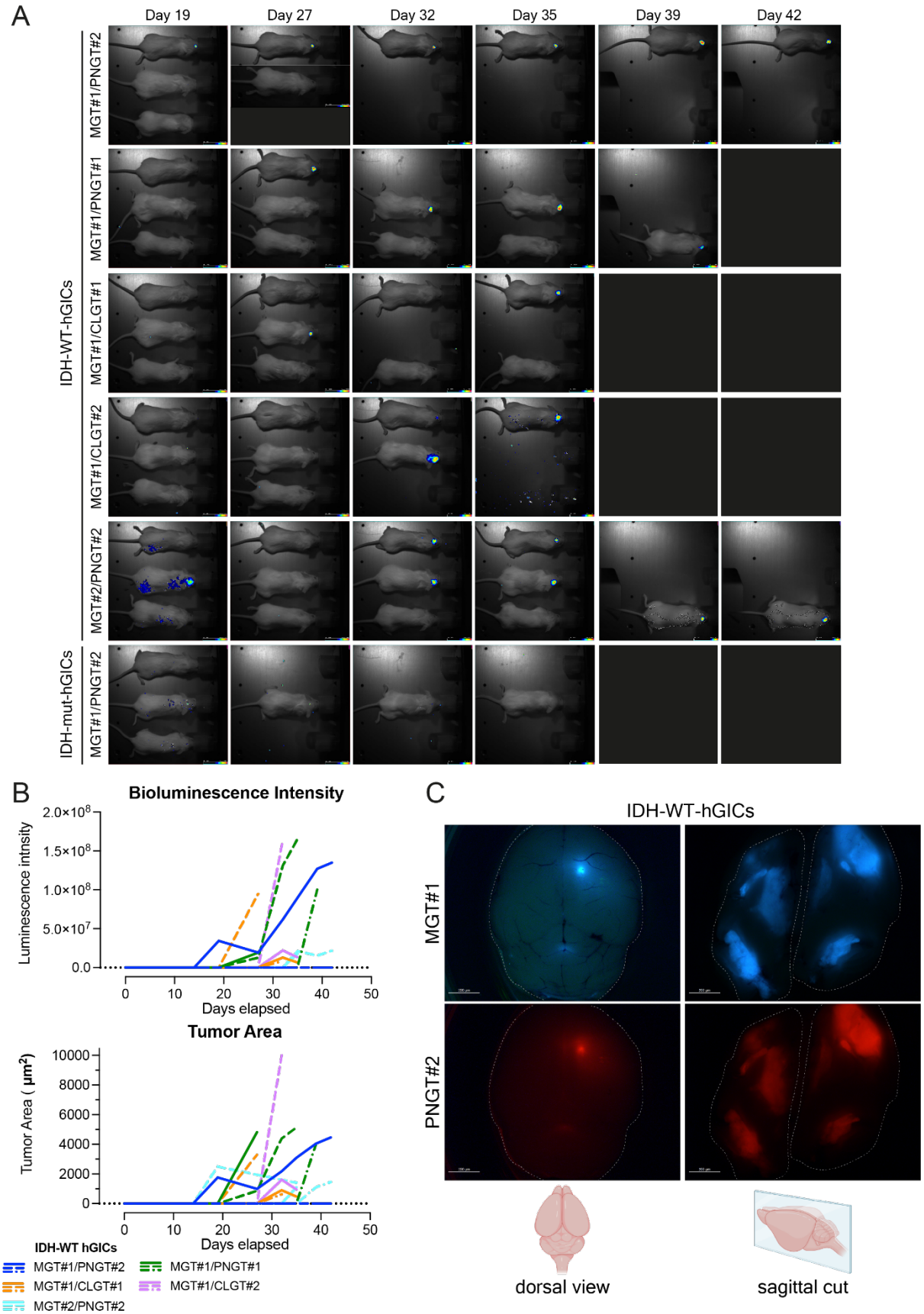
Figure 7 - Synthetic genetic tracing *in vivo* reveals Glioblastoma heterogeneity.

(**A**) Above, schematic representation of the experiment. Coronal forebrain images of IDH-wt-hGICs-MGT#1 xenograft in NSG mice at a humane end point are shown below (n=10). Lower left, HE staining; lower right, insets of mVenus, Tubulin, and DAPI counterstained tissue with uniform MGT#1-high invasive glioma front. (**B**) A representative lesion with a mix of high and low mesenchymal reporter expression. (**C-D**) Representative H2B-CFP expression (arrowhead) in MGT#1-positive and MGT#1-negative lesions. (**E**) A schematic illustration of the experiment is shown above. Below is a representative t-SNE map of reporter expression *in vitro* and *in vivo* for IDH-wt-hGICs using the indicated dual reporter combination. The gating approach is illustrated in (Ext. Fig. 7b). (**F**) Quantitative analysis of t-SNE data in (E). (**G**) *In vitro* and *in vivo* expression of the dual reporters for the specified pairs (n=3/group). The significance of differential expression of each *in vivo* reporter group as compared to its relative *in vitro* control is determined using an unpaired t-test (****P<0.0001, ns=not significant). (**H**) Bubble plot showing GSEA-adjusted p-values and comparisons for the specified Glioblastoma subtypes/states. (**I**) Volcano plot of the differential expression analysis of PNGT#2-high and MGT#1-high *in vivo*. A selection of genes is highlighted. (**J**) Ingenuity pathway top upstream regulator analysis of differential expression analysis in (I).

To understand the dynamics of GBM cell states *in vivo*, we generated dual-sLCR lines of IDH-WT- and IDH-mutant-hGICs with one mVenus-driven mesenchymal reporter (MGT#1 or MGT#2) together with one mCherry-driven non-mesenchymal reporter (PNGT#1, PNGT#2, CLGT#1 or CLGT#2). Upon orthotopic intracranial injection into NOG mice (n=18), all animals showed tumor formation starting from day 14-20, as monitored by IVIS (Ext. Fig. 7a A-B). At humane endpoint, with a median survival of 36 days (IDH-WT hGICs) or 26 days (IDH-mut hGICs) (Fig. 6B), animals presented with large tumors that in most cases also infiltrated into distant areas of the brain (Ext. Fig. 7a C).

After freshly isolating tumor cells, we applied t-Distributed Stochastic Neighbour Embedding (t-SNE) to categorize parallel *in vivo*/*in vitro* flow cytometry data; this consisted of hierarchical components including cell shape, granularity, viability dyes, mesenchymal and non-mesenchymal fate reporters (Ext. Fig. 7b). We observed *in vivo* a significant increase in cells with high expression of the mesenchymal reporter MGT#1 ($p<0.0001$) and a milder but still significant decrease in proportions of tumor cells with proneural or classical reporter expression ($p<0.01$) compared to their *in vitro* counterparts profiled simultaneously (Fig. 7E-G).

We next wanted to deduce the transcriptional identity of mesenchymally-committed cells *in vivo*. Therefore, we FAC-sorted hundreds of IDH-WT-hGICs from xenograft cell suspensions with high expression of mesenchymal (MGT#1) and non-mesenchymal reporters (PNGT#2 and CLGT#1) and profiled their gene expression signatures with RNA-seq alongside with the *in vitro* cultured counterparts. The profiles we generated from tumor cells with high MGT#1 expression *in vivo* were significantly enriched for the TCGA mesenchymal-GBM gene set as opposed to tumor cells sorted for high non-mesenchymal reporter expression, which in turn showed enrichment for the TCGA proneural-GBM and oligodendrocyte-like (Neftel-OPC) gene sets (Fig. 7H).


A
Day 19
Day 27
Day 32
Day 35
Day 39
Day 42
IDH-WT-hGICs
MGT#1/PNGT#2
MGT#1/PNGT#1
MGT#1/CLGT#1
MGT#1/CLGT#2
MGT#2/PNGT#2
IDH-mut-hGICs
MGT#1/PNGT#2
B
Bioluminescence Intensity
Luminescence intnsity
2.0×10^8
1.5×10^8
1.0×10^8
5.0×10^7
0.0
0
10
20
30
40
50
Days elapsed
Tumor Area
Tumor Area (µm²)
10000
8000
6000
4000
2000
0
Days elapsed
IDH-WT hGICs
MGT#1/PNGT#2
MGT#1/PNGT#1
MGT#1/CLGT#1
MGT#1/CLGT#2
MGT#2/PNGT#2
C
IDH-WT-hGICs
MGT#1
PNGT#2
dorsal view
sagittal cut

Extended Figure 7a - Synthetic genetic tracing *in vivo* reveals Glioblastoma heterogeneity.

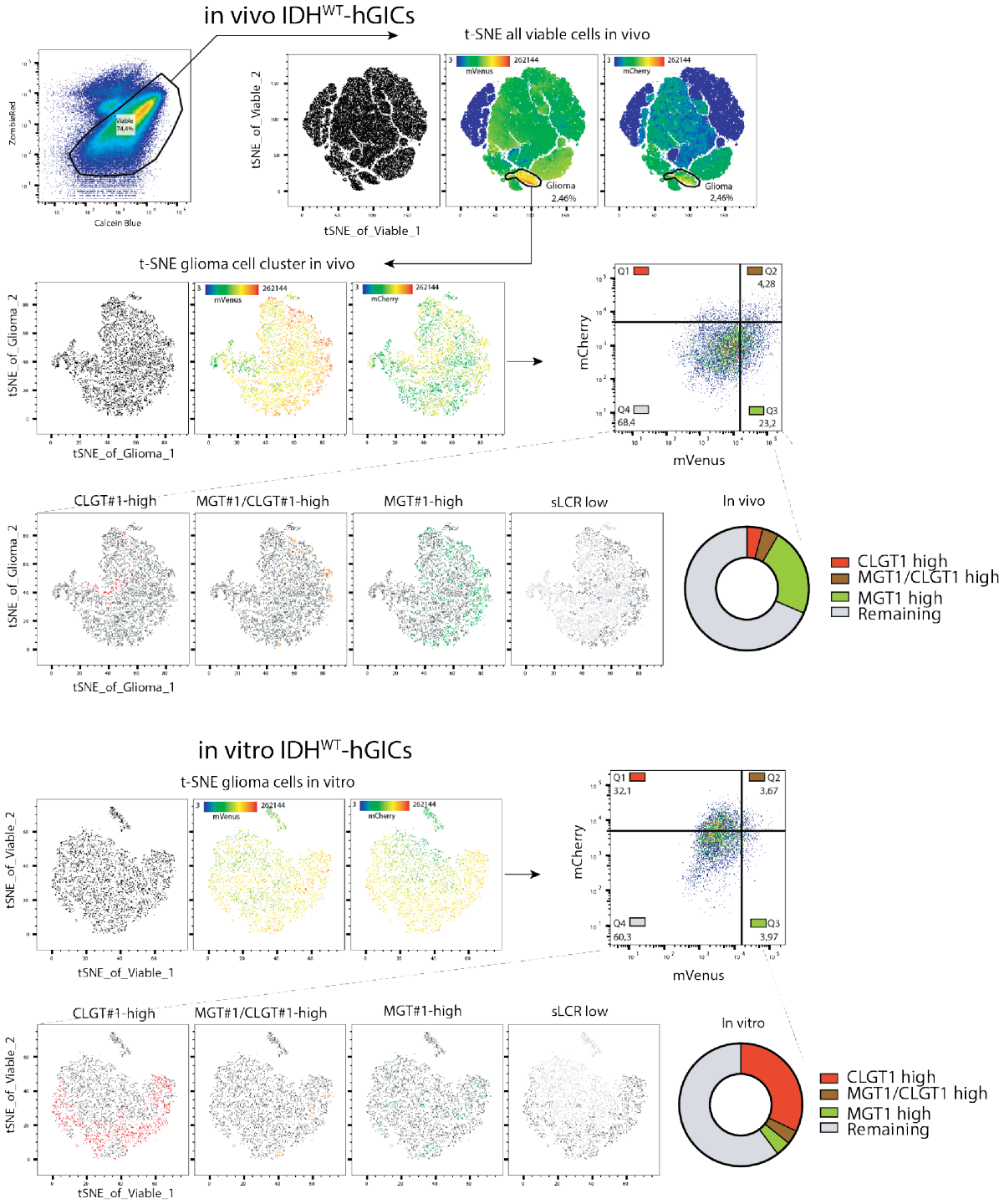
(**A**) Bioluminescence imaging using IVIS of NOG mice injected with hGICs expressing the specified sLCR-combinations. (**B**) Quantification of the bioluminescence intensity and area of the tumors shown in (A). (**C**) Fluorescence microscopy of freshly dissected mouse brains in dorsal and sagittal sections at humane endpoint. Tumors exhibit MGT#1-mVenus and PNGT#2-mCherry sLCR signals.

Likewise, instead of comparing the two opposing cell states found *in vivo* and focusing the analysis on the comparison of MGT#1-high *in vivo* against PNGT#2-high *in vitro*, we can observe a significant enrichment for a patient-derived mesenchymal-GBM gene set (Neftel-MES1). Finally, comparing the transcriptional profiles between MGT#1-high expressing tumor cells *in vivo* against *in vitro*, revealed an enrichment for the astrocyte-like (Neftel-AC) gene set to be a dominant feature of MGT#1-high cells *in vivo* (Fig. 7H).

One powerful aspect of our genetic tracing technology is the possibility to selectively isolate cells that highly express the reporter within an experimental or biological setting and profile their transcriptome to learn about the underlying molecular mechanisms governing the cell fate determination. We wanted to utilise this aspect in order to identify the hallmarks of proneural-to-mesenchymal transition *in vivo*. For this purpose, we performed differential gene expression analysis between MGT#1-high *in vivo* cells against cells with the most opposing cellular state (PNGT#2-high) *in vitro*. We found 954 genes in the MGT#1 *in vivo* cells, and 234 genes in the PNGT#2 subset *in vitro* to be specifically upregulated (padj<0.05; log2FC±1.5; Fig. 7I). As expected, we found mVenus (the MGT#1 reporter protein) to be among the most significantly upregulated genes (log2FC=5.34; padj=0) and mCherry (PNGT#2 reporter protein) among the list of depleted genes, yet not meeting the criteria for statistical significance (log2FC=-0.48; padj=0.42; Fig. 7I). Genes that marked MGT#1-high hGICs *in vivo* were found to be transcription factors of the early growth factor response (EGR1, EGR2, EGR3) as well as members of the Fos (FOS, FOSB, FOSL2) and Jun (JUN, JUNB) transcription factor families (Fig. 7I). The latter are known members to form the ubiquitous dimeric AP-1 transcription factor complexes, which are crucial for cells to rewire their gene expression profiles in response to various extracellular and cell-autonomous signaling cues (Angel & Karin, 1991; Karin et al., 1997; Eferl & Wagner, 2003).

Inferring potential upstream regulators via ingenuity pathway analysis revealed that the upregulated genes of MGT#1-cells *in vivo* are ranging downstream of several pro-inflammatory regulators, including TNFα (z-score=2.655, p=4.49-24), NFκB (z-score=1.915, p=3.05-13), and interferons (Fig. 7J). Further putative upstream regulators comprise targets of the synthetic retinoic acid tretinoin (z-score=2.99, p=3.31-22), the TGFB and VEGF pathways (z-score=2.51, p=3.28-26 and z-score=1.531, p=3.2-15, respectively; data not shown).

In conclusion, we were able to harness our synthetic reporters in order to trace a cell fate transition of rather homogenously proneural hGICs *in vitro* towards a more heterogeneous distribution *in vivo*, with a fraction of highly mesenchymally-committed tumor cells. Furthermore, selective isolation and transcriptomic profiling of distinct reporter-expressing cells allowed identification of molecular key players underlying the phenotypic plasticity.



Extended Figure 7b - Synthetic genetic tracing *in vivo* reveals Glioblastoma heterogeneity.

A gating strategy for FACS analysis of live and single IDH-wt-hGICs-MGT#1-CLGT#1 cells *in vitro* and *in vivo* is demonstrated. Freshly dissociated mouse brain tumor samples were pre-gated for viable cells before dimensionality reduction of all acquired parameters (FSC-H/W/A, SSC-H/W/A, positive and negative viability dyes, and mVenus/mCherry sLCR expression). By superimposing sLCR expression heatmaps on the generated t-SNE maps, we discovered and gated glioma cell clusters that separated from mouse cells. Further t-SNE dimensionality reduction of gated glioma cells was done to analyze clustering of the sLCR reporter distribution in individual *in vivo*-derived tumor cells and to compare it to *in vitro*-cultured cells used for transplantation that were simultaneously analysed. The quantification of sLCR-high cells was established by developing a four-quadrant gating scheme in mVenus vs. mCherry plots on *in vitro* cells to identify sLCR-high populations and then applying this gating scheme to t-SNE gated *in vivo* glioma cells.

2.4 The mesenchymal Glioblastoma identity is adaptive and reversible

Glioblastoma cells are highly plastic and can undergo cell fate transitions along certain axes of variation converging over single-cell transcriptomic states: Oligodendrocyte progenitor cell (OPC)-like, Neural progenitor cell (NPC)-like, Astrocyte (AC)-like and Mesenchymal (MES)-like (Neftel et al., 2019). A Projection of the previously defined TCGA-bulk signature subtypes over the single-cell identities suggests two major axis of variation with the MES subtype as a conserved entity, the Proneural subtype being a mixture of NPC- and OPC-like states and the Classical subtype consisting of AC-like and MES-like cells (Fig. 5F). The OPC-, NPC-, and AC-like states mirror normal neurodevelopmental pathways, whereas the MES-like state stands out with no clear relationship and may diverge from this paradigm (Neftel et al., 2019; Couturier et al., 2020).

Our hGICs were derived from human neural progenitor cells (Fig. 6A) and express *in vitro* high levels of the proneural reporter PNGT#2 (Fig. 6C-D). In contrast, MGT#1 is expressed at low levels *in vitro* and highly MGT#1-expressing cells only emerge in the context of a physiologically relevant tumor microenvironment (Fig. 7A-G). We therefore hypothesised that the proneural identity of our cells might be engraved and sustained by intrinsic mechanisms, whereas the acquisition of a mesenchymal state might be an adaptive reaction and requires external signaling.

To test this hypothesis, we subjected our reporter-bearing hGICs to a phenotypic screen, where we provided external signaling to the standard *in vitro* culture conditions with the addition of an array of cytokines and bioactive compounds (Fig. 8A). In comparison to the untreated condition, after 48h of compound exposure our mesenchymal reporter was strongly upregulated in IDH-WT-hGICs in response to TNFα signaling, human/bovine serum, LIF and to a lower extent to Activin A, highlighting these compounds as potential mediators of the mesenchymal state (Fig. 8B-C).

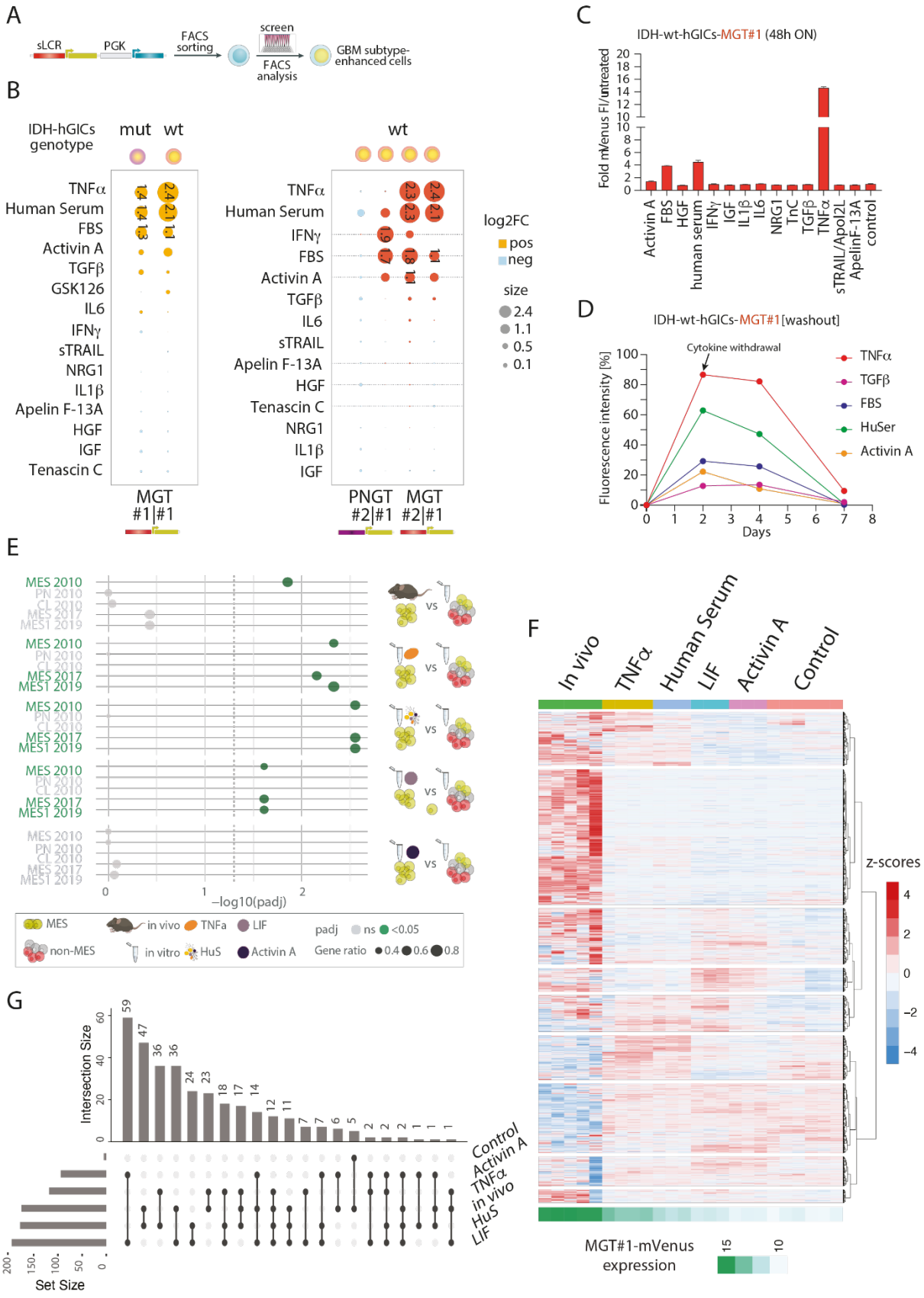

A
sLCR
PGK
FACS sorting
screen
FACS analysis
GBM subtype-enhanced cells
B
IDH-hGICs genotype
mut
wt
TNFα
Human Serum
FBS
Activin A
TGFβ
GSK126
IL6
IFNγ
sTRAIL
NRG1
IL1β
Apelin F-13A
HGF
IGF
Tenascin C
MGT #1|#1
PNGT #2|#1
MGT #2|#1
log2FC
pos
neg
size
2.4
1.1
0.5
0.1
C
IDH-wt-hGICs-MGT#1 (48h ON)
Fold mVenus FI/untreated
Activin A
FBS
HGF
human serum
IFNγ
IGF
IL1β
IL6
NRG1
TnC
TGFβ
TNFα
sTRAIL/Apol2L
ApelinF-13A
control
D
IDH-wt-hGICs-MGT#1[washout]
Cytokine withdrawal
Fluorescence intensity [%]
Days
TNFα
TGFβ
FBS
HuSer
Activin A
E
MES 2010
PN 2010
CL 2010
MES 2017
MES1 2019
vs
–log10(padj)
MES
non-MES
in vivo
in vitro
TNFa
HuS
LIF
Activin A
padj
ns
<0.05
Gene ratio
0.4
0.6
0.8
F
In vivo
TNFα
Human Serum
LIF
Activin A
Control
z-scores
MGT#1-mVenus expression
G
Intersection Size
Set Size
Control
Activin A
TNFα
in vivo
HuS
LIF

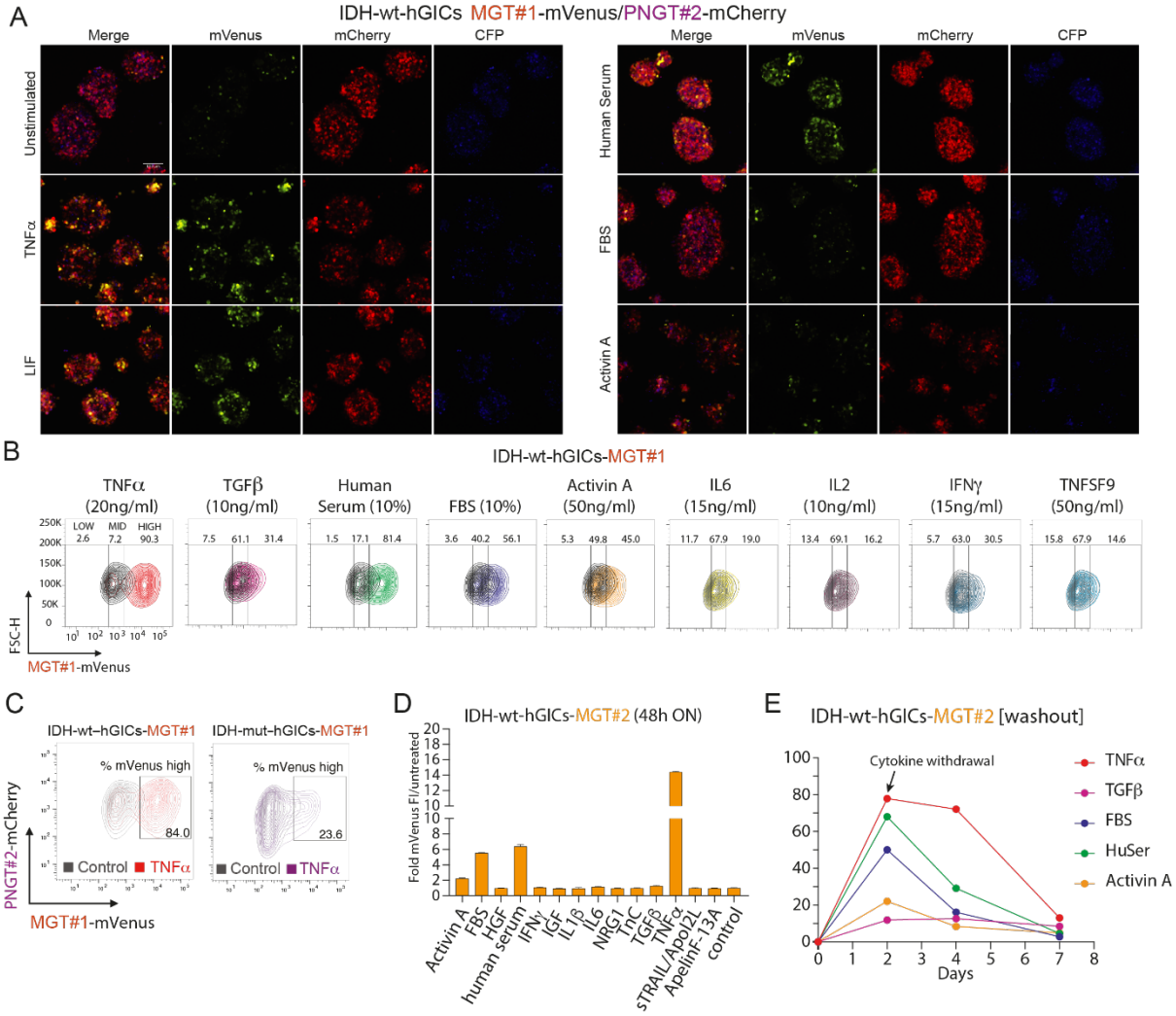
Figure 8 - The mesenchymal Glioblastoma identity is adaptive and reversible.

(**A**) Schematic representation of the phenotypic screening procedure with sLCRs. (**B**) A bubble plot depiction of the screening for the indicated factors that regulate MGT#1 in IDH-mut- and IDH-wt-hGICs (left) or MGT#1, MGT#2, PNGT#1 or PNGT#2 in IDH-wt-hGICs (right) (right). The size and color of the bubbles denote the extent and direction of the change. (**C**) Bar graph illustrating individual responses to the specified factors/sLCRs 48 hours after induction. (**D**) Representation of the MGT#1/2-mVenus gene's longitudinal expression in response to the indicated conditions beginning on day 0. (stimulation). The time-point for cytokine withdrawal is indicated by arrows. (**E**) Bubble plot showing GSEA-adjusted p-values for the specified GBM subtypes/states and comparisons between MES-inducing stimuli. (**F**) Heatmap representing the normalized values of all up-regulated genes for the highlighted comparisons. MGT#1-mVenus expression is indicated separately for each sample grouped according to condition (green). (**G**) Upset plot of all intersections for the specified MGT#1 activation stimuli, sorted according to the magnitude of the intersections. The circles in the matrix that are interconnected represent number of genes that are found in common.

The screening results were further validated via fluorescence live-cell imaging and FACS (Ext. Fig. 8a A+B). Of further note and in line with earlier observations, the IDH-mutant cells displayed an attenuated MGT#1-activation in response to external signaling (Fig. 8B and Ext. Fig. 8a C).

In contrast to the mesenchymal reporters, we were unable to discover signaling molecules in the same screen that consistently altered the expression of the proneural reporters, with the exception of interferon gamma and serum, which moderately increased PNGT#1 expression (Fig. 8B). Since proneural reporters have a high baseline-expression, which cannot easily be modulated by external signaling, we reasoned that the proneural identity is anchored in our cellular model either by the cell of origin or the Receptor-tyrosine-kinase signaling rather than dependent on external stimuli (e.g. from the microenvironment).

In contrast to that, the mesenchymal identity appeared to be an adaptive state, which is acquired in response to external signaling. Harnessing the power of our genetic tracing reporters, we next wanted to investigate whether the acquisition of a mesenchymal identity represents a stable transition or would also be reversible. After we instructed IDH-WT-hGICs to a mesenchymal state in presence of the previously identified triggers for a period of 48h (Fig. 8B), we have removed the cytokines from the culture medium and repeatedly measured reporter intensity over a time-course of 7 days. Within a time frame of five days, the intensity of the MGT#1 reporter faded back to baseline levels when cells were switched back to the basal medium (Fig. 8C-D). Importantly, these results were reproducible across two independently designed mesenchymal reporters (i.e. MGT#1-2; Ext. Fig. 8a D-E).



Extended Figure 8a - The mesenchymal Glioblastoma identity is adaptive and reversible.

(**A**) Representative confocal live cell imaging of MGT#1/PNGT#2 dual-sLCR IDH-wt-hGICs after 48h of exposure to indicated stimuli. (**B**) Representative FACS profiles demonstrating MGT#1 activation in IDH-wt-hGICs in response to the specified stimuli. (**C**) A representative FACS profile of MGT#1 and PNGT#2 expression in specified hGIC genotypes under the indicated conditions. (**D**) Bar plot illustrating the individual response of an independently designed MGT#2-mVenus sLCR to the indicated compounds 48 hours after induction. (**E**) Representation of the MGT#2-mVenus gene's longitudinal expression in response to the indicated conditions beginning on day 0 (stimulation). The time-point for cytokine withdrawal is indicated by arrows.

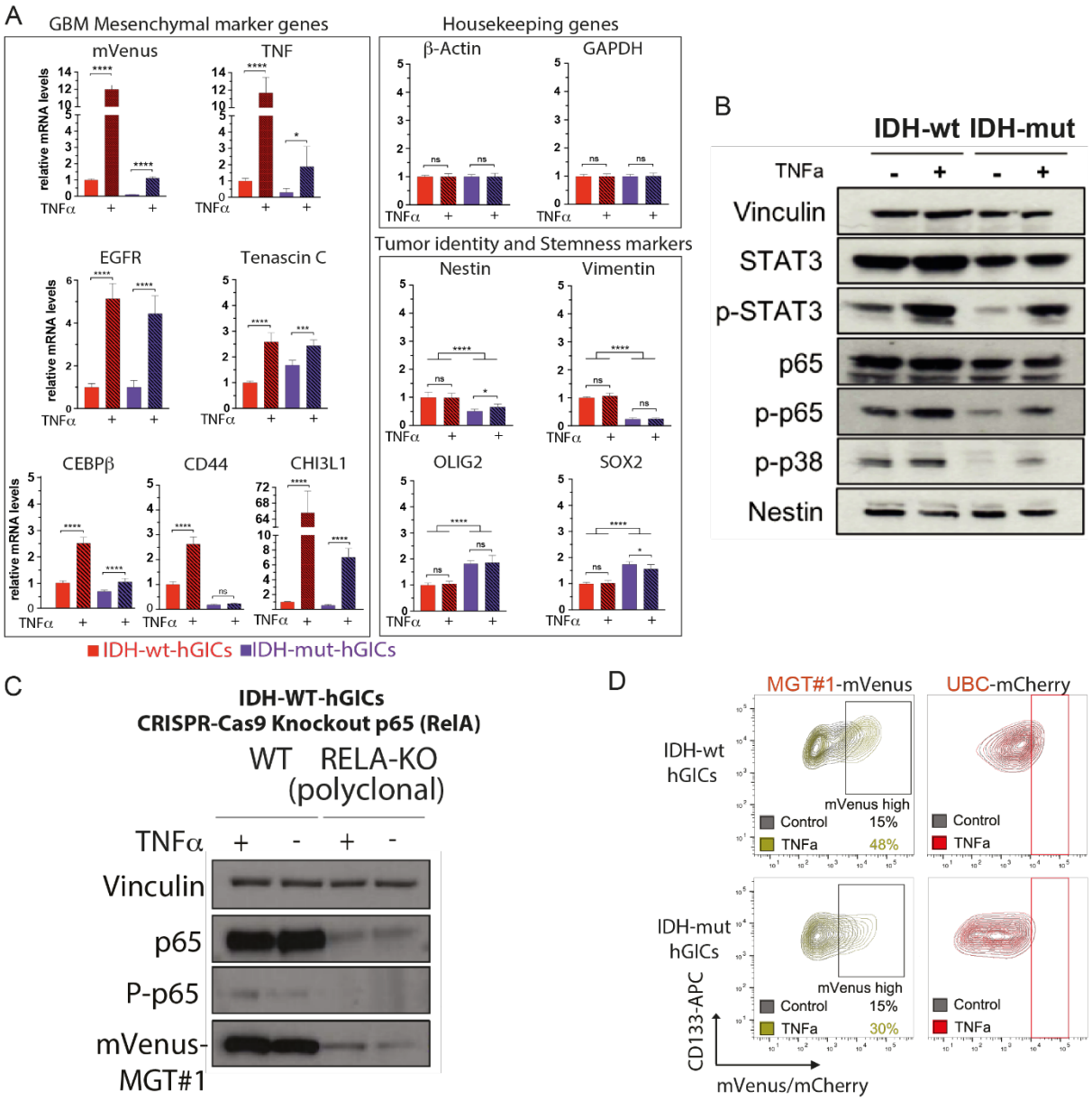
One of the top mesenchymal triggers in our screen turned out to be TNFα. TNFα is a well-described activator of the NFκB signaling pathway, which acts on genes regulating inflammation, cell survival, proliferation and differentiation (Hayden et al., 2014). This is in line with the presence of binding sites for the transcription factor NFκB in the cis-regulatory elements representing the MGT#1 reporter (Table 2 and Table 3), our identification of NFκB signaling being a regulator of the mesenchymal fate *in vivo* (Fig. 7J) and validating previous studies with patient-derived glioma cell lines (Bhat et al., 2013). To confirm that TNFα stimulation resulted in the upregulation of genes associated with the MES-GBM signature, we performed RT-qPCR in IDH-WT and -mutant

hGICs before and after treatment with TNFα. Besides a strong induction of transcripts for mVenus and TNF, we detected specific upregulation of prominent mesenchymal-GBM marker genes, such as CD44, CHI3L1, EGFR and Tenascin C (Ext. Fig. 8b A). Furthermore, on the protein level, TNFα led to the phosphorylation of the NFκB-subunit p65, p38-MAPK and STAT3, which is a well-described master regulator of mesenchymal transformation in malignant gliomas (Carro et al., 2009; Ext. Fig. 8b B). CRISPR-Cas9 mediated ablation of p65 resulted in reduced baseline protein levels of MGT#1-mVenus and drastically attenuated its induction by TNFα, underlining the crucial role of NFκB signaling for mesenchymal transition (Ext. Fig. 8b C). The measured induction of the sLCR in response to TNFα was specific to the MGT#1-reporter and not observable in a non-specific control-sLCR, which is expressing mCherry under an ubiquitously active polyubiquitin C (UBC) promoter (Ext. Fig. 8b D).

Next, we wanted to clarify whether also the various other elicitors of MGT#1 expression we found in our screen would independently cause the adoption of a mesenchymal gene expression signature. We FAC-sorted after 48h IDH-WT-hGICs that upregulated MGT#1 in response to TNFα, human Serum, FBS, LIF or Activin A and profiled their transcriptome using RNA-seq and compared those signatures to unsorted homeostatic hGICs and to MGT#1-high cells from xenografts. Except for the weak MGT#1-inducer Activin A and to a lower extent with LIF, all treatments promoted the acquisition of a signature that is enriching for mesenchymal gene sets found in human GBM patients (Fig. 8E; TCGA-MES 2010 [Verhaak et al., 2010]; TCGA-MES 2017 [Wang et al., 2017] and MES1 2019 [Neftel et al., 2019]). This data suggests that multiple routes over different stimuli can independently from each other converge over the acquisition of a more or less pronounced mesenchymal fate. Interestingly, when comparing all treated *in vitro* profiles to the one generated from MGT#1-high xenografts, we could observe that each compound generally had a milder transcriptional activation and each shared particular clusters of upregulated genes with MGT#1-high expressing cells *in vivo* (Fig. 8F). Notably, in this analysis, the transcriptional outputs of TNFα, human serum and LIF treated cells shared the greatest similarities with the profiles obtained from MGT#1-high cells *in vivo*, which prompts the hypothesis that mesenchymal transition *in vivo* might be driven by a combinatorial effect of multiple external stimuli (Fig. 8G).

To sum up, genetic tracing of cell fate transitions using our sLCRs revealed that the proneural Glioblastoma cell state appears to be a rather stable entity, which builds on intrinsic signaling, whereas the mesenchymal identity represents an adaptive and reversible state. Multiple independent roadways to mesenchymal transition *in vitro* were identified, each inducing a

characteristic gene expression profile, which share similarities to *in vivo* mesenchymal differentiation and are also correlating to human mesenchymal GBM patients.



