

Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

Caitlin Sojka

Date

DIVERGENCE FROM THE HUMAN ASTROCYTE DEVELOPMENTAL TRAJECTORY
IN GLIOBLASTOMA

By

Caitlin Sojka
Doctor of Philosophy

Graduate Division of Biological and Biomedical Science
Neuroscience

Steven Sloan, MD PhD
Advisor

Jennifer Spangle, PhD
Committee Member

Thomas Kukar, PhD
Committee Member

David Katz, PhD
Committee Member

Renee Read, PhD
Committee Member

Accepted:

Kimberly Jacob Arriola, Ph.D., MPH
Dean of the James T. Laney School of Graduate Studies

Date

DIVERGENCE FROM THE HUMAN ASTROCYTE DEVELOPMENTAL TRAJECTORY
IN GLIOBLASTOMA

By

Caitlin Sojka
B.S., California State University, Long Beach 2016

Advisor: Steven Sloan, MD PhD

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Neuroscience
2023

ABSTRACT

DIVERGENCE FROM THE HUMAN ASTROCYTE DEVELOPMENTAL TRAJECTORY IN GLIOBLASTOMA

By
Caitlin Sojka

Properties of early embryonic development are frequently recycled in cancer, including the acquisition of a highly plastic stem-like cell state, uncontrolled cell growth and proliferation, and adaptation to a harsh microenvironment. This mirroring of development is evident in glioma, where tumors reflect the differentiation hierarchies that underpin normal glial cell formation during neurodevelopment. Glioblastomas (GBM), the most severe glioma class, harbor cells that resemble immature progenitor populations, including oligodendrocyte, neural, and astrocyte precursor cells, and demonstrate the capacity to transition between these cell states. Given the close parallels between neurodevelopment and GBM cell programs, utilizing maps of normal glial lineages could inform us about how tumor cells hijack and progress along developmental trajectories to promote tumor survival. However, glial differentiation and maturation are challenging to delineate as it peaks between late gestational and early postnatal ages, a window of development that is hard to capture in humans due to limited primary tissue samples and suboptimal two-dimensional in vitro model systems. To overcome this hurdle, we leveraged human cortical organoids (hCOs), a three-dimensional in vitro model of the developing human cortex, to generate a comprehensive molecular timeline of human astrocyte maturation. We then projected this developmental trajectory onto GBM astrocyte-like tumor cells to identify how astrocyte development is recapitulated in GBM. We maintained hCOs in culture for nearly two years, profiling the chromatin- and transcriptome-level changes in hCO astrocytes at 10 discrete time points. In doing so, we found three molecularly distinct stages of maturation, including a novel intermediate stage that may serve as a key lineage determination cell state. This particular intermediate stage of maturation was consistently and highly expressed in astrocyte-like cells from GBM tumors, potentially serving as an “attractor” maturation state, where malignant tumor cells thrive. When looking at maturation signature across the diverse tumor cohort, we discovered that astrocyte-like cells from tumors harboring an IDH1 mutation were substantially more mature compared to IDH1-wildtype tumors, suggesting that the IDH1 mutation may directly or indirectly preserve astrocyte maturation state. We hypothesized that conserved mature molecular programs may be related to IDH1-mutant-associated DNA hydroxymethylation (5hmC) patterns and found an enrichment of 5hmC in differentially expressed maturation gene sets. Together, these experiments describe a novel astrocyte developmental program that is preferentially activated in human GBM revealing new facets of tumor biology and therapeutic targeting to explore.

DIVERGENCE FROM THE HUMAN ASTROCYTE DEVELOPMENTAL TRAJECTORY
IN GLIOBLASTOMA

By

Caitlin Sojka
B.S., California State University, Long Beach

Advisor: Steven Sloan, M.D., Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
Graduate Division of Biological and Biomedical Sciences
Neuroscience
2023

ACKNOWLEDGEMENTS

I would like to first thank my advisor, Steven Sloan, for his invaluable guidance and unwavering support throughout my time in graduate school. In 2018, Steven took a chance on a marine biologist-turned-neuroscientist and gave me the opportunity to pursue a project that spanned many challenging fields—neurodevelopment, glial biology, and cancer—none of which I had a background in, and he encouraged my desire to hone new skills in computational biology, which I did not have any experience with. Steven, your optimism and encouragement pushed me to try new things, think big, and to never be afraid of failure. Perhaps most importantly, you always made me feel heard, respected, and appreciated. I could not have asked for a better mentor to have by my side throughout this experience.

I would also like to thank two of my previous mentors, Bruno Pernet and Niamh Cawley, who helped me discover my love for scientific research and pushed me to pursue my passions, whatever they may be.

The work in this dissertation would not have been possible without our brilliant and fun lab. Every member of the Sloan Lab played a role in supporting my scientific, professional, and personal growth, and I could not have done this without their support. My fellow Sloan Lab grad students, thank you for keeping me sane and for always being in my corner. Anna Voss, I cannot thank you enough for your coding guidance, our daily debriefs on our walks home, and for keeping life loud and lively! Tarun Bhatia, Maureen Sampson, and Anson Sing, I sincerely appreciate your mentorship and guidance throughout this journey. Tarun, your intellectual and technical contributions have been instrumental in pushing this last leg of my project along. Alexia King, I am forever grateful for your mentorship and friendship. You kept me going through thick and thin, always entertained my curiosity, and you are one heck of a lab/hub manager!

I received lots of help from collaborators and members of the Human Genetics Department, including Pankaj Chopra, Bing Yao, Yangping Li, Victor Corces, and Hsiao-Lin Wang. Your hard work and expertise made this work possible!

I owe an enormous debt of gratitude to my family and friends for their unconditional love and support throughout my (very) long educational journey. You kept me grounded by reminding me that there is more to life than work. You lifted me up during challenging moments and were a constant source of motivation. I would not be where I am today without the encouragement and understanding of my parents, Jan Merna and Brian Sojka, and my brother, Nicholas Sojka. You were there every step of the way, even when my education took me to distant places. Finally, to my four-legged furry roommate, Tommi, you will never read this, but I could not have done this without your companionship and making sure I am awake on time for breakfast.

This work would not have been possible without an incredible support system. Your impact on my academic and personal life is immeasurable, and for that, I am deeply thankful.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION	1
1.1 Overview and organization.....	2
1.2 Glial development and maturation.....	3
1.2.1 The rise of glial cells: how, when, and where.....	3
1.2.1.1 Molecular regulators of the gliogenic switch	4
1.2.1.2 Extrinsic regulators of gliogenesis.....	6
1.2.1.3 Populating the CNS: where and when	8
1.2.2. Glial cell maturation	9
1.2.2.1 Astrocyte maturation	9
1.2.2.2 Oligodendrocyte maturation	11
1.2.3 Tipping the scales: making astrocytes vs oligodendrocytes.....	12
1.2.3.1 Evidence of a shared glial precursor.....	12
1.2.3.2 Diverging molecular programs.....	14
1.3 Cancer as an echo of glial development.....	17
1.3.1 Glioblastoma.....	19
1.3.2 Other diffuse and lower-grade gliomas	23
1.3.2.1 Diffuse glioma	24
1.3.2.2 Pilocytic astrocytoma.....	25
1.3.3 Diffuse midline glioma.....	26
1.4 How do glioma cells move across developmental time?	30

1.4.1 Where along the normal developmental trajectory do gliomas start?	30
1.4.2 When in developmental time do glioma cells reside and thrive during tumor progression?	32
1.4.3 When in developmental time are glioma cells capable of moving towards with intervention?.....	34
1.5 Summary and thesis objectives.....	37
CHAPTER 2: DIVERGENCE FROM THE HUMAN ASTROCYTE DEVELOPMENTAL TRAJECTORY IN GLIOBLASTOMA.....	40
2.1 Abstract	41
2.2 Introduction.....	42
2.3 Results.....	45
2.3.1 A molecular map of human astrocyte maturation in cortical organoids	45
2.3.2 Molecular differences between IDHwt tumor and margin astrocyte lineage cells	47
2.3.3 Projecting GBM astrocytes onto a normal human astrocyte maturation trajectory.....	50
2.3.4 Subtype-specific molecular signatures in GBM tumor astrocyte lineage cells	51
2.3.5 The contribution of DNA methylation to IDHwt and IDH1mt maturation differences.....	53
2.4 Discussion.....	86
2.5 Materials and methods	91
CHAPTER 3: DISCUSSION	110

3.1 Summary	111
3.2 Key findings and future directions	112
3.2.1 Human astrocyte maturation: filling in the gaps	112
3.2.2 What can the principles of astrocyte maturation tell us about GBM biology?	113
3.2.3 Can maturation TFs be harnessed for therapeutic intervention?	116
3.2.4 Revisiting the GBM cell of origin	119
3.2.5 Developmental differences across glioma subtypes	122
3.3 Conclusion	126
APPENDIX: ONGOING AND FUTURE FUNCTIONAL EXPERIMENTS	127
Abstract	128
Introduction	129
Results	130
Overexpressing TF candidates in hCSs	130
Mapping molecular maturation programs onto published fetal datasets	131
The effects of D2HG on astrocyte maturation	132
Discussion	141
Methods	144
REFERENCES	150

LIST OF FIGURES

Figure 1: Hypothesized glial differentiation trajectories	16
Figure 2: Updated 2021 WHO diffuse glioma classification	18
Figure 3: Representation of developmental lineages across glioma subtypes	29
Figure 4: How do glioma cells move across developmental time?	38
Figure 5: A molecular trajectory of human astrocyte maturation.....	57
Figure 6: Quality control of hCO astrocyte libraries.....	59
Figure 7: TF motif enrichment data in hCO-derived astrocytes	61
Figure 8: TF motif occurrence in hCO astrocytes.....	62
Figure 9: Diverging molecular profiles of GBM tumor and margin astrocyte lineage cells	64
Figure 10: GBM astrocyte lineage ATAC-seq library quality control	66
Figure 11: Molecular differences between IDHwt tumor and margin astrocytes	68
Figure 12: GBM astrocyte lineage cells deviate from the normal developmental trajectory at middle stages	70
Figure 13: sn-multiome quality control and cell-type identification	72
Figure 14: Measurements of maturation TFs in GBM.....	74
Figure 15: Subtype-specific molecular signatures in tumor astrocyte lineage cells.....	76
Figure 16: Additional genomic diagnosis-specific ATAC- and RNA-seq tumor astrocyte signatures	78
Figure 17: Gene expression comparisons between IDHwt, IDH1mt, and margin astrocyte lineage cells.....	80
Figure 18: DNA methylation and hydroxymethylation of astrocyte lineage cells in IDH1mt and IDHwt tumors	82
Figure 19: 5mC and 5hmC quantification in IDH1mt and IDHwt tumors.....	84
Figure 20: Vector maps for overexpressing maturation TFs	134
Figure 21: Approach for overexpressing maturation TFs.....	135
Figure 22: Test induction of RFX4 in hCSs	136
Figure 23: Mapping maturation modules onto fetal datasets	137
Figure 24: The effects of D2HG on astrocyte maturity	139

ABBREVIATIONS

5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
AS	Astrocyte
ATAC	Assay for Transposase-Accessible Chromatin
BBB	Blood brain barrier
CNS	Central nervous system
CNTF	Ciliary neurotrophic factor
CNV	Copy number variation
CSC	Cancer stem cell
CT-1	Cardiotrophin-1
DhMR	differentially hydroxymethylated regions
DIPG	Diffuse intrinsic pontine glioma
DMG	Diffuse midline glioma
DMR	differentially methylated regions
GBM	Glioblastoma
GEMM	Genetically engineered mouse model
GSC	Glioma stem cell
GW	Gestational week
hMe-Seal	hydroxymethylated-DNA selective chemical labeling
IDH	Isocitrate dehydrogenase
LGG	Low-grade glioma
LIF	Leukemia inhibitory factor
MeDIP	methylated-DNA immunoprecipitation
MES	Mesenchymal
NPC	Neural-progenitor
NSC	Neural stem cell
OL	Oligodendrocyte
OPC	Oligodendrocyte precursor cell
PA	Pilocytic astrocytoma
PECA	Paired Expression and Chromatin Accessibility
RA	Retinoic acid
RG	Radial glia
SC	Spinal cord
TCGA	The Cancer Genome Atlas
TET	Ten-eleven translocation
TF	Transcription factor
TG	Target gene
TME	Tumor microenvironment
TMZ	Temozolomide
WGCNA	Weighted correlation network analysis
WHO	World Health Organization

CHAPTER 1: INTRODUCTION

1.1 Overview and organization

One canonical feature of cancer is the hijacking of early developmental features, which drives a stem-like cell state and promotes tumorigenesis. This is evident across different types of gliomas where neoplastic cells closely resemble neurodevelopmental cell types. Given these parallels, glial developmental trajectories could be harnessed to uncover key aspects of tumor biology, including tumor cellular origins, mechanisms of resilience, and therapeutic targets. However, our understanding of healthy glial lineage progression remains fragmented, which is largely due to restrictions on primary human tissue samples. In this dissertation I seek to fill in some of these knowledge gaps by generating a comprehensive map of human astrocyte maturation using an *in vitro* model of human brain development and projecting this time course onto cells from primary glioblastomas (GBM), the most severe glioma subtype, to uncover how astrocyte maturation converges on GBM cell state.