Extended Figure 8b - The mesenchymal Glioblastoma identity is adaptive and reversible.

(**A**) Validation of the expression of the specified genes in response to Tumor Necrosis Factor alpha (TNFa) treatment in the indicated GICs using reverse transcription-quantitative PCR. ANOVA followed by Dunnet's post hoc test; n=3 biologically independent samples; ****=P<0.0001; mRNAs are normalized to b-Actin and IDH-wt-hGICs control cells. (**B-C**) Immunoblotting using the conditions and antibodies specified. (**D**) A representative FACS profile of the expression of selected markers under the specified condition. UBC-mCherry is a control reporter construct derived from a broadly expressed promoter (UBC).

2.5 *Genetic tracing of mesenchymal Glioblastoma demonstrates cell fate transition in response to ionizing radiation but not hypoxia*

The current standard-of-care for newly diagnosed GBM patients consists of maximal surgical resection followed by radio-chemotherapy (Stupp et al., 2005). Despite treatment, the tumor invariably recurs at a >90% rate, leaving patients with a poor prognosis and 2-year-survival of less than 20% (Omuro et al., 2013). Therapy-resistant recurrent tumors often present themselves in a more aggressive and invasive form with second-line treatments being sparse and largely ineffective (Weller et al., 2013; Birzu et al., 2020). Due to phenotypic and transcriptomic plasticity, a frequent treatment-induced switch in transcriptional identity towards the mesenchymal GBM signature was seen upon recurrence (Wang et al., 2017; Varn et al., 2021). The connection between mesenchymal transition and radiotherapy - which is a part of the standard-of-care - has been intensively reported on, yet a direct causal link remains to be demonstrated (Phillips et al., 2006; Bhat et al., 2013; Halliday et al., 2014; Behnan et al., 2019).

Given that our sLCRs permit the monitoring of cell fate changes, we sought to determine experimentally if ionizing radiation (IR) would establish a mesenchymal identity on our hGICs. We subjected MGT#1-bearing hGICs to various doses of IR with phosphorylation of the DNA damage marker γH2AX confirming the induction of double-strand breaks. After a time frame of 72h post irradiation, we measured a dose-dependent increase of MGT#1 expression, indicating the acquisition of a mesenchymal signature (Fig. 9A). To prove that a proneural-mesenchymal-transition (PMT) had occurred, we sorted MGT#1-high IDH-WT-hGICs after 48h post irradiation with 20Gy and performed transcriptional profiling via RNA-seq. We found 135 genes significantly up- and 31 genes down-regulated (log2FC±1.5; padj<0.05) and gene-set enrichment analysis of these genes confirmed that these cells underwent PMT, coupling it with the DNA damage-induced activation of the ATM signaling pathway for DNA repair (Fig. 9B-C).

Another critical driver of GBM transcriptional heterogeneity can also be found in the context of the tumor microenvironment. Glioblastoma cells are exposed to and reside in various anatomical features of a tumor (i.e. infiltrating tumor edge, cellular bulk tumor, pseudopalisading cells around necrotic areas and the perivascular niche), with each enriching for a characteristic transcriptional profile of tumor cells (Puchalski et al., 2018; Varn et al., 2021). Especially within necrotic or poorly vascularized areas of the tumor, GBM cells are exposed to hypoxic conditions due to low oxygen availability caused by the increased oxygen diffusion distance from the nearest blood-vessel supply. The master regulators of the hypoxia response are HIF transcription factors and through stabilization of HIF-1/2α, hypoxic cells were reported to be transcriptionally remodelled to produce a range of pro-angiogenic factors and attract inflammatory cells (Monteiro et al., 2017).

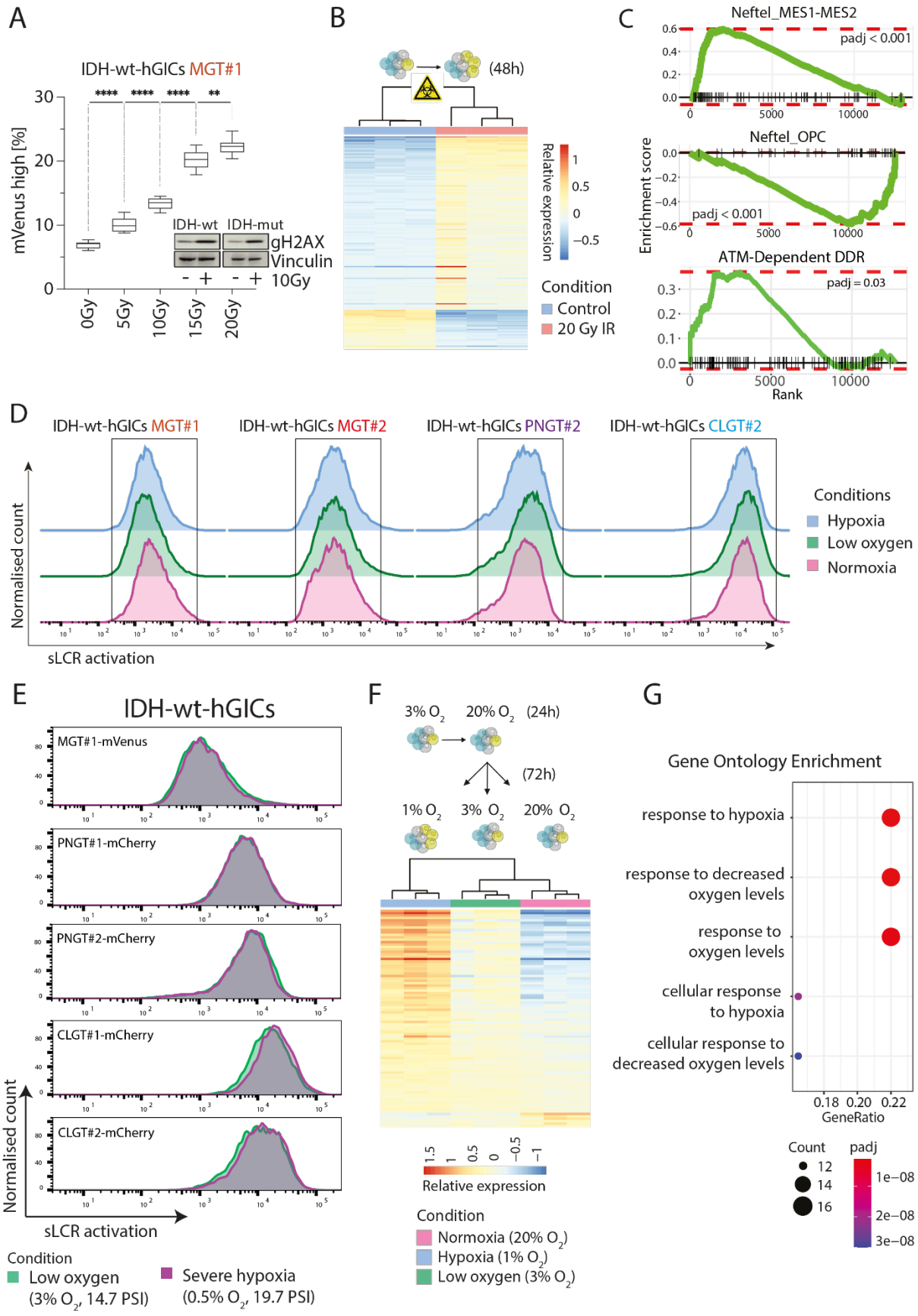

A
IDH-wt-hGICs MGT#1
mVenus high [%]
IDH-wt IDH-mut
gH2AX
Vinculin
10Gy
0Gy
5Gy
10Gy
15Gy
20Gy
B
(48h)
Relative expression
Condition
Control
20 Gy IR
C
Neftel_MES1-MES2
padj < 0.001
Neftel_OPC
padj < 0.001
ATM-Dependent DDR
padj = 0.03
Enrichment score
Rank
D
IDH-wt-hGICs MGT#1
IDH-wt-hGICs MGT#2
IDH-wt-hGICs PNGT#2
IDH-wt-hGICs CLGT#2
Normalised count
sLCR activation
Conditions
Hypoxia
Low oxygen
Normoxia
E
IDH-wt-hGICs
MGT#1-mVenus
PNGT#1-mCherry
PNGT#2-mCherry
CLGT#1-mCherry
CLGT#2-mCherry
Normalised count
sLCR activation
Condition
Low oxygen
(3% O2, 14.7 PSI)
Severe hypoxia
(0.5% O2, 19.7 PSI)
F
3% O2
20% O2
(24h)
(72h)
1% O2
3% O2
20% O2
Relative expression
Condition
Normoxia (20% O2)
Hypoxia (1% O2)
Low oxygen (3% O2)
G
Gene Ontology Enrichment
response to hypoxia
response to decreased oxygen levels
response to oxygen levels
cellular response to hypoxia
cellular response to decreased oxygen levels
GeneRatio
Count
padj

Figure 9 - Mesenchymal Glioblastoma genetic tracing reveals cell state change in response to ionizing radiation but not hypoxia.

(**A**) Activation of MGT#1 in response to increasing ionizing radiation (IR) doses at 72-hour time point. The inset displays immunoblotting with the identified antibodies and the corresponding conditions 1 hour after ionizing radiation delivery. (**B**) Heatmap of genes significantly differentially regulated by ionizing radiation (padj<0.05 and log2FC>1.5) in the MGT#1-high IDH-wt-hGICs fraction (pink, n=3) compared to non-irradiated control cells (blue, n=3). (**C**) Graphs of gene set enrichment analysis (GSEA) for the indicated gene sets. (**D**) Quantification of the relevant sLCRs by FACS under hypoxic (blue), low oxygen (green), and normoxic (pink) conditions. (**E**) Representative FACS profiles of sLCR expression in hGICs following 3 days of severe hypoxia (magenta) versus low oxygen (yellow) (green). (**F**) A schematic representation of the RNA-seq experimental design is shown above. Below, heatmap of the genes that were differently regulated between hypoxic (blue, n=3) and normoxic (pink, n=3) environments (padj<0.05 and log2FC>1.5). The color-coding of the heatmap is determined by the relative rlog-normalized gene expression values across samples. (**G**) Bubble plot of the top hypoxia-responsive gene sets. Significance and gene ratio are denoted by color and size, respectively.

It was demonstrated that through the HIF1α-ZEB1-axis, GBM cells are enhancing their invasive ability and undergo a mesenchymal transition (Joseph et al., 2015). Furthermore in single-cell-sequencing studies of GBM patients, the mesenchymal signature was found to be partially constituted by genes involved in hypoxia-response signaling (Neftel et al., 2019). To investigate whether exposure of glioma cells to hypoxic conditions alone is sufficient to induce a mesenchymal identity, we cultured sLCR-modified hGICs under low oxygen (3% O2) levels, hypoxia (1% O2), and ambient oxygen partial pressure (21% O2). The expression of the mesenchymal MGT#1 and MGT#2 reporters remained constant throughout all conditions, which also did not change if cells were exposed to levels of severe hypoxia (0.5% O2) for a duration of three days (Fig. 9D-E). We next transcriptionally profiled hGICs from low oxygen, hypoxic and normoxic conditions by RNA-seq, which revealed a set of differentially regulated genes (Fig. 9F). In line with the experimental setup and the behaviour of our reporters, we found an enrichment of gene sets related to response to low oxygen levels, which was not accompanied by an enrichment for a mesenchymal signature (Fig. 9G).

In summary, sLCRs allowed us to decouple the cause-effect relationship of physiological stimuli from the observed phenotypic consequences. Whereas ionizing radiation leads to the acquisition of a mesenchymal identity in a cell-intrinsic way, hypoxia alone does not appear to cause mesenchymal commitment and likely acts in a non-cell-autonomous fashion, possibly through the interplay with other mediators.

2.6 *Crosstalk between tumor and innate immune cells drives non-cell autonomous phenotypic plasticity.*

Mesenchymal trans-differentiation may occur either in a glioma cell-intrinsic or a microenvironment-driven modality. Microenvironmental factors such as the bidirectional interaction with non-neuronal cells, including glial cells, endothelial cells and pericytes, and in particular brain-resident microglia and infiltrating monocyte-derived macrophages, have been linked to the formation of GBMs (Venkatesh et al., 2019; Venkataramani et al., 2019). GBM tumors are composed of a range of subclonal populations of malignant cells in various biological states that coexist in the same tumor microenvironment, with mesenchymal tumors comprising more macrophages, neutrophils, and other immunoglobulins and a lower tumor content (Wang Q et al., 2017; Neftel et al., 2019; Wang LB et al., 2021). Cell-cell contact and paracrine pathways are involved in these interactions, which result in phenotypic adaptation of tumor and TME subpopulations in diverse tumor environments (Hambardzumyan et al., 2016). Understanding the impact of cell intrinsic and extrinsic variables on intratumoral variability in GBM patients is critical for designing and optimizing GBM multimodal therapeutic options.

The issue remains as to whether there is a causal link, how immune cells might modulate cancer cells at the molecular level, and if this can be exploited therapeutically. GAMs may not only be recruited by mesenchymal Glioblastoma cells, but they may also actively contribute to the specification of these cells. We co-cultured IDH-wild-type cells with an early-passage immortalized human microglia cell line to test this hypothesis (hMG cl.C20; Garcia-Mesa et al., 2017). The *in vitro* co-culture of hGIC tumor spheres and hMG cells was set up utilizing trans-well inserts, which allow cell-cell contact while keeping the two populations physically separated. Under these conditions, hMG upregulated MGT#1 expression to a level equivalent to TNF (Fig. 10A-D). Concomitantly, the hMG changed their morphology when co-cultured together with hGICs, supporting the occurence of a bidirectional crosstalk (Ext. Fig. 10A).

Next, we differentiated primary human CD34+ cells into immature cells known as myeloid-derived suppressor-like (MDSC-like) and the human monocytic cell line THP-1 into M1 or M2 macrophage-like cells to see whether non-brain innate immune cells may also direct mesenchymal differentiation in human GICs. When co-cultured with macrophage-like cells, MGT#1 was highly elevated, with M1 polarized cells causing the highest phenotypic commitment, whilst MDSC-like cells only prompted a minor MGT#1 activation (Fig. 10A-C). This is consistent with our finding that cytokines linked with innate immunity, such as TNF, evoke a more explicit MGT#1 activation than signaling molecules originating from adaptive immunity or stroma (IFN/IL-2 and IL-6, respectively; Fig. 8A-B).

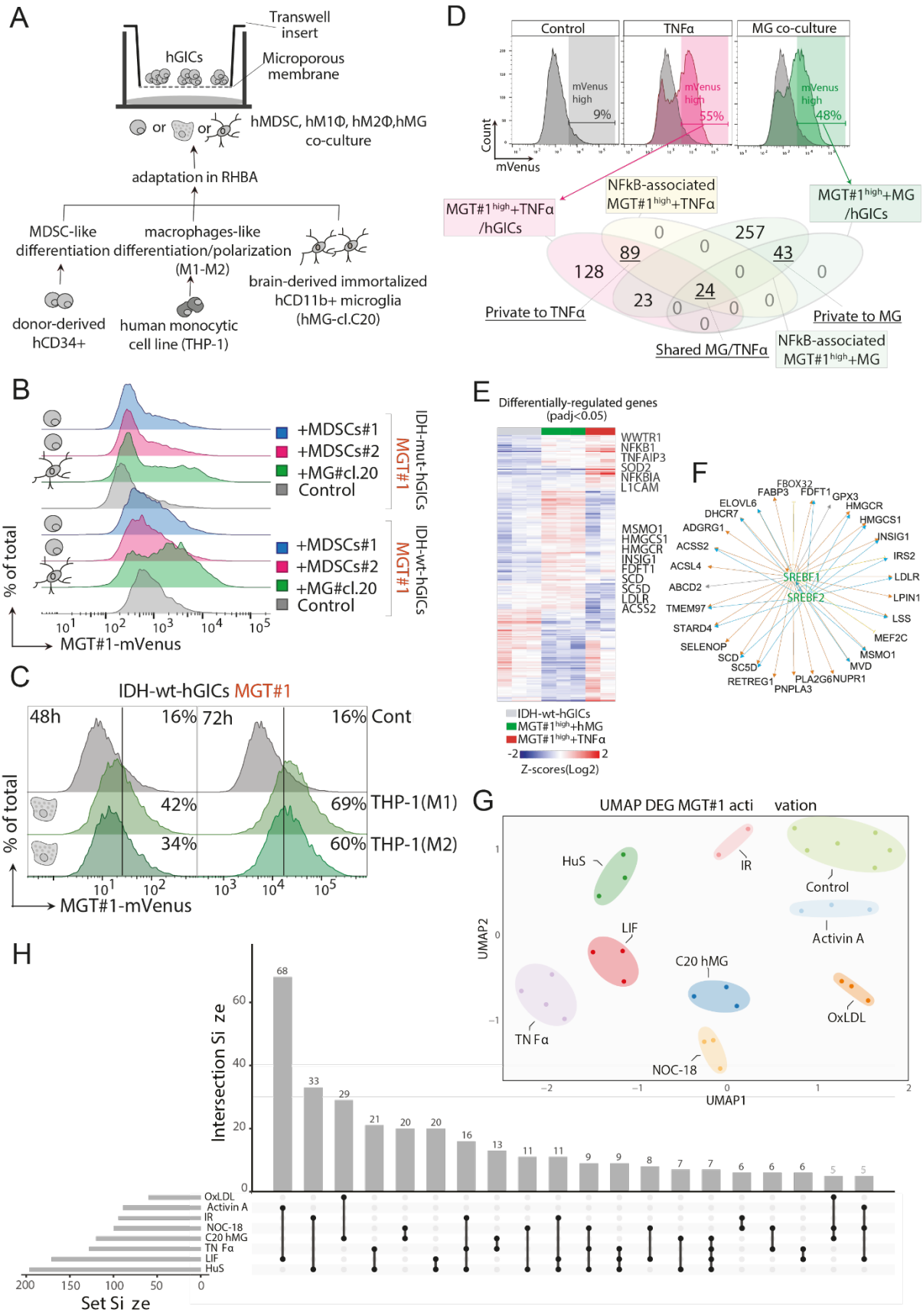
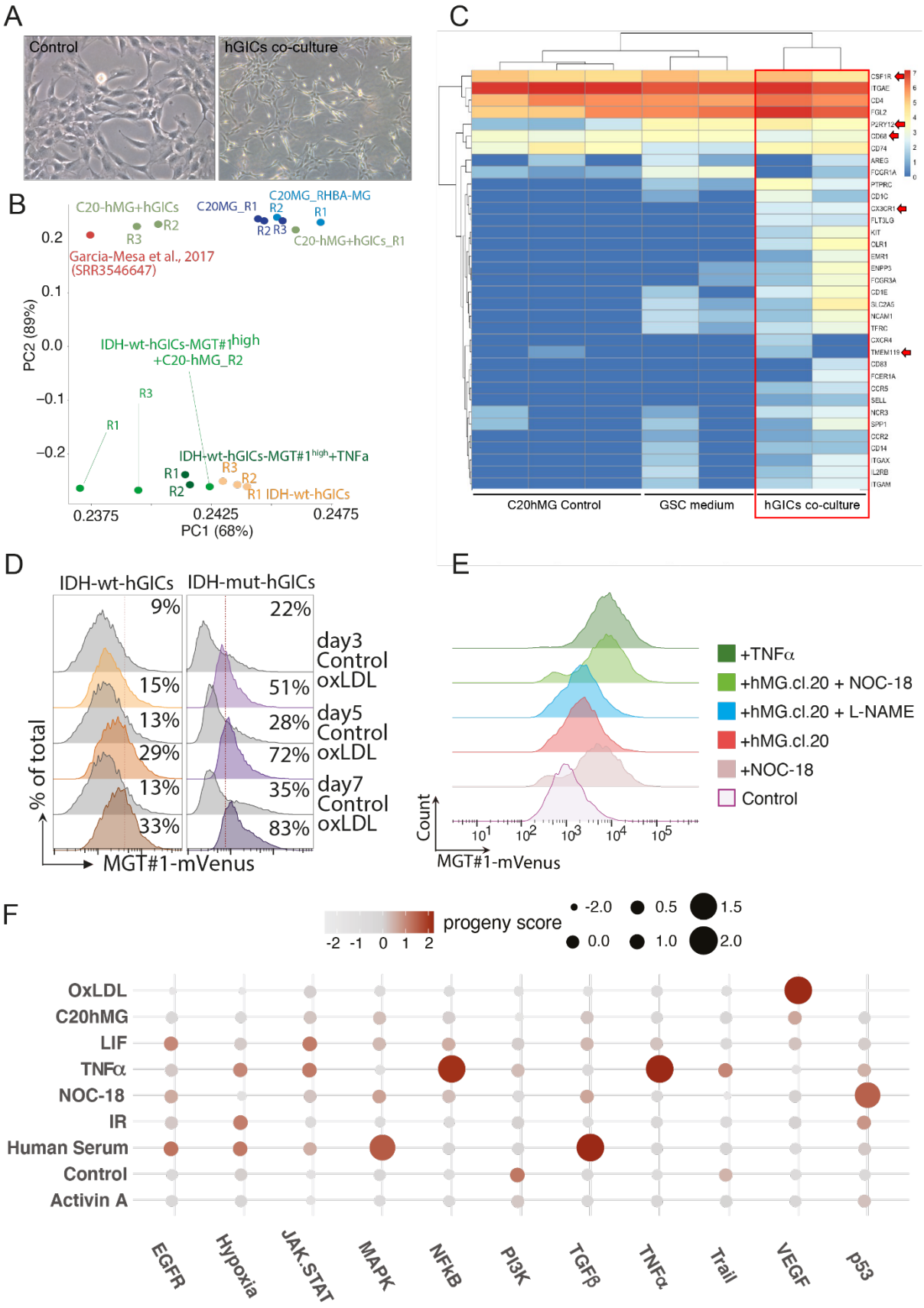

A
Transwell insert
hGICs
Microporous membrane
or or
hMDSC, hM1Φ, hM2Φ,hMG co-culture
adaptation in RHBA
MDSC-like differentiation
macrophages-like differentiation/polarization (M1-M2)
brain-derived immortalized hCD11b+ microglia (hMG-cl.C20)
donor-derived hCD34+
human monocytic cell line (THP-1)
B
+MDSCs#1
+MDSCs#2
+MG#cl.20
Control
IDH-mut-hGICs MGT#1
IDH-wt-hGICs MGT#1
% of total
MGT#1-mVenus
C
IDH-wt-hGICs MGT#1
48h
16%
72h
16%
Cont
42%
69%
THP-1(M1)
34%
60%
THP-1(M2)
% of total
MGT#1-mVenus
D
Control
TNFα
MG co-culture
mVenus high
9%
55%
48%
Count
mVenus
MGT#1high+TNFα /hGICs
NFkB-associated MGT#1high+TNFα
MGT#1high+MG /hGICs
128
89
23
24
257
43
Private to TNFα
Private to MG
Shared MG/TNFα
NFkB-associated MGT#1high+MG
E
Differentially-regulated genes (padj<0.05)
WWTR1
NFKB1
TNFAIP3
SOD2
NFKBIA
L1CAM
MSMO1
HMGCS1
HMGCR
INSIG1
FDFT1
SCD
SC5D
LDLR
ACSS2
IDH-wt-hGICs
MGT#1high+hMG
MGT#1high+TNFα
Z-scores(Log2)
F
FBOX32
FABP3
FDFT1
GPX3
ELOVL6
HMGCR
DHCR7
HMGCS1
ADGRG1
INSIG1
ACSS2
IRS2
ACSL4
SREBF1
LDLR
ABCD2
SREBF2
LPIN1
TMEM97
LSS
STARD4
MEF2C
SELENOP
SCD
MSMO1
SC5D
MVD
RETREG1
PLA2G6
NUPR1
PNPLA3
G
UMAP DEG MGT#1 acti vation
HuS
IR
Control
LIF
Activin A
C20 hMG
OxLDL
TN Fα
NOC-18
UMAP2
UMAP1
H
Intersection Si ze
68
33
29
21
20
20
16
13
11
11
9
9
8
7
7
6
6
6
5
5
OxLDL
Activin A
IR
NOC-18
C20 hMG
TN Fα
LIF
HuS
Set Si ze

Figure 10 - Innate immune cells drive non-cell autonomous mesenchymal commitment in tumor cells.

(**A**) Schematic illustration of hGICs co-cultured with immune cells in the absence of direct interaction (see Methods). (**B**) Representative FACS profiles of IDH-wt or IDH-mut-hGICs-MGT#1 cells cultivated alone or in co-culture with human microglia (hMG#cl.20) or human CD34+ myeloid-derived suppressor-like cells produced *in vitro* (MDSCs). (**C**) Representative FACS profiles of IDH-wt-hGICs-MGT#1 cells grown alone or in co-culture with human THP1-derived M1 or M2 macrophage-like cells for the specified time periods. (**D**) Representative FACS profiles and gating technique of IDH-wt-hGICs-MGT#1 cells cultured alone or in the presence of TNFa or co-cultured with hMG. Venn diagram of NFkB-related genes via Ingenuity Pathway is shown below. The DRGs are calculated in comparison to the control hGICs (log2FC>1, padj<0.05). (**E**) Heatmap of the DRGs associated with the specified conditions. Reads from RNA-seq were normalized to the number of transcripts per million, log2 transformed, and z-scored. The limma R package was used to determine statistical significance (control, n=3, hMG, n=3, TNFa, n=2; padj<0.05). (**F**) Ingenuity Upstream Regulator Analysis of genes up-regulated in IDH-wt-hGICs-MGT#1-high cells by hMG co-culture against TNFa. (**G**) Dimensional reduction using UMAP of the MGT#1 activation triggers expression profiles by aggregating all up-regulated genes (see Methods). (**H**) Upset plot of all intersections for the given MGT#1 activation cue comparison, sorted according to the magnitude of the intersections. The circles in the matrix that are interconnected represent number of genes that are found in common.

We utilized FACS to isolate and transcriptionally profile MGT#1-high expressing IDH-wt-hGICs following co-culture to discover how microglia promote a mesenchymal Glioblastoma state in a non-autonomous way. TNFα was used as a control because innate immune cells in mouse and human glioma can be a potential source for secretion of TNFα (Quail et al., 2016; Szulzewsky et al., 2016). Both cell types were transcriptionally altered as a result of the co-culture (Ext. Fig. 10B). As compared to their monocultured counterparts, C20hMG in co-culture with hGICs started to express more homeostatic core microglia markers (CSF1R, P2RY12, CX3CR1, TMEM119) and genes of glioma-associated microglia (e.g. SPP1, Sankowsi et al., 2019), suggesting mutual modulating effects of expression signatures towards a model system to study the cross-talk between innate immune cells and glioma cells (Ext. Fig. 10C). Microglia and TNFα both activate NFκB-related genes in hGICs, but each predominantly through a private collection of genes, as demonstrated by pathway analysis (Fig. 10D-E). Our approach offered no indication that TNFα induced the hMG-driven phenotype (data not shown). Instead, MGT#1-high expressing IDH-wt-hGICs triggered by microglia revealed a modification of the metabolic transcriptome that included genes involved in cholesterol production and the SREBP1/2 pathway (Fig. 10E-F).

Extracellular lipid droplets have long been known to exist in Glioblastoma (Manuelidis et al., 1961) and it has recently been shown that GBM cells display a dis-regulated cholesterol metabolism, are incapable of energy-intensive de novo cholesterol synthesis and rely on cholesterol uptake from the tumor microenvironment via low-density lipoprotein receptors (LDLRs) to fuel their growth, potentially granting them to direct the surplus of metabolic capacity into invasion and proliferation (Villa et al., 2016; Pimoradi et al., 2019).


A
Control
hGICs co-culture
B
C20-hMG+hGICs
C20MG_R1
C20MG_RHBA-MG
C20-hMG+hGICs_R1
Garcia-Mesa et al., 2017
(SRR3546647)
IDH-wt-hGICs-MGT#1high
+C20-hMG_R2
IDH-wt-hGICs-MGT#1high+TNFa
IDH-wt-hGICs
PC2 (89%)
PC1 (68%)
C
C20hMG Control
GSC medium
hGICs co-culture
D
IDH-wt-hGICs
IDH-mut-hGICs
9%
22%
15%
51%
13%
28%
29%
72%
13%
35%
33%
83%
day3 Control
oxLDL
day5 Control
oxLDL
day7 Control
oxLDL
% of total
MGT#1-mVenus
E
+TNFα
+hMG.cl.20 + NOC-18
+hMG.cl.20 + L-NAME
+hMG.cl.20
+NOC-18
Control
Count
MGT#1-mVenus
F
progeny score
-2 -1 0 1 2
-2.0
0.5
1.5
0.0
1.0
2.0
OxLDL
C20hMG
LIF
TNFα
NOC-18
IR
Human Serum
Control
Activin A
EGFR
Hypoxia
JAK.STAT
MAPK
NFkB
PI3K
TGFβ
TNFα
Trail
VEGF
p53

Extended Figure 10 - Innate immune cells drive non-cell autonomous mesenchymal commitment in tumor cells.

(**A**) Brightfield images of C20hMG monoculture or co-culture with hGICs. (**B**) Principal component analysis of the RNA-seq profiles indicated. The distances were determined using the average level of expression of selected human MG markers collected from (Gosselin et al., 2017). (**C**) Heatmap of the C20hMG expression profile of microglia-specific marker genes under specified culture conditions. The arrows denote canonical markers of human brain homeostatic microglia. (**D**) Representative FACS profiles of IDH-wt and IDH-mut-hGICs-MGT#1 treated for the given time with 15µg/ml oxLDL or control. (**E**) Representative FACS profiles of IDH-wt-hGICs-MGT#1 after 48 hours of treatment as indicated. In co-culture experiments, hMG cells were pre-treated with the indicated drugs for 24 hours prior to seeding of IDH-wt-hGICs-MGT#1. At the time of seeding IDH-wt-hGICs for co-culture, the drug treatments were renewed. (**F**) PROGENY was used to conduct a pathway analysis of condition-specific genes for each MGT#1 activator.

Oxysterols represent enzymatically modified forms of Cholesterol and regulate the activity of liver X receptors (LXRs), which in turn maintain cholesterol homeostasis through modulation of LDLR expression for Cholesterol uptake and ABCA1/ABCG1 Cholesterol efflux pumps (Ahmad et al., 2019). Consequently, we treated our cells with oxidized Low-Density Lipoprotein (oxLDL), which has been shown to activate the SREBP1/2 pathway through PPARγ. MGT#1 activation was triggered by OxLDL in both IDH wild-type and mutant models (Ext. Fig. 10D).

Furthermore, we evaluated whether the nitric oxide (NO) production route is capable of triggering mesenchymal trans-differentiation. Endothelial cell NO activity has previously been associated with glioma development and invasiveness, and upregulation of the NO pathway activity is a frequent hallmark of innate immune cell activation (Charles et al., 2010). NOC-18, a NO donor, similarly induced MGT#1 expression to levels equivalent to those induced by hMG or TNFα. In contrast, neither incubation with anti-LDLR antibodies nor the NOS inhibitor NG nitroarginine methyl ester (L-NAME) was able to prevent microglia from inducing MGT#1 (Ext. Fig. 10E). This suggests that innate immune cell-driven phenotypic alterations in glioma may be regulated by many concurrent processes. Following that, we conducted gene expression profiling of MGT#1-expressing IDH-wt-hGICs governed by oxLDL and NOC-18 and compared the results to all of the MGT#1-high expression profiles obtained earlier. In this way, we were able to explore common features and investigate how distinct upstream signaling inputs are incorporated into the transcriptional response identified by our reporter. Although each form of upstream signaling resulted in a distinct transcriptional output, we discovered that the microglia-induced mesenchymal transition shared characteristics with oxLDL, NOC-18, TNF, and - to a lesser degree - serum and LIF (Fig. 9G-H and Ext. Fig. 10F).

Collectively, our findings demonstrate a causal connection between the crosstalk of innate immune cells and glioma cells *in vitro* and mesenchymal trans-differentiation.

2.7 *Therapeutic implications of innate immune cell-induced phenotypic alterations*

Resistance to widely used and clinically tested GBM therapeutic agents has been found to be closely related to the development of mesenchymal GBM (Bhat et al., 2013; Sandmann et al., 2015; Segerman et al., 2016). GBM cells with a mesenchymal identity have been shown to be more resistant to radiotherapy, which is currently used to treat GBM patients, and a proneural-mesenchymal-transition in response to chemoradiotherapy has been described (Bhat et al., 2013; Halliday et al., 2014). Moreover, the tumor microenvironment plays an important role in overcoming selective forces during tumor growth and therapeutic stress adaptation. Tumor-associated macrophages and microglia (TAMs) are important constituents of the stroma whose interactions with glioma cells are linked to tumor aggressiveness and resistance to standard therapies (Hambardzumyan et al., 2016). It will be critical to understand the impact of human tumor microenvironmental characteristics on intratumoral variability in GBM in order to create and optimize multimodal therapeutic options for patients with GBM.

To this end, we sought to explore if the microglia-driven mesenchymal state described in our study might be correlated to and may have therapeutic implications for actual GBM patients. First, we looked for enrichment of the signatures we found in microglia co-cultured cells and hGICs treated with the previously reported pro-mesenchymal triggers in Glioblastoma patient expression profiles. Patients who were classified as mesenchymal based on TCGA transcriptional subtypes, also showed high enrichment for the microglia-driven signature, as well as signatures from cells treated with TNFa, human serum, IR, NOC-18, and OxLDL (Fig. 11A). Furthermore, both TNFα- and the microglia driven-signatures scored high for mesenchymal single Glioblastoma cells by ssGSEA (Fig. 11B). Microglia exposure, in particular, seemed to affect the expression of DNA damage and cell cycle genes in Glioblastoma cells (Fig. 11C-D).

In order for standard-of-care actions to succeed, Glioblastoma cells must engage DNA damage responses and proliferate (Stupp et al., 2009). In other words, the microglia-driven program discovered here may have major therapeutic implications, in a way that Glioblastoma cells exposed to microglia react differentially to therapies. To test our prediction, we FAC-sorted hGICs according to high or low MGT#1 expression and subjected them to a series of conventional and targeted chemotherapeutics. Remarkably, both IDH-wt- and IDH-mut-hGICs expressing MGT#1 demonstrated increased resistance to DNA damage response-based treatments when compared to their sLCR-low and naive counterparts (Olaparib, ATR inhibitor VE-821, Topotecan, Mitomycin C; Fig. 11E-F).

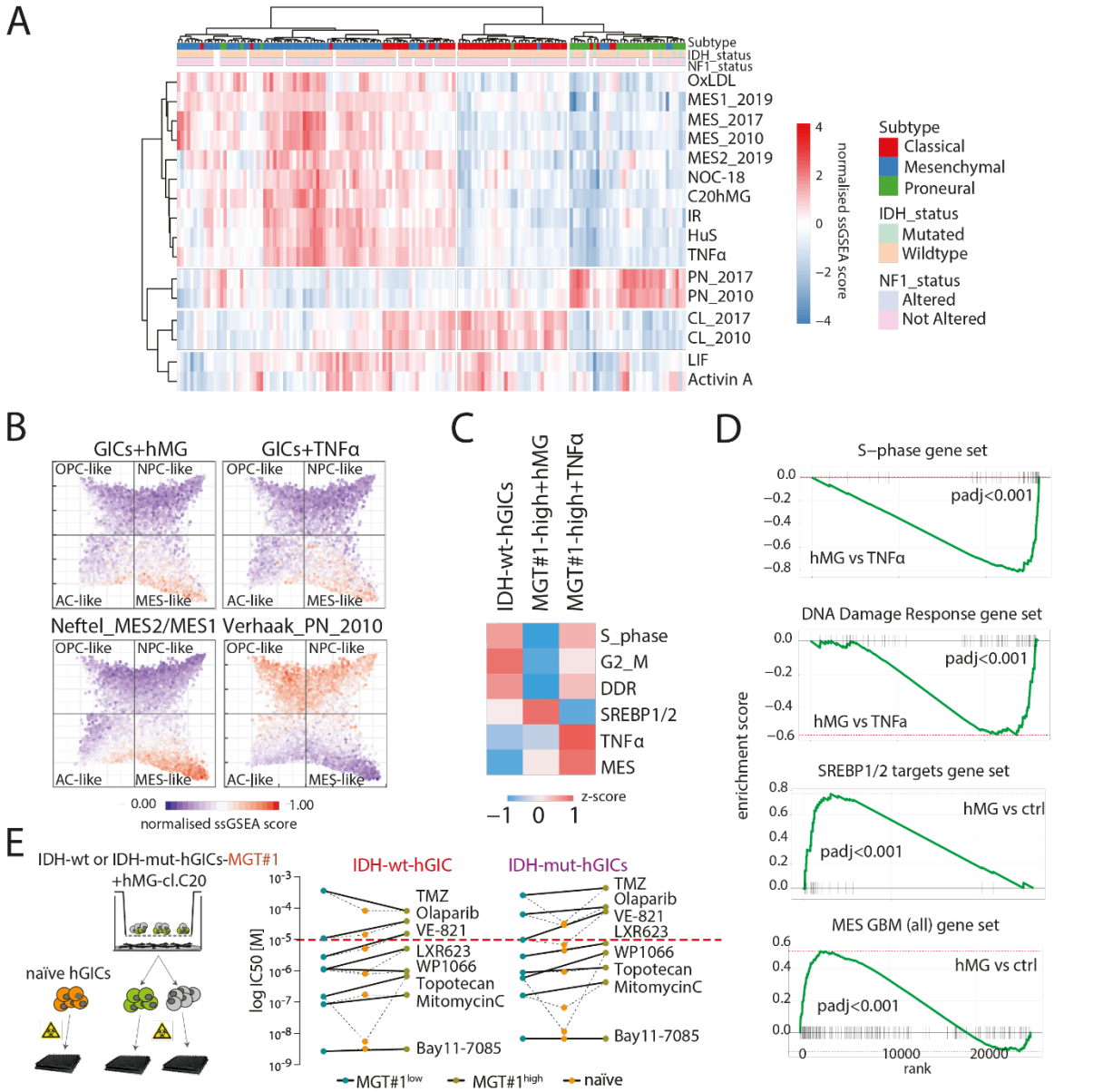

A
Subtype
IDH_status
NF1_status
OxLDL
MES1_2019
MES_2017
MES_2010
MES2_2019
NOC-18
C20hMG
IR
HuS
TNFα
PN_2017
PN_2010
CL_2017
CL_2010
LIF
Activin A
normalised ssGSEA score
Subtype
Classical
Mesenchymal
Proneural
IDH_status
Mutated
Wildtype
NF1_status
Altered
Not Altered
B
GICs+hMG
GICs+TNFα
OPC-like
NPC-like
AC-like
MES-like
Neftel_MES2/MES1
Verhaak_PN_2010
normalised ssGSEA score
C
IDH-wt-hGICs
MGT#1-high+hMG
MGT#1-high+TNFα
S_phase
G2_M
DDR
SREBP1/2
TNFα
MES
z-score
D
S–phase gene set
DNA Damage Response gene set
SREBP1/2 targets gene set
MES GBM (all) gene set
padj<0.001
hMG vs TNFα
hMG vs ctrl
enrichment score
rank
E
IDH-wt or IDH-mut-hGICs-MGT#1
+hMG-cl.C20
naïve hGICs
IDH-wt-hGIC
IDH-mut-hGICs
log IC50 [M]
TMZ
Olaparib
VE-821
LXR623
WP1066
Topotecan
MitomycinC
Bay11-7085
MGT#1low
MGT#1high
naïve
F
IDH-wt-hGICs-MGT#1
Viability [%]
log[M] Topotecan
log[M] Mitomycin C
log[M] Olaparib
log[M] Temozolomide
log[M] LXR623
log[M] Bay11-7085
log[M] VE-821
log[M] WP1066

Figure 11 - Therapeutic implications of phenotypic changes in glioma initiating cells driven by innate immune cells.

(**A**) Heatmap displaying the relative ssGSEA normalized score for the given gene sets in patients with TCGA Glioblastoma. Gene sets showing various GBM subtypes/states are included, as are up-regulated MGT#1 activation triggers (Fig. 10 G-H). Additionally, the IDH1 and NF1 mutation statuses, as well as the corresponding GBM subtypes, are highlighted. (**B**) ssGSEA-normalized scores for up-regulated MGT#1-high genes as shown in Fig. 10D. (see Methods). The cell states specified by (Neftel et al., 2019) are represented in each quadrant, and the original dot positions are preserved in the two-dimensional representation of GBM cell states (or meta-modules; Methods). (**C-D**) GSEA with differential scores for the indicated comparisons. Students t-test and Kolmogorov-Smirnov tests are employed separately to determine significance. (**E**) Left, schematic illustration of sLCR chemosensitivity profiling in the sLCR high and low states. On the right, the dot indicates the log[IC50] value for FAC-sorted MGT#1-high and MGT#1-low fractions of the specified genotypes in response to increasing drug concentrations. The dotted line indicates the 10μM concentration threshold, which is not achievable in brain tissue. (**F**) Concentration-dependent dose-response curves for MGT#1-high, -low, or naive IDH-wt-hGICs subjected to increasing concentrations of various chemicals as outlined in (E).

Our observations imply that when innate immune cells drive Glioblastoma cells to activate cholesterol production, the therapeutic advantage of reducing cell-intrinsic cholesterol levels may be lost. Importantly, the aforementioned differential response to chemotherapeutics was selective, since IDH-wt and IDH-mut-hGICs exhibited a comparable dose-response to targeted treatments such as BAY11-7085 (IκB inhibitor) and WP1066 (STAT3 inhibitor) regardless of MGT#1 expression status (Fig. 11E-F).

In conclusion, synthetic genetic tracing demonstrated a causal relationship between glioma cells undergoing a mesenchymal cell fate transition imposed by innate immune cells and the development of resistance to DNA-damage-based treatments.

2.8 Logical design of sLCRs for complex cellular states beyond GBM

To generate sLCRs in a more robust and automated way, we have developed a computational framework, which employs an algorithm for logical synthetic reporter design (LSD). From a pool of CREs which is derived from the two lists of differentially regulated signature genes and transcription factor genes, the algorithm scores, ranks and selects the top listed CREs in order to maximise the total amount and diversity of all identified TFBS for the final selection of the reporter (Fig. 12A). At this stage, we envision to introduce in the future ATAC-seq, ChIP-seq or other chromatin-accessibility profiles in order to delineate cis-regulatory DNA-boundary regions or filter the signature genes for cell-type specific TFBS accessibility.

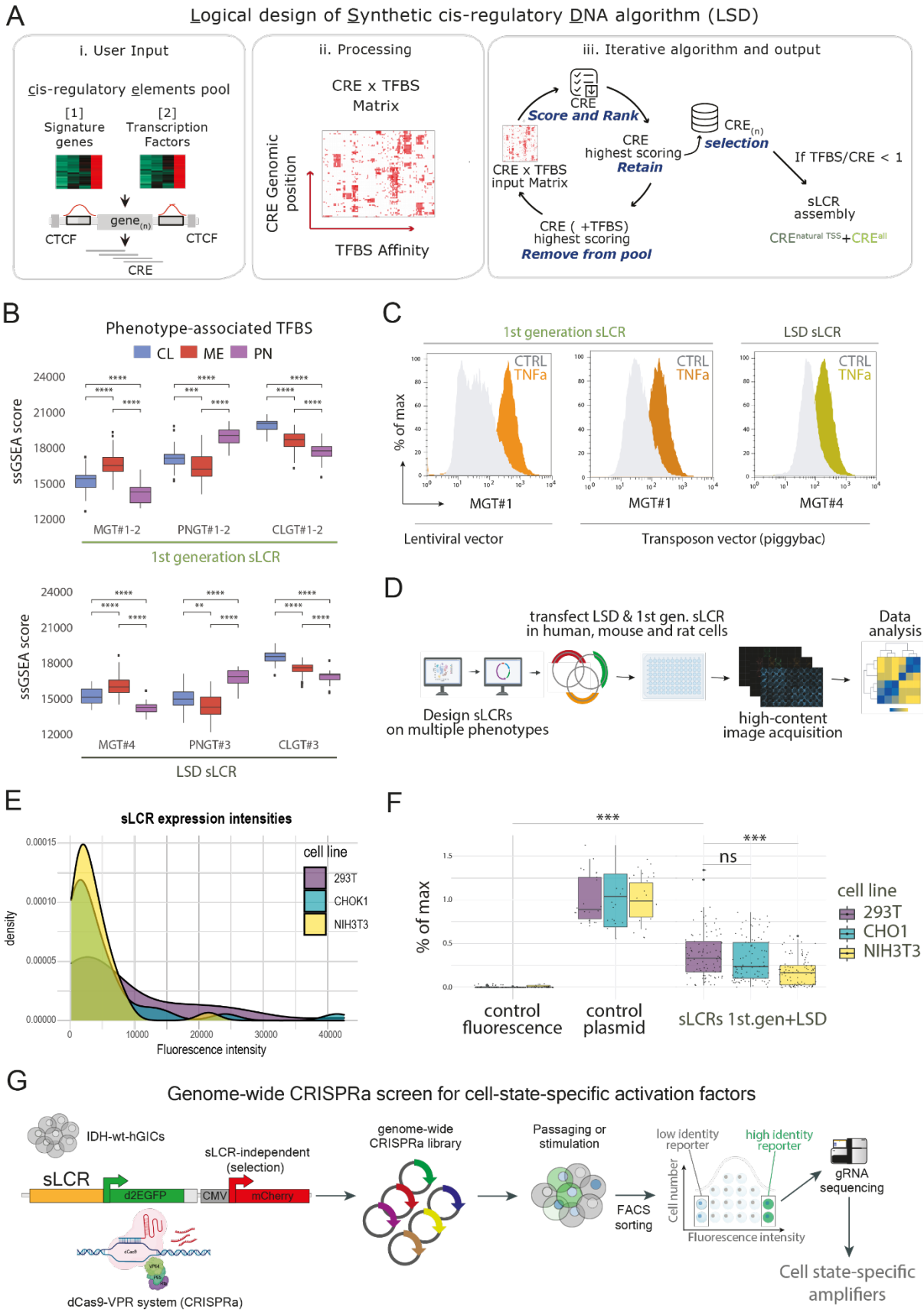

A
Logical design of Synthetic cis-regulatory DNA algorithm (LSD)
i. User Input
cis-regulatory elements pool
[1] Signature genes
[2] Transcription Factors
gene(n)
CTCF
CTCF
CRE
ii. Processing
CRE x TFBS Matrix
CRE Genomic position
TFBS Affinity
iii. Iterative algorithm and output
CRE
Score and Rank
CRE x TFBS input Matrix
CRE highest scoring
Retain
CRE(n)
selection
If TFBS/CRE < 1
sLCR assembly
CRE natural TSS+CRE all
CRE (+TFBS) highest scoring
Remove from pool
B
Phenotype-associated TFBS
CL ME PN
ssGSEA score
MGT#1-2
PNGT#1-2
CLGT#1-2
1st generation sLCR
MGT#4
PNGT#3
CLGT#3
LSD sLCR
C
1st generation sLCR
LSD sLCR
CTRL
TNFa
% of max
MGT#1
MGT#1
MGT#4
Lentiviral vector
Transposon vector (piggybac)
D
transfect LSD & 1st gen. sLCR in human, mouse and rat cells
Data analysis
Design sLCRs on multiple phenotypes
high-content image acquisition
E
sLCR expression intensities
cell line
293T
CHOK1
NIH3T3
density
Fluorescence intensity
F
% of max
ns
cell line
293T
CHO1
NIH3T3
control fluorescence
control plasmid
sLCRs 1st.gen+LSD
G
Genome-wide CRISPRa screen for cell-state-specific activation factors
IDH-wt-hGICs
sLCR
d2EGFP
sLCR-independent (selection)
CMV
mCherry
dCas9-VPR system (CRISPRa)
genome-wide CRISPRa library
Passaging or stimulation
FACS sorting
low identity reporter
high identity reporter
Cell number
Fluorescence intensity
gRNA sequencing
Cell state-specific amplifiers

Figure 12 - Automation via logical sLCR design algorithm streamlines unbiased reporter generation.

(**A**) Diagram illustrating the LSD sLCR design method. Steps (I-III) show sLCR assembly by iterative selection of the highest-ranked CREs from the CRExTFBS matrix (see Methods). Each iteration eliminates the CRE and TFBS with the highest score from the CRExTFBS matrix until there were no TFBS or CRE remaining in the CRExTFBS. Using the ranking list, N CREs are assembled with priority for the CRE nearest to a natural TSS. (**B**) ssGSEA comparison of Glioblastoma subtype-specific TF selection. Top panel presents the input to the first generation sLCR TF, while bottom panel evaluates the input to the LSD-sLCR TF. Boxplots provide the score for the expression profiles of Glioblastoma patients (TCGA-GBM) categorized by their identified transcriptional subtype. Wilcox.test (adj.pvalue < 0.05) was used to compare Glioblastoma subtypes for each TF gene-set. (**C**) FACS profile of TNFa-stimulated reporter activation of lentiviral- and transposon-mediated reporter integration of MGT#1 (1st generation sLCR) and MGT#4 (LSD-sLCR) in hGICs after 24h. Note similar induction for 1st gen. MGT#1 sLCR and LSD-generated MGT#4. (**D**) Schematic of the systematic screening of sLCRs designed on diverse phenotypic signatures and tested in three difference species. (**E**) Density plot for fluorescence intensities of sLCRs (n=28), transfected in human epithelial 293T, rat epithelial CHO-K1 and mouse fibroblast NIH3T3 cell lines. (**F**) Box plot of sLCR activities, defined as fluorescence normalized by controls and control plasmid transfection efficiency per cell line. Left, activity scores for negative untransfected controls (n=9) and positive control plasmids (n=3) including GFP, mCherry and iRFP670 channels. Right, aggregate plot shows relative activity of human sLCR transfected in human (293T) or non-human (CHO-K1, NIH3T3) cells. Each sLCR measurement was assessed in technical triplicates. Statistical significance was calculated with 2-way ANOVA followed by Dunnet post-hoc test. (**G**) Schematic representation of a CRISPR activation screen utilising sLCRs to identify cell state-specific modulators.

We generated the second generation of GBM subtype-specific sLCRs MGT#4, PNGT#3, and CLGT#3 using the same list of signature genes and a subset of subtype-specific TFs (high-expressed TF genes, > quantile 75 percent) derived from TCGA-GBM patients' RNA-seq expression profiles. Single-sample gene set enrichment analysis (ssGSEA) using lists of TFs associated with either subtype on GBM molecular profiles of TCGA patients suggests that both sLCR reporter systems hold the potential to be transcriptionally activated and thus being representative of their target phenotype (Fig. 12B).

Experimentally, we tested side-by-side the inducibility of LSD-generated and first generation mesenchymal-GBM reporters through TNFa mediated NFkB-pathway activation, which we have shown to be a well-characterised mesenchymal trigger. Both, the first generation MGT#1 and the LSD-sLCR MGT#4 were strongly induced upon 48h treatment with TNFa, independent of the delivery system used for genetic engineering (Fig. 12C).

Next, we wanted to test whether our pipeline allows the systematic generation of functional reporters for a broad range of cellular states and target phenotypes. For this purpose, we generated a total of 28 sLCR targeting multiple biological settings (e.g. ER stress response, cellular senescence, T-cell exhaustion, glioma-associated microglia reprogramming and SARS-CoV2 entry) using the LSD-algorithm. The signature gene and transcription factor lists, as well as the chosen fluorescent protein, selection markers and transgene-delivery systems varied significantly among this pool of reporters. We transiently transfected all sLCRs into human 293T, mouse NIH3T3 and

hamster CHO-K1 cells in order to assess their transcriptional activity and functionality in a setting of diverse tissue- and species-background (Fig. 12D). Most of the sLCRs displayed a low intensity of expression, although their overall activity was significantly above the background. Interestingly, although the average sLCR intensity and activity is higher in human 293T as compared to "off-target" cell lines, the non-human cells still displayed measurable reporter activity despite the fact that a different genome and TFBS annotation was used to construct them (Fig. 12E-F). This suggests the notion that sLCRs generated with our algorithm are functional, responsive to their targeted transcriptional signature and robust enough to show activity in a tissue- and species-independent manner without the need for a specific promoter, given that parts of the regulatory network associated to the target signature are sufficiently conserved.

The manifestion of a target phenotype is accompanied by the acquisition of a transcriptional signature, which can be measured and serves as input for sLCR generation. However, detailed understanding of the underlying transcriptional networks is not necessary for this purpose, yet often holds the key for the development of therapies targeting the cellular state of interest, as in the case of GBM. For this reason, sLCRs can be used to uncover cell state-specific amplifiers when combined with unbiased genetic screens, such as CRISPRa. Reporter-bearing cells, modified with the dCas9-VPR system and infected with a genome-wide CRISPRa library can be either maintained for an unbiased approach or selectively stimulated with putative modulators until reporter-expressing cells are FAC-sorted for sequencing and identification of enriched sgRNAs. Coupled with the use of patient-derived data, such type of screens may likely help to identify key amplifiers of the targeted cellular state and can likewise be used to counter-screen for molecules to disrupt its transcriptional activation (Fig. 12G).

3 Discussion

Large-scale genomic and transcriptome profiling of GBMs revealed an extraordinary degree of inter- and intratumoral heterogeneity and outlined critical features of GBM cells, such as intrinsic plasticity and reversible adaptation to variable microenvironmental settings. Multiple attempts to resolve this heterogeneity and establish a classifier for transcriptomic subtypes of GBMs converged on a fluid continuum of cellular states, in which tumor cells can readily interconvert along several axes of cellular properties, such as proliferation, stemness, differentiation, inflammation, neurodevelopmental programs, metabolism and the tumor microenvironment.

In this context, the resulting phenotypic plasticity has been considered an essential source of intra-tumoral heterogeneity, paving one potential way for treatment resistance and establishing a disease that is highly adaptive and persistent.. In the end, it is unlikely that treatment failure for GBM can be explained by a single mechanism of therapy resistance; rather, numerous simultaneous processes may be operating, with the most beneficial mechanism being selected for.

Single-cell technologies and longitudinal tracking of molecular features from primary and recurring tumors greatly expanded our understanding of these mechanisms. They provided valuable insights about the situation in the patient at the time of sampling but still lack the ability to follow in real-time the dynamic modalities of phenotypic adaption via cell fate transitions.

In an attempt to bridge this technological gap, we developed an approach towards designing synthetic locus control regions that translate multifactorial signatures representing cellular states into fluorescent tracing vectors that offer the possibility to selectively identify cellular states and follow cell fate transitions in real-time and without perturbations.

Employing lineage tracing reporters has generated a great advance in understanding of cell fate decisions and clarifying specific questions in disciplines like developmental or stem cell biology. In the field of cancer biology, lineage tracing has been proven a powerful tool and successfully applied to trace the contribution of individual tumor cells to tumor formation and relapse, although the scope remains limited to study very specific mechanisms (Chen et al., 2012; Driessens et al., 2012; Schepers et al., 2012).

A common strategy for lineage tracing reporters is the labelling of a phenotype-specific signature gene, whose expression is correlated with the cellular fate of interest. The validity of the underlying

dependency of single marker gene expression to target phenotype is an evident constraint in this design.

Therefore, this approach is first of all limited to cell types for which one reliable and highly expressed marker gene can be identified and generally targets only one cell type at a time at its terminal differentiation endpoint. Linking the destiny of reporter-functionality to the expression or absence of one particular gene on one hand comes with the risk of either false-positive labelling in off-target cells, which might ectopically express this gene or false-negative labelling in case the marker genes expression is repressed. On the other hand, losing out on intermediate cell states along the fate trajectory, which do not yet express sufficient levels of the chosen marker gene, greatly diminishes such reporters' ability to trace the longitudinal dynamics underlying cell fate commitment. Last but not least, designing a lineage tracing reporter founded on a single-gene-to-phenotype paradigm bears the possibility to rather be capturing the regulatory mechanism of this gene instead of the targeted cell fate.

With our approach, we intend to turn the full regulatory network of cis-regulatory elements within a given target-phenotype expression profile into a synthetic promoter-like sequence to drive our reporter, rather than a single gene. This enables us to generate reporters for cases in which no reliable single marker gene can be identified and, in general, to target distinct cellular entities as well as more complex signaling pathways (e.g. endoplasmic reticulum (ER) stress or senescence), which can be transient or context-dependent and have a characteristic expression profile.

However, we are acknowledging that certain challenges come along with the reporter design, which we are currently still further optimising. Firstly, the quality of the input transcriptomic profiles and specificity to the target phenotype are crucial for the pipeline to reliably identify signature genes and transcription factor genes (Fig. 5A). We expect a certain amount of unspecific marker genes to be tolerated by our method, although a larger number would possibly affect the reporter specificity, since an increase in the amount of unspecific CREs identified likewise increases the risk of such elements being included in the final output and skewing the reporter towards a false-positive induction.

A systematic assessment on the possibility of compensatory mechanisms for the robustness of input data and, more importantly, under which circumstances this would be the case needs to be conducted in the future. Second, in order to generate the list of cis-regulatory elements, their existence and the functional annotation of TFBS in publicly available databases as well as those experimentally derived or computationally inferred needs to be properly curated, which is in many cases still suboptimal and incomplete. Third, the experimental restraintson the length of a synthetic promoter due to technical feasibility of the transgene delivery system, time or budget limits may

restrict the choice of CREs that can be included. In this case, since not all regulatory elements from a typically large list (Ext. Fig. 5A) can be included, a selection based on reasonably defined parameters needs to be made, where weighting the importance of each feature can be difficult and may also depend on the experimental context. Ideally, this selection should be a maximal coverage of all important CREs and be representative of the whole cistrome of the target phenotype. The current development of unsupervised or minimally-supervised sLCR selection algorithms, in combination with methods for massively parallel reporter testing such as STARR-seq, will be helpful in the future to pre-screen for active regulatory sequences and guide sLCR design (Arnold et al., 2013; Muerdter et al., 2015).

In the case of GBM molecular subtypes, we succeeded in designing functional reporters, which reliably indicated their target phenotypes in silico and *in vitro* (Fig. 5F & Fig. 6). Our human glioma-initiating cell model is cultured under serum-free conditions, considered to be favourable for proliferation and maintenance of stemness of glioma stem cells (GSC; Singh et al., 2004; Lee et al., 2006). They express low levels of our mesenchymal reporter and display a strong inherent proneural identity *in vitro* under basal stem-cell culture conditions and in the absence of a tumor microenvironment. When provided with external signaling related to inflammation or differentiation cues, tumor cells rapidly and transiently acquired MES-features.

Our observations support a concept of GSC-multiplicity in which, rather than maintaining a stable hierarchy of tumor cell identities, glioma cells undergo a proneural-mesenchymal-transition and are also capable of reverting to their initial state (Fig. 4). The ability to directly trace the adaptive and, more importantly, reversible mode of this mechanism further underscores phenotypic plasticity of GBM cells fluidly interconverting between states and the power of our approach.

Of note, we observed that hGICs with the IDHR132H mutation appeared to have a generally attenuated response to pro-mesenchymal triggers, supporting the globally observed transcriptional reduction through hypermethylation of the IDH-mutant and its role as a biomarker for patient stratification (Han et al., 2020). Taken further, our observations of a generally dampened induction of a mesenchymal programme across conditions in IDH-mutant cells could provide experimental support for the general association of IDH-mutations to the proneural subtype (Noushmehr et al., 2010; Verhaak et al., 2010; Brennan et al., 2013).

Our reporters are well-suited to disentangle the individual contributions of important mechanisms for tumor initiation and progression to cell fate commitment and gain insights into their molecular

basis by systematically screening for cell fate transitions and characterizing their transcriptional responses.

The mesenchymal GBM correlates with a relatively higher aggressiveness compared to the other transcriptional subtypes, is often acquired upon recurrence and was linked to the application of radiochemotherapy (Phillips et al., 2006; Carro et al., 2010; Bhat et al., 2013; Segerman et al., 2016; Wang et al., 2017). There are numerous reports on phenotypic plasticity and proneural-to-mesenchymal transition (PMT), and it has been established that both external and intrinsic cellular mechanisms, which are either microenvironment-driven or -independent, play a role in mesenchymal transition (Bhat et al., 2013; Halliday et al., 2014; Behnan et al., 2019; Azam et al., 2020; Kim et al., 2021).

The various pro-mesenchymal cues we were able to identify and transcriptionally profile add along those lines. We discovered that a subset of cells undergo mesenchymal transition when exposed to a complex microenvironment *in vivo* and that the upstream regulators of this PMT are predominantly involved in inflammatory pathways and may be regulated by members of the AP-1 transcription factor complex. This is largely consistent with the mediators of mesenchymal cell fate transition we have identified in our phenotypic *in vitro* screen, and comparative analyses suggest that the signature obtained from *in vivo* mesenchymal tumor cells shared features with the transcriptional output from the triggers in the screen, which when combined appear to converge over the *in vivo* profile. This data indicate that the mesenchymal state that emerges *in vivo* is most likely a result of a combination of different external stimuli originating from distinct compartments of the tumor microenvironment, and that multiple effectors can orchestrate the acquisition of the same cellular state via multiple mechanisms.

In terms of cellular characteristics, pathological features, and transcriptome changes, PMT in GBM is similar to the epithelial-mesenchymal transition (EMT) observed in other solid tumors. EMT occurs, when epithelial tumor cells lose their properties (cell–cell adhesion and polarity) and gain more aggressive mesenchymal characteristics (Dongre & Weinberg, 2019). Supporting this notion, we observed that our mesenchymal-GBM MGT#1 reporter was able to discriminate mesenchymal identities in a tissue-independent manner and, in particular, trace the process of TGFb-induced EMT in lung cancer cells (Serresi et al., 2021). Combined with the fact that the mesenchymal GBM-subtype is the only one that does not resemble neurodevelopmental programmes and is (together with the proneural subtype) the most robustly identified molecular signature across studies and technologies, mesenchymal transition exhibits traits of a pan-cancer programme that

shares a common root of underlying molecular programmes, which we were able to extract and transform into a broadly applicable mesenchymal reporter with our method.

GBMs inevitably become resistant to treatment, which has been linked in part to the development of a mesenchymal signature in recurrent tumors. Numerous studies have been conducted on the association between proneural-mesenchymal-transformation and radiotherapy, which is currently a component of the standard treatment (Phillips et al., 2006; Bhat et al., 2013; Halliday et al. 2014; Minata et al., 2019; Behnan et al., 2019). Such radiotherapy-related alterations have been shown to contribute to a poor response of relapsing tumors and consequently poor patient prognosis.

Hypoxia, a hallmark of the tumor microenvironment in solid tumors including GBM, can also cause mesenchymal transition in a variety of malignancies via modulating hypoxia-inducible factors (HIF) and proteins (Tam et al., 2020). Hypoxic cells are transcriptionally remodelled by stabilization of HIF-1/2α, which activates pro-tumorigenic and -angiogenic factors such as VEGF/VEGFR, TGFβ, and PDGFR. Multiple studies have reported an enrichment of hypoxic signatures in mesenchymal GBM tumors (Patel et al., 2014; Neftel et al., 2019; Wang LB et al., 2021).

Genetic tracing using our mesenchymal reporter established a causal link between radiation-induced DNA damage and PMT in our hGICs. In contrast, as an example of how our method allows to decouple causality and correlation, hypoxia did not trigger the activation of our mesenchymal reporter in a cell-intrinsic manner. This suggests that the transcriptional programme of hypoxia-response, which we have detected in our experiments, might not directly drive mesenchymal transition. One possible explanation is that hypoxia rather acts in conjunction with further external factors in a non-cell autonomous manner, for instance via the recruitment of innate immune cells. On the other hand, the detection of hypoxia-related signatures might be a secondary effect of tumor cells adapting to a glycolytic metabolism, which has been strongly associated with the mesenchymal state and regulated by HIF1a (Del Rey et al., 2017; Miska et al., 2019; Garofano et al., 2021). The aforementioned observations on mesenchymal GBM raise a serious concern about the possible therapeutic risk associated with both existing standard-of-care and new treatment approaches and highlight the need to understand causal connections between molecular processes, biological features such as metabolic programmes and phenotypic alterations (Lasorella & Iavarone, 2021).

Importantly, using our sLCRs, we were able to establish another causal relationship between mesenchymal transformation and the interaction with innate immune cells. While many reports could detect a correlation between the mesenchymal subtype and the presence of a immune-cell related signature, direct evidence of an immune-cell mediated mesenchymal commitment has remained elusive (Wang et al., 2017; Neftel et al., 2019; Wang LB et al., 2021).

We were able to show that innate immune cell-like cells of the myeloid lineage, such as microglia and macrophages, engaged in a bi-directional crosstalk with tumor cells, which triggered the expression of our mesenchymal sLCR and adoption of a mesenchymal gene expression profile in a non-cell-autonomous way. This signature featured in part the transcriptional amplification of TNFa-mediated NFKB-pathway activation but also a distinct alteration in the metabolic pathway of cholesterol biosynthesis via SREBP1/2.

Most importantly, we were able to show that the microglia-driven mesenchymal commitment is similar to mesenchymal states found in GBM patients and might have therapeutic implications as it renders tumor cells more resistant to chemotherapy. We observed a selective resistance towards compounds targeting DNA-damage response and replication, which was consistent with a downregulation of these pathways on the transcriptional level detected in co-cultured mesenchymal glioma cells (Fig. 11C-F).

Cholesterol biosynthesis has been proposed in several studies as a potential target, as GBM cells were reported to be deficient in producing their own Cholesterol and relying on external sources for survival (Villa et al., 2016; Pirmoradi et al., 2019; Ahmad et al., 2019; Li et al., 2020). Small molecule inhibitors targeting Liver X receptors, which control Cholesterol levels in the cell by inhibiting uptake through reduced LDLR and promoting efflux through increased ABCA1 transporter expression, have been developed and tested with some success on GBM cells *in vitro* (Villa et al., 2016). In our experiments, strong effects have not been observed, yet the data points towards a potential decrease of sensitivity towards LXR-623 (an agonist of Liver X receptor) of mesenchymally-converted glioma cells from microglia co-culture.

To what extent the interaction with innate immune cells is driving a metabolic rewiring in our glioma cell model and whether these changes are directly associated with the mesenchymal programme and therapy-resistance we have observed still remains to be explored. Interestingly, a recent approach to subtype GBM according to biological pathways revealed a bifurcation into a neurodevelopmental/proliferative and a metabolic axis, which was subdivided into a mitochondric (MTC) and glycolytic/plurimetabolic (GPM) component. In this study, the GPM subtype showed significant alterations in glycolysis, carbohydrate and lipid metabolism and was related to the mesenchymal-GBM signature and immune-related functions (Garofano et al., 2021). This appears

in line with a second study, reporting on GSCs' falling along two major axes of variation, comprising of a neurodevelopmental trajectory and an injury response programme. The latter signature also showed overlap with the mesenchymal-GBM signatures, and a genome-wide CRISPR-Cas9 dropout screen found injury-response GSCs to be dependent on genes related to glycolysis and lipid metabolism (Richards et al., 2021).

We must acknowledge the limitations of our experimental strategy, which involves the use of immortalized immune-cell-like cultures that may not retain all of the physiological features of their primary counterparts, although we tried compensating this by the fact that we are using multiple models of immune-like cells of different origin and could show similar results between THP1-derived macrophages and immortalized human microglia (Fig. 10B-C). A recent study has confirmed our findings by showing that macrophages induce a mesenchymal transition of Glioblastoma cells *in vivo* and *in vitro*, likely through macrophage-derived Oncostatin M via STAT3 signaling (Hara et al., 2021). The external signaling-driven highlighted mechanism in this study is different to ours, although we clearly showed that multiple routes exist to the acquisition of a mesenchymal state and on the other hand, also targeting a single mechanism for abrogation of pro-mesenchymal modulation is insufficient (Ext. Fig. 8b). Furthermore, Hara and colleagues do not provide insights into the cell-intrinsic consequences of glioma cells, especially on the aspect of activated metabolic programmes (e.g. glycolysis and lipid metabolism), which clearly are important and cannot be disconnected from mesenchymal transdifferentiation (Garrofano et al., 2012).

Taken together, this calls for taking into account the mutual interactions of glioma cells with microenvironmental and immune cells and assembling multi-cellular test platforms in future drug screening and validation studies, which can ideally be coupled with our genetic tracing technology to monitor cell-fate transitions in real-time.

Our data demonstrate the usefulness of sLCRs to gain valuable insights into GBM biology and provides the ground for adopting this approach beyond that. In fact, we could show that the mesenchymal MGT#1-reporter, designed specifically on the mesenchymal GBM signature, appears to be equally reporting on EMT in lung cancer cells and differentiate between more or less mesenchymally committment cell lines of non-glioma origin (Ext. Fig. 6; Serresi et al., 2021) . This implies that the concept of sLCR design might be suitable for further biological settings where genetic tracing can generate insights into disease-related or developmental processes. To foster an unbiased and more systematic construction of sLCRs for a diverse range of target phenotypes, we have devised an automated algorithm for the selection CREs, maximising the regulatory potential

in an approach with reduced supervision. Reporters generated with this algorithm are in general functionally active in abscence of a specific promoter and show similar specificity as first-generation sLCRs in the case of mesenchymal GBM. Future experiments will be required to determine whether genetic tracing reporters hold the key to identifying the main regulators of cell state-specific transcriptional networks when coupled with genome-wide CRISPR-screens.

Despite an intensively growing body of research on the molecular and cellular characteristics of Glioblastoma (GBM), it remains one of the most difficult solid tumors to treat. Tumor recurrence is unavoidable, although rigorous treatment options for GBM patients exist, such as surgical resection, radiation, and chemotherapy. Since GBM displays a high level of genetic and cellular heterogeneity , as well as phenotypic plasticity, the question remains whether these properties are major contributors to its therapeutic resistance. To address this, we developed synthetic genetic tracing vectors based on human Glioblastoma subtype expression profiles to monitor cell fate transitions and investigate the fundamental molecular and cellular mechanisms driving Glioblastoma heterogeneity and therapy resistance. Using our novel technology, we were able to uncover and characterize intrinsic and non-cell autonomous determinants of cell fate commitment in glioma cells. In particular, we found the mesenchymal GBM programme emerges in the presence of signaling cues from an intact microenvironment and is regulated by inflammatory and differentiation signals *in vitro* and *in vivo*. We provide ground for the mesenchymal identity to be considered a cellular state rather than an entity, since acquisition of a mesenchymal signature is an adapative and reversible process, which can be achieved via multiple routes with partially overlapping transcriptional responses, such as various stimuli of external signaling or ionizing radiation. We established a causal association between the mesenchymal state and the interaction of glioma cells with cells from the tumor microenvironment. Mesenchymally-committed tumor cells from co-culture with innate immune cells display transcriptional alterations in cell cycle progression, DNA-damage response and cholesterol biosynthesis and were marked by selective resistance to targeted therapies. Overall, we are presenting sLCRs as an innovative approach to trace cell fate transitions in complex biological settings, such as uncovering a mechanistic role of cellular crosstalk between malignant and non-tumor cells in the context of therapy resistance and phenotypic plasticity. Beyond that, this method offers broad translational potential to be applied to other fields of research, such as developmental biology, infection biology or regenerative medicine. The presented framework represents the first generation of genetic tracing reporters and further sophistication in the process of selecting cis-regulatory elements by

including chromatin-accessibility profiles and the sLCR construction pipeline by implementation of algorithms and mathematical models will certainly enhance specificity and reliability of future reporter designs.

4 Material and Methods

4.1 GBM-sLCR generation

To narrow our focus on cell intrinsic gene signatures, we screened out genes with low expression in GBM stem-like cells (GSCs) from prior investigations and confirmed their putative intrinsic expression in a validated cohort of GSCs from others.

From the literature, we created a database of 1,818 motifs (position weight matrices, PWM) encoding known transcription factor binding preferences (Bucher et al., 1990; Berger et al., 2008; Badis et al., 2009; Jolma et al., 2010; Portales-Casamar et al., 2010). Pre-selected PWMs were based on subtype-specific TFs. The regions corresponding to DRGs were extracted from the hg19 genome (Refseq table acquired from the UCSC genome browser (accessed on October 5, 2012) and fragmented into 150bp and 50bp windows (hereafter referred to as CRE). The scanned region surrounding each signature gene was manually demarcated by two distal CTCF sites that were more than 10 kb from the TSS or TES. FIMO(PMID: 21330290) was used to identify high-affinity TF-binding sites in specific genomic areas using the —output-pthresh 1e-4 —no-qvalue option. When several matches for the same PWM were detected for each window, the adj. p-value 0.01 of the best match (multiple backgrounds) was used as a proxy for the TF's affinity towards that location.

The pairwise correlation heatmaps for TFBS in (Ext. Fig. 5B) employed the top 500 locations in terms of the -log10 score (p-value). With the top 100 scoring locations, heatmaps of genomic coordinates vs TFBS correlation were created, including the ones in (Fig. 5C).

4.2 Vectors generation

The sLCRs were initially synthesized at IDT, then at GenScript, and most recently at VectorBuilder. MGT#1-mVenus was cloned into the PacI-BsrGI segment of the Mammalian Expression, Lentiviral FUGW vector (a gift from David Baltimore; Addgene#14883). Additional alterations, such as the insertion of H2B-CFP (a donation from Elaine Fuchs, Addgene#25998), the substitution of mVenus for mCherry, and the replacement of MGT#1 with all other sLCRs, were performed via restriction enzyme digestion or Gibson cloning.

The Igk-mVenus-TM sequence was derived from (Ohinata et al., 2008). A NLS was used to modify the mCherry. The sLCRs vectors are a third-generation lentiviral system that was employed in

conjunction with the pCMV-G(Addgene#8454), pRSV-REV(Addgene#12253), and pMDLG/pRRE(Addgene#12251) vectors.

4.3 Cell lines

All lines used in this project were thawed from frozen batches and propagated for a limited number of passages (10-15x), and all lines were screened on a regular basis for contamination using the Mycoplasma Detection kit (Jena Bioscience11828383, PP-401L). Global expression profiling was used to authenticate all glioma initiating cells and GSC cell lines. Murine NIH3T3 cells were cultured in DMEM + 10% FBS + 1% Penicillin/Streptomycin. The hamster ovary-derived cell line CHO-K1 was cultured in DMEM-F12 + 10% FBS + 1% Penicillin/Streptomycin. All lines were cultured at 37°C and 5% CO2 in a humidified incubator

4.4 Human glioma cell lines

Our lab developed the IDH-wt-hGICs and IDH-mut-hGICs, which will be described in the following paragraph. IDH-mut-hGICs were produced by transfecting human NPCs with pLenti6.2/V5-IDH1-R132H (kindly donated by Hai Yan, Duke University, USA), TP53R173H and TP53R273H (point mutations incorporated into TP53 ccsbBroad304 07088 from the CCSB-Broad Lentiviral Expression Library), and pRS-Puro-s. Human NPC were transformed with the constructs pRSPURO-sh-PTEN(#1), pLKO.1-sh-TP53(TRCN0000003754), and pRS-shNF1 to generate IDH-wt-hGICs. Comprehensive genetic, transcriptional, and epigenetic characterisation, as well as in vivo tumor growth and phenotypic mimicry, have been established on these lines (unpublished data by the GG lab). Rainer Glass, LMU Munich, Germany, generously contributed the patient-derived glioma stem cell lines GBM2 (Binda et al., 2017; TCP#2), GBM14, and NCH421K (Campos et al., 2010). The lines GBM166 and GBM179 were kindly donated by Peter Dirks (University of Toronto, Canada (Pollard et al., 2009)) and the lines BLN-5 and BLN-7 (Schulze-Heuling et al., 2017) were kindly provided by Phillip Euskirchen, Charité Berlin, Germany.

All glioma lines were generated in vitro according to the protocol described (Gargiulo et al., 2013) with one modification. Along with EGF (20ng/ml; R&D, 236-EG), bFGF (20ng/ml; R&D, 233-FB), heparin (1μg/ml; Sigma, H3149), and 1% penicillin and streptomycin, PDGF-AA (20ng/ml; R&D, 221-AA) is added to RHB-A (Takara, Y40001). This composition of the medium shall be

referred to as RHB-A complete. hGICs were cultivated at 37°C in a 5% CO2–95% air incubator with 3% O2 and a humidified environment.