Chapter 1 provides an overview of glia (astrocyte and oligodendrocyte) development and maturation, how glial development is mirrored in glioma tumors, and how principles of glial development can reveal novel information about glioma biology and guide therapeutic exploration. Chapter 2 includes the rationale, methods, and results for experiments that investigate the specific astrocyte maturation program(s) that are recapitulated by GBM. In this chapter, I start by chronicling the transcriptomic and chromatin accessibility changes throughout astrocyte maturation using human cortical organoids (hCOs) as a model of normal human neurodevelopment. Next, I investigate how astrocyte maturation molecular programs are aberrantly reflected in GBM tumors, using bulk and single-nucleus datasets that I procured from an extensive cohort of

primary GBM tissue samples. Additionally, I assess how maturation signature varies according to tumor molecular diagnostic background, highlighting that astrocytes from tumors with IDH1 mutations are significantly more mature than IDH-wildtype tumors. Lastly, I explored how DNA hydroxymethylation, which is disrupted in IDH1-mutant tumors, may contribute to a more mature astrocyte state in IDH1-mutant tumors. Chapter 3 summarizes the key findings of this dissertation research, expands upon the broader implications of this work, and explores future directions.

1.2 Glial development and maturation

1.2.1 *The rise of glial cells: how, when, and where*

Neuroectodermal development involves multipotent neural stem cells (NSC) called radial glia (RG) giving rise to three key cell populations—neurons, astrocytes (AS), and oligodendrocytes (OL). RG first undergo symmetrical divisions during early gestation to expand their pool before dividing asymmetrically towards neurogenic fates during mid-gestation, followed by gliogenic fates (astrocytes and oligodendrocytes) at later gestational and early postnatal stages (1). This shift from neurogenesis to gliogenesis termed the “gliogenic switch”, occurs around 16 gestational weeks (GW) in humans and is mediated by both intrinsic and extrinsic factors that synergize to suppress neurogenesis, release molecular brakes impeding gliogenesis, and actively promote gliogenic commitment (2, 3). For the purpose of this dissertation, reference to “glia” specifically pertains to astrocyte and oligodendrocyte macroglia populations.

1.2.1.1 Molecular regulators of the gliogenic switch

During neurogenic phases of development, premature astrogenesis is primarily prevented through inhibition of the JAK/STAT pathway and more specifically, STAT3-mediated transcription of astrocyte genes, including GFAP and S100B (3-6). Pro-neuronal transcription factors (TF), such as NGN1 (7), and neurotrophins, like BDNF (8), inhibit STAT3-mediated astrogenesis while simultaneously promoting neurogenic pathways, like MEK-ERK signaling (9, 10). These mechanisms ensure a robust population of early immature neurons prior to the emergence of astrocytes and oligodendrocytes.

A key event that drives the shift towards gliogenesis is the remodeling of regulatory genomic regions into favorable states that promote the transcription of gliogenic genes. During astrogenesis, this occurs through synergistic activation of the JAK/STAT, BMP, and Notch signaling pathways, which modulate the landscape of DNA methylation, histone methylation, and acetylation (3, 11-14). The p300/CBP complex is an important component of the JAK/STAT pathway and has intrinsic acetyltransferase activity, including helping to induce H3K9 and H3K14 acetylation at the STAT3 binding site of the GFAP promoter (6). Around the time of the gliogenic switch, Polycomb group (PcG) proteins silence NGN1 activity, inducing the release of p300/CBP, which forms a co-activator complex with STAT3 at the promoter of astrocyte genes to activate expression (11, 12). Additionally, the binding of astrocytic TFs, such as NFIA, has been shown to displace DNMT1 from astrocyte-specific promoters, helping to facilitate an active gliogenic transcriptional state (15-18).

Transcription factors are powerful molecular regulators that initiate changes in cell state, differentiation, and maturation. Advancements in high-throughput sequencing coupled with new and robust methods for studying glia—such as sophisticated 2D and 3D model systems, improved glial purification methods, and more specific genetic targeting of glia (19, 20)—helped identify several TFs that contribute to gliogenesis. Two of the first TFs that were identified as key players in the induction of astrogenesis include NFIA and SOX9. NFIA is not only necessary and sufficient to induce astrocyte formation (21, 22) but also drives HES5 expression, a Notch pathway effector required for the inhibition of neurogenesis (23). Similarly, reduced Sox9 expression results in prolonged neurogenesis and delayed gliogenesis *in vitro* (24). Kang and colleagues later discovered that Sox9 not only induces NFIA expression but identified that the two TFs form a complex to facilitate transcription of astrocyte genes (25). Two additional Sox9 binding partners, NFIB and Zbtb20, also collectively induce cortical astrocyte differentiation in mice (26, 27). Several studies have subsequently identified key regulators of the SOX9-NFIA complex, including TFs PITX1, which promotes SOX9 expression (28), and Brn2, which plays a key role in SOX9-induction of NFIA (29). Together, this paints a picture of a complex network of TF activation required to promote the switch from neurogenesis to astrogenesis. Several studies have also investigated the role of TFs at later stages of astrocyte maturation, although this developmental window remains comparatively more elusive than the initiation of astrogenesis. Lattke and colleagues uncovered Rorb, Dbx2, Lhx2, and Fezf2 as potential regulators of astrocyte maturation in the developing mouse cortex (30). However, it is notable that the authors only observed substantial changes in the degree of maturation when TFs were simultaneously overexpressed, suggesting that when acting

alone, each of the four TFs are not sufficient to induce maturation. More likely, these TFs act synergistically or even in a complex to promote maturation.

The gliogenic switch entails a shift not only from neurogenic to astrogenic fates but also towards oligodendrocyte lineages. Several TFs are implicated in early oligodendrocyte precursor cell (OPC) development and maintenance. OLIG2, an important TF in the fate specification of OPCs, activates additional OL-lineage TFs, including SOX10 (31), which acts in combination with SOX9 to promote OPC maintenance and proliferation (32). Co-deletion of Sox9 and Sox10 reduces the density of Olig2-positive OPCs within the developing spinal cord, and the remaining Olig2-positive OPCs are deficient in *Pdgfra*, a signaling pathway that promotes OPC survival and proliferation (32, 33). In return, Sox10 helps maintain Olig2 expression in Sox10-expressing cells in a positive feedback loop, together supporting the maintenance of a robust OPC population (34). Additionally, OPCs express several TFs, including SOX5, SOX6, HES5, ID2, and ID4, which prevent OPC differentiation and maturation by inhibiting TFs like OLIG1/2 and SOX10 and the downstream transcription of key maturation genes (35-37).

1.2.1.2 Extrinsic regulators of gliogenesis

In addition to intrinsic regulators, multiple extrinsic cues are also important for promoting gliogenic commitment and downstream glial development (38). Some of the most well-documented intrinsic factors in astrogenesis are a trio of IL-6 cytokines—cardiotrophin-1 (CT-1), leukemia inhibitory factor (LIF), and ciliary neurotrophic factor (CNTF)—that promote astrocyte formation through activation of the JAK/STAT pathway

(39). CT-1, the most potent agonist of the three, is secreted by newly born neurons and is necessary for the neurogenic to astrogenic transition (40). LIF and CNTF are capable of inducing astrocyte formation (41), although may not be necessary for the gliogenic switch to occur (42). Newborn neurons also secrete ligands Jagged 1 and Delta-like 1, which contribute to the gliogenic switch through activation of the Notch signaling pathway (17). Multiple cytokines, including BMP2, BMP4, and TGF- β 1, have also been implicated in astrogenesis by promoting the formation of a Smad:p300/CBP:STAT complex that facilitates the transcription of astrocyte genes (43-47). FGF2 (48, 49) and retinoic acid (RA) (50) may act more broadly to promote astrogenesis by facilitating shifts in chromatin state to elicit transcription of astrocyte genes at the onset of the gliogenic switch. Additionally, there is evidence that ligands (TGF β 2, NLGN1, TSLP, DKK1, and BMP4) synergistically activate key signaling pathways to promote astrocyte development, suggesting that much like TFs, astrogenesis is likely orchestrated by a concert of extrinsic cues (51).

Extrinsic cues also play a substantial role in OL development, including PDGF- α , FGF-2, and IGF-1. PDGF- α is secreted by both neurons and astrocytes and helps maintain the OPC population by promoting OPC proliferation and preventing premature differentiation (32, 33, 52, 53). When the PDGF mitogen binds to and activates PDGF receptors, it triggers a reorganization of the actin filament structure, stimulating changes in cell growth and motility, a cascade that when hyperactivated can serve as an oncogenic program (54). The mitogen FGF-2 helps to maintain the expression of PDGFR α and blocks oligodendrocyte differentiation by downregulating major myelin proteins (55-57).

FGF-2 and PDGF α , in combination with IGF-1, also work together to promote OPC DNA synthesis and proliferation to continually replenish OPC populations (58, 59).

1.2.1.3 Populating the CNS: where and when

Astrocytes are thought to populate the rodent brain in two waves. The first of these waves occurs prenatally when RG give rise to astrocyte progenitors at the ventricular zone during the early stages of the gliogenic switch. However, astrocyte numbers increase significantly during the first few weeks after birth, when RG numbers are much lower (60). This suggests that a secondary mode of astrogenesis is occurring postnatally when astrocytes migrate and proliferate locally to populate the central nervous system (CNS) (61). While the exact mechanisms that underlie this second wave of astrogenesis are still largely unknown, evidence suggests that in the cerebral cortex, this drastic increase in astrocyte number is the product of symmetrical division of differentiated astrocytes (62).

Much like astrocytes, oligodendrocyte lineage cells also populate nervous system tissues in multiple stages. In the spinal cord (SC), OPCs first arise in the progenitor pMN domain, which first serves as a primary source of motor neurons, and then OPCs at later stages. In this region, Shh signaling induces basic helix–loop–helix (bHLH) transcription factor (TF) Olig2, a key lineage TF for both motor neurons and OPCs (63). Phosphorylation of Olig2 at Ser147 first promotes motor neuron specification, then, later in development, this site becomes dephosphorylated, allowing Olig2 to sequester neuronal TF NGN2 and drive a shift from motor neuron to OPC production (64). This OPC population then migrates dorsolaterally, innervating the remainder of the SC (65). Around E15.5, a second

wave of OPC production occurs dorsally in the SC, contributing to ~10% of the final population in the SC (66, 67). In the forebrain, the first wave of OPC production occurs around E12.5 arising from Nkx2.1-precursors in the ventral medial ganglionic eminences before migrating throughout the cerebral cortex around E16 (68, 69). A second wave of OPC production occurs around E14.5 from Gsh2-expressing cells in the lateral and/or caudal ganglionic eminences, followed by a third wave around birth in the postnatal cortex from Emx1-expressing cortical precursors (70).