4.5 Cancer cell lines

The MCF7 and MDA-231 cell lines were kindly donated by the Rene Bernards lab at the NKI Amsterdam, Netherlands (Life Technologies, 21875091). Both cell lines were cultured at 37°C in a 5% CO2–95% air incubator supplied with 10% FBS and 1% penicillin and streptomycin. A549 and H1944 cell lines (kindly donated by the Rene Bernards lab at NKI Amsterdam, Netherlands) were grown in RPMI media. Both cell lines were cultured at 37°C in a 5% CO2–95% air incubator supplied with 10% FBS and 1% penicillin and streptomycin.

4.6 Human Microglia cell line

Immortalized primary human Microglia C20 cells (kindly provided by David Alvarez-Carbonell, Case Western Reserve University, Cleveland, USA (Garcia-Mesa et al., 2017)) were cultured at 37°C in a 5% CO2 incubator in RHB-A medium supplemented with 1% FBS, 2.5mM Glutamine (Thermofisher; 35050038), 1μM Dexamethasone (Sigma; D1756), and 1% penicillin and streptomycin.

4.7 Human hematopoietic progenitor CD34 differentiation

SFEM II (StemCell Technologies, 09605), SCF, FLT3-L, TPO, and IL6 (all 100ng/ml; easyexperiments.com), UM171 (Selleck, 35nM), SR1 (Selleck, 0.75μM), and 19-deoxy-9-methylene-16,16-dimethyl PGE2 (Cayman, 10μM) were used to grow donor-derived CD34+ cells. Prior to co-culture, CD34+ cells were differentiated into immature MDSC-like cells by shifting to RHB-A media supplemented with human SCF(50ng/ml) and human GM-CSF(100ng/ml) for 7-12 days.

4.8 Human Monocyte cell line differentiation

The ATCC TIB-202 human monocytic THP-1 cells were obtained from S. Minucci (IEO Milan, Italy) and cultured in Roswell Park Memorial Institute media (RPMI 1640, Thermofisher) supplemented with 10% fetal bovine serum (Gibco, 10270106), 1mM pyruvate (Life

Technologies), and 2mM GlutaMAX (Life Technologies) (Thermofisher, 35050-038). After 48 hours incubation with 150nM phorbol 12-myristate 13-acetate (PMA, Cayman Chemicals; Cay10008014), THP-1 monocytes are differentiated into macrophages.
By incubating macrophages with 20ng/ml IFN (R&D system, 285-IF) and 10pg/ml LPS, macrophages were polarized into M1 macrophages (Sigma, L2630). The polarization of macrophage M2 was induced by incubation with 20ng/ml interleukin 4 (Sigma, A3134) and 20ng/ml interleukin 13. (PeproTech, 200-13).

4.9 Transfection/Transduction

Transfection and transduction have been extensively described previously (Gargiulo et al., 2014). Specifically, 12μg of DNA mix (lentivector, pCMV-G, pRSV-REV, and pMDLG/pRRE) was incubated for 15min at room temperature with FuGENE(Promega, E2311)-DMEM/F12(Life Technologies, 31331), and then added to the antibiotic-free medium covering the 293T cells. A first batch of viral supernatant was collected 40h after transfection. The titer was determined according to the manufacturer's instructions using the Lenti-X p24 Rapid Titer Kit (Takara, 631280). We inoculated target cells with virus particles in the appropriate complete medium supplemented with 2.5μg/ml protamine sulfate. The medium was replaced with the appropriate complete media after 12-14h of incubation with the viral supernatant.

4.10 Fluorescence-activated cell sorting (FACS)

Harvested single cell suspensions from transduced cell lines were resuspended in cold media and filtered into FACS tubes. Sorting was performed using BD FACSAria III or Fusion systems. Depending on the fluorophores to be sorted, the suitable laser-filter combinations were determined.
To exclude dead cells, events were typically gated according to their shape and granularity (FSC-A vs. SSC-A) and doublets were removed (FSC-A vs. FSC-H). Positive gates were set for populations with low to moderate levels of sLCR-dependent fluorophore expression using PGK-driven and constitutively transcribed H2B-CFP as a sorting reporter.

4.11 FACS analysis

FlowJo v10 was used for all analyses. To analyze the *in vivo* data in (Fig. 7E-G), freshly dissociated mouse brain tumor samples were pre-gated for viable cells prior to performing dimensionality reduction on all acquired parameters (FSC-H/W/A, SSC-H/W/A, positive and negative viability dies, and mVenus/mCherry sLCR expression) using the inbuilt auto t-distributed stochastic neighbor embedding (opt-SNE) (Perplexity 30, Iterations 1000). By superimposing sLCR expression heatmaps on the generated t-SNE maps, we discovered and gated glioma cell clusters that separated from mouse cells. Further t-SNE dimensionality reduction of gated glioma cells was done to analyze clustering of the sLCR reporter distribution in individual *in vivo*-derived tumor cells and to compare it to *in vitro*-cultured cells utilized for transplantation that were simultaneously assessed. The quantification of sLCR-high cells was accomplished by creating a four-quadrant gating scheme in mVenus vs. mCherry plots on *in vitro* cells to identify sLCR-high populations and using this gating scheme to t-SNE gated *in vivo* glioma cells (see Ext. Fig. 7b).

4.12 Phenotypic screening

Tumor cells were propagated in the manner described previously until they were screened. In 384 well plates (Corning), we seeded 15,000 cells/50µl in Gibco FluoroBrite DMEM medium supplemented with the required growth factors. The SPARK 20M Injector system was used to distribute cells in a 50µl suspension into each well (50µl injection volume; 100µl/s injection speed). Cells that were not adherent (e.g. hGICs) were further centrifuged at 1500rpm for 1h 30min at 37°C. Fluorescence was measured at the bottom of the plate using a SPARK 20M TECAN plate reader at 37°C in a 5% CO2 (added 3% O2 for hGICs) in a humidified cassette with the following settings for mVenus: monochromator, Ex 505nm/20nm, Em 535nm/7.5nm. Cell viability was determined in separate replicas using a 0.02 % AlamarBlue solution in FluoroBrite medium with the following settings: Fluorescence Monochromator, top reading, Ex 565nm/10nm, Em 592nm/10nm.

Automated aliquotation of DMSO-soluble compounds, such as GSK126, was performed using a D300e compound printer (TECAN), whereas cytokines were robotically aliquoted to each well using an Andrew pipetting robot (AndrewAlliance). PRISM7 was used to import the data (GraphPad). The fluorescence intensity of control dead cells was subtracted from all data as a background. Individual data were normalized to the control group mean and expressed as fold change.

4.13 Irradiation of hGICs

The XenX irradiator platform (XStrahl Life Sciences) was used to deliver the ionizing radiation, which is equipped with a 225 kV X-ray tube for targeted irradiation. hGICs cultivated in 6-well or 96-well plates were positioned in the beamline's focal plane and subjected to irradiation for a certain time period determined by an internal calculating program to reach the desired target dosage.

4.14 Induction of hypoxia

To evaluate the effects of hypoxia on IDH-wt-hGICs, cells were transferred for 24h from a typical 3% O2 culture to ambient O2-levels. Cells were then seeded at a density of 250,000 cells per well into 6-well plates and cultured in 1% O2, 3% O2, and ambient O2, respectively. In the presence of severe hypoxia (Fig. 9E), plates were cultivated in pressured incubators (Avatar, Xcellbio) at 0.5 percent O2 and 5 Psi (344 mbar) over Berlin's normal atmospheric pressure of 14.7Psi (1010-1030 mbar). Following 3 days of incubation, cells were collected into single cell suspensions using Accutase for FACS analysis of sLCR expression, and RNA was extracted according to the protocol for RNA sequencing.

4.15 Gene Knock-out using CRISPR/Cas9

The knockout of RelA (p65) was carried out using the Gene Knockout Kit v2 (Synthego). The sgRNAs were dissolved to a stock concentration of 30µM in nuclease-free 1xTE buffer. RNP complexes were produced by combining Cas9 nuclease and gRNAs in a 6:1 ratio. Each RNP complex was nucleofected into 250,000 IDH-wt-hGICs-MGT#1 cells using the 4D-Nucleofector Core Unit's CA-138 pulse program (Lonza). Protein lysates from knockout and wildtype cells were prepared approximately 7 days after electroporation for western blot analysis of knockout efficiency.

4.16 RT-qPCR

cDNA was produced using the using the SuperScript™ VILO™ MasterMix (Invitrogen, 11755050), commencing with 0.5–2.5µg RNA in a 20µl reaction and incubated at 25°C for 10 minutes, 42°C for 60min, and 85°C for 5min. RT-qPCR was run in a 384-well ViiA7 System with

10ng cDNA/well and 1x Power SYBR Green PCR Master Mix (Applied Biosystems, 4368702) in 10µl/well.

4.17 Immunoblot

Cell pellets were lysed in RIPA buffer (20mM Tris-HCl pH7.5, 150mM NaCl, 1mM EDTA, 1mM EGTA, 1% NP-40), which was supplemented with a 1x Protease inhibitor cocktail (Roche), 10mM NaPPi, 10mM NaF, and 1mM Sodium orthovanadate. If required, the lysates were sonicated, and electrophoresis was done using NuPAGE Bis-Tris precast gels in NuPAGE MOPS SDS Running Buffer (Life Technologies) (50mM MOPS, 50mM Tris Base, 0.1% SDS, 1mM EDTA). Protein was transferred onto Nitrocellulose membranes for 1h at 120mA in transfer buffer (25mM Tris-HCl pH 7.5, 192mM Glycine, 20% Methanol). Protein transfer was detected by staining for 5min with Ponceau Red, followed by two washes with TBS-T. Membranes were blocked for 1h at room temperature with 5% BSA in PBS. Primary antibodies were diluted in PBS+5% BSA and membranes incubated overnight at 4°C. Following three 5min TBS-T washes, membranes were incubated for 45min at room temperature with dilutions of suitable HRP-coupled secondary antibodies in PBS+5% BSA. After three 5min washes with TBS-T, ECL detection reagent (Sigma, RPN2209) was added and membranes were exposed to ECL Hyperfilms (Sigma, GE28-9068-37) for chemoluminescent signal detection.

4.18 Copy number normalised sLCR expression

Transduction of hGICs, patient-derived glioma stem cells (GSCs), lung adenocarcinoma, breast adenocarcinoma, and leukemia cell lines using the sLCRs MGT#1 and PNGT#2 was performed as previously described. To accomplish lentiviral copy number normalisation, gDNA was extracted according to the manufacturer's protocol using AMPure XP beads. The relative abundance of sLCR integration sites inside the genomes of target cells was determined using qPCR with mVenus(MGT#1) and mCherry(PNGT#2) specific primers and N2 primers targeting a genomic region on Chromosome 13 for input normalisation. 1ng of gDNA was amplified in quadruplicate using the appropriate primers and Power SYBR Green PCR Master Mix in a total reaction volume of 10µl. MGT#1 or PNGT#2 relative DNA quantities were normalized to N2 levels to determine the copy number abundance of each sLCR in each sample. qPCR was used to determine the expression levels of sLCRs in matching samples in quadruplicates using the One Step TB Green PrimeScript RT-PCR Kit II (Takara, RR086A) with an input of 2ng total RNA

using mVenus(MGT#1) and mCherry(PNGT#2) specific primers and GAPDH primers for normalisation. MGT#1 or PNGT#2 relative sLCR expression was estimated by normalising over GAPDH expression for each sLCR in each sample. Both qPCR normalisations were performed as two separate technical replicates, with data from both runs merged for final normalisation. The final normalisation, as shown in (Fig. 6C), was accomplished by first correcting GAPDH-normalised sLCR expression divided by N2-normalised copy number abundance and then normalising to IDH-wt-hGICs by calculating the fold-change increase in copy number normalised sLCR expression.

4.19 Intracranial orthotopic glioma xenograft

All mice experiments followed a protocol approved by the Institutional Animal Care and Use Committee and were done in compliance with European Union standards. Orthotopic glioma xenograft studies were performed using NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ(NSG) mice obtained from The Jackson Laboratory and maintained in a specific pathogen-free (SPF) environment. We utilized male and female mice ranging in age from 7 to 12 weeks. Orthotopic glioma xenograft experiments were carried out with changes to those previously described (Gargiulo et al., 2013; Gargiulo et al., 2014). Male and female NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ(NSG) mice aged 7-12 weeks were utilized for injections. To summarize, the mouse's head was immobilized within a stereotactic frame and the skull exposed via a minor skin incision. A tiny burr hole was drilled at the stereotactic coordinates of 1.0mm anterior and 2.0mm lateral of the bregma. Throughout the surgery, mice were maintained on a warming pad under 1.5-2% isoflurane mixed with ambient air and oxygen anaesthesia. Accutase-treated hGICs tumorspheres were resuspended as single cells at a concentration of 50,000 cells/µl. Typically, 4-5µl of hGICs were stereotactically injected into the corpus callosum at a depth of 3mm below the cortical surface. To prevent intracranial pressure and reflux, the injection flow was set to 0.4µl/min and the needle was removed after one minute of rest. Bone wax was used to seal the burr hole, and surgical clips were used to close the scalp. Mice were returned to their cages and evaluated daily for evidence of neurological problems until complete recovery. Additionally, mice were followed using IVIS imaging as needed until neurological symptoms developed and animals were to be euthanized. Brains were harvested immediately upon death and evaluated under a fluorescence microscope to determine the tumor boundary. Xenografted tumors were processed the same day for FACS analysis and sorting, or cells were frozen in medium + 10% DMSO until needed. Following that, the residual brain tissues were fixed as stated below. For the experiment

depicted in (Fig. 7E-J), 8-week-old female NOG mice were utilized in accordance with same experimental methods and welfare requirements authorized by the Berlin authorities (LaGeSo).

4.20 Tissue dissociation and brain tumor cell sorting

The dissection of a brain tumor was previously documented in detail (Gargiulo et al., 2013; Gargiulo et al., 2014). To summarize, the tissue was sliced with a scalpel and digested in Accutase/DNaseI (947µl Accutase, 50µl DNase I Buffer, 3µl DNase I) at 37 °C using C-tubes in a Miltenyi Biotec OctoMACS dissociator (program 37C_BTDK_1). Before RBC lysis, suspensions were filtered through a 100µm cell strainer and then a 70µm cell strainer (NH4Cl, 155 mM; KHCO3, 10 mM; EDTA, pH 7.4, 0.1 mM). After washing in cold PBS, viability and cell count were determined automatically using a TECAN SPARK20M and 0.4% Trypan Blue staining.
Typically, 200,000 cells/antibody were employed in 15ml Falcons when surface markers were evaluated. Staining volume was 50µl in RHB-A medium with primary antibody (e.g. CD133-APC; Miltenyi) for 30min on ice, in the dark. Two washes of PBS were used to eliminate unbound antibody. BD LSRFortessa was used to acquire data, whereas the BD Aria II/III or an Astrios Moflo was used to sort cells. Depending on the fluorophores being examined, the proper laser-filter combinations were determined. To exclude dead cells, events were typically gated on the basis of shape and granularity (FSC-SSC), and we employed either Calcein UltraBlue or DRAQ5 and ZombieRed as viability dyes (depending on the fluorophores being analyzed). FlowJo V10 was used to conduct the analysis.

4.21 Immunohistochemistry

Tissues and tumorspheres were fixed for 20min in 4% PFA. Following fixation, dehydration was accomplished by raising the concentration of EtOH from 70% to 100%, Xylene, and overnight Paraffin incubation. Paraffin-embedded samples were sliced using an HM 355S microtome (Thermo Scientific). Standard hematoxylin/eosin (HE) staining techniques were used, and slide images were taken using an automated microscope (Keyence).

4.22 Immunofluorescence

At room temperature, cells were grown on coverslips or spheroids spun down on glass, fixed with 4% paraformaldehyde (PFA, Sigma Aldrich, 16005) in PBS for 10min, washed three times

with PBS, permeabilized with 0.5% triton X-100 in PBS for 5min, blocked for 15min with 4% BSA (Roth, 3854.4). After staining with primary and secondary antibodies and 20µg/ml Hoechst33258 (Cayman, 16756-50), coverslips were mounted onto glass slides using nail polish and Vectashield (Linaris, H1000). Deparaffinization and citrate buffer antigen retrieval were conducted using conventional methods on paraffin-embedded tissues. Triton X-100 0.25% in PBS was used to permeabilize the membranes, and endogenous peroxidases were inhibited when necessary using 3% H2O2 in water. Generally, we used 5% normal goat serum to block (NGS). Primary antibodies were anti-GFP (Abcam, ab6556, 1:1,000), anti-MED1 (Abcam, ab64965, 1:500), and anti-Tubulin (BD T5168, 1:2,000), and secondary antibodies (1:200) were Alexa Fluor 647, A31573, A11055, and A31571 Alexa Fluor 488, and A31570 Alexa Fluor 555.

4.23 Imaging

The Zeiss LSM700 andLeica MZ10 microscopes were employed in this study. Confocal images in (Fig. 7A-D) were obtained using a Zeiss LSM 700 with the appropriate laser-filter combinations. Live-cell imaging for mVenus, CFP and mCherry fluorescence in (Fig. 6 & Ext. Fig.8a) was obtained with Zeiss LSM 700. Epifluorescence images of freshly isolated mouse brains in (Ext. Fig. 7a C) were acquired with a Leice MZ10. ImageJ or Photoshop were used to process the images.

4.24 Live-cell imaging for sLCR expression

Human lung and breast cancer cells transduced with the MGT#1 reporter and treated with GSK126 (5µM for 5 days), TGF-β1 (5ng/ml for 5 days), or vehicle were seeded out on 384-well plates in Gibco's FluoroBrite DMEM supplemented with 10% FBS for live-cell imaging.
Using 488nm excitation and a 470- to 540nm emission filter (confocal mode, 40x water objective), imaging was done on a high-content imaging platform (Operetta CLS, PerkinElmer).
Maximum projections of MGT#1 reporter intensities and DRAQ5-stained nuclei from 5µm z-stacks are displayed (Ext. Fig. 6B).

4.25 Transwell co-culture

Using hydrophilic PTFE 6-well cell culture inserts with a pore size of 0.4µm, co-cultures of hGICs and immortalized primary human Microglia C20 or human monocytes were established (Merck). Human microglia, CD34+ monocytes, or THP-1-derived M1/M2 macrophages were

seeded at 1.5x10e5 cells/well in respective media for 24h. After aspirating the media and washing the cells once with PBS, 1ml of RHB-A complete medium was added. 5x10e5 single hGICs were plated on the insert surface using a total volume of 1ml of RHB-A complete medium. After 48h of co-culture, hGICs and C20 human microglia were collected for further investigation.

4.26 Drug dose-response screening

Drug dose-response experiments were carried out according to the protocol outlined earlier (Serresi et al., 2018). Transduced hGICs were recovered from transwell co-culture assays and sorted into mVenus high and low populations using a BD FACSAria III. Cells were counted and seeded at a density of 7,000 cells/50µl/well in RHB-A complete media onto 384-well black walled plates using the SPARK20M Injector system (50µl injection volume; 100µl/s injection speed). Typically, drugs were dissolved in DMSO to a concentration of 10mM and dispensed using the D300e compound printer (TECAN) for dose-response studies with plate randomization and DMSO normalization. Cell viability was determined 72h after incubation with 10µl of 0.02% final AlamarBlue solution (Sigma, R7017) assay reagent using the following parameters: Fluorescence top reading, Monochromator, Ex 565nm/10nm, Em 592nm/10nm. PRISM7 was used to import the data (GraphPad). Fluorescence intensities from empty wells were subtracted from all results as a background. Individual data were adjusted to the mean of untreated positive and SDS-treated negative control conditions. The dose-response curve and IC50 values were determined using non-linear regression modeling (log(inhibitor) vs. normalized response — Variable slope).

4.27 RNA-seq Generation

RNA was extracted using the TRIzol Reagent (Invitrogen, 15596026), followed by isopropanol precipitation and purification using AMPure XP beads. The concentration of RNA was determined using the Qubit RNA HS Assay Kit (Invitrogen) and/or the One-Step TB Green PrimeScript RT-PCR Kit II (Takara Bio). RNA integrity was verified using the High Sensitivity RNA ScreenTape System (Agilent, 5067-5581). To create the GSC expression profiles required for the development of the GBM sLCR, the TruSeq Stranded Total RNA Library Prep (Illumina, 20020596) and Ribo-Zero Gold rRNA Removal Kit (Illumina, MRZG12324) kits were used according to the manufacturer's instructions. The resulting libraries were quantified using the Invitrogen Qubit dsDNA High Sensitivity kit and/or the KAPA Library Quantification Kit (Roche, 7960204001). The TapeStation High Sensitivity D1000 ScreenTapes kit was used to

determine the suitable library size distribution (Agilent). Pooled libraries were sequenced in either single-read 51bp or paired-end 100bp mode on the Illumina HiSeq2500 or HiSeq4000 platforms. Cutadapt (https://cutadapt.readthedocs.io/en/stable/) was used to trim Illumina adaptors from raw reads, and raw reads were aligned to the human genome (hg19 or hg38) using TopHat (Tophat2 v2.1.0; parameters: —library-type fr-firststrand -g 1 -p 8 -G ENSEMBL Annotation v82.gtf). The number of uniquely assigned reads for each gene was determined using HTSeq-count v0.6.1p1 (parameters: -m intersection-nonempty -a 10 -i gene id -s reverse -f bam). To produce log2 counts per million, reads were normalized and log2 converted (CPM).

The TruSeq Stranded Total RNA Library Prep Gold Kit (Illumina, 20020598) was used to generate the *in vitro* RNA-seq data (Fig. 8-11) according to the manufacturer's procedure with an input of 0.1-1μg total RNA. The *in vivo* RNA profiling was performed using the SMARTer Stranded Total RNA-Seq Kit v2 - Pico Input Mammalian (Takara Bio, 634413) (Fig. 6-7). The library was constructed according to the manufacturer's protocol using 0.2-10ng of total RNA as input. Quantity and quality controls were conducted in the manner stated before. Pooled libraries were sequenced in a 2x75bp or 2x100bp mode using the Illumina NextSeq500 or NovaSeq 6000 platforms. The package bcl2fastq was used to perform the demultiplexing (v2.20.0). STARv2.6.0c was used to align the reads to a customized genome (GRCh38 comprising sLCRs and reporter sequences). The number of uniquely assigned reads for each gene was determined using HTSeq (–m intersection-nonempty -a 10 -i gene id -s reverse -f bam).

4.28 RNA-seq Analysis

R v3.6 was used to conduct RNA-seq analysis on *in vivo* and *in vitro* high/low data sets (Fig. 6 and Fig. 7). Following the data processing phase, the quality of each sample was determined separately using the dupRadar v1.18 R package (default parameters). Then, using DESeq2 v1.24 on raw pre-filtered counts (>100 and >50), differential expression analyses were performed between particular sLCR activation, high/low, and *in vivo*/*in vitro*. Separate comparisons employed only MGT#1-high homogeneous samples (Fig. 7H). If log2FC > 1, padj < 0.05, and baseMean > 5, genes were called as differentially regulated (Fig. 7I). The *in vivo* MGT#1-high gene set comprises only the genes that were significantly up-regulated in the comparison of in vivo MGT#1-high and -low samples. DESeq2 v1.24 was used to execute differential expression studies of MGT#1 activation cues RNA-seq datasets (Fig. 8 & Fig. 10). After filtering for low-count genes, each condition (TNFa, Human Serum (huSer), ionizing radiation (IR), Activin A, NOC-18, oxidized LDL (OxLDL), and C20-human microglia co-culture) was compared separately to control samples

(filterByExpr function from the edgeR v3.26 R package). When necessary, batch correction using the sva v3.32 R package was used to account for technical discrepancies across sequencing runs. For the various representations, lists of up-regulated genes (log2FC > 1, padj < 0.05, and baseMean > 5) were considered gene sets (Fig. 8F-G, 10G-H, 11A, and Ext. Fig. 10F). The UpSetR v1.4 R package was used to identify common genes amongst MGT#1 activation signals (Fig. 8G and 10H). Notably, the control gene set was derived by comparing control samples to the remaining MGT#1-high samples. The umap function from the uwot R package was used to construct the UMAP dimensional reduction (Fig. 10G) (n neighbors=10, metric="manhattan", search k=100). The procedure was ran using the batch adjusted matrix (removeBatchEffect limma v3.46 R package function) filtered for each comparison by the sum of all up-regulated genes. The majority of the graphics in this research were created using ggplot2 v3.3.2, with the exception of the upset plots (above) and the heatmap in (Fig. 8F), which were created using the R package pheatmap v1.0.12 (Hierarchical clustering is based on Manhattan distance and the ward.D2 clustering method). The RNA-seq analysis in (Fig. 9) was performed using the DESeq2 pipeline on raw pre-filtered counts (>50) for the specified comparisons. The heatmaps in (Fig. 9B & F) illustrate the substantially differentially expressed genes (padj < 0.05, log2FC > 1.5). They were created using the pheatmap package. The color coding indicates the relative rlog-normalized values of gene expression across samples. To conduct and show the Gene Ontology enrichment analysis, the enrichGO and dotplot functions from the clusterProfiler R package v.3.16.1 were used (Fig. 9G).

SeqMonk was used to examine the transcriptome profiles in (Fig. 10D-F), and reads were normalized using the standard process, which included DNA contamination correction. DESeq2 differential gene expression analysis was performed using raw numbers. For visualization, the same pipeline with log transformation was utilized. The significance level was adjusted to 0.05 using normal SeqMonk settings, followed by independent intensity filtering. Quantification was carried out in the manner described previously. IPA was used to find the NFkB-related genes in MG versus hGICs and TNFa vs hGICs comparisons (Fig. 10E-F; Qiagen Bioinformatics). The MES GBM signatures were collected from the various publications, and Venny plots were created. For MES-GBM (log2FC > 0.5 fold with padj=0), PN-GBM (log2FC < -0.4 fold with padj=0), and SREBP (log2FC > 1 fold with padj=0), GSEA significance was calculated.

The interaction map in (Fig. 10F) was constructed using the Ingenuity upstream regulator tool from IPA to compare MGT#1-high TNFa and MGT#1-high C20MG co-culture. When comparing TNFa and hMG samples, an independent filter of padj < 0.05, log2FC > 1, and

log2Avg > 5 was applied to the DESeq2 data, resulting in the selection of only up-regulated genes that matched to the TNFa and MG signature gene sets.

4.29 Gene Set Enrichment Analysis

Gene-set enrichment analysis (GSEA) was performed to determine the enrichment of any of the GBM subtypes utilizing GBM cell-state public gene sets (Verhaak et al., 2010; Wang et al., 2017; Neftel et al., 2019) in comparisons of RNA-seq profiles from FACS sorted cells following MGT#1 activation (above). The enrichment was built using the runGSA function in the piano v2.0.2 R package (parameters: geneSetStat="page", signifMethod="geneSampling", and nPerm=1000). The graph was created by calculating the enriched positive gene set as -1*log10(padj(dist.dir.up) for the corresponding comparison (Fig. 6E-F, 7H, and 8E). GSEA was performed and displayed for the GBM signatures (Neftel et al., 2019) in (Fig. 9C) using the fast-pre-ranked gene set enrichment analysis (fgsea) R package v.1.14.0, with all genes included in the comparison. The genes were pre-ranked using the log-fold change obtained from differential expression analysis for the relevant comparisons, with a permutation count of 100,000 used to determine the enrichment significance. The graphics for this study were created using ggplot2 v3.3.2, except for the heatmap in (Fig. 6F), which was created using the R package pheatmap v1.0.12 (Hierarchical clustering is based on Manhattan distance and the ward.D2 clustering method). The single-sample Gene set enrichment analysis (ssGSEA) method was used in conjunction with gsva (ssgsea.norm=TRUE) from the GSVA v1.32.0 R package to generate a signature score matrix for GBM bulk (TCGA) and single-cell patient expression profiles (Neftel et al., 2019; Fig. 5F, 11A-C). The heatmap in (Fig. 11A) was created directly from this matrix (pheatmap v1.0.12 R package; hierarchical clustering is based on Manhattan distance and the ward.D2 method).

4.30 LSD algorithm

The LSD algorithm takes a list of PWMs, a list of marker genes of a target phenotype, and the reference genome of the organism of interest, and it generates a list of naturally-occurring, putative cis-regulatory elements used to assemble the synthetic-reporter. The algorithm can be divided into three steps. In step I, LSD generates a pool of potential CRE with a fixed length within user-defined regulatory landscapes (default is a 150bp window sliding with a 50bp step). In step II, LSD assigns TF-binding sites to the CRE pool using FIMO (default --output-pthresh 1e-

4 --no-qvalue), and creates a matrix of putative CREs x TFBS. In step III, LSD ranks and selects the minimal number of CREs representing the complete set of TFBS. For that purpose, it uses an algorithm to sort and select the best CRE based on the overall TFBS affinity and diversity among input TFs showing high affinity for the CRE. Starting from the ranked CREs, LSD selects the highest-ranking CRE defined by the sum of the affinity score (-log10(p-value)) and TFBS diversity (number of different TFBS). Subsequently, it removes the selected CRE and the corresponding TFBS from the CRE x TFBS matrix and repeats the selection. This continues until either none of the CRE or of the TFBS is left. In the ranking, priority is given to CREs proximal to known TSS, based on 5' CAGE data (ENCODE) in order to increase the chances of successful transcriptional firing using the same strategy as above. Finally, LSD returns an ordered list of the selected CREs, together with a representation of the TFBS scores (Fig. 12A). Several reporters were designed using the LSD method, using different signature gene and transcription factor lists as input. First-generation GBM-sLCRs were designed by manual integration of a selection of top-ranked CRE as described above. LSD GBM-sLCRs used as input the first generation sLCR gene signatures, and a selection of subtype-specific TFs (high-expressed TF genes, > quantile 75%) obtained from TCGA-GBM patients' RNA-seq expression profiles (RPKM-UQ). The framework to run the LSD algorithm is available at: https://gitlab.com/gargiulo_lab/sLCR_selection_framework.

4.31 sLCR activity transfection screening

To conduct assessment of sLCRs activites, we used three different cell lines of variable species-background. Cell lines were plated at a density of 3,000 cells (293T and CHO-K1) or 5,000 cells (NIH3T3) in their respective medium in black-walled 96-well-plates for optical imaging (Greiner, #655090) and allowed to settle overnight. For transfection on the following day, we used the Fugene HD reagent and determined optimised conditions according to the manufacturers protocol for each cell line in a 96-well-plate format in a pre-experiment. In brief, we found 100ng of DNA and varying ratios of Fugene:DNA ratios (293T 2.5:1; NIH3T3 4:1; CHO-K1 4:1) to yield sufficient transfection efficiencies in a total reaction volume of 100µl per well. Mastermixes of DNAs from 28 sLCR and three transfection control plasmids with Fugene reagent and cell line-specific medium were prepared accordingly in a separate 96-well U-bottom plate and transferred to the screening plates using an electronic multichannel pipette. Every plasmid was transfected as biological triplicates alongside six untransfected wells for each cell line. After 48h of incubation time, nuclei were stained for 4h with 2µM Hoechst 33258 and fluorescent live-cell imaging for Hoechst, GFP, mCherry and iRFP was conducted on a high-content imaging platform (Operetta

CLS, Perkin Elmer). We used the non-confocal mode and a 10x air objective to image the whole field of view of every fluorescent channel in every well in our screen, while temperature and CO2 control were active. Settings for LED power and detector exposure time were adjusted based on non-sLCR transfection controls and untransfected wells.

4.32 High-content screening analysis for sLCR activities

After filtering each fluorescent channel (sliding parabola 10px), we used the Harmony-Software building blocks to identify and count nuclei based on Hoechst-staining. Fluorescent cell-objects were identified based on GFP, mCherry or iRFP intensities. Fluorescent objects were filtered by applying a threshold for object size and mean intensities as well as number of objects were determined. Data with all relevant parameters was exported as csv files and analysed using R Studio. To account for differences in transfection efficiencies and to allow cross-comparison of sLCR expression among the three cell lines, we first calculated a transfection score as a proxy for efficiency. From the three non-sLCR transfection control plasmids (pMAX-GFP, UBC-mCherry, piRFP670) we established this score separately per cell line by determining the transfection_score = (control_fluorescent_objects_number / nuclei_number) for each control plasmid and calculated the combined mean from pMAX-GFP, UBC-mCherry and piRFP670. The value of this score represents the highest fluorescence intensity in the screen for each line and allows scaling of the sLCR plasmid activities between a value of 0 for untransfected controls and 1, as outlined in the following sentence. To assess the sLCR activity in each line, we calculated the sLCR_activity_score = (sLCR_fluorescent_objects_number / nuclei_number) and normalised this value by dividing through the previously established transfection score, that is setting the upper bar for the highest rate of fluorescent activity in each line and allows comparing sLCR activities across cell lines. Mean values of activity scores for each of the 28 sLCRs was calculated from the biological triplicates and data was plotted as box-plots using the ggplot2 package. Statistical testing was done through two-way ANOVA with multiple comparisons testing and Dunnett contrasts p-value adjustment.

5 Academic contributions

5.1 Publications

* denotes equal contribution of authors

[1] **M. J. Schmitt***, C. Company*, Y. Dramaretska*, I. Barozzi, A. Göhrig, S. Kertalli, M. Grossman, H. Naumann, J. Wierikx, D. Hulsman, R. Glass, M. Squatrito, M. Serresi and G. Gargiulo (2021). Phenotypic mapping of pathological crosstalk between Glioblastoma and innate immune cells by synthetic genetic tracing. **Cancer Discovery**, 2021 Mar;11(3):754-777

[2] M. Serresi, S. Kertalli*, L. Li*, **M. J. Schmitt**, Y. Dramaretska, J. Wierikx, D. Hulsman, G. Gargiulo (2021). Functional antagonism of chromatin modulators regulates epithelial-mesenchymal transition. **Science Advances**, 2021 Feb 24;7(9):eabd7974

[3] M. Serresi, B. Siteur, D. Hulsman, C. Company, **M. J. Schmitt**, C. Lieftink, et al. (2018). Ezh2 inhibition in Kras-driven lung cancer amplifies inflammation and associated vulnerabilities. **Journal Exp. Med.**, 2018 Dec 3;215(12):3115-3135

[4] P. C. M. van den Berk, C. Lancini, C. Company, M. Serresi, D. Hulsman, **M. J. Schmitt**, C. Pritchard, J-Y.Song, E. Tanger, M. Grossman, I. Huijbers, H.Jacobs, M. van Lohuizen, G. Gargiulo, and E. Citterio (2020). USP15 Deubiquitinase Safeguards Hematopoiesis and Genome Integrity in Hematopoietic Stem Cells and Leukemia Cells. **Cell Reports**, 2020 Dec 29;33(13):108533.

[5] J. Kühnemundt*, H. Leifeld*, F. Scherg*, **M. Schmitt***, L. Nelke, T. Schmitt, F. Baur, C. Göttlich, M. Fuchs, M. Kunz, M. Peindl, C. Brähler, C. Kronenthaler, J. Wischhusen, M. Prelog, H. Walles, T. Dandekar, G. Dandekar and S. L. Nietzer (2020). Modular micro-physiological human tumor/tissue models based on decellularized tissue for improved preclinical testing. **ALTEX**, 2020 Dec 11. doi: 10.14573/altex.2008141

5.2 Working Manuscripts

[1] C. Company*, **M. J. Schmitt***, Y. Dramaretska, S. Kertalli, B. Jiang, M. Serresi, I. Barozzi, G. Gargiulo (in preparation). Logical design of synthetic cis-regulatory DNA for genetic tracing of cell identities and state changes.

[2] M. Serresi, S. Kertalli*, Y. Dramaretska*, **M. J. Schmitt***, H. Naumann, M. Zschummel, G. Gargiulo (in preparation). Transcriptional amplification of oncogenic Kras drives lung cancer progression and metastasis.

[3] **M. J. Schmitt**, J. M. Faustino-Martins, Y. Dramaretska, M. Serresi, M. Gouti and G. Gargiulo (in preparation). Ex vivo modelling of human Glioblastoma using multicellular organoids.

[4] T. Breitenbach, **M. J. Schmitt**, T. Dandekar (in revision). Optimization of synthetic molecular reporters for a mesenchymal Glioblastoma transcriptional program by integer programming. **Bioinformatics**

6 Declaration of own contribution

The majority of this work was originally published in *Cancer Discovery* and is supplemented with additional published and unpublished results.

M. J. Schmitt*, C. Company*, Y. Dramaretska*, I. Barozzi, A. Göhrig, S. Kertalli, M. Grossman, H. Naumann, J. Wierikx, D. Hulsman, R. Glass, M. Squatrito, M. Serresi and G. Gargiulo (2021). Phenotypic mapping of pathological crosstalk between Glioblastoma and innate immune cells by synthetic genetic tracing. Cancer Discovery, 2021 Mar;11(3):754-777

The presented results in this thesis originate from the collaborative effort of multiple colleagues and therefore, I would like to highlight this fact and declare my personal contribution in order to properly credit the work of others.
I have designed and performed most of the wet-lab experiments (generation and characterization of sLCR cell lines; phenotypic screening; irradiation and hypoxia experiments; end-point processing, immunostaining and FACS analysis of *in vivo* experiments; co-culture experiments and drug dose-response) and contributed to data interpretation for most of the dry-lab data. Iros Barozzi (Imperial College, London) originally developed the sLCR algorithm and contributed to data interpretation. Carlos Company (MDC Berlin, Gargiulo Lab) implemented the automated algorithm and performed and/or contributed to most computational analyses and their interpretation. Yuliia Dramaretska (MDC Berlin, Gargiulo Lab) performed RNA extraction, library preparation and analysis of Next Generation Sequencing-experiments and contributed to data interpretation. Sonia Kertalli (MDC Berlin, Gargiulo Lab) contributed under my direct supervision as a master student to perform the phenotypic screen. Andreas Göhrig, Pilar Sanchez, Heike Naumann, Melanie Großmann, Ben Jiang (all MDC Berlin, Gargiulo Lab) and Danielle Hulsman (NKI Amsterdam, Netherlands) contributed to *in vitro* and *in vivo* experiments. Rainer Glass (LMU Munich, Germany) shared unpublished reagents. Massimo Squatrito (CNIO Madrid, Spain) contributed to expression and survival analyses and data interpretation. Michela Serresi (MDC Berlin, Gargiulo Lab) contributed to perform the experiments, to data interpretation and to supervision. Gaetano Gargiulo (MDC Berlin, Gargiulo Lab) developed the concept, designed and supervised the study, performed some of the experiments, interpreted the data and wrote the sLCR manuscript with inputs from others.

In detail, my own as well as the contribution of others will be outlined per individual figure in the following section.

To Figure 5:
IB originally developed the sLCR algorithm under input and supervision of GG. CC implemented the algorithm and used it to generated GBM-subtype specific sLCRs. CC together with GG sketched the outline of the method in (A). CC performed and plotted the analysis in (F). I have used the output from the pipeline of the top selected CRE to plot (D) and drawn the schematic with the final selection in (E).

To Extended Figure 5:
I have used the output from the pipeline to visualise the top selected CRE and conducted the pairwise-correlation plot of TFBS.

To Figure 6:
All generation, validation and maintenance of cell lines, acquired experimental data, analysis and plots in (A-E) have been done by myself with the following help:
Intracranial injection and maintenance of animals until brain harvesting have been performed by EPO, Berlin-Buch. MS and YD helped in generating qPCR-data for the sLCR activity screening in (C). The wet lab part including cell culture and FAC-sorting for the experiment outlined in (E) have been conducted by myself. YD extracted RNA and conducted RNA-sequencing. CC analysed and plotted the data in (F).

To Extended Figure 6:
MS and SK generated with my help sLCR-bearing cell lines, conducted GSK or TGFb treatment and generated qPCR-data. I performed live-cell imaging of sLCR-modified cells.

To Figure 7:
MG together with AJG and EPO, Berlin-Buch performed intracranial injection, monitoring and maintenance of animals until harvesting of brains. I performed harvesting of organs, immunofluorescence staining and imaging, dissociated tissues, FAC-sorted tumor cells, analysed and plotted the FACS data in (A-G). YD extracted RNA and conducted RNA-sequencing. CC and YD analysed and plotted the data in (H-J).

To Extended Figure 7a:
Intracranial injection, maintenance of animals and bioluminscence measurements until brain harvesting have been performed by EPO, Berlin-Buch. All analyses and plots have been created by myself.

To Extended Figure 7b:
From harvested organs, I dissociated tissues, decided the strategy for FAC-sorting, collected FACS data and devised the analysis strategy of dimensionality reduction.

To Figure 8:
I generated and maintained cell lines, catalogised compounds and developed, conducted and analysed the cytokine screening in (A-B). I treated cells, longitudinally monitored sLCR expression via FACS, analysed and plotted the data in (C-D). I treated and maintained cells, devised a strategy and conducted FAC-sorting for data shown in (E-G). YD extracted RNA and conducted RNA-sequencing. CC analysed and plotted the data in (E-G).

To Extended Figure 8a:
All experimental data, analyses and plots in this figure have been done by myself.

To Extended Figure 8b:
All experimental data, analyses and plots in this figure have been done by myself with the exception of the generation of p65-KO and western blot validation, which SK performed together with me and under my direct supervision as a master student. MS helped with Western Blots in (B).

To Figure 9:
I have generated and maintained cell lines, performed irradiation experiments, FACS analysis/sorting and Western Blot in (A-B). YD extracted RNA, conducted RNA-sequencing, analysed and plotted the data in (B-C). I have generated and maintained cell lines under normoxia or hypoxia for data presented in (D-G). I conducted FACS experiments, analysis and visualisation in (D-E). YD extracted RNA, conducted RNA-sequencing, analysed and plotted the data in (F-G).

To Figure 10:

I have generated and maintained all cell lines presented in this figure, with the exception of CD34+ MDSC, which have been generated by GG. I have developed and conducted the co-culture experiments, performed FACS experiments, analysed and plotted the data from (A-D). I have generated, maintained and treated cell lines as well as conducted FAC-sorting for data presented in (E-H). YD extracted RNA and conducted RNA-sequencing. CC analysed and plotted the data in (E-G). GG generated the Venn-diagramm in (D).

To Extended Figure 10:

I have generated, maintained and treated all cell lines and conducted co-culture experiments presented in this figure. CC performed principle component analysis (B), C20-hMG marker expression heatmap (C) and PROGENy pathway analysis in (F).

To Figure 11:

MSq, GG and CC processed, analysed and plotted the data in (A-C). I have generated, maintained and treated cell lines as well as conducted FAC-sorting for data presented in (B-D). YD extracted RNA, conducted RNA-sequencing, analysed and plotted the data in (C-D). I have conducted co-culture experiments, performed FAC-sorting, drug screening and viability assays, as well as analysed and plotted all data in (E-F).

To Figure 12:

CC developed and implemented with input from GG the LSD concept and pipeline as sketched in (A). CC performed the ssGSEA enrichment analysis and comparison of 1st generation with LSD-sLCRs in (B). I have generated hGICs with lentiviral and PiggyBac MGT#1 and MGT#4, respectively, and conducted experimental validation via stimulation with TNFa and performed FACS experiments, while GG has visualized results in (C). I have conceptualised the experimental outline of the sLCR screen, developed the image analysis pipeline and analysed and visualized the data of the screening results (E-F). BJ has performed the transfection of plasmids and imaging of the screening plates. GG has sketched the schematic in (D).

7 References

[1] Aalipour, A., Chuang, H. Y., Murty, S., D'Souza, A. L., Park, S. M., Gulati, G. S., Patel, C. B., Beinat, C., Simonetta, F., Martinić, I., Gowrishankar, G., Robinson, E. R., Aalipour, E., Zhian, Z., & Gambhir, S. S. (2019). Engineered immune cells as highly sensitive cancer diagnostics. Nature biotechnology, 37(5), 531–539. https://doi.org/10.1038/s41587-019-0064-8

[2] Agnihotri S, Zadeh G. Metabolic reprogramming in Glioblastoma: the influence of cancer metabolism on epigenetics and unanswered questions. Neuro Oncol. 2016 Feb;18(2):160-72. doi: 10.1093/neuonc/nov125. Epub 2015 Jul 14. PMID: 26180081; PMCID: PMC4724176.

[3] Ahmad, F., Sun, Q., Patel, D., & Stommel, J. M. (2019). Cholesterol Metabolism: A Potential Therapeutic Target in Glioblastoma. Cancers, 11(2), 146.

[4] Aldape, K., Zadeh, G., Mansouri, S., Reifenberger, G., & von Deimling, A. (2015). Glioblastoma: pathology, molecular mechanisms and markers. Acta neuropathologica, 129(6), 829–848.

[5] Andor N, Graham TA, Jansen M, Xia LC, Aktipis CA, Petritsch C, Ji HP, Maley CC. Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. Nat Med. 2016 Jan;22(1):105-13. doi: 10.1038/nm.3984. Epub 2015 Nov 30. PMID: 26618723; PMCID: PMC4830693.

[6] Angel P, Karin M. The role of Jun, Fos and the AP-1 complex in cell-proliferation and transformation. Biochim Biophys Acta. 1991 Dec 10;1072(2-3):129-57. doi: 10.1016/0304-419x(91)90011-9. PMID: 1751545.

[7] Anido J, Sáez-Borderías A, Gonzàlez-Juncà A, Rodón L, Folch G, Carmona MA, Prieto-Sánchez RM, Barba I, Martínez-Sáez E, Prudkin L, Cuartas I, Raventós C, Martínez-Ricarte F, Poca MA, García-Dorado D, Lahn MM, Yingling JM, Rodón J, Sahuquillo J, Baselga J, Seoane J. TGF-β Receptor Inhibitors Target the CD44(high)/Id1(high) Glioma-Initiating

Cell Population in Human Glioblastoma. Cancer Cell. 2010 Dec 14;18(6):655-68. doi: 10.1016/j.ccr.2010.10.023. PMID: 21156287.

[8] Arnold CD, Gerlach D, Stelzer C, Boryń ŁM, Rath M, Stark A. Genome-wide quantitative enhancer activity maps identified by STARR-seq. Science. 2013 Mar 1;339(6123):1074-7. doi: 10.1126/science.1232542.

[9] Azam Z, To ST, Tannous BA. Mesenchymal Transformation: The Rosetta Stone of Glioblastoma Pathogenesis and Therapy Resistance. Adv Sci (Weinh). 2020 Sep 28;7(22):2002015. doi: 10.1002/advs.202002015. PMID: 33240762; PMCID: PMC7675056.

[10] Badis G, Berger MF, Philippakis AA, Talukder S, Gehrke AR, Jaeger SA, Chan ET, Metzler G, et al. Diversity and complexity in DNA recognition by transcription factors. Science. 2009 Jun 26;324(5935):1720-3.

[11] Balss, J., Meyer, J., Mueller, W., Korshunov, A., Hartmann, C., & von Deimling, A. (2008). Analysis of the IDH1 codon 132 mutation in brain tumors. Acta neuropathologica, 116(6), 597–602.

[12] Bao S, Wu Q, McLendon RE, Hao Y, Shi Q, Hjelmeland AB, Dewhirst MW, Bigner DD, Rich JN. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. Nature. 2006 Dec 7;444(7120):756-60. doi: 10.1038/nature05236. Epub 2006 Oct 18. PMID: 17051156.

[13] Baron CS, van Oudenaarden A. Unravelling cellular relationships during development and regeneration using genetic lineage tracing. Nat Rev Mol Cell Biol. 2019 Dec;20(12):753-765. doi: 10.1038/s41580-019-0186-3.

[14] Barthel FP, Johnson KC, Varn FS, Moskalik AD, Tanner G, Kocakavuk E, Anderson KJ, Abiola O, Aldape K, Alfaro KD, Alpar D, Amin SB, Ashley DM, Bandopadhayay P, Barnholtz-Sloan JS, Beroukhim R, Bock C, Brastianos PK, Brat DJ, Brodbelt AR, Bruns AF, Bulsara KR, Chakrabarty A, Chakravarti A, Chuang JH, Claus EB, Cochran EJ, Connelly J, Costello JF, Finocchiaro G, Fletcher MN, French PJ, Gan HK, Gilbert MR, Gould PV, Grimmer MR, Iavarone A, Ismail A, Jenkinson MD, Khasraw M, Kim H, Kouwenhoven MCM, LaViolette PS, Li M, Lichter P, Ligon KL, Lowman AK, Malta TM, Mazor T,

McDonald KL, Molinaro AM, Nam DH, Nayyar N, Ng HK, Ngan CY, Niclou SP, Niers JM, Noushmehr H, Noorbakhsh J, Ormond DR, Park CK, Poisson LM, Rabadan R, Radlwimmer B, Rao G, Reifenberger G, Sa JK, Schuster M, Shaw BL, Short SC, Smitt PAS, Sloan AE, Smits M, Suzuki H, Tabatabai G, Van Meir EG, Watts C, Weller M, Wesseling P, Westerman BA, Widhalm G, Woehrer A, Yung WKA, Zadeh G, Huse JT, De Groot JF, Stead LF, Verhaak RGW; GLASS Consortium. Longitudinal molecular trajectories of diffuse glioma in adults. Nature. 2019 Dec;576(7785):112-120. doi: 10.1038/s41586-019-1775-1. Epub 2019 Nov 20. PMID: 31748746; PMCID: PMC6897368.

[15] Bayik D, Lathia JD. Cancer stem cell-immune cell crosstalk in tumour progression. Nat Rev Cancer. 2021 Aug;21(8):526-536. doi: 10.1038/s41568-021-00366-w. Epub 2021 Jun 8. PMID: 34103704; PMCID: PMC8740903.

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

[17] Bell CC, Gilan O. Principles and mechanisms of non-genetic resistance in cancer. Br J Cancer. 2020 Feb;122(4):465-472. doi: 10.1038/s41416-019-0648-6. Epub 2019 Dec 13. PMID: 31831859; PMCID: PMC7028722.

[18] Beppu T, Kamada K, Yoshida Y, Arai H, Ogasawara K, Ogawa A. Change of oxygen pressure in Glioblastoma tissue under various conditions. J Neurooncol. 2002 May;58(1):47-52. doi: 10.1023/a:1015832726054. PMID: 12160140.

[19] Berger MF, Badis G, Gehrke AR, et al. Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. Cell. 2008;133(7):1266-1276.

[20] Bhaduri, A., Di Lullo, E., Jung, D., Müller, S., Crouch, E. E., Espinosa, C. S., Ozawa, T., Alvarado, B., Spatazza, J., Cadwell, C. R., Wilkins, G., Velmeshev, D., Liu, S. J., Malatesta, M., Andrews, M. G., Mostajo-Radji, M. A., Huang, E. J., Nowakowski, T. J., Lim, D. A., Diaz, A., … Kriegstein, A. R. (2020). Outer Radial Glia-like Cancer Stem Cells Contribute to Heterogeneity of Glioblastoma. Cell stem cell, 26(1), 48–63.e6.

[21] Bhat KPL, Balasubramaniyan V, Vaillant B, Ezhilarasan R, Hummelink K, Hollingsworth F, Wani K, Heathcock L, James JD, Goodman LD, Conroy S, Long L, Lelic N, Wang S, Gumin J, Raj D, Kodama Y, Raghunathan A, Olar A, Joshi K, Pelloski CE, Heimberger A, Kim SH, Cahill DP, Rao G, Den Dunnen WFA, Boddeke HWGM, Phillips HS, Nakano I, Lang FF, Colman H, Sulman EP, Aldape K. Mesenchymal differentiation mediated by NF-κB promotes radiation resistance in Glioblastoma. Cancer Cell. 2013 Sep 9;24(3):331-46. doi: 10.1016/j.ccr.2013.08.001. Epub 2013 Aug 29. PMID: 23993863; PMCID: PMC3817560.

[22] Binda E, Visioli A, Giani F, Trivieri N, Palumbo O, Restelli S, et al. Wnt5a Drives an Invasive Phenotype in Human Glioblastoma Stem-like Cells. Cancer Res. 2017 Feb 15;77(4):996-1007. Dec 23. Erratum in: Cancer Res. 2017 Jul 15;77(14):3962.

[23] Birzu C, French P, Caccese M, Cerretti G, Idbaih A, Zagonel V, Lombardi G. Recurrent Glioblastoma: From Molecular Landscape to New Treatment Perspectives. Cancers (Basel). 2020 Dec 26;13(1):47. doi: 10.3390/cancers13010047. PMID: 33375286; PMCID: PMC7794906.

[24] Boumahdi S, de Sauvage FJ. The great escape: tumour cell plasticity in resistance to targeted therapy. Nat Rev Drug Discov. 2020 Jan;19(1):39-56. doi: 10.1038/s41573-019-0044-1. Epub 2019 Oct 10. PMID: 31601994.

[25] Bradner JE, Hnisz D, Young RA. Transcriptional Addiction in Cancer. Cell. 2017 Feb 9;168(4):629-643. doi: 10.1016/j.cell.2016.12.013. PMID: 28187285; PMCID: PMC5308559.

[26] Brennan CW, Verhaak RG, McKenna A, Campos B, Noushmehr H, Salama SR, Zheng S, Chakravarty D, Sanborn JZ, Berman SH, Beroukhim R, Bernard B, Wu CJ, Genovese G, Shmulevich I, Barnholtz-Sloan J, Zou L, Vegesna R, Shukla SA, Ciriello G, Yung WK, Zhang W, Sougnez C, Mikkelsen T, Aldape K, Bigner DD, Van Meir EG, Prados M, Sloan A, Black KL, Eschbacher J, Finocchiaro G, Friedman W, Andrews DW, Guha A, Iacocca M, O'Neill BP, Foltz G, Myers J, Weisenberger DJ, Penny R, Kucherlapati R, Perou CM, Hayes DN, Gibbs R, Marra M, Mills GB, Lander E, Spellman P, Wilson R, Sander C, Weinstein J, Meyerson M, Gabriel S, Laird PW, Haussler D, Getz G, Chin L; TCGA Research Network. The somatic genomic landscape of Glioblastoma. Cell. 2013 Oct

10;155(2):462-77. doi: 10.1016/j.cell.2013.09.034. Erratum in: Cell. 2014 Apr 24;157(3):753. PMID: 24120142; PMCID: PMC3910500.

[27] Broekman ML, Maas SLN, Abels ER, Mempel TR, Krichevsky AM, Breakefield XO. Multidimensional communication in the microenvirons of Glioblastoma. Nat Rev Neurol. 2018 Aug;14(8):482-495. doi: 10.1038/s41582-018-0025-8. PMID: 29985475; PMCID: PMC6425928.

[28] Bucher P. Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. J Mol Biol. 1990;212(4):563-578.

[29] Cabrera MC, Hollingsworth RE, Hurt EM. Cancer stem cell plasticity and tumor hierarchy. World J Stem Cells. 2015 Jan 26;7(1):27-36. doi: 10.4252/wjsc.v7.i1.27. PMID: 25621103; PMCID: PMC4300934.

[30] Campos B, Wan F, Farhadi M, Ernst A, Zeppernick F, Tagscherer KE, et al. Differentiation therapy exerts antitumor effects on stem-like glioma cells. Clin Cancer Res. 2010 May 15;16(10):2715-28.

[31] Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human Glioblastoma genes and core pathways. Nature. 2008 Oct 23;455(7216):1061-8. doi: 10.1038/nature07385. Epub 2008 Sep 4. Erratum in: Nature. 2013 Feb 28;494(7438):506. PMID: 18772890; PMCID: PMC2671642.

[32] Capper D, Jones DTW, Sill M, Hovestadt V, Schrimpf D, Sturm D, Koelsche C, Sahm F, Chavez L, Reuss DE, Kratz A, Wefers AK, Huang K, Pajtler KW, Schweizer L, Stichel D, Olar A, Engel NW, Lindenberg K, Harter PN, Braczynski AK, Plate KH, Dohmen H, Garvalov BK, Coras R, Hölsken A, Hewer E, Bewerunge-Hudler M, Schick M, Fischer R, Beschorner R, Schittenhelm J, Staszewski O, Wani K, Varlet P, Pages M, Temming P, Lohmann D, Selt F, Witt H, Milde T, Witt O, Aronica E, Giangaspero F, Rushing E, Scheurlen W, Geisenberger C, Rodriguez FJ, Becker A, Preusser M, Haberler C, Bjerkvig R, Cryan J, Farrell M, Deckert M, Hench J, Frank S, Serrano J, Kannan K, Tsirigos A, Brück W, Hofer S, Brehmer S, Seiz-Rosenhagen M, Hänggi D, Hans V, Rozsnoki S, Hansford JR, Kohlhof P, Kristensen BW, Lechner M, Lopes B, Mawrin C, Ketter R, Kulozik A, Khatib Z, Heppner F, Koch A, Jouvet A, Keohane C, Mühleisen H, Mueller W, Pohl U, Prinz M,

Benner A, Zapatka M, Gottardo NG, Driever PH, Kramm CM, Müller HL, Rutkowski S, von Hoff K, Frühwald MC, Gnekow A, Fleischhack G, Tippelt S, Calaminus G, Monoranu CM, Perry A, Jones C, Jacques TS, Radlwimmer B, Gessi M, Pietsch T, Schramm J, Schackert G, Westphal M, Reifenberger G, Wesseling P, Weller M, Collins VP, Blümcke I, Bendszus M, Debus J, Huang A, Jabado N, Northcott PA, Paulus W, Gajjar A, Robinson GW, Taylor MD, Jaunmuktane Z, Ryzhova M, Platten M, Unterberg A, Wick W, Karajannis MA, Mittelbronn M, Acker T, Hartmann C, Aldape K, Schüller U, Buslei R, Lichter P, Kool M, Herold-Mende C, Ellison DW, Hasselblatt M, Snuderl M, Brandner S, Korshunov A, von Deimling A, Pfister SM. DNA methylation-based classification of central nervous system tumours. Nature. 2018 Mar 22;555(7697):469-474. doi: 10.1038/nature26000. Epub 2018 Mar 14. PMID: 29539639; PMCID: PMC6093218.

[33] Carro MS, Lim WK, Alvarez MJ, Bollo RJ, Zhao X, Snyder EY, Sulman EP, Anne SL, Doetsch F, Colman H, Lasorella A, Aldape K, Califano A, Iavarone A. The transcriptional network for mesenchymal transformation of brain tumours. Nature. 2010 Jan 21;463(7279):318-25. doi: 10.1038/nature08712. Epub 2009 Dec 23. PMID: 20032975; PMCID: PMC4011561.

[34] Ceccarelli M, Barthel FP, Malta TM, Sabedot TS, Salama SR, Murray BA, Morozova O, Newton Y, Radenbaugh A, Pagnotta SM, Anjum S, Wang J, Manyam G, Zoppoli P, Ling S, Rao AA, Grifford M, Cherniack AD, Zhang H, Poisson L, Carlotti CG Jr, Tirapelli DP, Rao A, Mikkelsen T, Lau CC, Yung WK, Rabadan R, Huse J, Brat DJ, Lehman NL, Barnholtz-Sloan JS, Zheng S, Hess K, Rao G, Meyerson M, Beroukhim R, Cooper L, Akbani R, Wrensch M, Haussler D, Aldape KD, Laird PW, Gutmann DH; TCGA Research Network, Noushmehr H, Iavarone A, Verhaak RG. Molecular Profiling Reveals Biologically Discrete Subsets and Pathways of Progression in Diffuse Glioma. Cell. 2016 Jan 28;164(3):550-63. doi: 10.1016/j.cell.2015.12.028. PMID: 26824661; PMCID: PMC4754110.

[35] Chaligne R, Gaiti F, Silverbush D, Schiffman JS, Weisman HR, Kluegel L, Gritsch S, Deochand SD, Gonzalez Castro LN, Richman AR, Klughammer J, Biancalani T, Muus C, Sheridan C, Alonso A, Izzo F, Park J, Rozenblatt-Rosen O, Regev A, Suvà ML, Landau DA. Epigenetic encoding, heritability and plasticity of glioma transcriptional cell states. Nat Genet. 2021 Oct;53(10):1469-1479. doi: 10.1038/s41588-021-00927-7. Epub 2021 Sep 30. PMID: 34594037; PMCID: PMC8675181.

[36] Charles NA, Holland EC, Gilbertson R, Glass R, Kettenmann H. The brain tumor microenvironment. Glia. 2012 Mar;60(3):502-14. doi: 10.1002/glia.21264. PMID: 22379614.

[37] Chédeville AL, Madureira PA. The Role of Hypoxia in Glioblastoma Radiotherapy Resistance. Cancers (Basel). 2021 Feb 1;13(3):542. doi: 10.3390/cancers13030542. PMID: 33535436; PMCID: PMC7867045.

[38] Chen J, Li Y, Yu TS, McKay RM, Burns DK, Kernie SG, Parada LF. A restricted cell population propagates Glioblastoma growth after chemotherapy. Nature. 2012 Aug 23;488(7412):522-6. doi: 10.1038/nature11287. PMID: 22854781; PMCID: PMC3427400.

[39] Chinot OL, Wick W, Mason W, Henriksson R, Saran F, Nishikawa R, Carpentier AF, Hoang-Xuan K, Kavan P, Cernea D, Brandes AA, Hilton M, Abrey L, Cloughesy T. Bevacizumab plus radiotherapy-temozolomide for newly diagnosed Glioblastoma. N Engl J Med. 2014 Feb 20;370(8):709-22. doi: 10.1056/NEJMoa1308345. PMID: 24552318.

[40] Cohen AL, Holmen SL, Colman H. IDH1 and IDH2 mutations in gliomas. Curr Neurol Neurosci Rep. 2013 May;13(5):345. doi: 10.1007/s11910-013-0345-4. PMID: 23532369; PMCID: PMC4109985.

[41] Colegio OR, Chu NQ, Szabo AL, Chu T, Rhebergen AM, Jairam V, Cyrus N, Brokowski CE, Eisenbarth SC, Phillips GM, Cline GW, Phillips AJ, Medzhitov R. Functional polarization of tumour-associated macrophages by tumour-derived lactic acid. Nature. 2014 Sep 25;513(7519):559-63. doi: 10.1038/nature13490. Epub 2014 Jul 13. PMID: 25043024; PMCID: PMC4301845.

[42] Conklin, EG. The organization and cell lineage of the ascidian egg. J. Acad. Nat. Sci. Phila. 13, 1–119 (1905).

[43] Couturier CP, Ayyadhury S, Le PU, Nadaf J, Monlong J, Riva G, Allache R, Baig S, Yan X, Bourgey M, Lee C, Wang YCD, Wee Yong V, Guiot MC, Najafabadi H, Misic B, Antel J, Bourque G, Ragoussis J, Petrecca K. Single-cell RNA-seq reveals that Glioblastoma recapitulates a normal neurodevelopmental hierarchy. Nat Commun. 2020 Jul 8;11(1):3406. doi: 10.1038/s41467-020-17186-5. Erratum in: Nat Commun. 2020 Aug 7;11(1):4041. PMID: 32641768; PMCID: PMC7343844.

[44] Dagogo-Jack I, Shaw AT. Tumour heterogeneity and resistance to cancer therapies. Nat Rev Clin Oncol. 2018 Feb;15(2):81-94. doi: 10.1038/nrclinonc.2017.166. Epub 2017 Nov 8. PMID: 29115304.

[45] Darmanis S, Sloan SA, Croote D, Mignardi M, Chernikova S, Samghababi P, Zhang Y, Neff N, Kowarsky M, Caneda C, Li G, Chang SD, Connolly ID, Li Y, Barres BA, Gephart MH, Quake SR. Single-Cell RNA-Seq Analysis of Infiltrating Neoplastic Cells at the Migrating Front of Human Glioblastoma. Cell Rep. 2017 Oct 31;21(5):1399-1410. doi: 10.1016/j.celrep.2017.10.030. PMID: 29091775; PMCID: PMC5810554.

[46] Das PK, Pillai S, Rakib MA, Khanam JA, Gopalan V, Lam AKY, Islam F. Plasticity of Cancer Stem Cell: Origin and Role in Disease Progression and Therapy Resistance. Stem Cell Rev Rep. 2020 Apr;16(2):397-412. doi: 10.1007/s12015-019-09942-y. PMID: 31965409.

[47] Del Rey MJ, Valín Á, Usategui A, García-Herrero CM, Sánchez-Aragó M, Cuezva JM, Galindo M, Bravo B, Cañete JD, Blanco FJ, Criado G, Pablos JL. Hif-1α Knockdown Reduces Glycolytic Metabolism and Induces Cell Death of Human Synovial Fibroblasts Under Normoxic Conditions. Sci Rep. 2017 Jun 16;7(1):3644. doi: 10.1038/s41598-017-03921-4. PMID: 28623342; PMCID: PMC5473902.

[48] Dirkse, A., Golebiewska, A., Buder, T., Nazarov, P. V., Muller, A., Poovathingal, S., Brons, N., Leite, S., Sauvageot, N., Sarkisjan, D., Seyfrid, M., Fritah, S., Stieber, D., Michelucci, A., Hertel, F., Herold-Mende, C., Azuaje, F., Skupin, A., Bjerkvig, R., Deutsch, A., … Niclou, S. P. (2019). Stem cell-associated heterogeneity in Glioblastoma results from intrinsic tumor plasticity shaped by the microenvironment. Nature communications, 10(1), 1787.

[49] Doetsch F, Caillé I, Lim DA, García-Verdugo JM, Alvarez-Buylla A. Subventricular zone astrocytes are neural stem cells in the adult mammalian brain. Cell. 1999 Jun 11;97(6):703-16. doi: 10.1016/s0092-8674(00)80783-7. PMID: 10380923.

[50] Dongre A, Weinberg RA. New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. Nat Rev Mol Cell Biol. 2019 Feb;20(2):69-84. doi: 10.1038/s41580-018-0080-4. PMID: 30459476.

[51] Driessens G, Beck B, Caauwe A, Simons BD, Blanpain C. Defining the mode of tumour growth by clonal analysis. Nature. 2012 Aug 23;488(7412):527-30. doi: 10.1038/nature11344.

[52] Eferl R, Wagner EF. AP-1: a double-edged sword in tumorigenesis. Nat Rev Cancer. 2003 Nov;3(11):859-68. doi: 10.1038/nrc1209. PMID: 14668816.

[53] Engler JR, Robinson AE, Smirnov I, Hodgson JG, Berger MS, Gupta N, James CD, Molinaro A, Phillips JJ. Increased microglia/macrophage gene expression in a subset of adult and pediatric astrocytomas. PLoS One. 2012;7(8):e43339. doi: 10.1371/journal.pone.0043339. Epub 2012 Aug 22. PMID: 22937035; PMCID: PMC3425586.

[54] Fack F, Espedal H, Keunen O, Golebiewska A, Obad N, Harter PN, Mittelbronn M, Bähr O, Weyerbrock A, Stuhr L, Miletic H, Sakariassen PØ, Stieber D, Rygh CB, Lund-Johansen M, Zheng L, Gottlieb E, Niclou SP, Bjerkvig R. Bevacizumab treatment induces metabolic adaptation toward anaerobic metabolism in Glioblastomas. Acta Neuropathol. 2015 Jan;129(1):115-31. doi: 10.1007/s00401-014-1352-5. Epub 2014 Oct 17. PMID: 25322816; PMCID: PMC4282692.

[55] Friedman HS, Prados MD, Wen PY, Mikkelsen T, Schiff D, Abrey LE, Yung WK, Paleologos N, Nicholas MK, Jensen R, Vredenburgh J, Huang J, Zheng M, Cloughesy T. Bevacizumab alone and in combination with irinotecan in recurrent Glioblastoma. J Clin Oncol. 2009 Oct 1;27(28):4733-40. doi: 10.1200/JCO.2008.19.8721. Epub 2009 Aug 31. PMID: 19720927.

[56] Galbraith K, Snuderl M. DNA methylation as a diagnostic tool. Acta Neuropathol Commun. 2022 May 8;10(1):71. doi: 10.1186/s40478-022-01371-2. PMID: 35527288; PMCID: PMC9080136.

[57] Galli R, Binda E, Orfanelli U, Cipelletti B, Gritti A, De Vitis S, Fiocco R, Foroni C, Dimeco F, Vescovi A. Isolation and characterization of tumorigenic, stem-like neural precursors from human Glioblastoma. Cancer Res. 2004 Oct 1;64(19):7011-21. doi: 10.1158/0008-5472.CAN-04-1364. PMID: 15466194.

[58] Garcia-Mesa Y, Jay TR, Checkley MA, Luttge B, Dobrowolski C, Valadkhan S, et al. Immortalization of primary microglia: a new platform to study HIV regulation in the central nervous system. J Neurovirol. Springer International Publishing; 2017;23:47–66.

[59] Gargiulo G, Cesaroni M, Serresi M, de Vries NA, Hulsman D, Bruggeman S, et al. In vivo RNAi screen for BMI1 targets identifies TGF-β/BMP-ER stress pathways as key regulators of neural- and malignant glioma-stem cell homeostasis. Cancer Cell. 2013;23:660–76.

[60] Garofano L, Migliozzi S, Oh YT, D'Angelo F, Najac RD, Ko A, Frangaj B, Caruso FP, Yu K, Yuan J, Zhao W, Di Stefano AL, Bielle F, Jiang T, Sims P, Suvà ML, Tang F, Su XD, Ceccarelli M, Sanson M, Lasorella A, Iavarone A. Pathway-based classification of Glioblastoma uncovers a mitochondrial subtype with therapeutic vulnerabilities. Nat Cancer. 2021 Feb;2(2):141-156. doi: 10.1038/s43018-020-00159-4. Epub 2021 Jan 11. PMID: 33681822; PMCID: PMC7935068.

[61] Gilbert MR, Dignam JJ, Armstrong TS, Wefel JS, Blumenthal DT, Vogelbaum MA, Colman H, Chakravarti A, Pugh S, Won M, Jeraj R, Brown PD, Jaeckle KA, Schiff D, Stieber VW, Brachman DG, Werner-Wasik M, Tremont-Lukats IW, Sulman EP, Aldape KD, Curran WJ Jr, Mehta MP. A randomized trial of bevacizumab for newly diagnosed Glioblastoma. N Engl J Med. 2014 Feb 20;370(8):699-708. doi: 10.1056/NEJMoa1308573. PMID: 24552317; PMCID: PMC4201043.

[62] Gimple RC, Bhargava S, Dixit D, Rich JN. Glioblastoma stem cells: lessons from the tumor hierarchy in a lethal cancer. Genes Dev. 2019 Jun 1;33(11-12):591-609. doi: 10.1101/gad.324301.119. PMID: 31160393; PMCID: PMC6546059.

[63] Gimple RC, Yang K, Halbert ME, Agnihotri S, Rich JN. Brain cancer stem cells: resilience through adaptive plasticity and hierarchical heterogeneity. Nat Rev Cancer. 2022 Jun 16. doi: 10.1038/s41568-022-00486-x. Epub ahead of print. PMID: 35710946.

[64] Gosselin D, Skola D, Coufal NG, Holtman IR, Schlachetzki JCM, Sajti E, Jaeger BN, et al. An environment-dependent transcriptional network specifies human microglia identity. Science. 2017 Jun 23;356(6344):eaal3222.

[65] Gupta PB, Pastushenko I, Skibinski A, Blanpain C, Kuperwasser C. Phenotypic Plasticity: Driver of Cancer Initiation, Progression, and Therapy Resistance. Cell Stem Cell. 2019 Jan 3;24(1):65-78. doi: 10.1016/j.stem.2018.11.011. Epub 2018 Dec 13. PMID: 30554963; PMCID: PMC7297507.

[66] Halliday J, Helmy K, Pattwell SS, Pitter KL, LaPlant Q, Ozawa T, Holland EC. In vivo radiation response of proneural glioma characterized by protective p53 transcriptional program and proneural-mesenchymal shift. Proc Natl Acad Sci U S A. 2014 Apr 8;111(14):5248-53. doi: 10.1073/pnas.1321014111. Epub 2014 Mar 24. PMID: 24706837; PMCID: PMC3986190.

[67] Hambardzumyan D, Bergers G. Glioblastoma: Defining Tumor Niches. Trends Cancer. 2015 Dec;1(4):252-265. doi: 10.1016/j.trecan.2015.10.009. PMID: 27088132; PMCID: PMC4831073.

[68] Hambardzumyan D, Gutmann DH, Kettenmann H (2016) The role of microglia and macrophages in glioma maintenance and progression. Nat Neurosci 19:20–27.

[69] Han, S., Liu, Y., Cai, S. J., Qian, M., Ding, J., Larion, M., Gilbert, M. R., & Yang, C. (2020). IDH mutation in glioma: molecular mechanisms and potential therapeutic targets. British journal of cancer, 122(11), 1580–1589.

[70] Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011 Mar 4;144(5):646-74. doi: 10.1016/j.cell.2011.02.013. PMID: 21376230.

[71] Hata AN, Niederst MJ, Archibald HL, Gomez-Caraballo M, Siddiqui FM, Mulvey HE, Maruvka YE, Ji F, Bhang HE, Krishnamurthy Radhakrishna V, Siravegna G, Hu H, Raoof S, Lockerman E, Kalsy A, Lee D, Keating CL, Ruddy DA, Damon LJ, Crystal AS, Costa C, Piotrowska Z, Bardelli A, Iafrate AJ, Sadreyev RI, Stegmeier F, Getz G, Sequist LV, Faber AC, Engelman JA. Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. Nat Med. 2016 Mar;22(3):262-9. doi: 10.1038/nm.4040. Epub 2016 Feb 1. PMID: 26828195; PMCID: PMC4900892.

[72] Hara T, Chanoch-Myers R, Mathewson ND, Myskiw C, Atta L, Bussema L, Eichhorn SW, Greenwald AC, Kinker GS, Rodman C, Gonzalez Castro LN, Wakimoto H, Rozenblatt-

Rosen O, Zhuang X, Fan J, Hunter T, Verma IM, Wucherpfennig KW, Regev A, Suvà ML, Tirosh I. Interactions between cancer cells and immune cells drive transitions to mesenchymal-like states in Glioblastoma. Cancer Cell. 2021 Jun 14;39(6):779-792.e11. doi: 10.1016/j.ccell.2021.05.002. Epub 2021 Jun 3. PMID: 34087162; PMCID: PMC8366750.

[73] Harris AL. Hypoxia--a key regulatory factor in tumour growth. Nat Rev Cancer. 2002 Jan;2(1):38-47. doi: 10.1038/nrc704. PMID: 11902584.

[74] Hayden MS, Ghosh S. Regulation of NF-κB by TNF family cytokines. Semin Immunol. 2014 Jun;26(3):253-66. doi: 10.1016/j.smim.2014.05.004. Epub 2014 Jun 21. PMID: 24958609; PMCID: PMC4156877.

[75] Höving AL, Windmöller BA, Knabbe C, Kaltschmidt B, Kaltschmidt C, Greiner JFW. Between Fate Choice and Self-Renewal-Heterogeneity of Adult Neural Crest-Derived Stem Cells. Front Cell Dev Biol. 2021 Apr 8;9:662754. doi: 10.3389/fcell.2021.662754. PMID: 33898464; PMCID: PMC8060484.

[76] Jain KK. A Critical Overview of Targeted Therapies for Glioblastoma. Front Oncol. 2018 Oct 15;8:419. doi: 10.3389/fonc.2018.00419. PMID: 30374421; PMCID: PMC6196260.

[77] Jain RK, di Tomaso E, Duda DG, Loeffler JS, Sorensen AG, Batchelor TT. Angiogenesis in brain tumours. Nat Rev Neurosci. 2007 Aug;8(8):610-22. doi: 10.1038/nrn2175. PMID: 17643088.

[78] Johnson BE, Mazor T, Hong C, Barnes M, Aihara K, McLean CY, Fouse SD, Yamamoto S, Ueda H, Tatsuno K, Asthana S, Jalbert LE, Nelson SJ, Bollen AW, Gustafson WC, Charron E, Weiss WA, Smirnov IV, Song JS, Olshen AB, Cha S, Zhao Y, Moore RA, Mungall AJ, Jones SJM, Hirst M, Marra MA, Saito N, Aburatani H, Mukasa A, Berger MS, Chang SM, Taylor BS, Costello JF. Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. Science. 2014 Jan 10;343(6167):189-193. doi: 10.1126/science.1239947. Epub 2013 Dec 12. PMID: 24336570; PMCID: PMC3998672.