OL lineage progression consists of multiple distinct differentiation stages—OPCs, pre-OLs, pre-myelinating OLs, and myelinating OLs—that are largely defined by their myelination capacity. In humans, OPCs arise in the ganglionic eminence around 10 GW, populating the cortex by 15 GW (71). During mid-gestation (from 18-27 GW), OPCs and pre-OLs are the predominant developmental stage of the OL lineage with the first pre-myelinating OLs beginning to emerge (72). OL differentiation ramps up during the latter half of gestation (from 28-41 GW), marked by a rapid increase in pre-myelinating OLs and the emergence of mature myelinating OLs (72). These shifts in developmental state are denoted by morphological and transcriptomic changes, revealing distinct markers of each stage (63).

1.2.2. *Glial cell maturation*

1.2.2.1 Astrocyte maturation

After populating the CNS, astrocytes undergo a profound maturation process, evidenced by changes in gene expression, morphology, and function. In the first month of rodent

postnatal development astrocyte processes transition dramatically from filopodial processes that overlap with neighboring astrocytes to dense elaborate branching. These cells also occupy spatially segregated non-overlapping domains, a process referred to as “tiling” (73-76). In addition to morphological changes, prenatal astrocyte precursor cells (APCs) also have unique transcriptomic profiles in comparison with postnatal and adult astrocytes. Recent studies using mouse models (30) and primary human fetal tissue samples (77, 78) have identified thousands of differentially expressed genes (DEGs) between APCs and postnatal astrocytes, highlighting differences in physiology and function between these two maturation states. For instance, immature astrocytes express high levels of proliferation genes TOP2A and MKI67, consistent with a developmental window when these cells are populating the CNS. Conversely, mature astrocytes are enriched for transcripts encoding physiologically specialized functions like gap junction genes (GJA1 and GJB6), water channels (AQP4), and synaptogenic genes (SPARCL1).

Unsurprisingly, immature astrocytes serve different roles in the developing brain compared to mature astrocytes. During early development, astrocytes help neurons populate the brain by promoting neuronal generation, migration, and axon-path finding (79, 80). Astrocyte functions then shift to support the wave of synaptogenesis that occurs between the third trimester and 3 weeks postpartum (81-83) by guiding synapse formation, elimination, and stabilization. Pioneering work from the Barres lab demonstrated that immature astrocytes secrete thrombospondins (TSP1 and TSP2) and glypicans (Gpc4 and Gpc6) to induce the formation of structurally mature and functional synapses, respectively (84-86). Similar studies also demonstrate that astrocytes play a role in synapse pruning by phagocytosing synapses through the MEGF10 and MERTK

pathways (87) and by inducing the expression of neuronal C1q, a complement cascade protein, as a mechanism of “tagging” synapses for microglial engulfment (88). While CNS vasculature appears prior to the onset of astrogenesis, around E15 in mice (89), there is evidence suggesting that astrocyte development also contributes to blood-brain barrier (BBB) formation (90, 91) and eventually nearly all CNS astrocytes will contain endfeet that ensheath blood vessels to help regulate CNS blood flow, nutrient uptake, and waste excretion (89, 92). While immature astrocytes help guide CNS construction, mature astrocyte functions shift towards a more homeostatic state. For instance, astrocytes play a crucial role in supporting the fluctuating metabolic demands of the CNS, such as transferring energy substrates like glutamine and lactate to neurons to feed neuronal activity (93, 94), which accounts for 80–90% of the total energy consumed by the brain (95). In the maturing brain, astrocytes also help modulate neuronal signaling by recycling neurotransmitters and secreting neuroactive molecules called gliotransmitters (96).

1.2.2.2 Oligodendrocyte maturation

Similar to astrocytes, the OL lineage also demonstrates morphological, transcriptomic, and functional changes throughout maturation. Structurally, OPCs closely resemble NPCs, with a bipolar morphology and a small number of processes that emanate from opposing regions of the soma (97, 98). As OPCs differentiate into postmitotic pre-OLs and pre-myelinating OLs, they expand their total surface area by engaging with neighboring axons, losing their bipolarity, and acquiring filamentous myelin outgrowths (63, 98). This change in morphology coincides with a cascade of TFs binding to regulatory sites of myelination-promoting genes (99). For instance, during early differentiation, Olig2 is

recruited to SOX10 and myelin regulatory factor (Myrf) enhancers, activating their expression (100-102). The activation of Myrf and additional TFs, including Nkx2-2, Olig1, Ascl1, YY1, Zfhx1b, and Sox10, is necessary for proper OL differentiation into mature myelinating cells (63). Later in OL development, Olig2 and Brg1 are recruited to the enhancers of cell morphogenesis regulators, such as Cdc42 and Rac1, guiding cytoskeleton reorganization, an important step in the progression toward myelinating OLs (103). During this shift from pre-myelinating OLs to mature myelinating OLs, myelin structural proteins, including proteolipid protein (PLP), myelin-associated glycoprotein (MAG), and myelin basic protein (MBP), are upregulated coinciding with increased myelin ensheathment of axons (104, 105).

1.2.3 *Tipping the scales: making astrocytes vs oligodendrocytes*

1.2.3.1 Evidence of a shared glial precursor

For many years, OPCs were classified by their high expression of a gene called neural-gli antigen 2 (NG2). It was also widely accepted that these cells are the fate-restricted progenitors for myelinating oligodendrocytes (106). However, there is now a growing debate about whether NG2 are multipotent and can also give rise to neurons and astrocytes (106, 107). Several experiments demonstrated that when purified from rat optic nerve and embryonic spinal cord and cultured in serum-containing media, NG2 cells demonstrated the capacity to differentiate into both oligodendrocytes and “type-2 astrocytes” (108, 109). However, there was little evidence of this phenomenon in vivo, so these observations were deemed an in vitro artifact. More recent transplant studies in which purified OPCs were engrafted back into the brains of mice suggest NG2 cells can

give rise to both astrocyte and oligodendrocyte-lineage cells, indicating they have the capacity to be bipotent, but perhaps are restricted to the oligodendrocyte lineage by in vivo internal and external regulatory cues (110, 111). There is also evidence that the fate of this glial progenitor may be age- and/or region-dependent, as NG2 cells demonstrated a greater propensity for an astrocytic lineage during embryonic development and in the gray matter of the ventral forebrain and spinal cord (112, 113).

Greater access to primary tissue specimens has brought to light new and diverse progenitor populations in the developing human brain, including uniquely hominid features (114-116). Recently, several single-cell RNA-seq papers have uncovered a bipotent glial progenitor in the developing human brain by examining the transcriptome of primary human fetal tissue at a single-cell resolution. These datasets document changes across the developing human cortex and spinal cord revealing an intermediate progenitor cell type that is EGFR+/OLIG2+/OLIG1+/ASCL1+ that may give rise to both astrocyte and oligodendrocyte lineage cells (117-119). In both the cortex and spinal cord, EGFR-positive cells are split into two groups—those enriched for astrocyte markers (SOX9 and AQP4) and a separate population expressing canonical oligodendrocyte markers (SOX10, PDGFRA, and PCDH15). These data suggest that at some point this bipotent glial precursor may diverge towards either an astrocyte or oligodendrocyte trajectory (117, 119). Notably, one of these studies also identified a similar population of multipotent intermediate progenitor cells (mIPC) that are EGFR-negative. This population, enriched for both neuronal and radial glial markers (RBFOX1, ADGRV1, and NRG), suggests that EGFR may serve as a marker for progenitors committed specifically to the glial fate (117).

1.2.3.2 Diverging molecular programs

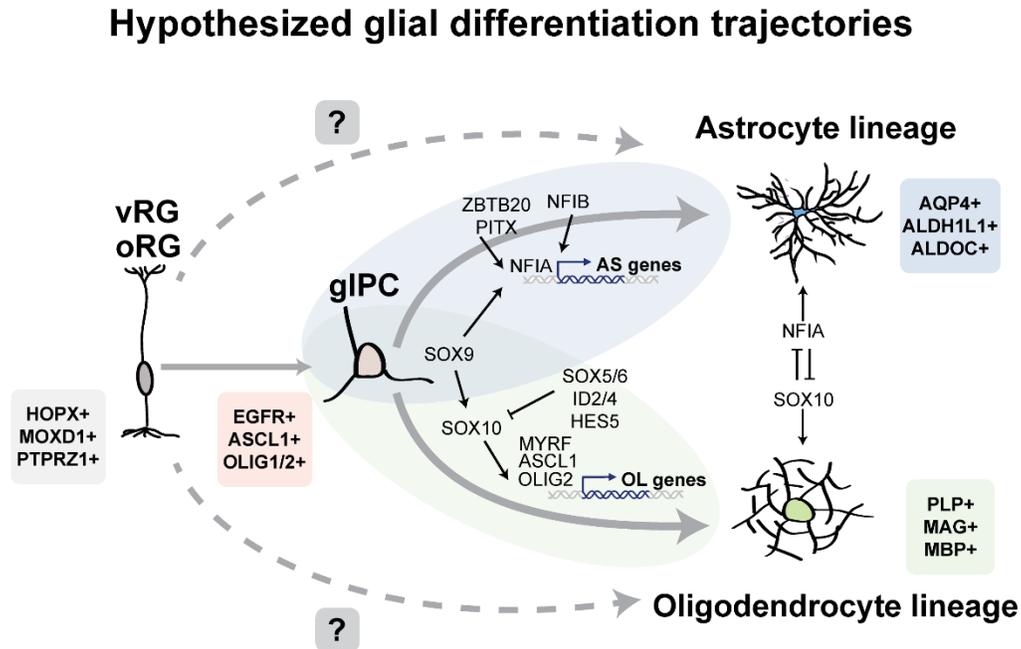
Assuming astrocytes and oligodendrocytes emanate from a shared precursor, it would be critical that molecular regulators are positioned at the right time and place to assure appropriate proportions and developmental timing of each glial lineage. Transcription factors are one such class of lineage fate determinants that can simultaneously promote one lineage trajectory and repress another. This is evident during gliogenesis with such strong overlap in the molecular programs that drive AS and OL lineages; however, these shared drivers of development behave in unique and in some cases opposing ways to promote one cellular fate or the other.

This concept is perhaps most evident when evaluating the role of SOX9 and its binding partners in determining glial fate specification. SOX9 appears to be an important component of both astrocyte and oligodendrocyte development (24, 25, 32, 120) as Sox9 knockout in the developing spinal cord inhibits both astrogenesis and oligogenesis (24). However, it serves contrasting roles in each lineage because of differences in when, where, and with whom it binds. Studies in the developing rodent spinal cord indicate that glial genes are prebound by Sox3 in NSCs. During the initial wave of astrogenesis, genomic sites marked by Sox3 are targeted by Sox9, specifically at regions enriched for Nfi binding motifs (121). Together, Sox9 and Nfi TFs facilitate the transcription of astrocyte genes to drive early astrogenesis (25). In oligogenesis, Sox9 is prebound at multiple oligodendrocyte genes, which are then targeted by Sox10 to facilitate oligodendrocyte development (24, 121). Unlike in astrocyte development, SOX9 expression appears to

peak during the OPC lineage commitment phase of OL development but then drops off during later stages of maturation (122), suggesting that it serves different roles in astrocyte and oligodendrocyte developmental progression.

Not only does SOX9 display differential binding and functional properties in AS and OL lineages, but there is also evidence that the binding partners of SOX9 in one lineage may directly antagonize SOX9 binding partners of a diverging lineage. Work by Glasgow et al using chick and mouse models demonstrated that induction of NFIA and SOX10 exhibit antagonizing effects on each other. Expression of SOX10 impedes NFIA-induced expression of AS genes and reciprocally, NFIA inhibits SOX10-induction of OL genes (123). The same study provided evidence suggesting that Olig2 may play a key role in this NFIA/SOX10 relationship by reinforcing the interaction between SOX10/NFIA, promoting a lineage-fate-decision stage (123). This suggests that while NFIA and SOX10 promote their respective lineages by interacting with SOX9, they also suppress competing lineages by interfering with each other's ability to transcribe specific glial gene sets, thereby tipping the scales toward one glial lineage or the other.

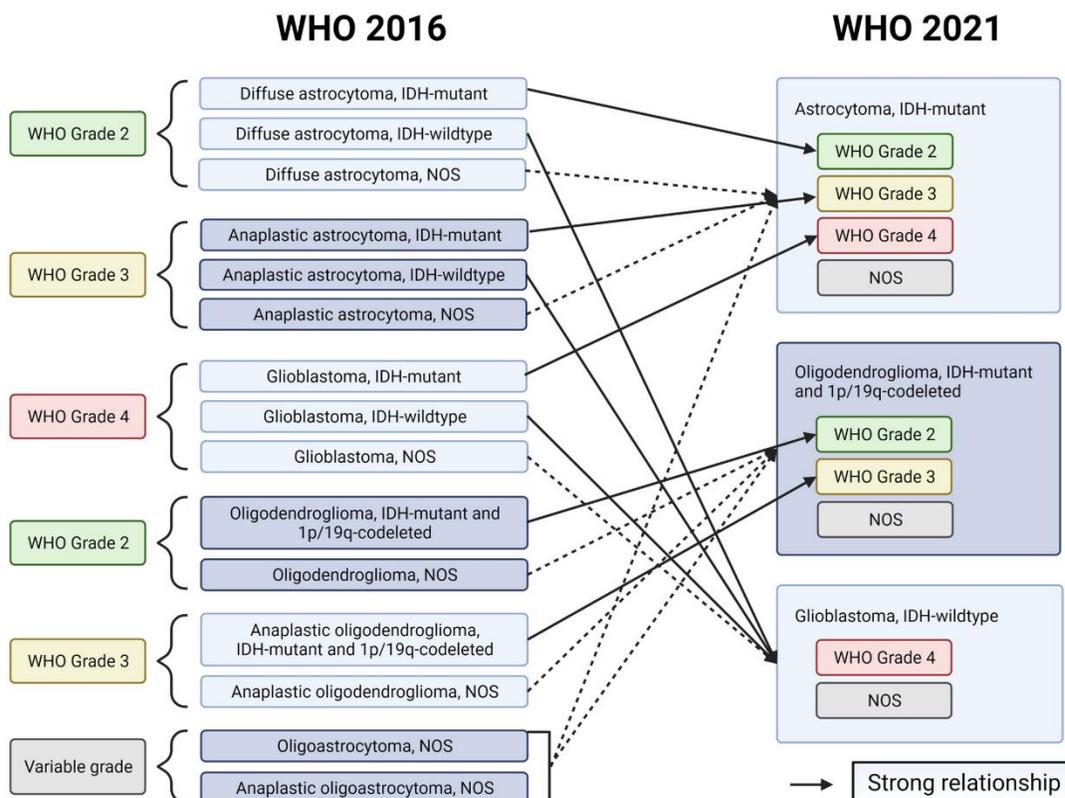
The prospect of a shared AS/OL precursor cell and a precarious scale of AS/OL fate has important implications for glioma research, where malignant cells resemble AS- and OL-like cell types and in some cases, have the capacity to differentiate between the two glial fates.

Figure 1: Hypothesized glial differentiation trajectories**Fig. 1: Hypothesized glial differentiation trajectories**

Schematized representation of proposed glial differentiation trajectories, whereby RG are hypothesized to either (solid arrows) give rise to a shared glial intermediate progenitor that can generate both astrocytes and oligodendrocytes or (dashed arrows) directly generate astrocyte and oligodendrocyte lineages. Colored boxes include cell markers and TF drivers and inhibitors are listed next to respective lineage types. Outer radial glia (oRG), ventricular radial glia (vRG), glial intermediate progenitor cell (gIPC).

1.3 Cancer as an echo of glial development

Cancer echoes many early developmental principles, including rapid cell proliferation, the activation of nascent developmental signaling pathways, a high degree of cellular plasticity, and a susceptibility to local environmental cues. Brain tumors in particular are a prime example of this developmental mimicry. Advancements in single-cell sequencing datasets confirm that brain tumors, especially glioblastomas, exhibit cellular heterogeneity comprised of hierarchies reflective of early neurodevelopment with a cancer stem cell-like (CSC) population at the apex. CSC populations have been identified in several different malignancies, including various glioma types, where they are referred to as glioma stem cells (GSCs). While it remains unclear what type of cell(s) these represent and if there is a pan-GSC marker, there are several defining features of GSCs across glioma types, including their high expression of embryonic stem cell genes and self-renewal capabilities (124-126). GSCs also demonstrate the ability to adapt to the tumor microenvironment and differentiate into multiple lineage types, reminiscent of the NSC population within the embryonic brain (124-126). This population of cells is believed to be the source of tumor propagation (125, 126) and is capable of evading immune surveillance and treatments such as chemotherapy and radiation (127-129). Additionally, functional studies in *Drosophila* and rodent models show that key early neurodevelopmental signaling cascades, such as Wnt, Notch, and Hedgehog pathways, are implicated in the tumorigenesis of CNS malignancies, suggesting that brain tumors recycle early developmental blueprints for generating and maintaining progenitor populations (130-134). Given the parallels between neurodevelopmental principles and brain tumor biology, a better understanding of the internal and external regulators of normal early developmental trajectories could inform how we study and treat gliomas.

Figure 2: Updated 2021 WHO diffuse glioma classification**Fig. 2: Updated 2021 WHO diffuse glioma classification**

Schematized explanation of how diffuse glioma classification is different between the WHO 2016 and WHO 2021 reports. Dashed lines indicate how a classification is most likely to now be defined and solid lines indicate strong relationships between the previous (WHO 2016) and current (WHO 2021) classifications. Image reproduced from *Whitfield and Huse, 2022*, with copyright permission from the publisher.

1.3.1 Glioblastoma

Glioblastoma (GBM) is classified by the World Health Organization (WHO) as a grade IV glioma. These tumors are the most aggressive and common primary CNS malignancy, accounting for approximately 16% of all primary CNS neoplasms (135). For primary (de novo) GBMs, which account for 80% of all GBM tumors, the median age of diagnosis is 62 (136), whereas secondary GBMs, which develop from lower-grade astrocytoma or oligodendrogliomas, are more frequent in younger patients (mean age 45 years) (136, 137). The typical treatment course for patients with GBM consists of maximal safe surgical resection followed by radiotherapy and temozolomide (TMZ) chemotherapy (138). Unfortunately, due to the diffuse, heterogeneous, and resilient nature of GBM, these tumors are nearly impossible to entirely eradicate, and the prognosis remains bleak with a median survival of 15 months (139, 140).

Glioblastoma was the first cancer included in The Cancer Genome Atlas (TCGA), a collaborative effort between the National Cancer Institute and the National Human Genome Research Institute to catalog high-throughput sequencing datasets from large cohorts of human tumors. The first wave of datasets included in this project highlighted inter-tumoral transcriptional heterogeneity across GBMs, classifying tumors into four transcriptional subtypes—proneural, neural, mesenchymal, and classical—where each subtype was enriched for cell type-specific gene signatures and oncogenic events (141, 142). However, studies incorporating multi-region sampling across individual GBM tumors demonstrated that many transcriptional subtypes exist within different regions of the same tumor (143). This finding was confirmed and delineated by a wave of GBM single-cell transcriptomic studies, which depict ample intra-tumoral cellular

heterogeneity that was not captured through bulk sequencing approaches. Together, these studies revealed an additional layer of complexity to this specific neoplasm and uncovered previously unknown mechanisms of tumor resilience and treatment resistance.

Generating transcriptomic profiles with single-cell resolution facilitated several pioneering discoveries in the GBM field. The first of which was the ability to distinguish between neoplastic and non-neoplastic cells in the tumor bulk using predicted copy number variations for each individual cell. This capability led to the finding that the “neural” transcriptional tumor subtype was an artifact of non-cancerous neuronal populations driving that particular gene signature (144). More recent studies have adopted a more nuanced approach to GBM transcriptional classification. Neftel et al demonstrated that GBM malignant tumor cells generally fall into four transcriptional subtypes that reflect- (1) neural-progenitor-like (NPC-like), (2) oligodendrocyte-progenitor-like (OPC-like), (3) astrocyte-like (AC-like), and (4) mesenchymal-like (MES-like) states, where any given tumor possesses cells that exist in all of these states in varying ratios (145). Pseudotime analysis suggested that malignant cells exist along a stemness hierarchy, where a small population of malignant tumor cells that closely resemble multipotent NSCs sit at the apex, and the remaining majority of tumor cells exist in transcriptional paths along the four differentiation trajectories (145).

It is important to note that these transcriptomic analyses capture cell states at a single moment in time. Functional studies where cells of a specific GSC population have been engrafted into patient-derived xenografts demonstrate that GSC state is anything but

stagnant, and that cell state is not driven solely by genetics, but also by microenvironment (145-147). For example, Neftel et al found that regardless of the cell population used to initiate the xenograft—AC-like, NPC-like, or MES-like—the resulting tumor presented all three cell states in comparable frequencies (145). Several key tumor microenvironmental (TME) niches influence tumor cell biology: the perivascular niche, which provides cues to maintain stemness and induces pathways that support migration and DNA repair (148-151); the hypoxic niche, which promotes GSC maintenance, proliferation, and therapy resistance, mainly through hypoxia-inducible factor 1 (HIF-1) and HIF-2 (152-155); and the invasive edge, where GSCs migrate along vasculature and white matter tracts (156). Additionally, evidence suggests that therapeutic intervention, itself, can induce a shift in GSC state to a phenotype more conducive for evading harsh treatment strategies (157, 158). However, extrinsic cues alone do not dictate GSC state. Neftel et al also found that frequencies of each state are associated with genetic alterations in CDK4, PDGFRA, EGFR, and NF1 that appear to bias cell identity towards a particular state (145). Thus, while GSCs and normal developmental cell types share the capacity to respond to environmental cues, it is the combination of oncogenic mutations, general genomic instability, and disruption of chromatin regulators that permits GSCs to override normal systems of checks and balances.

Many of the genetic aberrations in GBM are in genes that play critical roles in normal glial development. One powerful approach for studying the tumorigenic potential of different driver gene candidates is with genetically engineered mouse models (GEMMs). GEMM approaches including the Cre-lox recombinase system, virus-mediated gene delivery (RCAS–TVA), and transposon-based insertional mutagenesis are particularly attractive

options as they offer temporal and cell-type control over tumor initiation (159, 160). These approaches illustrate that loss-of-function of GBM-associated tumor suppressor genes (TP53, PTEN, NF1) or gain-of-function of oncogenes (EGFR, PDGFR, RAS, AKT) can induce dedifferentiation of quiescent glia (161-165), restrict progenitors to an immature state (166, 167), and may even promote the inter-conversion between glial types (168). Likewise, neurodevelopmental TFs can act as oncogenes by inappropriately activating developmental programs that drive tumorigenesis (169-171). This was verified in work from the Bernstein lab, which demonstrated that activation of ASCL1, a driver of both neuro- and gliogenesis, can sustain GSC proliferation via WNT pathway activation, secondary to the repression of the WNT inhibitor DKK1 (169). This team also uncovered several additional neurodevelopmental TFs including POU3F2, SOX2, SALL2, and OLIG2, that activate GSC regulatory networks and push differentiated GBM cells from non-tumorigenic to tumor propagating GSCs (170). Olig2 has also demonstrated the capacity to dictate GSC subtype, as a loss of Olig2 causes a shift from a proneural transcriptional subtype towards a more astrocytic phenotype, including downregulation of PDGFR and concomitant upregulation of EGFR (172).