[79] Johnson KC, Anderson KJ, Courtois ET, Gujar AD, Barthel FP, Varn FS, Luo D, Seignon M, Yi E, Kim H, Estecio MRH, Zhao D, Tang M, Navin NE, Maurya R, Ngan CY, Verburg N, de Witt Hamer PC, Bulsara K, Samuels ML, Das S, Robson P, Verhaak RGW. Single-

cell multimodal glioma analyses identify epigenetic regulators of cellular plasticity and environmental stress response. Nat Genet. 2021 Oct;53(10):1456-1468. doi: 10.1038/s41588-021-00926-8. Epub 2021 Sep 30. PMID: 34594038; PMCID: PMC8570135.

[80] Jolma A, Kivioja T, Toivonen J, et al. Multiplexed massively parallel SELEX for characterization of human transcription factor binding specificities. Genome Res. 2010;20(6):861-873.

[81] Joseph JV, Conroy S, Pavlov K, Sontakke P, Tomar T, Eggens-Meijer E, Balasubramaniyan V, Wagemakers M, den Dunnen WF, Kruyt FA. Hypoxia enhances migration and invasion in Glioblastoma by promoting a mesenchymal shift mediated by the HIF1α-ZEB1 axis. Cancer Lett. 2015 Apr 1;359(1):107-16. doi: 10.1016/j.canlet.2015.01.010. Epub 2015 Jan 12. PMID: 25592037.

[82] Jovčevska I, Kočevar N, Komel R. Glioma and Glioblastoma - how much do we (not) know? Mol Clin Oncol. 2013 Nov;1(6):935-941. doi: 10.3892/mco.2013.172. Epub 2013 Aug 26. PMID: 24649273; PMCID: PMC3916171.

[83] Kaffes I, Szulzewsky F, Chen Z, Herting CJ, Gabanic B, Velázquez Vega JE, Shelton J, Switchenko JM, Ross JL, McSwain LF, Huse JT, Westermark B, Nelander S, Forsberg-Nilsson K, Uhrbom L, Maturi NP, Cimino PJ, Holland EC, Kettenmann H, Brennan CW, Brat DJ, Hambardzumyan D. Human Mesenchymal Glioblastomas are characterized by an increased immune cell presence compared to Proneural and Classical tumors. Oncoimmunology. 2019 Aug 22;8(11):e1655360. doi: 10.1080/2162402X.2019.1655360. PMID: 31646100; PMCID: PMC6791439.

[84] Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. J Clin Invest. 2009 Jun;119(6):1420-8. doi: 10.1172/JCI39104. Erratum in: J Clin Invest. 2010 May 3;120(5):1786. PMID: 19487818; PMCID: PMC2689101.

[85] Karin M, Liu Zg, Zandi E. AP-1 function and regulation. Curr Opin Cell Biol. 1997 Apr;9(2):240-6. doi: 10.1016/s0955-0674(97)80068-3. PMID: 9069263.

[86] Kazmi F, Soon YY, Leong YH, Koh WY, Vellayappan B. Re-irradiation for recurrent Glioblastoma (GBM): a systematic review and meta-analysis. J Neurooncol. 2019 Mar;142(1):79-90. doi: 10.1007/s11060-018-03064-0. Epub 2018 Dec 6. PMID: 30523605.

[87] Kim J, Lee IH, Cho HJ, Park CK, Jung YS, Kim Y, Nam SH, Kim BS, Johnson MD, Kong DS, Seol HJ, Lee JI, Joo KM, Yoon Y, Park WY, Lee J, Park PJ, Nam DH. Spatiotemporal Evolution of the Primary Glioblastoma Genome. Cancer Cell. 2015 Sep 14;28(3):318-28. doi: 10.1016/j.ccell.2015.07.013. PMID: 26373279.

[88] Kim Y, Varn FS, Park SH, Yoon BW, Park HR, Lee C, Verhaak RGW, Paek SH. Perspective of mesenchymal transformation in Glioblastoma. Acta Neuropathol Commun. 2021 Mar 24;9(1):50. doi: 10.1186/s40478-021-01151-4. PMID: 33762019; PMCID: PMC7992784.

[89] Kim, T.H., Abdullaev, Z.K., Smith, A.D., Ching, K.A., Loukinov, D.I., Green, R.D., Zhang, M.Q., Lobanenkov, V.V., and Ren, B. (2007). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell 128, 1231–1245.

[90] Kretzschmar K, Watt FM. Lineage tracing. Cell. 2012 Jan 20;148(1-2):33-45. doi: 10.1016/j.cell.2012.01.002.

[91] Labrie M, Brugge JS, Mills GB, Zervantonakis IK. Therapy resistance: opportunities created by adaptive responses to targeted therapies in cancer. Nat Rev Cancer. 2022 Jun;22(6):323-339. doi: 10.1038/s41568-022-00454-5. Epub 2022 Mar 9. PMID: 35264777; PMCID: PMC9149051.

[92] Lai A, Tran A, Nghiemphu PL, Pope WB, Solis OE, Selch M, Filka E, Yong WH, Mischel PS, Liau LM, Phuphanich S, Black K, Peak S, Green RM, Spier CE, Kolevska T, Polikoff J, Fehrenbacher L, Elashoff R, Cloughesy T. Phase II study of bevacizumab plus temozolomide during and after radiation therapy for patients with newly diagnosed Glioblastoma multiforme. J Clin Oncol. 2011 Jan 10;29(2):142-8. doi: 10.1200/JCO.2010.30.2729. Epub 2010 Dec 6. PMID: 21135282; PMCID: PMC3058273.

[93] Lan X, Jörg DJ, Cavalli FMG, Richards LM, Nguyen LV, Vanner RJ, Guilhamon P, Lee L, Kushida MM, Pellacani D, Park NI, Coutinho FJ, Whetstone H, Selvadurai HJ, Che C, Luu B, Carles A, Moksa M, Rastegar N, Head R, Dolma S, Prinos P, Cusimano MD, Das S,

Bernstein M, Arrowsmith CH, Mungall AJ, Moore RA, Ma Y, Gallo M, Lupien M, Pugh TJ, Taylor MD, Hirst M, Eaves CJ, Simons BD, Dirks PB. Fate mapping of human Glioblastoma reveals an invariant stem cell hierarchy. Nature. 2017 Sep 14;549(7671):227-232. doi: 10.1038/nature23666. Epub 2017 Aug 30. PMID: 28854171; PMCID: PMC5608080.

[94] Lasorella A, Iavarone A. The making of the Glioblastoma classification. Br J Cancer. 2021 Jul;125(1):4-6. doi: 10.1038/s41416-021-01360-7. Epub 2021 Mar 25. PMID: 33767415; PMCID: PMC8257735.

[95] Lathia, J. D., Mack, S. C., Mulkearns-Hubert, E. E., Valentim, C. L., & Rich, J. N. (2015). Cancer stem cells in Glioblastoma. Genes & development, 29(12), 1203–1217. https://doi.org/10.1101/gad.261982.115

[96] Lau J, Ilkhanizadeh S, Wang S, Miroshnikova YA, Salvatierra NA, Wong RA, Schmidt C, Weaver VM, Weiss WA, Persson AI. STAT3 Blockade Inhibits Radiation-Induced Malignant Progression in Glioma. Cancer Res. 2015 Oct 15;75(20):4302-11. doi: 10.1158/0008-5472.CAN-14-3331. Epub 2015 Aug 17. PMID: 26282165; PMCID: PMC4609277.

[97] Lee J, Kotliarova S, Kotliarov Y, Li A, Su Q, Donin NM, Pastorino S, Purow BW, Christopher N, Zhang W, Park JK, Fine HA. Tumor stem cells derived from Glioblastomas cultured in bFGF and EGF more closely mirror the phenotype and genotype of primary tumors than do serum-cultured cell lines. Cancer Cell. 2006 May;9(5):391-403. doi: 10.1016/j.ccr.2006.03.030. PMID: 16697959.

[98] Lee, J. C., Vivanco, I., Beroukhim, R., Huang, J. H., Feng, W. L., DeBiasi, R. M., Yoshimoto, K., King, J. C., Nghiemphu, P., Yuza, Y., Xu, Q., Greulich, H., Thomas, R. K., Paez, J. G., Peck, T. C., Linhart, D. J., Glatt, K. A., Getz, G., Onofrio, R., Ziaugra, L., … Mellinghoff, I. K. (2006). Epidermal growth factor receptor activation in Glioblastoma through novel missense mutations in the extracellular domain. PLoS medicine, 3(12), e485.

[99] Le Rhun E, Preusser M, Roth P, Reardon DA, van den Bent M, Wen P, Reifenberger G, Weller M. Molecular targeted therapy of Glioblastoma. Cancer Treat Rev. 2019 Nov;80:101896. doi: 10.1016/j.ctrv.2019.101896. Epub 2019 Sep 11. PMID: 31541850.

[100] Li Y, Ali S, Clarke J, Cha S. Bevacizumab in Recurrent Glioma: Patterns of Treatment Failure and Implications. Brain Tumor Res Treat. 2017 Apr;5(1):1-9. doi: 10.14791/btrt.2017.5.1.1. Epub 2017 Apr 30. PMID: 28516072; PMCID: PMC5433944.

[101] Li, D., Li, S., Xue, A. Z., Smith Callahan, L. A., & Liu, Y. (2020). Expression of SREBP2 and cholesterol metabolism related genes in TCGA glioma cohorts. Medicine, 99(12), e18815. https://doi.org/10.1097/MD.0000000000018815

[102] Liberti MV, Locasale JW. The Warburg Effect: How Does it Benefit Cancer Cells? Trends Biochem Sci. 2016 Mar;41(3):211-218. doi: 10.1016/j.tibs.2015.12.001. Epub 2016 Jan 5. Erratum in: Trends Biochem Sci. 2016 Mar;41(3):287. Erratum in: Trends Biochem Sci. 2016 Mar;41(3):287. PMID: 26778478; PMCID: PMC4783224.

[103] Liu C, Sage JC, Miller MR, Verhaak RG, Hippenmeyer S, Vogel H, Foreman O, Bronson RT, Nishiyama A, Luo L, Zong H. Mosaic analysis with double markers reveals tumor cell of origin in glioma. Cell. 2011 Jul 22;146(2):209-21. doi: 10.1016/j.cell.2011.06.014.

[104] Liu G, Yuan X, Zeng Z, Tunici P, Ng H, Abdulkadir IR, Lu L, Irvin D, Black KL, Yu JS. Analysis of gene expression and chemoresistance of CD133+ cancer stem cells in Glioblastoma. Mol Cancer. 2006 Dec 2;5:67. doi: 10.1186/1476-4598-5-67. PMID: 17140455; PMCID: PMC1697823.

[105] Liu, K., Jin, H., & Zhou, B. (2020). Genetic lineage tracing with multiple DNA recombinases: A user's guide for conducting more precise cell fate mapping studies. The Journal of biological chemistry, 295(19), 6413–6424.

[106] Louis DN, Ohgaki H, Wiestler OD, Cavenee WK, Burger PC, Jouvet A, Scheithauer BW, Kleihues P. The 2007 WHO classification of tumours of the central nervous system. Acta Neuropathol. 2007 Aug;114(2):97-109. doi: 10.1007/s00401-007-0243-4. Epub 2007 Jul 6. Erratum in: Acta Neuropathol. 2007 Nov;114(5):547. PMID: 17618441; PMCID: PMC1929165.

[107] Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, Hawkins C, Ng HK, Pfister SM, Reifenberger G, Soffietti R, von Deimling A, Ellison DW. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. Neuro Oncol. 2021

Aug 2;23(8):1231-1251. doi: 10.1093/neuonc/noab106. PMID: 34185076; PMCID: PMC8328013.

[108] Maley CC, Aktipis A, Graham TA, Sottoriva A, Boddy AM, Janiszewska M, Silva AS, Gerlinger M, Yuan Y, Pienta KJ, Anderson KS, Gatenby R, Swanton C, Posada D, Wu CI, Schiffman JD, Hwang ES, Polyak K, Anderson ARA, Brown JS, Greaves M, Shibata D. Classifying the evolutionary and ecological features of neoplasms. Nat Rev Cancer. 2017 Oct;17(10):605-619. doi: 10.1038/nrc.2017.69. Epub 2017 Sep 15. PMID: 28912577; PMCID: PMC5811185.

[109] Mao P, Joshi K, Li J, Kim SH, Li P, Santana-Santos L, Luthra S, Chandran UR, Benos PV, Smith L, Wang M, Hu B, Cheng SY, Sobol RW, Nakano I. Mesenchymal glioma stem cells are maintained by activated glycolytic metabolism involving aldehyde dehydrogenase 1A3. Proc Natl Acad Sci U S A. 2013 May 21;110(21):8644-9. doi: 10.1073/pnas.1221478110. Epub 2013 May 6. PMID: 23650391; PMCID: PMC3666732.

[110] Marine JC, Dawson SJ, Dawson MA. Non-genetic mechanisms of therapeutic resistance in cancer. Nat Rev Cancer. 2020 Dec;20(12):743-756. doi: 10.1038/s41568-020-00302-4. Epub 2020 Oct 8. PMID: 33033407.

[111] Marques C, Unterkircher T, Kroon P, Oldrini B, Izzo A, Dramaretska Y, Ferrarese R, Kling E, Schnell O, Nelander S, Wagner EF, Bakiri L, Gargiulo G, Carro MS, Squatrito M. NF1 regulates mesenchymal Glioblastoma plasticity and aggressiveness through the AP-1 transcription factor FOSL1. Elife. 2021 Aug 17;10:e64846. doi: 10.7554/eLife.64846. PMID: 34399888; PMCID: PMC8370767.

[112] Marusyk A, Janiszewska M, Polyak K. Intratumor Heterogeneity: The Rosetta Stone of Therapy Resistance. Cancer Cell. 2020 Apr 13;37(4):471-484. doi: 10.1016/j.ccell.2020.03.007. PMID: 32289271; PMCID: PMC7181408.

[113] McGranahan N, Swanton C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. Cell. 2017 Feb 9;168(4):613-628. doi: 10.1016/j.cell.2017.01.018. PMID: 28187284.

[114] Minata M, Audia A, Shi J, Lu S, Bernstock J, Pavlyukov MS, Das A, Kim SH, Shin YJ, Lee Y, Koo H, Snigdha K, Waghmare I, Guo X, Mohyeldin A, Gallego-Perez D, Wang J, Chen D, Cheng P, Mukheef F, Contreras M, Reyes JF, Vaillant B, Sulman EP, Cheng SY, Markert JM, Tannous BA, Lu X, Kango-Singh M, Lee LJ, Nam DH, Nakano I, Bhat KP. Phenotypic Plasticity of Invasive Edge Glioma Stem-like Cells in Response to Ionizing Radiation. Cell Rep. 2019 Feb 12;26(7):1893-1905.e7. doi: 10.1016/j.celrep.2019.01.076. PMID: 30759398; PMCID: PMC6594377.

[115] Miska J, Lee-Chang C, Rashidi A, Muroski ME, Chang AL, Lopez-Rosas A, Zhang P, Panek WK, Cordero A, Han Y, Ahmed AU, Chandel NS, Lesniak MS. HIF-1α Is a Metabolic Switch between Glycolytic-Driven Migration and Oxidative Phosphorylation-Driven Immunosuppression of Tregs in Glioblastoma. Cell Rep. 2019 Apr 2;27(1):226-237.e4. doi: 10.1016/j.celrep.2019.03.029. PMID: 30943404; PMCID: PMC6461402.

[116] Mitchell K, Troike K, Silver DJ, Lathia JD. The evolution of the cancer stem cell state in Glioblastoma: emerging insights into the next generation of functional interactions. Neuro Oncol. 2021 Feb 25;23(2):199-213. doi: 10.1093/neuonc/noaa259. PMID: 33173943; PMCID: PMC7906055.

[117] Monteiro AR, Hill R, Pilkington GJ, Madureira PA. The Role of Hypoxia in Glioblastoma Invasion. Cells. 2017 Nov 22;6(4):45. doi: 10.3390/cells6040045. PMID: 29165393; PMCID: PMC5755503.

[118] Muerdter F, Boryń ŁM, Arnold CD. STARR-seq - principles and applications. Genomics. 2015 Sep;106(3):145-150. doi: 10.1016/j.ygeno.2015.06.001.

[119] Nathanson DA, Gini B, Mottahedeh J, Visnyei K, Koga T, Gomez G, Eskin A, Hwang K, Wang J, Masui K, Paucar A, Yang H, Ohashi M, Zhu S, Wykosky J, Reed R, Nelson SF, Cloughesy TF, James CD, Rao PN, Kornblum HI, Heath JR, Cavenee WK, Furnari FB, Mischel PS. Targeted therapy resistance mediated by dynamic regulation of extrachromosomal mutant EGFR DNA. Science. 2014 Jan 3;343(6166):72-6. doi: 10.1126/science.1241328. Epub 2013 Dec 5. PMID: 24310612; PMCID: PMC4049335.

[120] Neftel C, Laffy J, Filbin MG, Hara T, Shore ME, Rahme GJ, et al. (2019). An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. Cell. 2019 Aug 8;178(4):835-849.e21.

[121] Nicholson JG, Fine HA. Diffuse Glioma Heterogeneity and Its Therapeutic Implications. Cancer Discov. 2021 Mar;11(3):575-590. doi: 10.1158/2159-8290.CD-20-1474. Epub 2021 Feb 8. PMID: 33558264.

[122] Noushmehr H, Weisenberger DJ, Diefes K, Phillips HS, Pujara K, Berman BP, Pan F, Pelloski CE, Sulman EP, Bhat KP, Verhaak RG, Hoadley KA, Hayes DN, Perou CM, Schmidt HK, Ding L, Wilson RK, Van Den Berg D, Shen H, Bengtsson H, Neuvial P, Cope LM, Buckley J, Herman JG, Baylin SB, Laird PW, Aldape K; Cancer Genome Atlas Research Network. Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. Cancer Cell. 2010 May 18;17(5):510-22. doi: 10.1016/j.ccr.2010.03.017. Epub 2010 Apr 15. PMID: 20399149; PMCID: PMC2872684.

[123] Noy R, Pollard JW. Tumor-associated macrophages: from mechanisms to therapy. Immunity. 2014 Jul 17;41(1):49-61. doi: 10.1016/j.immuni.2014.06.010. Erratum in: Immunity. 2014 Nov 20;41(5):866. PMID: 25035953; PMCID: PMC4137410.

[124] Ohinata Y, Sano M, Shigeta M, Yamanaka K, Saitou M. A comprehensive, non-invasive visualization of primordial germ cell development in mice by the Prdm1-mVenus and Dppa3-ECFP double transgenic reporter. Reproduction. 2008 Oct;136(4):503-14.

[125] Omuro, A., & DeAngelis, L. M. (2013). Glioblastoma and other malignant gliomas: a clinical review. JAMA, 310(17), 1842–1850.

[126] Oronsky B, Reid TR, Oronsky A, Sandhu N, Knox SJ. A Review of Newly Diagnosed Glioblastoma. Front Oncol. 2021 Feb 5;10:574012. doi: 10.3389/fonc.2020.574012. PMID: 33614476; PMCID: PMC7892469.

[127] Osswald M, Jung E, Sahm F, Solecki G, Venkataramani V, Blaes J, Weil S, Horstmann H, Wiestler B, Syed M, Huang L, Ratliff M, Karimian Jazi K, Kurz FT, Schmenger T, Lemke D, Gömmel M, Pauli M, Liao Y, Häring P, Pusch S, Herl V, Steinhäuser C, Krunic D, Jarahian M, Miletic H, Berghoff AS, Griesbeck O, Kalamakis G, Garaschuk O, Preusser M,

Weiss S, Liu H, Heiland S, Platten M, Huber PE, Kuner T, von Deimling A, Wick W, Winkler F. Brain tumour cells interconnect to a functional and resistant network. Nature. 2015 Dec 3;528(7580):93-8. doi: 10.1038/nature16071. Epub 2015 Nov 4. PMID: 26536111.

[128] Parada LF, Dirks PB, Wechsler-Reya RJ. Brain Tumor Stem Cells Remain in Play. J Clin Oncol. 2017 Jul 20;35(21):2428-2431. doi: 10.1200/JCO.2017.73.9540. Epub 2017 Jun 22. PMID: 28640710; PMCID: PMC5516484.

[129] Parsons, D. W., Jones, S., Zhang, X., Lin, J. C., Leary, R. J., Angenendt, P., Mankoo, P., Carter, H., Siu, I. M., Gallia, G. L., Olivi, A., McLendon, R., Rasheed, B. A., Keir, S., Nikolskaya, T., Nikolsky, Y., Busam, D. A., Tekleab, H., Diaz, L. A., Jr, Hartigan, J., … Kinzler, K. W. (2008). An integrated genomic analysis of human Glioblastoma multiforme. Science (New York, N.Y.), 321(5897), 1807–1812.

[130] Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, Cahill DP, Nahed BV, Curry WT, Martuza RL, Louis DN, Rozenblatt-Rosen O, Suvà ML, Regev A, Bernstein BE. Single-cell RNA-seq highlights intratumoral heterogeneity in primary Glioblastoma. Science. 2014 Jun 20;344(6190):1396-401.

[131] Pavlyukov MS, Yu H, Bastola S, Minata M, Shender VO, Lee Y, Zhang S, Wang J, Komarova S, Wang J, Yamaguchi S, Alsheikh HA, Shi J, Chen D, Mohyeldin A, Kim SH, Shin YJ, Anufrieva K, Evtushenko EG, Antipova NV, Arapidi GP, Govorun V, Pestov NB, Shakhparonov MI, Lee LJ, Nam DH, Nakano I. Apoptotic Cell-Derived Extracellular Vesicles Promote Malignancy of Glioblastoma Via Intercellular Transfer of Splicing Factors. Cancer Cell. 2018 Jul 9;34(1):119-135.e10. doi: 10.1016/j.ccell.2018.05.012. Epub 2018 Jun 21. PMID: 29937354; PMCID: PMC6048596.

[132] Peereboom DM, Shepard DR, Ahluwalia MS, Brewer CJ, Agarwal N, Stevens GH, Suh JH, Toms SA, Vogelbaum MA, Weil RJ, Elson P, Barnett GH. Phase II trial of erlotinib with temozolomide and radiation in patients with newly diagnosed Glioblastoma multiforme. J Neurooncol. 2010 May;98(1):93-9. doi: 10.1007/s11060-009-0067-2. Epub 2009 Dec 4. PMID: 19960228.

[133] Piao Y, Liang J, Holmes L, Henry V, Sulman E, de Groot JF. Acquired resistance to anti-VEGF therapy in Glioblastoma is associated with a mesenchymal transition. Clin Cancer Res. 2013 Aug 15;19(16):4392-403. doi: 10.1158/1078-0432.CCR-12-1557. Epub 2013 Jun 26. PMID: 23804423.

[134] Piao Y, Liang J, Holmes L, Zurita AJ, Henry V, Heymach JV, de Groot JF. Glioblastoma resistance to anti-VEGF therapy is associated with myeloid cell infiltration, stem cell accumulation, and a mesenchymal phenotype. Neuro Oncol. 2012 Nov;14(11):1379-92. doi: 10.1093/neuonc/nos158. Epub 2012 Sep 10. PMID: 22965162; PMCID: PMC3480262.

[135] Piccirillo SG, Reynolds BA, Zanetti N, Lamorte G, Binda E, Broggi G, Brem H, Olivi A, Dimeco F, Vescovi AL. Bone morphogenetic proteins inhibit the tumorigenic potential of human brain tumour-initiating cells. Nature. 2006 Dec 7;444(7120):761-5. doi: 10.1038/nature05349. PMID: 17151667.

[136] Pine AR, Cirigliano SM, Nicholson JG, Hu Y, Linkous A, Miyaguchi K, Edwards L, Singhania R, Schwartz TH, Ramakrishna R, Pisapia DJ, Snuderl M, Elemento O, Fine HA. Tumor Microenvironment Is Critical for the Maintenance of Cellular States Found in Primary Glioblastomas. Cancer Discov. 2020 Jul;10(7):964-979. doi: 10.1158/2159-8290.CD-20-0057. Epub 2020 Apr 6. PMID: 32253265.

[137] Pirmoradi, L., Seyfizadeh, N., Ghavami, S., Zeki, A. A., & Shojaei, S. (2019). Targeting cholesterol metabolism in Glioblastoma: a new therapeutic approach in cancer therapy. Journal of investigative medicine : the official publication of the American Federation for Clinical Research, 67(4), 715–719. https://doi.org/10.1136/jim-2018-000962

[138] Pollard SM, Yoshikawa K, Clarke ID, Danovi D, Stricker S, Russell R, et al. Glioma stem cell lines expanded in adherent culture have tumor-specific phenotypes and are suitable for chemical and genetic screens. Cell Stem Cell. 2009;4:568–80.

[139] Poon CC, Sarkar S, Yong VW, Kelly JJP. Glioblastoma-associated microglia and macrophages: targets for therapies to improve prognosis. Brain. 2017 Jun 1;140(6):1548-1560. doi: 10.1093/brain/aww355. PMID: 28334886.

[140] Portales-Casamar E, Thongjuea S, Kwon AT, Arenillas D, Zhao X, Valen E, et al. JASPAR 2010: the greatly expanded open-access database of transcription factor binding profiles. Nucleic Acids Res. 2010 Jan;38(Database issue):D105-10.

[141] Prager BC, Bhargava S, Mahadev V, Hubert CG, Rich JN. Glioblastoma Stem Cells: Driving Resilience through Chaos. Trends Cancer. 2020 Mar;6(3):223-235. doi: 10.1016/j.trecan.2020.01.009. Epub 2020 Feb 3. PMID: 32101725; PMCID: PMC8779821.

[142] Prager BC, Xie Q, Bao S, Rich JN. Cancer Stem Cells: The Architects of the Tumor Ecosystem. Cell Stem Cell. 2019 Jan 3;24(1):41-53. doi: 10.1016/j.stem.2018.12.009. PMID: 30609398; PMCID: PMC6350931.

[143] Puchalski RB, Shah N, Miller J, Dalley R, Nomura SR, Yoon JG, Smith KA, Lankerovich M, Bertagnolli D, Bickley K, Boe AF, Brouner K, Butler S, Caldejon S, Chapin M, Datta S, Dee N, Desta T, Dolbeare T, Dotson N, Ebbert A, Feng D, Feng X, Fisher M, Gee G, Goldy J, Gourley L, Gregor BW, Gu G, Hejazinia N, Hohmann J, Hothi P, Howard R, Joines K, Kriedberg A, Kuan L, Lau C, Lee F, Lee H, Lemon T, Long F, Mastan N, Mott E, Murthy C, Ngo K, Olson E, Reding M, Riley Z, Rosen D, Sandman D, Shapovalova N, Slaughterbeck CR, Sodt A, Stockdale G, Szafer A, Wakeman W, Wohnoutka PE, White SJ, Marsh D, Rostomily RC, Ng L, Dang C, Jones A, Keogh B, Gittleman HR, Barnholtz-Sloan JS, Cimino PJ, Uppin MS, Keene CD, Farrokhi FR, Lathia JD, Berens ME, Iavarone A, Bernard A, Lein E, Phillips JW, Rostad SW, Cobbs C, Hawrylycz MJ, Foltz GD. An anatomic transcriptional atlas of human Glioblastoma. Science. 2018 May 11;360(6389):660-663. doi: 10.1126/science.aaf2666. PMID: 29748285; PMCID: PMC6414061.

[144] Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. Nat Med. 2013 Nov;19(11):1423-37. doi: 10.1038/nm.3394. PMID: 24202395; PMCID: PMC3954707.

[145] Raizer JJ, Abrey LE, Lassman AB, Chang SM, Lamborn KR, Kuhn JG, Yung WK, Gilbert MR, Aldape KA, Wen PY, Fine HA, Mehta M, Deangelis LM, Lieberman F, Cloughesy TF, Robins HI, Dancey J, Prados MD; North American Brain Tumor Consortium. A phase II trial of erlotinib in patients with recurrent malignant gliomas and nonprogressive Glioblastoma multiforme postradiation therapy. Neuro Oncol. 2010 Jan;12(1):95-103. doi: 10.1093/neuonc/nop015. Epub 2009 Dec 14. PMID: 20150372; PMCID: PMC2940554.

[146] Raymond E, Brandes AA, Dittrich C, Fumoleau P, Coudert B, Clement PM, Frenay M, Rampling R, Stupp R, Kros JM, Heinrich MC, Gorlia T, Lacombe D, van den Bent MJ; European Organisation for Research and Treatment of Cancer Brain Tumor Group Study. Phase II study of imatinib in patients with recurrent gliomas of various histologies: a European Organisation for Research and Treatment of Cancer Brain Tumor Group Study. J Clin Oncol. 2008 Oct 1;26(28):4659-65. doi: 10.1200/JCO.2008.16.9235. PMID: 18824712; PMCID: PMC2653126.

[147] Richards, L.M., Whitley, O.K.N., MacLeod, G. et al. Gradient of Developmental and Injury Response transcriptional states defines functional vulnerabilities underpinning Glioblastoma heterogeneity. Nat Cancer 2, 157–173 (2021).

[148] Ringel F, Pape H, Sabel M, Krex D, Bock HC, Misch M, Weyerbrock A, Westermaier T, Senft C, Schucht P, Meyer B, Simon M; SN1 study group. Clinical benefit from resection of recurrent Glioblastomas: results of a multicenter study including 503 patients with recurrent Glioblastomas undergoing surgical resection. Neuro Oncol. 2016 Jan;18(1):96-104. doi: 10.1093/neuonc/nov145. Epub 2015 Aug 4. PMID: 26243790; PMCID: PMC4677413.

[149] Rivera LB, Bergers G. Location, location, location: macrophage positioning within tumors determines pro- or antitumor activity. Cancer Cell. 2013 Dec 9;24(6):687-9. doi: 10.1016/j.ccr.2013.11.014. PMID: 24332035; PMCID: PMC4243836.

[150] Rutledge WC, Kong J, Gao J, Gutman DA, Cooper LA, Appin C, Park Y, Scarpace L, Mikkelsen T, Cohen ML, Aldape KD, McLendon RE, Lehman NL, Miller CR, Schniederjan MJ, Brennan CW, Saltz JH, Moreno CS, Brat DJ. Tumor-infiltrating lymphocytes in Glioblastoma are associated with specific genomic alterations and related to transcriptional class. Clin Cancer Res. 2013 Sep 15;19(18):4951-60. doi: 10.1158/1078-0432.CCR-13-0551. Epub 2013 Jul 17. PMID: 23864165; PMCID: PMC3865611.

[151] Sa JK, Chang N, Lee HW, Cho HJ, Ceccarelli M, Cerulo L, Yin J, Kim SS, Caruso FP, Lee M, Kim D, Oh YT, Lee Y, Her NG, Min B, Kim HJ, Jeong DE, Kim HM, Kim H, Chung S, Woo HG, Lee J, Kong DS, Seol HJ, Lee JI, Kim J, Park WY, Wang Q, Sulman EP, Heimberger AB, Lim M, Park JB, Iavarone A, Verhaak RGW, Nam DH. Transcriptional regulatory networks of tumor-associated macrophages that drive malignancy in

mesenchymal Glioblastoma. Genome Biol. 2020 Aug 26;21(1):216. doi: 10.1186/s13059-020-02140-x. PMID: 32847614; PMCID: PMC7448990.

[152] Sandmann T, Bourgon R, Garcia J, Li C, Cloughesy T, Chinot OL, Wick W, Nishikawa R, Mason W, Henriksson R et al (2015) Patients with proneural Glioblastoma may derive overall survival benefit from the addition of bevacizumab to first-line radiotherapy and temozolomide: retrospective analysis of the AVAglio trial. J Clin Oncol 33:2735–2744.

[153] Sankowski R, Böttcher C, Masuda T, Geirsdottir L, Sagar, Sindram E, Seredenina T, Muhs A, Scheiwe C, Shah MJ, Heiland DH, Schnell O, Grün D, Priller J, Prinz M. Mapping microglia states in the human brain through the integration of high-dimensional techniques. Nat Neurosci. 2019 Dec;22(12):2098-2110. doi: 10.1038/s41593-019-0532-y. Epub 2019 Nov 18. PMID: 31740814.

[154] Sasaki, M., Knobbe, C. B., Itsumi, M., Elia, A. J., Harris, I. S., Chio, I. I., Cairns, R. A., McCracken, S., Wakeham, A., Haight, J., Ten, A. Y., Snow, B., Ueda, T., Inoue, S., Yamamoto, K., Ko, M., Rao, A., Yen, K. E., Su, S. M., & Mak, T. W. (2012). D-2-hydroxyglutarate produced by mutant IDH1 perturbs collagen maturation and basement membrane function. Genes & development, 26(18), 2038–2049.

[155] Schaettler MO, Richters MM, Wang AZ, Skidmore ZL, Fisk B, Miller KE, Vickery TL, Kim AH, Chicoine MR, Osbun JW, Leuthardt EC, Dowling JL, Zipfel GJ, Dacey RG, Lu HC, Johanns TM, Griffith OL, Mardis ER, Griffith M, Dunn GP. Characterization of the Genomic and Immunologic Diversity of Malignant Brain Tumors through Multisector Analysis. Cancer Discov. 2022 Jan;12(1):154-171. doi: 10.1158/2159-8290.CD-21-0291. Epub 2021 Oct 5. PMID: 34610950.

[156] Schepers AG, Snippert HJ, Stange DE, van den Born M, van Es JH, van de Wetering M, Clevers H. Lineage tracing reveals Lgr5+ stem cell activity in mouse intestinal adenomas. Science. 2012 Aug 10;337(6095):730-5. doi: 10.1126/science.1224676.

[157] Schmitt MW, Loeb LA, Salk JJ. The influence of subclonal resistance mutations on targeted cancer therapy. Nat Rev Clin Oncol. 2016 Jun;13(6):335-47. doi: 10.1038/nrclinonc.2015.175. Epub 2015 Oct 20. PMID: 26483300; PMCID: PMC4838548.

[158] Schulze Heuling E, Knab F, Radke J, Eskilsson E, Martinez-Ledesma E, Koch A, et al. Prognostic Relevance of Tumor Purity and Interaction with MGMT Methylation in Glioblastoma. Mol Cancer Res. 2017 May;15(5):532-540.

[159] Scott JG, Bauchet L, Fraum TJ, Nayak L, Cooper AR, Chao ST, Suh JH, Vogelbaum MA, Peereboom DM, Zouaoui S, Mathieu-Daudé H, Fabbro-Peray P, Rigau V, Taillandier L, Abrey LE, DeAngelis LM, Shih JH, Iwamoto FM. Recursive partitioning analysis of prognostic factors for Glioblastoma patients aged 70 years or older. Cancer. 2012 Nov 15;118(22):5595-600. doi: 10.1002/cncr.27570. Epub 2012 Apr 19. PMID: 22517216; PMCID: PMC3402652.

[160] Segerman A, Niklasson M, Haglund C, Bergström T, Jarvius M, Xie Y, Westermark A, Sönmez D, Hermansson A, Kastemar M, Naimaie-Ali Z, Nyberg F, Berglund M, Sundström M, Hesselager G, Uhrbom L, Gustafsson M, Larsson R, Fryknäs M, Segerman B, Westermark B. Clonal Variation in Drug and Radiation Response among Glioma-Initiating Cells Is Linked to Proneural-Mesenchymal Transition. Cell Rep. 2016 Dec 13;17(11):2994-3009. doi: 10.1016/j.celrep.2016.11.056. PMID: 27974212.

[161] Sengupta S, George RE. Super-Enhancer-Driven Transcriptional Dependencies in Cancer. Trends Cancer. 2017 Apr;3(4):269-281. doi: 10.1016/j.trecan.2017.03.006. Epub 2017 Apr 12. PMID: 28718439; PMCID: PMC5546010.

[162] Serganova, I., & Blasberg, R. G. (2019). Molecular Imaging with Reporter Genes: Has Its Promise Been Delivered?. Journal of nuclear medicine : official publication, Society of Nuclear Medicine, 60(12), 1665–1681.

[163] Serresi M, Kertalli S, Li L, Schmitt MJ, Dramaretska Y, Wierikx J, Hulsman D, Gargiulo G. Functional antagonism of chromatin modulators regulates epithelial-mesenchymal transition. Sci Adv. 2021 Feb 24;7(9):eabd7974. doi: 10.1126/sciadv.abd7974. PMID: 33627422; PMCID: PMC7904264.

[164] Serresi M, Siteur B, Hulsman D, Company C, Schmitt MJ, Lieftink C, et al. Ezh2 inhibition in Kras-driven lung cancer amplifies inflammation and associated vulnerabilities. J Exp Med. Rockefeller University Press; 2018;215:3115–35.

[165] Sharma SV, Lee DY, Li B, Quinlan MP, Takahashi F, Maheswaran S, McDermott U, Azizian N, Zou L, Fischbach MA, Wong KK, Brandstetter K, Wittner B, Ramaswamy S, Classon M, Settleman J. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. Cell. 2010 Apr 2;141(1):69-80. doi: 10.1016/j.cell.2010.02.027. PMID: 20371346; PMCID: PMC2851638.

[166] Shergalis A, Bankhead A 3rd, Luesakul U, Muangsin N, Neamati N. Current Challenges and Opportunities in Treating Glioblastoma. Pharmacol Rev. 2018 Jul;70(3):412-445. doi: 10.1124/pr.117.014944. PMID: 29669750; PMCID: PMC5907910.

[167] Shibue T, Weinberg RA. EMT, CSCs, and drug resistance: the mechanistic link and clinical implications. Nat Rev Clin Oncol. 2017 Oct;14(10):611-629. doi: 10.1038/nrclinonc.2017.44. Epub 2017 Apr 11. PMID: 28397828; PMCID: PMC5720366.

[168] Shime H, Yabu M, Akazawa T, Kodama K, Matsumoto M, Seya T, Inoue N. Tumor-secreted lactic acid promotes IL-23/IL-17 proinflammatory pathway. J Immunol. 2008 Jun 1;180(11):7175-83. doi: 10.4049/jimmunol.180.11.7175. PMID: 18490716.