Another clear example of the convergence between developmental and oncogenic programs is shown by the redundancy of EGFR activity in development and gliomagenesis. EGFR activity is essential during normal gliogenesis (118, 173-175) and both EGFR amplification or constitutively activating mutations (EGFRvIII) are amongst the most common molecular features of GBM, occurring in about 50% of all cases (176, 177). The tumor biology of EGFR signaling is highly nuanced. Liu et al demonstrated that the most common EGFR mutation, EGFRvIII, remodels the enhancer regulatory

landscape of GBM, inducing two TFs that are key regulators of astrocyte development, SOX9, and FOXG1. Together, these EGFR-dependent TFs work collaboratively to induce oncogenic programs, including c-MYC target genes and EGFR-regulated genes (178). Interestingly, a recent study from the Deneen lab showed that one of these EGFR targets, SOX9, behaves differently in diverse brain tumor subtypes, which each exhibit unique epigenomic states and drive divergent roles in tumorigenesis (179). Thus, while the activation of developmental programs is a shared biological phenomenon in glioma, the same molecular perturbation can induce opposing outcomes within different cellular contexts.

1.3.2 *Other diffuse and lower-grade gliomas*

While GBM is more prevalent in adulthood and has a higher incidence in the supratentorial compartment, low-grade gliomas (LGG) and pediatric high-grade gliomas have a higher incidence in the posterior fossa. Diffuse gliomas consist of astrocytomas (WHO grades II, III, and IV) and oligodendrogliomas (WHO grades II and III), most of which are classified as having IDH1/2 mutations, but oligodendroglioma is distinguished by a chromosomal 1p/19 co-deletion (180, 181). LGGs also include WHO grade I astrocytomas, the most prevalent of which are pilocytic astrocytomas (PAs); however, these tumors are not typically diffusely infiltrating and generally have a more favorable prognosis (180, 181). Across various glioma subtypes, there are abundant molecular, histopathological, and prognostic variations; however, one shared feature is the reflection of early neurodevelopmental cell types and molecular programs.

1.3.2.1 Diffuse glioma

Diffuse oligodendrogliomas and astrocytomas demonstrate a recycling of early glial differentiation programs to fuel immature developmental cell states. For instance, the regulatory chromatin architecture that is present in normal gliogenesis and the binding of astrocytic TFs like SOX9, NFIA, and BRN2, is shared by models of diffuse glioma and promoted tumorigenesis (29). Several of these potent glial fate determinants even demonstrate the capacity to regulate glioma subtype specification reminiscent of the early developmental decision to bias towards AS versus OL lineages. This was perhaps most clearly demonstrated by experiments overexpressing NFIA in a mouse model of oligodendroglioma, which shifted tumor histopathology to more closely reflect astrocytomas (123). In addition to intrinsic regulators, extrinsic cues also play a role in driving glioma phenotype. PDGF, a potent mitogen involved in generating and maintaining OPCs in the developing brain, induces tumors that reflect oligodendroglioma biology (182, 183).

While diffuse astrocytoma and oligodendroglioma are characterized by unique histological features, genomic perturbations, and markers of gliogenic regulation, surprisingly, neurodevelopmental lineages are reflected quite consistently between the two tumor types. Single-cell transcriptomic work by Venteicher et al highlighted the similarities between the two classes of IDH-mutant glioma, demonstrating that both harbor three groups of malignant tumor cells—a relatively small proliferative NSC-like population, and two populations of nonproliferating cells that resemble AS and OL lineages (184). Interestingly, the primary differences between astrocytomas and oligodendrogliomas are related to genetic events and tumor microenvironmental niches

(184). When focusing on the cellular heterogeneity within oligodendrogliomas, Tirosh et al found that CNV-subclones within these tumors span all three different transcriptional states—NSC-like, OL-like, and AS-like—suggesting that factors beyond genetic events were contributing to the observed developmental hierarchy (185). This finding is supported by experiments where PDGF exposure yields an inconsistent tumor phenotype between WHO grade II oligodendroglioma and a mixed oligoastrocytoma profile that expresses both GFAP and Vimentin (182, 183).

1.3.2.2 Pilocytic astrocytoma

Unlike grade II and III astrocytomas and oligodendrogliomas, PAs do not progress to higher-grade gliomas and most commonly arise in the optic pathway, brainstem, and cerebellum (186). In comparison to higher-grade diffuse gliomas, PAs are more genomically simple, with most exhibiting only a single-driver alteration activating the MAPK pathway (186). PAs in the cerebellum commonly develop sporadically and display a somatic rearrangement where the BRAF gene kinase domain is fused to the KIAA1549 gene (referred to as KIAA1549:BRAF) (186). An additional PA subtype is present in children with Neurofibromatosis type 1 (NF1) tumor predisposition syndrome, who typically experience tumors in optic pathways (186).

Akin to other glioma, the developmental origins of PAs are largely still unknown, although, there is some evidence implicating OPC and astrocyte populations, specifically. For instance, NF1-deficient astrocytes display hyperactive mTOR signaling and a greater proliferative capacity, a phenotype that is observed in patient NF1 PA tumors (187, 188).

Conversely, limited evidence suggests that ectopic expression of the KIAA1549:BRAF fusion protein does increase proliferation of NSCs in vitro and in vivo; however, there is more uncertainty about the cell of origin in KIAA1549:BRAF fusion PAs as a result of limited experimental models (189, 190). However, recent single-cell RNA-seq studies have begun unraveling the cellular hierarchies within these tumors, providing new evidence that an OPC-like progenitor population enriched for MAPK signaling may give rise to a much larger group of AC-like cells with diminished MAPK signaling activity (191, 192). In comparison to GBM and other diffuse astrocytomas/oligodendrogliomas, the NSC signature is noticeably absent from PA cells, suggesting that PAs may be driven by a more developmentally committed OPC-like cell (191, 192).

1.3.3 Diffuse midline glioma

Diffuse midline gliomas (DMGs) are a primarily pediatric and extremely aggressive glioma subtype with a median survival of about one-year post-diagnosis (193). These tumors are regionally specific to midline structures occurring in the thalamus, midbrain, cerebellum, or pons; the latter of which are called diffuse intrinsic pontine gliomas (DIPG) (194). A major breakthrough in understanding the biology of these tumors was the finding that many of these tumors contain a lysine27-to-methionine (K27M) mutation in histone 3 (H3). In H3K27-altered DMGs, H3K27M suppresses EZH2, the catalytic subunit of polycomb repressive complex 2 (PRC2). Polycomb activity is involved in a variety of epigenetic regulatory processes, including trimethylation of Lys-27 on histone 3 (H3K27me3) (195, 196), which leads to genome-wide dysregulation of gene repression and cell differentiation (196, 197).

Single-cell transcriptomic profiling of these tumors has uncovered a similar developmental hierarchy in H3K27M-glioma to other diffuse gliomas; however, there are several noteworthy differences (198). H3K27M-gliomas contain a substantially larger pool of undifferentiated cells, consistent with the more aggressive nature of this tumor (198). Undifferentiated cells in DMG more closely resembled OPC lineages (198), unlike the putative GSCs in IDH-mutant diffuse gliomas that reflect an NSC identity. Additionally, in contrast to IDH-mutant glioma, DMGs exhibit minimal signatures of differentiated OL-like cells, and only a small percentage of differentiated AC-like cells (198).

More recent work from Jessa et al and Liu et al implemented a barrage of single-cell genomic, epigenomic, and chromatin profiling approaches to dissect region- and age-related developmental signatures in DMG. Jessa et al profiled cells across DMGs that harbor the H3K27M mutation in different histone variants (H3.1 and H3.3) and demonstrated that while K27-gliomas appear to maintain a developmentally conserved chromatin regulatory architecture, differences between H3.1 and H3.3 samples point to distinct OPC developmental origins (199). Specifically, the molecular profile of H3.1K27M ACVR1-mutant pontine gliomas resembled early ventral NKX6-1+/SHH-dependent brainstem OPCs, whereas the H3.3K27M signature was more closely aligned with later dorsal PAX3+/BMP-dependent progenitors (199). Liu et al observed the presence of a stem-like OPC population across all H3K27M-gliomas, regardless of age or tumor location (200). Remarkably, the team identified location-specific OPC subpopulations, where pontine tumors were enriched for a more immature pre-OPC-like signature in

comparison with thalamic tumor OPC signatures, corroborating the findings of Jessa et al that pontine K27-gliomas may arise from an OPC population of earlier origins (200). Together, these two studies suggest that DMG arising in different brain regions may descend from distinct cells of origin, but likely undergo similar developmental pressures that shape a shared DMG cellular hierarchy.

Figure 3: Representation of developmental lineages across glioma subtypes

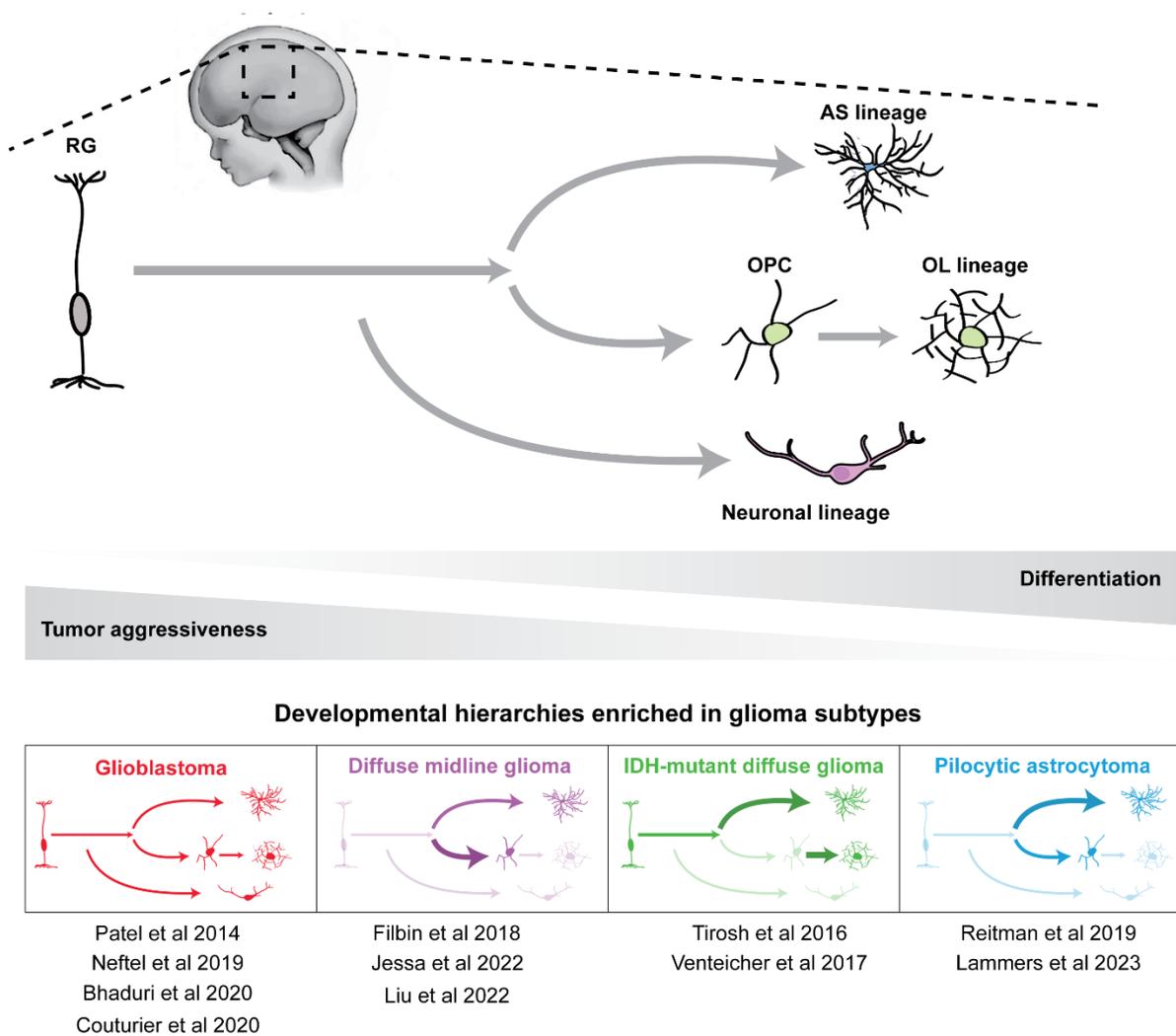


Fig. 3: Representation of developmental lineages across glioma subtypes

Schematic depicting the normal neurodevelopmental cell hierarchy and how these cell states are (over)represented across glioma subtypes. The glioma subtypes are arranged in order of relative aggressiveness, with the most aggressive and dedifferentiated (glioblastoma) on the far left. Relative enrichment of neurodevelopmental cell states are represented by arrow thickness. Relevant single-cell publications that support these findings are listed below respective tumor types.

1.4 How do glioma cells move across developmental time?

While early glial cell types and developmental hierarchies are recapitulated in most gliomas, we are still learning how to use normal developmental trajectories to better understand glioma biology and potential therapeutic interventions. More detailed and complete glial maturation atlases can provide insight into the initial, present, and future developmental stages that glioma cells progress through. This can be thought of as (initial) where along the normal developmental trajectory do gliomas *start* (i.e., cell of origin); (present) when in developmental time do glioma cells *reside* and thrive during tumor progression; and (future) when in developmental time are glioma cells capable of *moving towards* with intervention.

1.4.1 Where along the normal developmental trajectory do gliomas start?

The first of these questions is a long-standing enigma: which cell types have the capacity to give rise to gliomas? An important note here is that there is certainly a difference between cells with gliomagenic capacity and the reality of which cells tend to originate tumors *in vivo*. There are two prevailing theories addressing glioma cellular origin(s), a question that remains highly controversial in the field. The first is a scenario where a differentiated somatic cell stochastically gains a combination of oncogenic and/or tumor suppressor mutations, through a variety of possible mechanisms including replication errors or DNA damage, transforming quiescent cells into a stem-like state. Critics of this theory argue that it is unlikely that a mature non-proliferative cell with a limited lifespan could accumulate the perfect cocktail of mutations to induce such tumorigenic potential. However, recent work investigating this possibility in the context of chronic inflammation lends credence to this hypothesis.

By analyzing published astrocyte reactivity datasets, Simpson Ragdale et al identified an upregulation of the canonical tumor suppressor gene TP53 in injured astrocytes. Interestingly, although these cells are reactive, they do not exhibit dedifferentiated features. This suggests that TP53 may actively prevent reactive cells from returning to a more stem-like state (162). However, in the context of glioma, where TP53 is frequently mutated, this putative dedifferentiation brake might be absent, permitting injury-induced dedifferentiation. Supporting this hypothesis, TP53 knockout in reactive astrocytes destabilized astrocyte fate, such that adult astrocytes were capable of dedifferentiating 8-10 months after injury when inflammation reached a chronic state (162). Over time, age-exacerbated inflammation coupled with EGF secretion from periwound astrocytes, induced mTOR-dependent reacquisition of early neurodevelopmental TF programs, including Sox2, Olig2, and Ascl1 activation (162). Notably, TP53 loss or injury alone was insufficient to induce adult astrocyte dedifferentiation (162). This suggests that TP53 mutation may lift the restraint on fate commitment, while chronic inflammation could serve as a “second hit” to induce dedifferentiation at later time points.

The second cell-of-origin theory proposes that GSCs arise when an endogenous quiescent stem cell in the brain acquires oncogenic mutations. Importantly, the exact identity of this stem-like cell is still up for debate and likely varies across glioma subtypes. Given the cellular heterogeneity of GBM, many hypothesize that the GSC origin in GBM is adult NSCs, which developmentally have the capacity to generate each of the transcriptomic subtypes documented by Neftel et al (AS-like, OPC-like, and NPC-like). Indeed, a multitude of evidence supports this theory. For instance, GBM tumor cells share many

properties with NSCs, including high expression of NSC markers (201, 202); they can form neurospheres that have a similar structure to those derived from adult human subventricular zone cells (203, 204); and Nestin-positive tumor cells are critical for tumor growth and chemotherapy resistance (127). Alternatively, others hypothesize that a lineage-committed precursor, such as an OPC or astrocyte precursor cell, is a more likely culprit, given that GSCs express markers for these cell types as well (205-208) and there is evidence that both lineages possess tumor propagating potential (163, 205, 209, 210). However, OPCs are unique in that they are the major dividing cell population in the adult CNS, vastly outnumbering NSCs, which are only present in ventricular zone niches and the dentate gyrus (211, 212). Far less evidence exists on whether astrocyte progenitor cells exist in the adult brain and if so, exhibit the same proliferative potential as OPCs in the adult CNS (213). It is also conceivable, and highly likely, that there are discrete cells of origin for different GSC subtypes or for the same GSC subtype across different brain regions.

1.4.2 When in developmental time do glioma cells reside and thrive during tumor progression?

While it is evident that gliomas reflect multiple early developmental cell types, it is unclear if there are specific maturation stages of glial development reflected in glioma tumors. This is largely due to our fragmented understanding of glial maturation, which in humans primarily occurs between the third trimester of gestation and the first postnatal month, a brief but critical period when access to primary human tissue samples is greatly restricted. Much of what we do know about glial development and maturation is derived from murine model systems and a limited number of second trimester primary fetal tissue

samples. While informative, there are major temporal gaps during this developmental window in human samples, which bears many neurodevelopmental differences from rodents (77, 214-217). Curating more comprehensive developmental timelines of glial lineages will help inform whether glioma cells are stalled at particular developmental stages and which molecular programs could be leveraged to coerce maturation towards a quiescent state.

One viable option for building comprehensive timelines of human glial maturation is to leverage human in vitro model systems, such as human brain organoids. There are now numerous robust protocols for forming and culturing human stem cell-induced 3D organoids that are patterned to reflect various regions of the CNS, including forebrain (218, 219), midbrain (220, 221), hindbrain (222, 223), and spinal cord (224, 225). This platform recapitulates many key features of human neurodevelopment including complex cellular composition, intricate tissue architecture, and functionally active neurons (219-221, 226-229). Additionally, long-term culture of human brain organoids depicts maturing astrocyte (230, 231) and oligodendrocyte (232-234) lineages with transcriptomic profiles that reflect pre- and postnatal stages of human brain development. Altogether, this makes organoids an ideal system for chronicling elusive windows of development at a high temporal resolution to capture all phases of glial maturation (235).

1.4.3 *When in developmental time are glioma cells capable of moving towards with intervention?*

Given the parallels between neurodevelopment and glioma biology, one might reasonably hypothesize that malignant glioma cells are susceptible to the same maturation cues that coerce quiescence in normal glial development. This is the rationale behind differentiation therapy, which explores therapeutic options to coerce tumor cells through developmental time by exploiting extrinsic and intrinsic factors that regulate normal cell differentiation and maturation. This approach has proven to be highly successful for myeloid malignancies, particularly acute myeloid leukemia (AML), leading to drastically improved patient survival rates (236). Differentiation therapy is a particularly attractive option for glioma treatment as well, given that GSCs have been shown to adapt to and evade current cytotoxic treatment regimens (127-129). Of course, a major challenge remains in identifying the optimal glial maturation cues to target.

One avenue under active investigation is targeting signaling pathways that are critical in initiating gliogenesis (BMP, Wnt, Notch, STAT3, MAPK/ERK, and TGF- β) and that are frequently hijacked in glioma progression (237). Several of these pathways appear to have particularly potent impacts on GSC differentiation when targeted through BMP4 and RA treatment, both of which are important in early astrogenesis (3) and are currently being tested in clinical trials (238). Studies treating GSCs with all-trans RA show a reduction of cell growth and proliferation, induction of lineage-specific differentiation markers, and decreased neurosphere-forming capacity through inhibition of Notch signaling (239, 240). Notably, when Notch signaling was constitutively activated, RA-induced differentiation was repressed, suggesting that RA treatment acts at least in part through

this pathway to abrogate tumorigenicity (240). Other groups found that when treated with BMP4, tumor sphere self-renewal, and GSC marker expression was significantly repressed (241, 242). However, most of these initial experiments were completed using in vitro cell lines and further testing indicated that GSCs from different patients respond inconsistently to BMP4 treatment (243, 244). Interestingly, GSCs with high expression of EGFR appear to be more vulnerable to TMZ after BMP4-induced differentiation (245).

In addition to signaling cascades, several TF candidates show promise in promoting GSC differentiation. TF activity is a key component of driving normal glial development, where the timing, binding partners, and binding location dictate lineage progression. Unsurprisingly, many glial development TFs are inappropriately activated in glioma and one proposed approach is to inhibit this aberrant TF activity in GSCs (238, 246-248). One key TF that promotes stemness in GSCs is SOX2 (169). Fang et al demonstrated that DNA-dependent protein kinase (DNA-PK) is essential for SOX2 stability as well as GSC maintenance (249). Inhibiting DNA-PK lead to prolonged survival of tumor-bearing mice and sensitized glioblastoma xenografts to radiotherapy (249). Likewise, STAT3, a key player in the gliogenic switch, is highly upregulated in gliomas, is associated with glioma EGFR amplification, and contributes to GSC proliferation and migration, thus making it a high-priority target for inhibition (246, 250). Multiple research groups have identified approaches for suppressing STAT3 activity in GSCs, resulting in increased GSC sensitivity to subsequent chemo and radiation therapy (250-252).

While there is accumulating evidence that malignant glioma cells are receptive to glial developmental cues, there are several technical challenges and caveats to differentiation

therapies that must be considered. As illustrated through rigorous testing of BMP4 treatment, it is extremely challenging to identify targetable molecular programs that overcome inter- and intra- tumoral heterogeneity boundaries (243, 244). Not only do tumor cells from different patients exhibit varying genetic backgrounds that respond inconsistently to BMP4 treatment, but it is likely that GSCs exposed to separate TME niches may also respond uniquely to treatment. For instance, while BMP4 treatment appears to be effective for inducing a pro-astrocytic differentiation program in GSCs, this did not hold true in the context of a hypoxic environment, where GSCs maintained their stem-like properties (253). Another obstacle is defining the benchmarks for successful GSC differentiation. As demonstrated by in vitro experiments overexpressing TFs to induce normal glial development (30), different TFs will likely induce unique epigenetic and transcriptomic changes. Deciphering which set(s) of changes are the benchmark for successful GSC differentiation will be a critical step in identifying effective differentiation candidates. This will be especially challenging without more detailed molecular maps of normal maturation in glial lineages. Lastly, even if GSCs respond to differentiation cues to progress through developmental time, these changes may be transient. In fact, Caren et al demonstrated that GSCs are capable of reverting to a stem-like state following BMP-treatment as a result of incomplete chromatin accessibility changes that permit aberrant SOX TF binding (254). This suggests that the most effective differentiation method will need to induce large-scale chromatin architecture shifts that are comparable to what occurs in normal glial maturation (247).

1.5 Summary and thesis objectives

Evidently, there is a clear overlap between early glia and adult glioma developmental programs, particularly in GBM, the most severe glioma subtype. Thus, utilizing blueprints of normal glial lineage trajectories could reveal new information about glioma occurrence, growth, and resilience. However, a lack of robust human model systems and primary tissue samples has left gaps in our understanding of normal glial maturation. In this dissertation, we leveraged human cortical organoids (hCOs) to first create a comprehensive molecular timeline of normal human astrocyte maturation, one of the primary glial cell types in hCOs. We followed this up with extensive transcriptomic and chromatin accessibility profiling of a diverse cohort of primary GBM tissue samples to identify specific windows of astrocyte maturation that are aberrantly activated in malignant tumor cells. The work presented in this dissertation seeks to address four primary questions: (1) what are the major molecular shifts—at the chromatin- and transcriptome- level—across normal human astrocyte maturation?; (2) which specific stage(s) of our complete astrocyte maturation time course are reflected in GBM; (3) is astrocyte maturation signature consistently represented across tumors with varying molecular backgrounds; (4) how might one distinctive molecular diagnostic mark—IDH1 mutations—preserve astrocyte maturity in a malignant tumor? Together, this research reveals the junction between astrocyte maturation and GBM cell state, providing valuable insight into glioma tumor biology.