[169] Singh SK, Hawkins C, Clarke ID, Squire JA, Bayani J, Hide T, Henkelman RM, Cusimano MD, Dirks PB. Identification of human brain tumour initiating cells. Nature. 2004 Nov 18;432(7015):396-401. doi: 10.1038/nature03128. PMID: 15549107.

[170] Singh, S. K., Clarke, I. D., Terasaki, M., Bonn, V. E., Hawkins, C., Squire, J., & Dirks, P. B. (2003). Identification of a cancer stem cell in human brain tumors. Cancer research, 63(18), 5821–5828.

[171] Son MJ, Woolard K, Nam DH, Lee J, Fine HA. SSEA-1 is an enrichment marker for tumor-initiating cells in human Glioblastoma. Cell Stem Cell. 2009 May 8;4(5):440-52. doi: 10.1016/j.stem.2009.03.003. PMID: 19427293; PMCID: PMC7227614.

[172] Sottoriva, A., Spiteri, I., Piccirillo, S. G., Touloumis, A., Collins, V. P., Marioni, J. C., Curtis, C., Watts, C., & Tavaré, S. (2013). Intratumor heterogeneity in human Glioblastoma reflects cancer evolutionary dynamics. Proceedings of the National Academy of Sciences of the United States of America, 110(10), 4009–4014.

[173] Stupp R, Hegi ME, Mason WP, van den Bent MJ, Taphoorn MJ, Janzer RC, Ludwin SK, Allgeier A, Fisher B, Belanger K, Hau P, Brandes AA, Gijtenbeek J, Marosi C, Vecht CJ, Mokhtari K, Wesseling P, Villa S, Eisenhauer E, Gorlia T, Weller M, Lacombe D, Cairncross JG, Mirimanoff RO; European Organisation for Research and Treatment of Cancer Brain Tumour and Radiation Oncology Groups; National Cancer Institute of Canada Clinical Trials Group. 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. 2009 May;10(5):459-66. doi: 10.1016/S1470-2045(09)70025-7. Epub 2009 Mar 9. PMID: 19269895.

[174] Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJ, Belanger K, Brandes AA, Marosi C, Bogdahn U et al (2005) Radiotherapy plus concomitant and adjuvant temozolomide for Glioblastoma. N Engl J Med 352:987–996.

[175] Sturm, D., Witt, H., Hovestadt, V., Khuong-Quang, D. A., Jones, D. T., Konermann, C., Pfaff, E., Tönjes, M., Sill, M., Bender, S., Kool, M., Zapatka, M., Becker, N., Zucknick, M., Hielscher, T., Liu, X. Y., Fontebasso, A. M., Ryzhova, M., Albrecht, S., Jacob, K., … Pfister, S. M. (2012). Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of Glioblastoma. Cancer cell, 22(4), 425–437.

[176] Suchorska B, Weller M, Tabatabai G, Senft C, Hau P, Sabel MC, Herrlinger U, Ketter R, Schlegel U, Marosi C, Reifenberger G, Wick W, Tonn JC, Wirsching HG. Complete resection of contrast-enhancing tumor volume is associated with improved survival in recurrent Glioblastoma-results from the DIRECTOR trial. Neuro Oncol. 2016 Apr;18(4):549-56. doi: 10.1093/neuonc/nov326. Epub 2016 Jan 27. PMID: 26823503; PMCID: PMC4799687.

[177] Sulston JE, Schierenberg E, White JG, Thomson JN. The embryonic cell lineage of the nematode Caenorhabditis elegans. Dev Biol. 1983 Nov;100(1):64-119. doi: 10.1016/0012-1606(83)90201-4.

[178] Suvà ML, Tirosh I. Single-Cell RNA Sequencing in Cancer: Lessons Learned and Emerging Challenges. Mol Cell. 2019 Jul 11;75(1):7-12. doi: 10.1016/j.molcel.2019.05.003. PMID: 31299208.

[179] Suvà ML, Tirosh I. The Glioma Stem Cell Model in the Era of Single-Cell Genomics. Cancer Cell. 2020 May 11;37(5):630-636. doi: 10.1016/j.ccell.2020.04.001. PMID: 32396858.

[180] Suzuki, H., Aoki, K., Chiba, K., Sato, Y., Shiozawa, Y., Shiraishi, Y., Shimamura, T., Niida, A., Motomura, K., Ohka, F., Yamamoto, T., Tanahashi, K., Ranjit, M., Wakabayashi, T., Yoshizato, T., Kataoka, K., Yoshida, K., Nagata, Y., Sato-Otsubo, A., Tanaka, H., … Ogawa, S. (2015). Mutational landscape and clonal architecture in grade II and III gliomas. Nature genetics, 47(5), 458–468. https://doi.org/10.1038/ng.3273

[181] Szulzewsky F, Arora S, de Witte L, Ulas T, Markovic D, Schultze JL, Holland EC, Synowitz M, Wolf SA, Kettenmann H. Human Glioblastoma-associated microglia/monocytes express a distinct RNA profile compared to human control and murine samples. Glia. 2016 Aug;64(8):1416-36. doi: 10.1002/glia.23014. PMID: 27312099.

[182] Tam SY, Wu VWC, Law HKW. Hypoxia-Induced Epithelial-Mesenchymal Transition in Cancers: HIF-1α and Beyond. Front Oncol. 2020 Apr 8;10:486. doi: 10.3389/fonc.2020.00486. PMID: 32322559; PMCID: PMC7156534.

[183] Tamimi AF, Juweid M. Epidemiology and Outcome of Glioblastoma. In: De Vleeschouwer S, editor. Glioblastoma [Internet]. Brisbane (AU): Codon Publications; 2017 Sep 27. Chapter 8. PMID: 29251870.

[184] Thakkar JP, Dolecek TA, Horbinski C, Ostrom QT, Lightner DD, Barnholtz-Sloan JS, Villano JL. Epidemiologic and molecular prognostic review of Glioblastoma. Cancer Epidemiol Biomarkers Prev. 2014 Oct;23(10):1985-96. doi: 10.1158/1055-9965.EPI-14-0275. Epub 2014 Jul 22. PMID: 25053711; PMCID: PMC4185005.

[185] Thankamony AP, Saxena K, Murali R, Jolly MK, Nair R. Cancer Stem Cell Plasticity - A Deadly Deal. Front Mol Biosci. 2020 Apr 30;7:79. doi: 10.3389/fmolb.2020.00079. PMID: 32426371; PMCID: PMC7203492.

[186] Thurman, R.E., Day, N., Noble, W.S., and Stamatoyannopoulos, J.A. (2007). Identification of higher-order functional domains in the human ENCODE regions. Genome Res 17, 917–927.

[187] Tomaszewski W, Sanchez-Perez L, Gajewski TF, Sampson JH. Brain Tumor Microenvironment and Host State: Implications for Immunotherapy. Clin Cancer Res. 2019 Jul 15;25(14):4202-4210. doi: 10.1158/1078-0432.CCR-18-1627. Epub 2019 Feb 25. PMID: 30804019; PMCID: PMC6635001.

[188] Touat M, Li YY, Boynton AN, Spurr LF, Iorgulescu JB, Bohrson CL, Cortes-Ciriano I, Birzu C, Geduldig JE, Pelton K, Lim-Fat MJ, Pal S, Ferrer-Luna R, Ramkissoon SH, Dubois F, Bellamy C, Currimjee N, Bonardi J, Qian K, Ho P, Malinowski S, Taquet L, Jones RE, Shetty A, Chow KH, Sharaf R, Pavlick D, Albacker LA, Younan N, Baldini C, Verreault M, Giry M, Guillerm E, Ammari S, Beuvon F, Mokhtari K, Alentorn A, Dehais C, Houillier C, Laigle-Donadey F, Psimaras D, Lee EQ, Nayak L, McFaline-Figueroa JR, Carpentier A, Cornu P, Capelle L, Mathon B, Barnholtz-Sloan JS, Chakravarti A, Bi WL, Chiocca EA, Fehnel KP, Alexandrescu S, Chi SN, Haas-Kogan D, Batchelor TT, Frampton GM, Alexander BM, Huang RY, Ligon AH, Coulet F, Delattre JY, Hoang-Xuan K, Meredith DM, Santagata S, Duval A, Sanson M, Cherniack AD, Wen PY, Reardon DA, Marabelle A, Park PJ, Idbaih A, Beroukhim R, Bandopadhayay P, Bielle F, Ligon KL. Mechanisms and therapeutic implications of hypermutation in gliomas. Nature. 2020 Apr;580(7804):517-523. doi: 10.1038/s41586-020-2209-9. Epub 2020 Apr 15. PMID: 32322066; PMCID: PMC8235024.

[189] Turner DL, Cepko CL. A common progenitor for neurons and glia persists in rat retina late in development. Nature. 1987 Jul 9-15;328(6126):131-6. doi: 10.1038/328131a0.

[190] van den Bent M. J. (2010). Interobserver variation of the histopathological diagnosis in clinical trials on glioma: a clinician's perspective. Acta neuropathologica, 120(3), 297–304.

[191] van den Bent MJ, Brandes AA, Rampling R, Kouwenhoven MC, Kros JM, Carpentier AF, Clement PM, Frenay M, Campone M, Baurain JF, Armand JP, Taphoorn MJ, Tosoni A, Kletzl H, Klughammer B, Lacombe D, Gorlia T. Randomized phase II trial of erlotinib versus temozolomide or carmustine in recurrent Glioblastoma: EORTC brain tumor group study 26034. J Clin Oncol. 2009 Mar 10;27(8):1268-74. doi: 10.1200/JCO.2008.17.5984. Epub 2009 Feb 9. PMID: 19204207; PMCID: PMC2667826.

[192] Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg effect: the metabolic requirements of cell proliferation. Science. 2009 May 22;324(5930):1029-33. doi: 10.1126/science.1160809. PMID: 19460998; PMCID: PMC2849637.

[193] VanHorn S, Morris SA. Next-Generation Lineage Tracing and Fate Mapping to Interrogate Development. Dev Cell. 2021 Jan 11;56(1):7-21. doi: 10.1016/j.devcel.2020.10.021.

[194] Varn, F. S., Johnson, K. C., Wade, T. E., Malta, T. M., Sabedot, T. S., Barthel, F. P., ... & GLASS Consortium. (2021). Longitudinal analysis of diffuse glioma reveals cell state dynamics at recurrence associated with changes in genetics and the microenvironment. bioRxiv.

[195] Végran F, Boidot R, Michiels C, Sonveaux P, Feron O. Lactate influx through the endothelial cell monocarboxylate transporter MCT1 supports an NF-κB/IL-8 pathway that drives tumor angiogenesis. Cancer Res. 2011 Apr 1;71(7):2550-60. doi: 10.1158/0008-5472.CAN-10-2828. Epub 2011 Feb 7. PMID: 21300765.

[196] Venkataramani V, Tanev DI, Strahle C, Studier-Fischer A, Fankhauser L, Kessler T, Körber C, Kardorff M, Ratliff M, Xie R, Horstmann H, Messer M, Paik SP, Knabbe J, Sahm F, Kurz FT, Acikgöz AA, Herrmannsdörfer F, Agarwal A, Bergles DE, Chalmers A, Miletic H, Turcan S, Mawrin C, Hänggi D, Liu HK, Wick W, Winkler F, Kuner T. Glutamatergic synaptic input to glioma cells drives brain tumour progression. Nature. 2019 Sep;573(7775):532-538. doi: 10.1038/s41586-019-1564-x. Epub 2019 Sep 18. PMID: 31534219.

[197] Venkatesh HS, Morishita W, Geraghty AC, Silverbush D, Gillespie SM, Arzt M, Tam LT, Espenel C, Ponnuswami A, Ni L, Woo PJ, Taylor KR, Agarwal A, Regev A, Brang D, Vogel H, Hervey-Jumper S, Bergles DE, Suvà ML, Malenka RC, Monje M. Electrical and synaptic integration of glioma into neural circuits. Nature. 2019 Sep;573(7775):539-545. doi: 10.1038/s41586-019-1563-y. Epub 2019 Sep 18. PMID: 31534222; PMCID: PMC7038898.

[198] Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, Miller CR, Ding L, Golub T, Mesirov JP, Alexe G, Lawrence M, O'Kelly M, Tamayo P, Weir BA, Gabriel S, Winckler W, Gupta S, Jakkula L, Feiler HS, Hodgson JG, James CD, Sarkaria JN, Brennan C, Kahn A, Spellman PT, Wilson RK, Speed TP, Gray JW, Meyerson M, Getz G, Perou

CM, Hayes DN; Cancer Genome Atlas Research Network. Integrated genomic analysis identifies clinically relevant subtypes of Glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell. 2010 Jan 19;17(1):98-110. doi: 10.1016/j.ccr.2009.12.020. PMID: 20129251; PMCID: PMC2818769.

[199] Villa, G. R., Hulce, J. J., Zanca, C., Bi, J., Ikegami, S., Cahill, G. L., Gu, Y., Lum, K. M., Masui, K., Yang, H., Rong, X., Hong, C., Turner, K. M., Liu, F., Hon, G. C., Jenkins, D., Martini, M., Armando, A. M., Quehenberger, O., Cloughesy, T. F., … Mischel, P. S. (2016). An LXR-Cholesterol Axis Creates a Metabolic Co-Dependency for Brain Cancers. Cancer cell, 30(5), 683–693. https://doi.org/10.1016/j.ccell.2016.09.008

[200] Vogt, W. Gestaltungsanalyse am amphibienkeim mit örtlicher vitalfärbung. W. Roux Archiv f. Entwicklungsmechanik, 120 (1929), pp. 384-706

[201] Wagner DE, Klein AM. Lineage tracing meets single-cell omics: opportunities and challenges. Nat Rev Genet. 2020 Jul;21(7):410-427. doi: 10.1038/s41576-020-0223-2.

[202] Wang J, Cazzato E, Ladewig E, Frattini V, Rosenbloom DI, Zairis S, Abate F, Liu Z, Elliott O, Shin YJ, Lee JK, Lee IH, Park WY, Eoli M, Blumberg AJ, Lasorella A, Nam DH, Finocchiaro G, Iavarone A, Rabadan R. Clonal evolution of Glioblastoma under therapy. Nat Genet. 2016 Jul;48(7):768-76. doi: 10.1038/ng.3590. Epub 2016 Jun 6. PMID: 27270107; PMCID: PMC5627776.

[203] Wang L, Babikir H, Müller S, Yagnik G, Shamardani K, Catalan F, Kohanbash G, Alvarado B, Di Lullo E, Kriegstein A, Shah S, Wadhwa H, Chang SM, Phillips JJ, Aghi MK, Diaz AA. The Phenotypes of Proliferating Glioblastoma Cells Reside on a Single Axis of Variation. Cancer Discov. 2019 Dec;9(12):1708-1719. doi: 10.1158/2159-8290.CD-19-0329. Epub 2019 Sep 25. PMID: 31554641; PMCID: PMC7161589.

[204] Wang LB, Karpova A, Gritsenko MA, Kyle JE, Cao S, Li Y, Rykunov D, Colaprico A, Rothstein JH, Hong R, Stathias V, Cornwell M, Petralia F, Wu Y, Reva B, Krug K, Pugliese P, Kawaler E, Olsen LK, Liang WW, Song X, Dou Y, Wendl MC, Caravan W, Liu W, Cui Zhou D, Ji J, Tsai CF, Petyuk VA, Moon J, Ma W, Chu RK, Weitz KK, Moore RJ, Monroe ME, Zhao R, Yang X, Yoo S, Krek A, Demopoulos A, Zhu H, Wyczalkowski MA, McMichael JF, Henderson BL, Lindgren CM, Boekweg H, Lu S, Baral J, Yao L, Stratton

KG, Bramer LM, Zink E, Couvillion SP, Bloodsworth KJ, Satpathy S, Sieh W, Boca SM, Schürer S, Chen F, Wiznerowicz M, Ketchum KA, Boja ES, Kinsinger CR, Robles AI, Hiltke T, Thiagarajan M, Nesvizhskii AI, Zhang B, Mani DR, Ceccarelli M, Chen XS, Cottingham SL, Li QK, Kim AH, Fenyö D, Ruggles KV, Rodriguez H, Mesri M, Payne SH, Resnick AC, Wang P, Smith RD, Iavarone A, Chheda MG, Barnholtz-Sloan JS, Rodland KD, Liu T, Ding L; Clinical Proteomic Tumor Analysis Consortium. Proteogenomic and metabolomic characterization of human Glioblastoma. Cancer Cell. 2021 Apr 12;39(4):509-528.e20. doi: 10.1016/j.ccell.2021.01.006. Epub 2021 Feb 11. PMID: 33577785; PMCID: PMC8044053.

[205] Wang X, Prager BC, Wu Q, Kim LJY, Gimple RC, Shi Y, Yang K, Morton AR, Zhou W, Zhu Z, Obara EAA, Miller TE, Song A, Lai S, Hubert CG, Jin X, Huang Z, Fang X, Dixit D, Tao W, Zhai K, Chen C, Dong Z, Zhang G, Dombrowski SM, Hamerlik P, Mack SC, Bao S, Rich JN. Reciprocal Signaling between Glioblastoma Stem Cells and Differentiated Tumor Cells Promotes Malignant Progression. Cell Stem Cell. 2018 Apr 5;22(4):514-528.e5. doi: 10.1016/j.stem.2018.03.011. PMID: 29625067; PMCID: PMC5947947.

[206] Wang Z, Zhang H, Xu S, Liu Z, Cheng Q. The adaptive transition of Glioblastoma stem cells and its implications on treatments. Signal Transduct Target Ther. 2021 Mar 23;6(1):124. doi: 10.1038/s41392-021-00491-w. PMID: 33753720; PMCID: PMC7985200.

[207] Wang, Q., Hu, B., Hu, X., Kim, H., Squatrito, M., Scarpace, L., deCarvalho, A. C., Lyu, S., Li, P., Li, Y., Barthel, F., Cho, H. J., Lin, Y. H., Satani, N., Martinez-Ledesma, E., Zheng, S., Chang, E., Sauvé, C. G., Olar, A., Lan, Z. D., … Verhaak, R. (2017). Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. Cancer cell, 32(1), 42–56.e6. https://doi.org/10.1016/j.ccell.2017.06.003

[208] Weil S, Osswald M, Solecki G, Grosch J, Jung E, Lemke D, Ratliff M, Hänggi D, Wick W, Winkler F. Tumor microtubes convey resistance to surgical lesions and chemotherapy in gliomas. Neuro Oncol. 2017 Oct 1;19(10):1316-1326. doi: 10.1093/neuonc/nox070. PMID: 28419303; PMCID: PMC5596180.

[209] Weisblat DA, Sawyer RT, Stent GS. Cell lineage analysis by intracellular injection of a tracer enzyme. Science. 1978 Dec 22;202(4374):1295-8. doi: 10.1126/science.725606. PMID: 725606.

[210] Weller M, Cloughesy T, Perry JR, Wick W. Standards of care for treatment of recurrent Glioblastoma--are we there yet? Neuro Oncol. 2013 Jan;15(1):4-27. doi: 10.1093/neuonc/nos273. Epub 2012 Nov 7. PMID: 23136223; PMCID: PMC3534423.

[211] Weller M, Wick W, Aldape K, Brada M, Berger M, Pfister SM, Nishikawa R, Rosenthal M, Wen PY, Stupp R, Reifenberger G. Glioma. Nat Rev Dis Primers. 2015 Jul 16;1:15017. doi: 10.1038/nrdp.2015.17. PMID: 27188790.

[212] Wen PY, Yung WK, Lamborn KR, Dahia PL, Wang Y, Peng B, Abrey LE, Raizer J, Cloughesy TF, Fink K, Gilbert M, Chang S, Junck L, Schiff D, Lieberman F, Fine HA, Mehta M, Robins HI, DeAngelis LM, Groves MD, Puduvalli VK, Levin V, Conrad C, Maher EA, Aldape K, Hayes M, Letvak L, Egorin MJ, Capdeville R, Kaplan R, Murgo AJ, Stiles C, Prados MD. Phase I/II study of imatinib mesylate for recurrent malignant gliomas: North American Brain Tumor Consortium Study 99-08. Clin Cancer Res. 2006 Aug 15;12(16):4899-907. doi: 10.1158/1078-0432.CCR-06-0773. PMID: 16914578.

[213] Whitman, CO. Memoirs: the embryology of clepsine. J. Cell Sci. s2-18, 215–315 (1878).

[214] Wick W, Osswald M, Wick A, Winkler F. Treatment of Glioblastoma in adults. Ther Adv Neurol Disord. 2018 Jul 25;11:1756286418790452. doi: 10.1177/1756286418790452. PMID: 30083233; PMCID: PMC6071154.

[215] Wick W, Platten M, Wick A, Hertenstein A, Radbruch A, Bendszus M, Winkler F. Current status and future directions of anti-angiogenic therapy for gliomas. Neuro Oncol. 2016 Mar;18(3):315-28. doi: 10.1093/neuonc/nov180. Epub 2015 Oct 12. PMID: 26459812; PMCID: PMC4767238.

[216] Wilson TA, Karajannis MA, Harter DH. Glioblastoma multiforme: State of the art and future therapeutics. Surg Neurol Int. 2014 May 8;5:64. doi: 10.4103/2152-7806.132138. PMID: 24991467; PMCID: PMC4078454.

[217] Wood MD, Reis GF, Reuss DE, Phillips JJ. Protein Analysis of Glioblastoma Primary and Posttreatment Pairs Suggests a Mesenchymal Shift at Recurrence. J Neuropathol Exp Neurol. 2016 Oct;75(10):925-935. doi: 10.1093/jnen/nlw068. Epub 2016 Aug 18. PMID: 27539476; PMCID: PMC6281078.

[218] Xie R, Kessler T, Grosch J, Hai L, Venkataramani V, Huang L, Hoffmann DC, Solecki G, Ratliff M, Schlesner M, Wick W, Winkler F. Tumor cell network integration in glioma represents a stemness feature. Neuro Oncol. 2021 May 5;23(5):757-769. doi: 10.1093/neuonc/noaa275. PMID: 33320195; PMCID: PMC8099480.

[219] Yabo YA, Niclou SP, Golebiewska A. Cancer cell heterogeneity and plasticity: A paradigm shift in Glioblastoma. Neuro Oncol. 2021 Dec 21:noab269. doi: 10.1093/neuonc/noab269. Epub ahead of print. PMID: 34932099.

[220] Yan, H., Parsons, D. W., Jin, G., McLendon, R., Rasheed, B. A., Yuan, W., Kos, I., Batinic-Haberle, I., Jones, S., Riggins, G. J., Friedman, H., Friedman, A., Reardon, D., Herndon, J., Kinzler, K. W., Velculescu, V. E., Vogelstein, B., & Bigner, D. D. (2009). IDH1 and IDH2 mutations in gliomas. The New England journal of medicine, 360(8), 765–773. https://doi.org/10.1056/NEJMoa0808710

[221] Yin J, Oh YT, Kim JY, Kim SS, Choi E, Kim TH, Hong JH, Chang N, Cho HJ, Sa JK, Kim JC, Kwon HJ, Park S, Lin W, Nakano I, Gwak HS, Yoo H, Lee SH, Lee J, Kim JH, Kim SY, Nam DH, Park MJ, Park JB. Transglutaminase 2 Inhibition Reverses Mesenchymal Transdifferentiation of Glioma Stem Cells by Regulating C/EBPβ Signaling. Cancer Res. 2017 Sep 15;77(18):4973-4984. doi: 10.1158/0008-5472.CAN-17-0388. Epub 2017 Jul 28. PMID: 28754668.

8 Appendix

Oligonucleotides	***Resource***	***Supplier***
RELA sgRNA	GAU-CUCCACAUAGGGGCCAG	Synthego
Recombinant DNA	***Resource***	***Supplier***
pCMV-VSV-G	Plasmid #8454	Addgene
pRSV-REV	Plasmid #12253	Addgene
pMDLG/pRRE	Plasmid #12251	Addgene
pHR-TRE3G-KRAB-dCas9-P2A-mCherry	Plasmid #60954	Addgene
LV-CFP	Plasmid #25998	Addgene
pHR-Tet3G	(61)	NA
Human CRISPR Knockout Pooled Library (Brunello)	Plasmid #73179	Addgene
Human Kinome CRISPRi library	NKI	NA
FH1-MGT#1-mVenus	This paper	Addgene ID #164094
FH1-MGT#1-mCherry	This paper	Addgene ID #164095
FH1-MGT#2-mVenus	This paper	Addgene ID #164096
FH1-MGT#2-mCherry	This paper	Addgene ID #164097
FH1-PNGT#1-mCherry	This paper	Addgene ID #164098
FH1-PNGT#2-mCherry	This paper	Addgene ID #164099
FH1-MGT#1-mVenus-PGK-H2B-CFP	This paper	Addgene ID #164100
FH1-MGT#2-mVenus-PGK-H2B-CFP	This paper	Addgene ID #164101
FH1-MGT#2-mCherry-PGK-H2B-CFP	This paper	Addgene ID #164102
FH1-PNGT#1-mCherry-PGK-H2B-CFP	This paper	Addgene ID #164103
FH1-PNGT#2-mCherry-PGK-H2B-CFP	This paper	Addgene ID #164104
FH1-CLGT#1-mCherry-PGK-H2B-CFP	This paper	Addgene ID #164105

FH1-CLGT#2-mCherry-PGK-H2B-CFP	This paper	Addgene ID #164106
pB-Supertransposase	PB210PA-1	System Biosciences
Antibodies	***Product code***	***Supplier***
Alexa Fluor 647	A31573	Thermofisher
	A11055	
	A31571	
Alexa Fluor 488	A21206	Thermofisher
Alexa Fluor 555	A31570	Thermofisher
GFP	ab6556	Abcam
Vinculin	NA	NA
p-Stat3 y705	9145L	Cell Signaling
Stat3	sc-482x	Santa Cruz
p-NFKB p65	3033P	Cell Signaling
NFKB p65	86299	Abcam
p-p38 t180 d3f9	45115	Cell Signaling
p-p38	9211s	Millipore
Nestin	611658	BD Biosciences
MED1	ab64965	Abcam
BRD4	ab128874	Abcam
Tubulin	T5168	BD Biosciences
p-yH2AX Ser 139	05-636	Millipore
K27me3	07-449	Millipore
H3 total	1791	Abcam
E-Cadherin	31950	Cell Signaling
Vimentin	5741s	Cell Signaling
Goat Anti-Mouse IgG (H L) - HRP	626520	Invitrogen

Goat Anti-Rabbit IgG (H L) - HRP	G21234	Invitrogen
CD133/2 (293C)-APC	130-090-851	Miltenyi
Celcein UltraBlue	21908	AAT Bioquest
ZombieRed	423110	Biolegend
DRAQ5	62251	Thermofisher
Hoechst 33258	16756	Cayman Chemicals

Compounds	***Product code***	***Supplier***
All-trans retinoic acid (ATRA)	R2625	Sigma
IL6	206-IL	R&D Systems
LPS	ALX-581	Enzo
TNFα	210-TA	R&D Systems
TGFb	240-B	R&D Systems
IFNg	285-IF	R&D Systems
Tenascin C	MBS230239	Mybiosource
HGF	294-HG	R&D Systems
IGF	50356,1	Biomol
FBS	10270106	Gibco
GSK126	NA	Cancer Epigenetics
Activin A	BV-P1078	Enzo
NRG1	97642,1	Biomol
IL1b	CYT-094	Biotrend
LIF	GW-TL-633-100	FF Technology e.K.
Human Serum	088HSER	TEBU-Bio
Topotecan	sc-204919A	SCBT
Olaparib	SEL-S1060	SelleckChem
Bay11-7085	B5681	Sigma
WP1066	S2796	SelleckChem

WAY-242623	PZ0257	SelleckChem
Mitomycin C	sc-3514	SCBT
Temozolomide	T2577	SelleckChem
TAK733	S2617	SelleckChem
VE-821	Cay17587	CaymanChem
Ku-60019	Cay17502	CaymanChem
NU 7441	Cay14881	CaymanChem
IKK-16 (IKK Inhibitor VII)	S2882	SelleckChem
RHB-A	Y40001	Takara
BD Matrigel GFR-Red, Phenol-red-free	734-1101	VWR
Matrigel-Matrix, Phenol-red-free	734-0272	VWR
RPMI 1640 Medium	21875091	Life Technologies (Gibco)
DMEM, high glucose, pyruvate	41966052	Life Technologies (Gibco)
FluoroBrite Medium	A1896701	Thermofisher
Trypsin-EDTA (0.05%), phenol red	25300054	Life Technologies (Gibco)
DMEM/F-12, GlutaMAX Supplement	31331093	Life Technologies (Gibco)
MEM α, nucleosides, GlutaMAX Supplement	733-1693	lonza
ALPHA MEM EAGLE W/O L-GLUTAMINE, DEOXYRI	32571028	Life Technologies (Gibco)
FuGENE HD Transfection Reagent	E2312	Promega
Dead Cell Removal Kit	130-090-101	Miltenyi Biotec
rhFGF	233-FB-025/CF	RD System
PDGFaa recombinant human protein	221-AA-050	R&D Systems
Recombinant Human EGF protein	236-EG-200	R&D Systems
Recombinant Human Interleukin-6 (IL-6) Protein	GW-TL-512-100	FF Technology e.K.
Recombinant Human Interleukin-22 (IL-22) Protein	GW-TL-631-50	FF Technology e.K.
Recombinant Human Stem Cell Factor (SCF) Protein	GW-TL-504-100	FF Technology e.K.

Recombinant Human FLT3 Ligand Protein	GW-TL-505-100	FF Technology e.K.
Recombinant Human Leukemia Inhibitory Factor (LIF) Protein	GW-TL-633-100	FF Technology e.K.
Recombinant Human Epidermal Growth Factor (EGF) Protein	GW-TL-613-1000	FF Technology e.K.
Recombinant Human PDGFbb Protein	GW-TL-644-100	FF Technology e.K.
Recombinant Human GM-CSF Protein	GW-TL-302-A100	FF Technology e.K.
Recombinant Human Thrombopoietin (TPO) (carrier-free)	BLD-597402	Biozol Diagnostica (Biolegend)
Y-27632 ROCK Inhibitor	S1049	SelleckChem
UM171	S7608	SelleckChem
StemRegenin 1 (SR1)	S2858	SelleckChem
P3 Primary Cell 4D-Nucleofector® X Kit S	V4XP-3032	lonza
Live/Dead Fixable Aqua Dead Cell Stain Kit	L34965	ThermoFisher
NOC-18	BV-2492	Enzo Life Sciences
NG-nitro-L-arginine methyl ester (L-NAME)	S2877	Selleck Chemicals
Low Density Lipoprotein from Human Plasma, oxidized (OxLDL)	L34357	Life Technologies
Anti-LDL Receptor antibody	ab52818	Abcam

Software and Algorithms	***Resource Version***	***Link***
bcl2fastq v2.20.0	v2.20.0	Illumina
R	v3.5; v3.6	https://cran.r-project.org/
Trim Galore	v0.6.2	https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
bwa	v0.7.17-r1188	http://bio-bwa.sourceforge.net/
Tophat2	v2.1.0	https://ccb.jhu.edu/software/tophat/index.shtml
Cutadapt	v1.7.1	https://cutadapt.readthedocs.io/en/v1.7.1/
HTSEq	v0.6.1p1	https://htseq.readthedocs.io/en/master/

dupRadar	v1.18	http://bioconductor.org/packages/release/bioc/html/dupRadar.html
STAR	v2.6.0c	https://github.com/alexdobin/STAR
GlioVis	NA	http://gliovis.bioinfo.cnio.es/
UCSC Genome browser hg19	NA	https://genome.ucsc.edu/
Venny	v2.1	https://bioinfogp.cnb.csic.es/tools/venny/
SeqMonk	v1.47	https://www.bioinformatics.babraham.ac.uk/
GraphPad Prism	v.8	https://www.graphpad.com/scientific-software/prism/
Flowjo	v.10	https://www.flowjo.com/
IPA	NA	https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/
GSVA	v1.32.0	https://bioconductor.org/packages/release/bioc/html/GSVA.html
piano	v2.0.2	https://bioconductor.org/packages/release/bioc/html/piano.html
fgsea	v1.14.0	https://bioconductor.org/packages/release/bioc/html/fgsea.html
Ggplot2	v3.3.2	https://cran.r-project.org/web/packages/ggplot2/index.html
UpSetR	v1.4	https://cran.r-project.org/web/packages/UpSetR/index.html
uwot	v0.1.8	https://cran.r-project.org/web/packages/uwot/index.html
pheatmap	v1.0.12	https://cran.r-project.org/web/pa

		ckages/pheatmap/index.html
DESeq2	v1.24	http://bioconductor.org/packages/release/bioc/html/DESeq2.html
clusterProfiler	v3.16.1	https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html
caRpools	v0.83	https://cran.r-project.org/web/packages/caRpools/
sva	v3.32	https://www.bioconductor.org/packages/release/bioc/html/sva.html
MEME suite tools	v5.0.2	http://meme-suite.org/
PROGENy	v1.6	https://bioconductor.org/packages/release/bioc/html/progeny.html
sva	v3.32.1	https://bioconductor.org/packages/release/bioc/html/GSVA.html
sLCR original pipeline	v1	https://gitlab.com/gargiulo_lab/publications/slcr_publication

Deposited Data	***Resource***	***Accession Code***
Raw and analyzed data	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE136751	GEO: GSE136751
TCGA GBM RNA-seq	http://cancergenome.nih.gov	phs000178.v3.p3
GSCs RNA-seq	https://www.ncbi.nlm.nih.gov/geo/	GSE119834
GSCs RNA-seq	https://www.ncbi.nlm.nih.gov/geo/	GSE67089
GSCs RNA-seq	https://www.ncbi.nlm.nih.gov/geo/	GSE8049

Table 1 - List of reagents, software and deposited data.

ID	MGT#1	MGT#2	CLGT#1	CLGT#2	PNGT#1	PNGT#2
PHENOTYPE	GBM-MES (Verhaak et al. 2010)	GBM-MES (Verhaak et al. 2010)	GBM-CL (Verhaak et al. 2010)	GBM-CL (Verhaak et al. 2010)	GBM-PN (Verhaak et al. 2010)	GBM-PN (Verhaak et al. 2010)
FACS SIGNAL IDH-MUT GICS	weak	weak	strong	strong	strong	strong
FACS SIGNAL IDH-WT GICS	weak	weak	strong	strong	strong	strong
TFBS RATIO	47,50	51,25	47,37	54,39	52,35	48,32
SIGNATURE-GENES	27	27	32	32	46	46
TF GENES	62	62	34	34	60	60
TFBS MOTIFS	160	160	114	114	149	149
CRE	5,51	6,91	7,40	7,41	7,09	7,76
SEQUENCE LENGTH	827	1037	1110	1112	1064	1164

Table 2 – Input data and features of sLCR reporters

	MGT#1
TFBS motifs	MA0002.1_RUNX1; MA0004.1_Arnt; MA0050.1_IRF1; MA0052.1_MEF2A; MA0061.1_NF-kappaB; MA0073.1_RREB1; MA0079.1_SP1; MA0080.1_SPI1; MA0083.1_SRF; MA0101.1_REL; MA0105.1_NFKB1; MA0112.1_ESR1; MA0117.1_Mafb; MA0002.2_RUNX1; MA0112.2_ESR1; MA0151.1_ARID3A; MA0163.1_PLAG1; MA0080.2_SPI1; MA0079.2_SP1; BU0001_Arid3a_primary; BU0002_Arid5a_primary; BU0023_Gata6_primary; BU0035_Irf5_primary; BU0038_Jundm2_primary; BU0041_Mafb_primary; BU0058_Sfpi1_primary; BU0075_Sp100_primary; BU0078_Srf_primary; ETS0002_h-ELF1; ETS0005_h-ELF4; ETS0008_h-Elk3; ETS0020_h-ETV7; ETS0022_h-Fli1; ETS0027_h-SPI1; HOMEO0040_Hlx1_2350.1; HOMEO0063_Hoxb6_3428.2; HOMEO0146_Pou2f2_3748.1; HOMEO0156_Prrx1_3442.1; TA0011_RUNX3_monomer; TA0013_POU2F2_monomer; TA0026_PRDM1_dimer; HC_AHR_si; HC_ARI3A_do; HC_ARI3A_f1; HC_ARNT_f1; HC_BACH1_si; HC_BATF_si; HC_CEBPB_f1; HC_CEBPD_f1; HC_ELF1_f1; HC_ELK3_f1; HC_ESR1_do; HC_ETV7_si; HC_FLI1_f1; HC_FOSL1_f2; HC_FOSL2_f1; HC_GATA6_f2; HC_HXB6_f1; HC_IKZF1_f1; HC_IRF1_si; HC_IRF5_f1; HC_IRF8_si; HC_JUNB_f1; HC_KLF6_si; HC_MAFB_f1; HC_MAF_f1; HC_MBD2_si; HC_MEF2A_f1; HC_NFKB1_f1; HC_NFKB2_f1; HC_PLAG1_f1; HC_PLAG1_si; HC_PO2F2_si; HC_PRDM1_f1; HC_PRRX1_f1; HC_RELB_si; HC_REL_do; HC_RREB1_si; HC_RUNX1_f1; HC_RUNX2_f1; HC_RUNX3_si; HC_SP1_f1; HC_SP1_f2; HC_SPI1_si; HC_SRF_do; HC_STA5A_do; HC_STAT2_f1; HC_STAT6_do; HC_TEAD3_si; HC_TFE3_f1; HC_VDR_f1; HC_VDR_f2; HC_ZBT7A_f1; HC_ZBT7B_si; HC_ZFHX3_f1; HC_low_KLF5_si; HC_low_OBF1_f1; TA0055_PRDM1_full_monomer_dimer; TA0059_SP1_DBD_monomer; TA0067_ZBTB7A_DBD_monomer; TA0068_ZBTB7B_full_monomer; TA0111_ELF1_DBD_monomer; TA0112_ELF1_full_monomer; TA0115_ELF4_full_monomer; TA0121_ELK3_DBD_monomer; TA0140_Elk3_DBD_monomer; TA0142_FLI1_DBD_monomer; TA0143_FLI1_DBD_monomer; TA0144_FLI1_full_monomer; TA0145_FLI1_full_monomer; TA0153_SPI1_full_monomer; TA0179_IRF5_full_dimer; TA0180_IRF5_full_monomer; TA0183_IRF8_DBD_dimer; TA0184_IRF8_full_dimer; TA0186_MEF2A_DBD_dimer; TA0189_SRF_DBD_dimer; TA0190_SRF_full_dimer; TA0219_NFKB1_DBD_dimer; TA0220_NFKB2_DBD_dimer; TA0239_POU2F2_DBD_monomer; TA0240_POU2F2_DBD_dimer; TA0261_Pou2f2_DBD_dimer; TA0262_Pou2f2_DBD_monomer; TA0275_RUNX2_DBD_dimer; TA0276_RUNX2_DBD_dimer; TA0277_RUNX2_DBD_monomer; TA0278_RUNX3_DBD_dimer; TA0279_RUNX3_DBD_monomer; TA0280_RUNX3_DBD_dimer; TA0281_RUNX3_full_monomer; TA0294_TBX19_DBD_dimer; TA0318_TEAD1_full_dimer; TA0319_TEAD3_DBD_dimer; TA0379_TFE3_DBD_dimer; TA0387_CEBPB_DBD_dimer; TA0388_CEBPB_full_dimer; TA0389_CEBPD_DBD_dimer; TA0398_Cebpb_DBD_dimer; TA0399_Creb3l2_DBD_dimer; TA0400_Creb3l2_DBD_dimer; TA0407_JDP2_DBD_dimer; TA0408_JDP2_DBD_dimer; TA0409_JDP2_full_dimer; TA0410_JDP2_full_dimer; TA0411_Jdp2_DBD_dimer; TA0412_Jdp2_DBD_dimer; TA0413_MAFF_DBD_dimer; TA0419_Mafb_DBD_monomer; TA0420_Mafb_DBD_dimer; TA0421_Mafb_DBD_dimer; TA0659_PRRX1_DBD_monomer; TA0660_PRRX1_full_monomer; TA0661_PRRX1_full_dimer; TA0680_VENTX_DBD_monomer; TA0681_VENTX_DBD_dimer; TA0689_ESR1_DBD_dimer; TA0762_VDR_full_dimer; TA0763_Vdr_DBD_dimer; TA0847_TFEC_DBD_dimer
TF genes	AHR; ARID3A; ARID5A; ARNT; BACH1; BATF; CEBPB; CEBPD; CREB3L2; ELF1; ELF4; ELK3; ESR1; ETV7; FLI1; FOSL1; FOSL2; GATA6; HLX; HOXB6; IKZF1; IRF1; IRF5; IRF8; JDP2; JUNB; KLF5; KLF6; MAF; MAFB; MAFF; MBD2; MEF2A; NFKB1; NFKB2; PLAG1; POU2AF1; POU2F2; PRDM1; PRRX1; REL; RELB; RREB1; RUNX1; RUNX2; RUNX3; SP1; SP100; SPI1; SRF; STAT2; STAT5A; STAT6; TBX19; TEAD3; TFE3; TFEC; VDR; VENTX; ZBTB7A; ZBTB7B; ZFHX3
Signature genes	KYNU; BNC2; DOCK2; F13A1; NRP1; LHFPL2; PTPRC; AIM1; CCL2; ALOX5; PRDM1; IL7R; AMPD3; MAPK13; TES; CYBB; RAB11FIP1; IL15RA; LCP1; ST14; NCF2; PLAU; SERPINE1; TLR1; C3AR1; IRAK1; PTPN7

Table 3 – Signature genes, TF genes and selected TFBS motifs of MGT#1.