Figure 4: How do glioma cells move across developmental time?

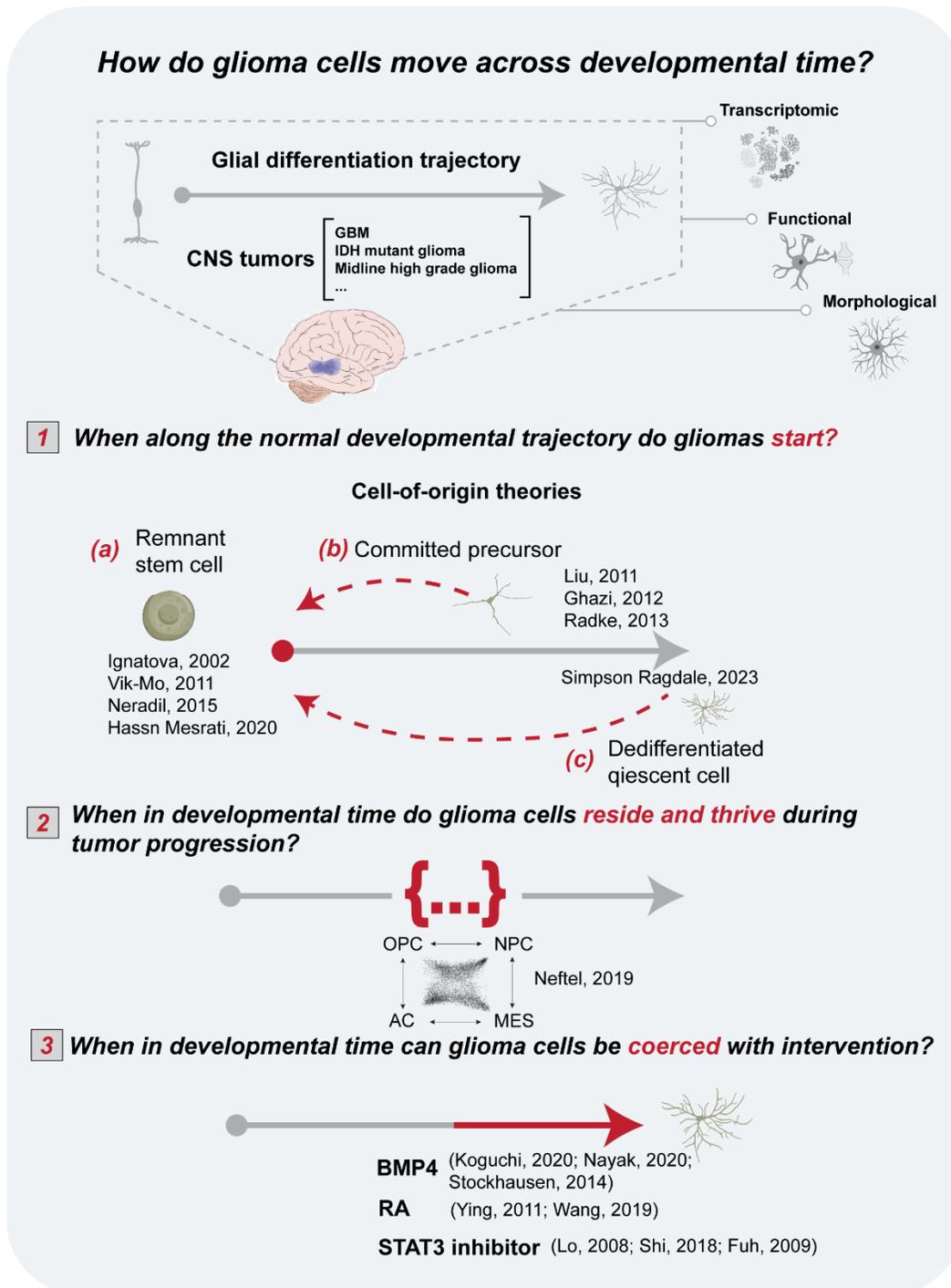


Fig. 4: How do glioma cells move across developmental time?

Schematized summary of how normal glial differentiation trajectories (generated from transcriptomic, functional, and morphological information) can inform how glioma cells progress through developmental time—where along the normal trajectory do tumor cells (1) start, (2) thrive during tumorigenesis, and (3) are capable of being coerced to?

Citations are listed for primary evidence related to each respective question.

CHAPTER 2: DIVERGENCE FROM THE HUMAN ASTROCYTE DEVELOPMENTAL TRAJECTORY IN GLIOBLASTOMA

Caitlin Sojka¹, Hsiao-Lin V. Wang^{1,2}, Yangping Li¹, Pankaj Chopra¹, Tarun Bhatia¹, Anson Sing¹, Anna Voss¹, Alexia King¹, Feng Wang¹, Kevin Joseph³, Vidhya Ravi³, Jeffrey Olson⁴, Kimberly Hoang⁴, Edjah Nduom⁴, Victor Corces^{1,2}, Bing Yao¹, Steven A. Sloan^{1,2*}

¹Department of Human Genetics, Emory University School of Medicine, Atlanta, GA, USA.

²Emory Center for Neurodegenerative Disease, Emory University School of Medicine, Atlanta, GA, USA.

³Department of Neurosurgery, Medical Center and Faculty of Medicine, University of Freiburg, Germany

⁴Department of Neurosurgery, Emory University School of Medicine, Atlanta, GA, USA.

2.1 Abstract

Glioblastoma (GBM) is defined by heterogeneous and resilient cell populations that closely reflect neurodevelopmental cell types. While it is clear that GBM echoes early and immature cell states, identifying the specific developmental programs disrupted in these tumors has been hindered by a lack of high-resolution trajectories of CNS developmental lineages. Here, we delineate the course of human astrocyte maturation to find where GBM astrocyte populations diverge. We generated a transcriptomic and chromatin accessibility map of human astrocyte maturation using cortical organoids maintained in culture for nearly two years. We chronicled a multi-phase developmental process orchestrated by a series of transcription factor and gene regulatory networks including a novel and molecularly distinct intermediate stage of human astrocyte maturation that separates proliferating progenitor from quiescent mature states. This intermediate stage serves as the site of developmental deviation in IDH-wildtype neoplastic astrocyte lineage cells. Interestingly, IDH1-mutant tumor astrocytes are the exception to this developmental perturbation, as they reflect a substantially more mature signature than IDH-wildtype astrocytes. We propose that this maturation preservation is likely a consequence of IDH1mt-associated epigenetic dysregulation and identified biased DNA hydroxymethylation (5hmC) in maturation genes as a possible protective mechanism. Together, this study illustrates a novel cellular state aberration in GBM astrocyte lineage cells and presents new developmental targets for experimental and therapeutic exploration.

2.2 Introduction

The hijacking of early cell development programs is a common oncogenic feature to promote tumor survival (255-257). Recent single-cell transcriptomic studies provide a better understanding of tumor cell heterogeneity, revealing subpopulations of tumor cells that show close parallels to normal developmental lineages (184, 185, 258-260). This is apparent in Glioblastoma (GBM), the most severe grade of astrocytoma, where tumor cells bear a resemblance to three major neurodevelopmental cell types: neural precursor cells (NPC-like), oligodendrocyte precursor cells (OPC-like), and astrocytes (AC-like) (142, 145, 261-266). These cells, along with a fourth mesenchymal cell state (MES-like), exist on a proliferative spectrum, populate patient tumors in varying ratios and demonstrate the capacity to transition into one of the other cell states (145). Together, this cellular heterogeneity and plasticity make it especially challenging to study and treat GBM.

Utilizing maps of normal developmental lineages has greatly improved our understanding of multiple brain tumors including medulloblastomas and diffuse midline gliomas (200, 267). Given that GBM also echoes features of early neurodevelopment, this crossover approach between normal and oncogenic development could serve a similar purpose and guide studies on how these tumors emerge, transform, and resist treatment. However, while we now have a plethora of transcriptomic datasets characterizing the cells that comprise GBM tumors (142, 145, 261-263, 265), we still lack a sufficiently detailed understanding of human CNS lineages to identify how and where GBM cells deviate from normal developmental programs.

Astrocytes, one of the most abundant cell populations in GBM neoplasms, are among the most challenging neurodevelopmental trajectories to define. Our current understanding of astrocyte development is fragmented because the majority of astrocyte maturation occurs between late gestational and early postnatal ages (268-271) when access to primary human tissue samples is restricted. Rodent models and a limited number of second trimester primary fetal tissue samples have demonstrated that there are large-scale transcriptional (30, 77, 272-275), morphological (61, 74, 275-277), and functional (61, 268, 277) differences between fetal/embryonic and postnatal astrocytes, suggesting that these cells undergo a profound maturation process during this elusive developmental window. Additionally, transcriptomic and epigenomic studies have uncovered several transcription factors (TF) capable of inducing astrogenesis (21, 22, 25) that may also be disrupted in GBM (123, 178, 278, 279). However, far less is known about the molecular dynamics that occur throughout human astrocyte maturation and the TFs and gene sets that promote a quiescent, non-proliferative state during later stages of development.

To overcome this hurdle, we used human cortical organoids (hCOs) to capture a critical longitudinal window of neurodevelopment and create a detailed molecular timeline of astrogenesis spanning nearly 2 years in culture. This system recapitulates many features of human cortical development (280, 281), including astrocyte maturation (230), making it a practical and informative approach for studying this time-sensitive process at a high temporal resolution. Here, we leverage both transcriptomics (RNA-seq) and chromatin accessibility (ATAC-seq) to achieve two important advantages. First, the redundancies across these platforms minimize the chance of transcriptional noise

blurring transient developmental states. Second, the synergy of these tools allows for the identification of candidate TFs that may drive core developmental processes at each individual stage of astrocyte maturation.

Using these approaches, we mapped dynamic TF binding and developmental gene programs across normal human astrocyte maturation and investigated how these networks are disrupted in GBM astrocytes. We first show that human astrocytes undergo multiple maturation stages and highlight specific TFs and gene sets unique to each developmental epoch. When projecting this time course onto GBM astrocyte lineage cells, we identified an intermediate maturation stage where astrocytes from IDH-wildtype (IDHwt) tumors deviate from the typical course of development. This maturation stage represents a previously undefined developmental state where astrocyte progenitors transition into a non-proliferative phase. We also provide evidence for how astrocytes from IDH1-mutant (IDH1mt) tumors bypass this developmental roadblock to reach more mature stages via disrupted IDH1-dependent DNA hydroxymethylation of maturation genes. This study offers a new window into an elusive neurodevelopmental process. Clarifying this trajectory identified specific regulatory mechanisms that are perturbed in GBM astrocytes and provides new opportunities for guiding experimental and therapeutic approaches targeting astrocyte maturation.

2.3 Results

2.3.1 *A molecular map of human astrocyte maturation in cortical organoids*

To temporally define the molecular changes across normal astrocyte maturation, we formed human cortical organoids (hCOs) from one male and one female hiPSC line using previously established methods (218). We sampled astrocytes across ten time points spanning the earliest onset of astrogenesis in hCO cultures (80 days), when astrocytes reflect a proliferative immature signature, through 550 days in culture, a timepoint that we previously demonstrated well-reflected quiescent adult human astrocytes (230) (**Fig. 5a**). Astrocytes within organoid sections across these developmental stages exhibited evidence of accumulating morphological complexity (**Fig. 5b**). At each time point, we purified hCO-derived astrocytes using a modified version of a previously established immunopanning protocol (77) and examined the changes of transcription factor binding and gene expression genome-wide, using ATAC- and RNA-seq, respectively (**Fig. 6; Table S1a**).