	MGT#2
TFBS motifs	MA0002.1_RUNX1; MA0004.1_Arnt; MA0050.1_IRF1; MA0052.1_MEF2A; MA0061.1_NF-kappaB; MA0073.1_RREB1; MA0079.1_SP1; MA0080.1_SPI1; MA0083.1_SRF; MA0101.1_REL; MA0105.1_NFKB1; MA0112.1_ESR1; MA0117.1_Mafb; MA0002.2_RUNX1; MA0112.2_ESR1; MA0151.1_ARID3A; MA0163.1_PLAG1; MA0080.2_SPI1; MA0079.2_SP1; BU0001_Arid3a_primary; BU0002_Arid5a_primary; BU0023_Gata6_primary; BU0035_Irf5_primary; BU0038_Jundm2_primary; BU0041_Mafb_primary; BU0058_Sfpi1_primary; BU0075_Sp100_primary; BU0078_Srf_primary; ETS0002_h-ELF1; ETS0005_h-ELF4; ETS0008_h-Elk3; ETS0020_h-ETV7; ETS0022_h-Fli1; ETS0027_h-SPI1; HOMEO0040_Hlx1_2350.1; HOMEO0063_Hoxb6_3428.2; HOMEO0146_Pou2f2_3748.1; HOMEO0156_Prrx1_3442.1; TA0011_RUNX3_monomer; TA0013_POU2F2_monomer; TA0026_PRDM1_dimer; HC_AHR_si; HC_ARI3A_do; HC_ARI3A_f1; HC_ARNT_f1; HC_BACH1_si; HC_BATF_si; HC_CEBPB_f1; HC_CEBPD_f1; HC_ELF1_f1; HC_ELK3_f1; HC_ESR1_do; HC_ETV7_si; HC_FLI1_f1; HC_FOSL1_f2; HC_FOSL2_f1; HC_GATA6_f2; HC_HXB6_f1; HC_IKZF1_f1; HC_IRF1_si; HC_IRF5_f1; HC_IRF8_si; HC_JUNB_f1; HC_KLF6_si; HC_MAFB_f1; HC_MAF_f1; HC_MBD2_si; HC_MEF2A_f1; HC_NFKB1_f1; HC_NFKB2_f1; HC_PLAG1_f1; HC_PLAG1_si; HC_PO2F2_si; HC_PRDM1_f1; HC_PRRX1_f1; HC_RELB_si; HC_REL_do; HC_RREB1_si; HC_RUNX1_f1; HC_RUNX2_f1; HC_RUNX3_si; HC_SP1_f1; HC_SP1_f2; HC_SPI1_si; HC_SRF_do; HC_STA5A_do; HC_STAT2_f1; HC_STAT6_do; HC_TEAD3_si; HC_TFE3_f1; HC_VDR_f1; HC_VDR_f2; HC_ZBT7A_f1; HC_ZBT7B_si; HC_ZFHX3_f1; HC_low_KLF5_si; HC_low_OBF1_f1; TA0055_PRDM1_full_monomer_dimer; TA0059_SP1_DBD_monomer; TA0067_ZBTB7A_DBD_monomer; TA0068_ZBTB7B_full_monomer; TA0111_ELF1_DBD_monomer; TA0112_ELF1_full_monomer; TA0115_ELF4_full_monomer; TA0121_ELK3_DBD_monomer; TA0140_Elk3_DBD_monomer; TA0142_FLI1_DBD_monomer; TA0143_FLI1_DBD_monomer; TA0144_FLI1_full_monomer; TA0145_FLI1_full_monomer; TA0153_SPI1_full_monomer; TA0179_IRF5_full_dimer; TA0180_IRF5_full_monomer; TA0183_IRF8_DBD_dimer; TA0184_IRF8_full_dimer; TA0186_MEF2A_DBD_dimer; TA0189_SRF_DBD_dimer; TA0190_SRF_full_dimer; TA0219_NFKB1_DBD_dimer; TA0220_NFKB2_DBD_dimer; TA0239_POU2F2_DBD_monomer; TA0240_POU2F2_DBD_dimer; TA0261_Pou2f2_DBD_dimer; TA0262_Pou2f2_DBD_monomer; TA0275_RUNX2_DBD_dimer; TA0276_RUNX2_DBD_dimer; TA0277_RUNX2_DBD_monomer; TA0278_RUNX3_DBD_dimer; TA0279_RUNX3_DBD_monomer; TA0280_RUNX3_DBD_dimer; TA0281_RUNX3_full_monomer; TA0294_TBX19_DBD_dimer; TA0318_TEAD1_full_dimer; TA0319_TEAD3_DBD_dimer; TA0379_TFE3_DBD_dimer; TA0387_CEBPB_DBD_dimer; TA0388_CEBPB_full_dimer; TA0389_CEBPD_DBD_dimer; TA0398_Cebpb_DBD_dimer; TA0399_Creb3l2_DBD_dimer; TA0400_Creb3l2_DBD_dimer; TA0407_JDP2_DBD_dimer; TA0408_JDP2_DBD_dimer; TA0409_JDP2_full_dimer; TA0410_JDP2_full_dimer; TA0411_Jdp2_DBD_dimer; TA0412_Jdp2_DBD_dimer; TA0413_MAFF_DBD_dimer; TA0419_Mafb_DBD_monomer; TA0420_Mafb_DBD_dimer; TA0421_Mafb_DBD_dimer; TA0659_PRRX1_DBD_monomer; TA0660_PRRX1_full_monomer; TA0661_PRRX1_full_dimer; TA0680_VENTX_DBD_monomer; TA0681_VENTX_DBD_dimer; TA0689_ESR1_DBD_dimer; TA0762_VDR_full_dimer; TA0763_Vdr_DBD_dimer; TA0847_TFEC_DBD_dimer
TF genes	AHR; ARID3A; ARID5A; ARNT; BACH1; BATF; CEBPB; CEBPD; CREB3L2; ELF1; ELF4; ELK3; ESR1; ETV7; FLI1; FOSL1; FOSL2; GATA6; HLX; HOXB6; IKZF1; IRF1; IRF5; IRF8; JDP2; JUNB; KLF5; KLF6; MAF; MAFB; MAFF; MBD2; MEF2A; NFKB1; NFKB2; PLAG1; POU2AF1; POU2F2; PRDM1; PRRX1; REL; RELB; RREB1; RUNX1; RUNX2; RUNX3; SP1; SP100; SPI1; SRF; STAT2; STAT5A; STAT6; TBX19; TEAD3; TFE3; TFEC; VDR; VENTX; ZBTB7A; ZBTB7B; ZFHX3
Signature genes	KYNU; BNC2; DOCK2; F13A1; NRP1; LHFPL2; PTPRC; AIM1; CCL2; ALOX5; PRDM1; IL7R; AMPD3; MAPK13; TES; CYBB; RAB11FIP1; IL15RA; LCP1; ST14; NCF2; PLAU; SERPINE1; TLR1; C3AR1; IRAK1; PTPN7

Table 4 – Signature genes, TF genes and selected TFBS motifs of MGT#2.

	PNGT#1
TFBS motifs	MA0032.1.FOXC1; MA0069.1.Pax6; MA0073.1.RREB1; MA0077.1.SOX9; MA0090.1.TEAD1; MA0143.1.Sox2; MA0161.1.NFIC; BU0033.Irf3_primary; BU0069.Sox21_primary; BU0082.Tcf3_primary; BU0089.Tcfe2a_primary; ETS0010.h-ERF; ETS0017.h-ETV4; HOMEO0104_Meis1_2335.1; HOMEO0134_Pax6_3838.3; HOMEO0149_Pou3f2_2824.1; HOMEO0172_Tgif2_3451.1; TA0007_GLI2_monomer; TA0028_GLI2; HC_ARNT2_si; HC_BMAL1_f1; HC_ERR3_f1; HC_ETV4_f1; HC_FOXC1_f1; HC_FOXO1_si; HC_GLI2_f1; HC_HES1_f1; HC_HNF4G_f1; HC_IRF3_f1; HC_JUND_f1; HC_MEIS1_f2; HC_NR2F6_f1; HC_PAX6_f1; HC_PO3F2_si; HC_RFX2_f1; HC_RREB1_si; HC_SMAD1_si; HC_SOX2_f1; HC_SOX9_f1; HC_TBX2_f1; HC_TEAD1_f1; HC_TEAD3_si; HC_TFE2_f2; HC_low_NOTC1_si; HC_low_PRGC1_si; TA0040_GLI2_DBD_monomer; TA0041_GLI2_DBD_monomer; TA0123_ERF_DBD_monomer; TA0135_ETV4_DBD_monomer; TA0166_SOX9_DBD_monomer; TA0177_IRF3_full_trimer; TA0192_MEIS1_DBD_monomer; TA0206_TGIF2_DBD_dimer; TA0231_PAX6_DBD_monomer; TA0245_POU3F2_DBD_monomer; TA0246_POU3F2_DBD_monomer; TA0264_RFX2_DBD_dimer; TA0265_RFX2_DBD_dimer; TA0272_Rfx2_DBD_dimer; TA0273_Rfx2_DBD_dimer; TA0282_GMEB2_DBD_dimer; TA0283_GMEB2_DBD_dimer; TA0284_GMEB2_DBD_dimer; TA0310_TBX2_full_dimer; TA0311_TBX2_full_monomer; TA0317_TEAD1_full_monomer; TA0318_TEAD1_full_dimer; TA0319_TEAD3_DBD_dimer; TA0320_TEAD3_DBD_monomer; TA0337_ARNTL_DBD_dimer; TA0374_TCF3_DBD_dimer; TA0401_Creb5_DBD_dimer; TA0433_FOXC1_DBD_dimer; TA0434_FOXC1_DBD_dimer; TA0435_FOXC1_DBD_monomer; TA0443_FOXG1_DBD_dimer; TA0444_FOXG1_DBD_multimer; TA0456_FOXO1_DBD_monomer; TA0457_FOXO1_DBD_dimer; TA0458_FOXO1_DBD_multimer; TA0469_Foxc1_DBD_dimer; TA0470_Foxc1_DBD_monomer; TA0471_Foxg1_DBD_dimer; TA0472_Foxg1_DBD_multimer; TA0473_Foxg1_DBD_monomer; TA0694_ESRRG_full_dimer; TA0695_ESRRG_full_dimer; TA0696_ESRRG_full_monomer; TA0707_NR2E1_full_monomer; TA0708_NR2E1_full_dimer; TA0712_NR2F6_DBD_dimer; TA0713_NR2F6_DBD_dimer; TA0714_NR2F6_full_dimer; TA0720_Nr2e1_DBD_monomer; TA0721_Nr2e1_DBD_dimer; TA0722_Nr2f6_DBD_dimer; TA0723_Nr2f6_DBD_dimer; TA0769_GMEB2_DBD_dimer; TA0800_SOX21_DBD_dimer; TA0801_SOX21_DBD_dimer; TA0802_SOX21_DBD_dimer; TA0803_SOX21_DBD_dimer; TA0804_SOX2_DBD_dimer; TA0805_SOX2_DBD_dimer; TA0806_SOX2_DBD_dimer; TA0807_SOX2_full_dimer; TA0808_SOX2_full_dimer; TA0809_SOX2_full_dimer; TA0822_SOX9_full_dimer; TA0823_SOX9_full_dimer; TA0824_SOX9_full_dimer; TA0825_SOX9_full_dimer; TA0826_SOX9_full_dimer; TA0827_SOX9_full_dimer
TF genes	BPTF; CDC5L; CTCF; E2F2; E2F3; ETV1; FOXM1; FUBP1; GBX2; GTF2I; HIC2; HLTF; HMBOX1; HOXD1; HOXD3; HSF2; INSM1; KLF12; MAZ; MECP2; MNX1; MYB; MYBL1; MYBL2; NFIB; NOTCH1; NR0B1; NR2C2; OLIG2; ONECUT2; PARP1; PBX1; POU2F1; PSIP1; RBPJ; SATB1; SMARCC1; SOX10; SOX11; SOX12; SOX2; SOX4; SOX5; SP4; TBX5; TCF12; TCF4; TCF7L1; TFAP2A; TFAP4; THRA; VAX2; YBX1; ZBTB33; ZEB1; ZNF238; ZNF281; ZNF589; ZNF652; ZSCAN16
Signature genes	AGPAT4; ALCAM; AMOTL2; BEX4; BRD3; C1QL1; CAMTA1; CHD7; CLASP2; DNAJB5; DPF1; EYA1; GSTA4; HDAC2; HN1; ICK; KCND2; KIF1A; KLF12; MAGEH1; MLLT11; MNX1; NCAM1; NFIB; NOL4; NR0B1; OLIG2; PATZ1; PELI1; PSIP1; RBPJ; RNF144A; SATB1; SNAP91; SOX11; SOX4; STMN1; TCEAL2; TCF4; TSPYL4; VAX2; WASF1; ZBTB5; ZNF184; ZNF643; ZNF711

Table 5 – Signature genes, TF genes and selected TFBS motifs of PNGT#1.

	PNGT#2
TFBS motifs	MA0032.1.FOXC1; MA0069.1.Pax6; MA0073.1.RREB1; MA0077.1.SOX9; MA0090.1.TEAD1; MA0143.1.Sox2; MA0161.1.NFIC; BU0033.Irf3_primary; BU0069.Sox21_primary; BU0082.Tcf3_primary; BU0089.Tcfe2a_primary; ETS0010.h-ERF; ETS0017.h-ETV4; HOMEO0104_Meis1_2335.1; HOMEO0134_Pax6_3838.3; HOMEO0149_Pou3f2_2824.1; HOMEO0172_Tgif2_3451.1; TA0007_GLI2_monomer; TA0028_GLI2; HC_ARNT2_si; HC_BMAL1_f1; HC_ERR3_f1; HC_ETV4_f1; HC_FOXC1_f1; HC_FOXO1_si; HC_GLI2_f1; HC_HES1_f1; HC_HNF4G_f1; HC_IRF3_f1; HC_JUND_f1; HC_MEIS1_f2; HC_NR2F6_f1; HC_PAX6_f1; HC_PO3F2_si; HC_RFX2_f1; HC_RREB1_si; HC_SMAD1_si; HC_SOX2_f1; HC_SOX9_f1; HC_TBX2_f1; HC_TEAD1_f1; HC_TEAD3_si; HC_TFE2_f2; HC_low_NOTC1_si; HC_low_PRGC1_si; TA0040_GLI2_DBD_monomer; TA0041_GLI2_DBD_monomer; TA0123_ERF_DBD_monomer; TA0135_ETV4_DBD_monomer; TA0166_SOX9_DBD_monomer; TA0177_IRF3_full_trimer; TA0192_MEIS1_DBD_monomer; TA0206_TGIF2_DBD_dimer; TA0231_PAX6_DBD_monomer; TA0245_POU3F2_DBD_monomer; TA0246_POU3F2_DBD_monomer; TA0264_RFX2_DBD_dimer; TA0265_RFX2_DBD_dimer; TA0272_Rfx2_DBD_dimer; TA0273_Rfx2_DBD_dimer; TA0282_GMEB2_DBD_dimer; TA0283_GMEB2_DBD_dimer; TA0284_GMEB2_DBD_dimer; TA0310_TBX2_full_dimer; TA0311_TBX2_full_monomer; TA0317_TEAD1_full_monomer; TA0318_TEAD1_full_dimer; TA0319_TEAD3_DBD_dimer; TA0320_TEAD3_DBD_monomer; TA0337_ARNTL_DBD_dimer; TA0374_TCF3_DBD_dimer; TA0401_Creb5_DBD_dimer; TA0433_FOXC1_DBD_dimer; TA0434_FOXC1_DBD_dimer; TA0435_FOXC1_DBD_monomer; TA0443_FOXG1_DBD_dimer; TA0444_FOXG1_DBD_multimer; TA0456_FOXO1_DBD_monomer; TA0457_FOXO1_DBD_dimer; TA0458_FOXO1_DBD_multimer; TA0469_Foxc1_DBD_dimer; TA0470_Foxc1_DBD_monomer; TA0471_Foxg1_DBD_dimer; TA0472_Foxg1_DBD_multimer; TA0473_Foxg1_DBD_monomer; TA0694_ESRRG_full_dimer; TA0695_ESRRG_full_dimer; TA0696_ESRRG_full_monomer; TA0707_NR2E1_full_monomer; TA0708_NR2E1_full_dimer; TA0712_NR2F6_DBD_dimer; TA0713_NR2F6_DBD_dimer; TA0714_NR2F6_full_dimer; TA0720_Nr2e1_DBD_monomer; TA0721_Nr2e1_DBD_dimer; TA0722_Nr2f6_DBD_dimer; TA0723_Nr2f6_DBD_dimer; TA0769_GMEB2_DBD_dimer; TA0800_SOX21_DBD_dimer; TA0801_SOX21_DBD_dimer; TA0802_SOX21_DBD_dimer; TA0803_SOX21_DBD_dimer; TA0804_SOX2_DBD_dimer; TA0805_SOX2_DBD_dimer; TA0806_SOX2_DBD_dimer; TA0807_SOX2_full_dimer; TA0808_SOX2_full_dimer; TA0809_SOX2_full_dimer; TA0822_SOX9_full_dimer; TA0823_SOX9_full_dimer; TA0824_SOX9_full_dimer; TA0825_SOX9_full_dimer; TA0826_SOX9_full_dimer; TA0827_SOX9_full_dimer
TF genes	BPTF; CDC5L; CTCF; E2F2; E2F3; ETV1; FOXM1; FUBP1; GBX2; GTF2I; HIC2; HLTF; HMBOX1; HOXD1; HOXD3; HSF2; INSM1; KLF12; MAZ; MECP2; MNX1; MYB; MYBL1; MYBL2; NFIB; NOTCH1; NR0B1; NR2C2; OLIG2; ONECUT2; PARP1; PBX1; POU2F1; PSIP1; RBPJ; SATB1; SMARCC1; SOX10; SOX11; SOX12; SOX2; SOX4; SOX5; SP4; TBX5; TCF12; TCF4; TCF7L1; TFAP2A; TFAP4; THRA; VAX2; YBX1; ZBTB33; ZEB1; ZNF238; ZNF281; ZNF589; ZNF652; ZSCAN16
Signature genes	AGPAT4; ALCAM; AMOTL2; BEX4; BRD3; C1QL1; CAMTA1; CHD7; CLASP2; DNAJB5; DPF1; EYA1; GSTA4; HDAC2; HN1; ICK; KCND2; KIF1A; KLF12; MAGEH1; MLLT11; MNX1; NCAM1; NFIB; NOL4; NR0B1; OLIG2; PATZ1; PELI1; PSIP1; RBPJ; RNF144A; SATB1; SNAP91; SOX11; SOX4; STMN1; TCEAL2; TCF4; TSPYL4; VAX2; WASF1; ZBTB5; ZNF184; ZNF643; ZNF711

Table 6 – Signature genes, TF genes and selected TFBS motifs of PNGT#1.

	CLGT#1
TFBS motifs	MA0003.1.TFAP2A; MA0070.1.PBX1; MA0087.1.Sox5; MA0100.1.Myb; MA0103.1.ZEB1; MA0109.1.Hltf; MA0139.1.CTCF; MA0143.1.Sox2; MA0155.1.INSM1; MA0442.1.SOX10; BU0008.E2F2_primary; BU0009.E2F3_primary; BU0045.Myb_primary; BU0046.Mybl1_primary; BU0061.Sox11_primary; BU0062.Sox12_primary; BU0071.Sox4_primary; BU0072.Sox5_primary; BU0076.Sp4_primary; BU0085.Tcfap2a_primary; BU0097.Zfp281_primary; ETS0014.h-ETV1; HOMEO0036_Gbx2_3110.1; HOMEO0041_Hlxb9_3422.1; HOMEO0042_Hmbox1_2674.1; HOMEO0076_Hoxd1_3448.1; HOMEO0081_Hoxd3_1742.2; HOMEO0136_Pbx1_3203.1; HOMEO0145_Pou2f1_3081.2; HOMEO0177_Vax2_3500.1; TA0019_HSF2_dimer-trimer; TA0021_TCF4_dimer; TA0030_TCF4; HC_AP2A_f2; HC_BPTF_si; HC_CDC5L_si; HC_CTCF_f2; HC_E2F2_f1; HC_E2F3_si; HC_FOXM1_f1; HC_FUBP1_f1; HC_HLTF_f1; HC_HSF2_si; HC_HTF4_f1; HC_INSM1_f1; HC_ITF2_f1; HC_KAISO_f1; HC_MAZ_f1; HC_MECP2_f1; HC_MYBB_f1; HC_MYB_f1; HC_NR0B1_si; HC_NR2C2_f1; HC_ONEC2_si; HC_PBX1_do; HC_PO2F1_f1; HC_SMRC1_f1; HC_SOX10_si; HC_SOX2_f1; HC_SOX4_f1; HC_SOX5_f1; HC_SP4_f1; HC_SUH_f1; HC_TBX5_si; HC_TFAP4_si; HC_THA_f1; HC_THA_f2; HC_YBOX1_f2; HC_ZEB1_do; HC_ZN238_f1; HC_ZN589_f1; HC_low_ETV1_si; HC_low_GTF2I_f1; HC_low_NOTC1_si; HC_low_PARP1_si; HC_low_PSIP1_f1; HC_low_SATB1_f1; TA0032_CTCF_full_monomer; TA0045_HIC2_DBD_monomer; TA0054_Klf12_DBD_monomer; TA0061_SP4_full_monomer; TA0074_ZNF238_DBD_monomer; TA0075_ZNF238_full_monomer; TA0085_Zfp652_DBD_monomer; TA0101_ONECUT2_DBD_monomer; TA0103_E2F2_DBD_dimer; TA0104_E2F2_DBD_dimer; TA0105_E2F3_DBD_dimer; TA0106_E2F3_DBD_dimer; TA0107_E2F3_DBD_dimer; TA0132_ETV1_DBD_monomer; TA0167_TCF7L1_full_monomer; TA0172_HSF2_DBD_trimer; TA0207_MYBL1_DBD_dimer; TA0208_MYBL1_DBD_dimer; TA0209_MYBL1_DBD_monomer; TA0210_MYBL1_DBD_dimer; TA0211_MYBL2_DBD_dimer; TA0212_MYBL2_DBD_dimer; TA0213_MYBL2_DBD_monomer; TA0214_MYBL2_DBD_dimer; TA0222_NFIB_full_dimer; TA0237_POU2F1_DBD_monomer; TA0238_POU2F1_DBD_dimer; TA0314_TBX5_DBD_monomer; TA0315_TBX5_DBD_dimer; TA0322_TFAP2A_DBD_dimer; TA0323_TFAP2A_DBD_dimer; TA0324_TFAP2A_DBD_dimer; TA0334_Tcfap2a_DBD_dimer; TA0335_Tcfap2a_DBD_dimer; TA0336_Tcfap2a_DBD_dimer; TA0369_OLIG2_DBD_dimer; TA0370_OLIG2_full_dimer; TA0375_TCF4_DBD_dimer; TA0376_TCF4_full_dimer; TA0377_TFAP4_DBD_dimer; TA0378_TFAP4_full_dimer; TA0533_GBX2_DBD_monomer; TA0534_GBX2_DBD_dimer; TA0535_GBX2_full_monomer; TA0541_Gbx2_DBD_monomer; TA0544_HMBOX1_DBD_monomer; TA0589_Hoxd3_DBD_monomer; TA0621_MNX1_DBD_monomer; TA0679_VAX2_DBD_monomer; TA0706_NR2C2_DBD_dimer; TA0758_THRA_FL_dimer; TA0777_ZNF435_full_dimer; TA0782_E2F2_DBD_dimer; TA0786_SOX10_full_dimer; TA0787_SOX10_full_dimer; TA0788_SOX10_full_dimer; TA0789_SOX10_full_dimer; TA0790_SOX10_full_dimer; TA0804_SOX2_DBD_dimer; TA0805_SOX2_DBD_dimer; TA0806_SOX2_DBD_dimer; TA0807_SOX2_full_dimer; TA0808_SOX2_full_dimer; TA0809_SOX2_full_dimer; TA0810_SOX4_DBD_dimer; TA0832_Sox10_DBD_dimer; TA0833_Sox10_DBD_dimer; TA0834_Sox10_DBD_dimer; TA0835_Sox11_DBD_dimer; TA0862_TFAP2A_DBD_dimer; TA0863_TFAP2A_DBD_dimer; TA0864_TFAP2A_DBD_dimer
TF genes	ARNT2; ARNTL; CREB5; ERF; ESRRG; ETV4; FOXC1; FOXG1; FOXO1; GLI2; GMEB2; HES1; HNF4G; IRF3; JUND; MEIS1; NFIC; NOTCH1; NR2E1; NR2F6; PAX6; POU3F2; PPARGC1A; RFX2; RREB1; SMAD1; SOX2; SOX21; SOX9; TBX2; TCF3; TEAD1; TEAD3; TGIF2
Signature genes	ABI2; ACSL3; B3GALT1; CC2D1A,CDC130; DAG1; DENND2A; FZR1; GNAS; HMG20B; IRS2; JUND; KIAA0355; MLC1; NES; PLCG1; POLRMT; POMT2; POU3F2; PTN; PTPRA; QTRT1; SALL1; SAMD4A; SCRN1; SEMA6D; SIN3B; SLC20A2; TLE1; TMEM161A; UPF1; VPS16; ZNF264

Table 7 – Signature genes, TF genes and selected TFBS motifs of CLGT#1.

	CLGT#2
TFBS motifs	MA0003.1.TFAP2A; MA0070.1.PBX1; MA0087.1.Sox5; MA0100.1.Myb; MA0103.1.ZEB1; MA0109.1.Hltf; MA0139.1.CTCF; MA0143.1.Sox2; MA0155.1.INSM1; MA0442.1.SOX10; BU0008.E2F2_primary; BU0009.E2F3_primary; BU0045.Myb_primary; BU0046.Mybl1_primary; BU0061.Sox11_primary; BU0062.Sox12_primary; BU0071.Sox4_primary; BU0072.Sox5_primary; BU0076.Sp4_primary; BU0085.Tcfap2a_primary; BU0097.Zfp281_primary; ETS0014.h-ETV1; HOMEO0036_Gbx2_3110.1; HOMEO0041_Hlxb9_3422.1; HOMEO0042_Hmbox1_2674.1; HOMEO0076_Hoxd1_3448.1; HOMEO0081_Hoxd3_1742.2; HOMEO0136_Pbx1_3203.1; HOMEO0145_Pou2f1_3081.2; HOMEO0177_Vax2_3500.1; TA0019_HSF2_dimer-trimer; TA0021_TCF4_dimer; TA0030_TCF4; HC_AP2A_f2; HC_BPTF_si; HC_CDC5L_si; HC_CTCF_f2; HC_E2F2_f1; HC_E2F3_si; HC_FOXM1_f1; HC_FUBP1_f1; HC_HLTF_f1; HC_HSF2_si; HC_HTF4_f1; HC_INSM1_f1; HC_ITF2_f1; HC_KAISO_f1; HC_MAZ_f1; HC_MECP2_f1; HC_MYBB_f1; HC_MYB_f1; HC_NR0B1_si; HC_NR2C2_f1; HC_ONEC2_si; HC_PBX1_do; HC_PO2F1_f1; HC_SMRC1_f1; HC_SOX10_si; HC_SOX2_f1; HC_SOX4_f1; HC_SOX5_f1; HC_SP4_f1; HC_SUH_f1; HC_TBX5_si; HC_TFAP4_si; HC_THA_f1; HC_THA_f2; HC_YBOX1_f2; HC_ZEB1_do; HC_ZN238_f1; HC_ZN589_f1; HC_low_ETV1_si; HC_low_GTF2I_f1; HC_low_NOTC1_si; HC_low_PARP1_si; HC_low_PSIP1_f1; HC_low_SATB1_f1; TA0032_CTCF_full_monomer; TA0045_HIC2_DBD_monomer; TA0054_Klf12_DBD_monomer; TA0061_SP4_full_monomer; TA0074_ZNF238_DBD_monomer; TA0075_ZNF238_full_monomer; TA0085_Zfp652_DBD_monomer; TA0101_ONECUT2_DBD_monomer; TA0103_E2F2_DBD_dimer; TA0104_E2F2_DBD_dimer; TA0105_E2F3_DBD_dimer; TA0106_E2F3_DBD_dimer; TA0107_E2F3_DBD_dimer; TA0132_ETV1_DBD_monomer; TA0167_TCF7L1_full_monomer; TA0172_HSF2_DBD_trimer; TA0207_MYBL1_DBD_dimer; TA0208_MYBL1_DBD_dimer; TA0209_MYBL1_DBD_monomer; TA0210_MYBL1_DBD_dimer; TA0211_MYBL2_DBD_dimer; TA0212_MYBL2_DBD_dimer; TA0213_MYBL2_DBD_monomer; TA0214_MYBL2_DBD_dimer; TA0222_NFIB_full_dimer; TA0237_POU2F1_DBD_monomer; TA0238_POU2F1_DBD_dimer; TA0314_TBX5_DBD_monomer; TA0315_TBX5_DBD_dimer; TA0322_TFAP2A_DBD_dimer; TA0323_TFAP2A_DBD_dimer; TA0324_TFAP2A_DBD_dimer; TA0334_Tcfap2a_DBD_dimer; TA0335_Tcfap2a_DBD_dimer; TA0336_Tcfap2a_DBD_dimer; TA0369_OLIG2_DBD_dimer; TA0370_OLIG2_full_dimer; TA0375_TCF4_DBD_dimer; TA0376_TCF4_full_dimer; TA0377_TFAP4_DBD_dimer; TA0378_TFAP4_full_dimer; TA0533_GBX2_DBD_monomer; TA0534_GBX2_DBD_dimer; TA0535_GBX2_full_monomer; TA0541_Gbx2_DBD_monomer; TA0544_HMBOX1_DBD_monomer; TA0589_Hoxd3_DBD_monomer; TA0621_MNX1_DBD_monomer; TA0679_VAX2_DBD_monomer; TA0706_NR2C2_DBD_dimer; TA0758_THRA_FL_dimer; TA0777_ZNF435_full_dimer; TA0782_E2F2_DBD_dimer; TA0786_SOX10_full_dimer; TA0787_SOX10_full_dimer; TA0788_SOX10_full_dimer; TA0789_SOX10_full_dimer; TA0790_SOX10_full_dimer; TA0804_SOX2_DBD_dimer; TA0805_SOX2_DBD_dimer; TA0806_SOX2_DBD_dimer; TA0807_SOX2_full_dimer; TA0808_SOX2_full_dimer; TA0809_SOX2_full_dimer; TA0810_SOX4_DBD_dimer; TA0832_Sox10_DBD_dimer; TA0833_Sox10_DBD_dimer; TA0834_Sox10_DBD_dimer; TA0835_Sox11_DBD_dimer; TA0862_TFAP2A_DBD_dimer; TA0863_TFAP2A_DBD_dimer; TA0864_TFAP2A_DBD_dimer
TF genes	ARNT2; ARNTL; CREB5; ERF; ESRRG; ETV4; FOXC1; FOXG1; FOXO1; GLI2; GMEB2; HES1; HNF4G; IRF3; JUND; MEIS1; NFIC; NOTCH1; NR2E1; NR2F6; PAX6; POU3F2; PPARGC1A; RFX2; RREB1; SMAD1; SOX2; SOX21; SOX9; TBX2; TCF3; TEAD1; TEAD3; TGIF2
Signature genes	ABI2; ACSL3; B3GALT1; CC2D1A,CDC130; DAG1; DENND2A; FZR1; GNAS; HMG20B; IRS2; JUND; KIAA0355; MLC1; NES; PLCG1; POLRMT; POMT2; POU3F2; PTN; PTPRA; QTRT1; SALL1; SAMD4A; SCRN1; SEMA6D; SIN3B; SLC20A2; TLE1; TMEM161A; UPF1; VPS16; ZNF264

Table 8 – Signature genes, TF genes and selected TFBS motifs of CLGT#2.

9 Acknowledgments

First of all, I would like to express my deepest gratitude for my supervisor, Dr. Gaetano Gargiulo, for not only giving me the opportunity to be a part of his lab and research, but also for the constant guidance and support in developing my skills as a scientist. His close, yet independence-focused supervision, constructive feedback and ideas have always been a steady source of help to advance my projects and succeed.

Not a less amount of gratitude is owed to Dr. Michela Serresi, who built and introduced me to the lab when it was just starting and was always a secure anchor in the lab for help and advice. Besides scientific discussions, also our conversations over everything and nothing did a great part for a nice atmosphere and great time that I have spent in the lab.

I would also like to especially thank my colleagues Carlos Company and Yuliia Dramaretska, without whom almost none of the presented work could have been realized. I enjoyed very much the large degree of collaboration in which our lab was operating and due to the constant and also interdisciplinary exchang, I learned a lot during the past 6 years.
I am very glad to not only had the honor of considering you my coworkers, but also my friends together with great memories of coffee breaks, late hours in the lab, Friday night beer hours or weekend activities.

My gratitude of course also goes out to all the other past and present members of the GG lab and collaborators, who also greatly contributed to our joint projects and generally a productive and friendly atmosphere in the lab.

I thank my PhD committee members, Prof. Dr. Markus Landthaler and Massimo Squatrito for their willingness to serve in my committee and the provided support and guidance during my thesis.

Last but not least, I would like to thank the most important people in my life – my partner Olya and my family and friends – for all their support and understanding during my PhD. It has been an intensive time for me with lots of ups and downs and I have not always had the time and energy to reach out as often as I could, but knowing that I always could have counted on your support and that you were covering my back was a priceless gift to me!

10 Selbstständigkeitserklärung

Hiermit erkläre ich, die Dissertation selbstständig und nur unter Verwendung der angegebenen Hilfen und Hilfsmittel angefertigt zu haben. Der Eigenanteil, sowie kollaborative Daten anderer an der vorgelegten Dissertation ist im Detail in Abschnitt (6) dargelegt.

Ich habe mich anderwärts nicht um einen Doktorgrad beworben und besitze keinen entsprechenden Doktorgrad. Ich erkläre, dass ich die Dissertation oder Teile davon nicht bereits bei einer anderen wissenschaftlichen Einrichtung eingereicht habe und dass sie dort weder angenommen noch abgelehnt wurde. Ich erkläre die Kenntnisnahme der dem Verfahren zugrunde liegenden Promotionsordnung der Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin vom 05. März 2015, veröffentlicht im Amtlichen Mitteilungsblatt der Humboldt-Universität zu Berlin Nr. 12/2015.

Weiterhin erkläre ich, dass keine Zusammenarbeit mit gewerblichen Promotionsberaterinnen/Promotionsberatern stattgefunden hat und dass die Grundsätze der Humboldt-Universität zu Berlin zur Sicherung guter wissenschaftlicher Praxis eingehalten wurden.

Datum, Ort	Unterschrift