We observed two distinct waves of astrocyte maturation, where astrocytes appear to cluster in three groups that we define as “early” (d80-d150), “middle” (d200-d350), and “late” (d400-d550) developmental stages (**Fig. 5c and 5d**). To delineate the transcriptional and chromatin changes that define these maturation stages we performed a weighted gene co-expression network analysis (WGCNA) on the RNA-seq and ATAC-seq data across all 10 time points. This revealed transcriptional and chromatin accessibility peak sets that uniquely define five modules — “early”, “middle”, and “late”, as well as transitional “early/middle” and “middle/late” signatures that are maintained throughout two of the three maturation stages (**Fig. 5e; Tables S1b and**

S1c). Early and late developmental stages included markers associated with proliferative immature (MKI67, TOP2A, ID3, NES) and quiescent mature (GJA1, AQP4, ALDH1L1) astrocytes (30, 77), respectively, as well as thousands of largely unexplored contributors to the astrocyte maturation process (**Fig. 5f; Table S1b**). Interestingly, the “middle” stage demonstrated a unique enrichment of key fate-determining transcription factors including SOX9, OLIG1/2, LHX2, POU3F3 (BRN2), EMX1, and SOX2 (**Table S1b**).

When paired together, ATAC-seq and RNA-seq datasets are complementary tools for predicting transcription factors (TFs) that drive transcriptional changes and entire lineage transitions (231, 272). Thus, we sought to integrate our ATAC- and RNA-seq data to predict candidate TF regulators of human astrocyte maturation. We utilized PECA2 (282, 283), an algorithm to (1) score the potential of all TFs to regulate putative target genes (TGs), (2) assess how these TF-TG regulatory networks change across astrocyte maturation, and (3) identify top TF hits that are most likely to regulate our module target genes (**Fig. 5g-1**). After binning our maturation samples into “early”, “middle”, and “late” groups, we used PECA2 to compare TF-TG network differences across our developmental groups and identify candidate TF drivers of astrocyte maturation at each stage (**Fig. 5g**).

We identified 29 candidate TFs (**Table S1d**) that demonstrate temporally graded changes in motif enrichment, occurrence (**Fig. 5m, 7, and 8**), and expression (**Fig. 5m**). Notably, while subsets of these TFs exhibited highly correlated gene expression and motif accessibility, the ATAC- and RNA-seq measurements did not always change

congruently, a discordance documented by other groups that could result from trans-acting influences or regulation of RNA polymerase (284, 285). Our candidate astrocyte maturation TFs included several previously linked to astrogenesis (LHX2, NHLH1, PRRX1, ASCL1) and gliomagenesis (RFX4, POU3F4, MYCN, PAX3), as well as many TFs not yet associated with astrocyte development. Interestingly, some of these TFs share overlapping roles with neuron and oligodendrocyte development (EOMES, SOX21, OLIG1/2, SOX8) (30, 231, 272, 286-290). We next asked which target genes these TFs are predicted to modulate. For each candidate TF, we identified the top target genes (based on a *trans*-regulation score) and quantified their distribution across our five previously defined maturation modules. Reassuringly, TFs active at early time points (d80-150) exhibited regulatory potential directed predominantly towards “early” module TGs, while TFs active in middle (d200-350) and late (d400-550) developmental stages were enriched for TGs belonging to more mature modules (**Fig. 5n**).

2.3.2 *Molecular differences between IDHwt tumor and margin astrocyte lineage cells*

Similar to normal neurodevelopment, astrocytes within GBM tumors also exist on a differentiation spectrum (145). Therefore, we next asked what molecular and genomic features distinguish malignant astrocyte lineage cells from parenchymal astrocytes in surrounding margin tissue. Using fresh primary GBM tissue samples, we performed bulk ATAC- and RNA-seq on purified astrocytes acquired from 13 IDHwt (grade IV GBM) tumor core and 11 paired non-contrast enhancing margin tissues (resected upon surgical approach) (**Fig. 9a and 10a-c; Table S2a**). For each sample, we immunopanned to enrich for astrocyte lineage cells (**Fig. 10d**), as previous studies have

demonstrated that antibodies directed against HepaCAM can capture both normal and neoplastic astrocytes (77, 291, 292). We observed clear chromatin- and transcriptome-level separation between tumor core and margin astrocyte populations (**Fig. 9b and 9c**), a partition that remained when we exclusively assessed expression of a 200-gene panel of previously described immature and mature human astrocyte markers (**Fig. 11c; Table S2b**) (77).

At the chromatin level, IDHwt tumor astrocytes and margin astrocytes each exhibited >2000 unique peaks ($FDR < 0.05$, $|\log_2FC| > 2$), with the majority present in non-coding regions (**Fig. 11a and 11b; Table S2c**). We found that differential peak accessibility appeared to be closely associated with the degree of sample stemness, which we calculated using the normalized cumulative expression of previously defined stem-like markers (**Fig. 9d, Table S2d**). At the transcriptomic level, differential expression analysis ($FDR < 0.05$, $|\log_2FC| > 2$) between IDHwt tumor and margin astrocytes revealed a significant overrepresentation of immature astrocyte genes upregulated in tumor astrocyte lineage cells ($p = 1 \times 10^{-12}$). This is in contrast to margin astrocytes, which showed a clear upregulation of mature astrocyte genes ($p = 7.7 \times 10^{-14}$) (**Fig. 9e-f, Table S2e**). Together, these data suggest that the chromatin and transcriptomic differences between IDHwt tumor and margin astrocytes are largely rooted in developmental processes.

While bulk sequencing of purified astrocytes provides the read depth to investigate even subtle molecular differences, an important caveat is a possibility that immunopanning might not capture all astrocyte lineage cells. To account for this, we performed single-

nucleus multiome (snRNA-seq and snATAC-seq) sequencing on a subset of the flash-frozen tissue samples from the above sample collection (**Fig. 9g**). After quality control and filtering, we obtained 27,036 total nuclei from three tumor and two margin tissue samples (**Fig. 13b; Table S3a**). We annotated clusters using gene activity scores of known cell type-specific markers (**Table S3c**) and identified all primary CNS cell types, including microglia, endothelial cells, oligodendrocytes, inhibitory and excitatory neurons, and astrocytes (**Fig. 9h and 13a; Table S3d**). Astrocyte lineage populations were defined by high gene activity scores for a combination of pre-defined astrocyte markers (**Table S3e**) and we focused on these clusters for downstream analyses (**Fig. 9h and 9i**). Notably, several clusters exhibited signatures of multiple glial and progenitor cell lineages (**Fig. 13d**), a testament to the cellular complexity of GBM tumors. Next, we classified cells as either neoplastic or non-neoplastic based on a combination of factors, including tissue source, transcriptomic marker expression, and CNV enrichment (**Fig. 9i, 13c, and 13d**). While neoplastic and non-neoplastic astrocytes largely segregated by whether they were collected from tumor core or margin resections, we observed sporadic exceptions to this trend, likely indicative of infiltrating tumor cells within the margin samples. When examining the gene activity score for immature and mature human astrocyte genes (**Table S2b**), we again observed a clear enrichment of mature astrocyte genes in non-neoplastic astrocytes and an enrichment of the immature astrocyte gene set in neoplastic cells (**Fig. 9j**). However, this panel of immature and mature astrocyte genes represents the extreme limits of the developmental spectrum and we hypothesized that our high-resolution astrocyte trajectory would provide a more specific indication of where neoplastic astrocyte lineage cells align or deviate from normal development.

2.3.3 *Projecting GBM astrocytes onto a normal human astrocyte maturation trajectory*

We next projected our GBM astrocyte data onto our newly defined human astrocyte maturation timeline (**Fig. 12a**). We began by correlating the expression of maturation module genes in our bulk IDHwt tumor core and margin astrocyte data. This analysis indicated that tumor astrocytes predominantly share a molecular signature with astrocytes from early-middle time points, unlike margin astrocytes, which more closely resemble hCO astrocytes from middle-late time points (**Fig. 11d and 12b**).

To further investigate this developmental bifurcation, we examined the enrichment of each maturation module (**Table S3e**) across our single nuclei astrocyte lineage cell clusters. This analysis again demonstrated a profound enrichment for “early”, “early/middle”, and “middle” gene modules within neoplastic astrocyte lineage cells (**Fig. 12c**). Conversely, non-neoplastic cells predominantly exhibited a clear enrichment of “middle/late” and “late” maturation modules, consistent with a more mature identity (**Fig. 12c**). Patient-to-patient heterogeneity is a defining hallmark of GBM, so we expanded this analysis to see if these findings held true across a larger and more diverse cohort. We analyzed sc-RNA-seq data collected from 110 separate individuals (>300k cells) with IDHwt GBM tumors (293). As we observed in the bulk and single nucleus data, we again saw a strong enrichment of the “middle” signature in the majority of neoplastic astrocyte lineage cells (**Fig. 12d**). The “early/middle” transition signature was specifically restricted to highly proliferative neoplastic astrocyte lineage cells, and the “middle/late” and “late” signatures were limited almost

exclusively to non-neoplastic parenchymal astrocytes (**Fig. 12d**). To investigate the spatial organization of our maturation gene signatures, we overlaid our modules onto published spatial transcriptomics datasets (294). There, we found distinct spatial arrangements, with “early” and “middle” genes enriched in the tumor parenchyma, including proliferative pseudopalisading regions. Conversely, the “middle/late” gene signature was more apparent in areas separating the tumor and infiltrative regions (**Fig. 12e**).

Astrocyte maturation TF candidates also displayed unique enrichment within GBM astrocyte lineage cells. We calculated motif enrichment for our 29 PECA-predicted TF candidates (**Table S1d**) across single nuclei from astrocyte lineage cells and found several TFs exhibited biased motif accessibility within neoplastic cells — POU3F2, POU3F3, POU3F4, PAX3, OLIG1, OLIG2, HEY1, and MYCN — many of which have been previously implicated in GBM (**Fig. 12f-g, 14a**) (170, 286, 287, 295-301). Other TFs were enriched in specific populations of neoplastic and/or non-neoplastic cells, occasionally from one individual patient (**Fig. 14a**). This may contribute to different patterns of module gene enrichment across clusters.

2.3.4 *Subtype-specific molecular signatures in GBM tumor astrocyte lineage cells*

Over the course of tissue collection, we procured tumor samples harboring a variety of genetic abnormalities. These included IDHwt, IDH1-mutant (IDH1mt), and several recurrent tumors. Our cohort of IDHwt tumors contained various combinations of common GBM-associated aberrations in genes like EGFR, PDGFRA, and PTEN (**Table**

S2a). To account for how these variables may contribute to differences in maturation state, we sought to identify molecular and transcriptomic signatures associated with each feature. To make these statistical comparisons, we binned our samples into six categories for which there was a minimum of at least two samples per group- (1) IDH1mt, (2) EGFR, (2) PTEN, (4) EGFR+PDGFRA, (5) EGFR+PDGFRA+PTEN, (6) recurrent (**Table S4a**).

At the ATAC-seq level, we found 11,134 peaks specific ($FDR < 0.05$) to astrocytes from one or more of our six categories, with 5,730 peaks unique to only one of the six categories (**Fig. 15a and 16a; Table S4b**). Within these specific peak signatures, we identified enrichment of unique sets of TF motifs affiliated with various biological functions, including stem cell maintenance, immune response, and early forebrain development (**Fig. 15b; Table S4c**). We performed a similar analysis with the RNA-seq data and found category-specific gene sets involved in a variety of metabolic, developmental, and homeostatic processes (**Fig. 15c and 16c; Table S4d**). However, when looking at each feature individually (ex: yes/no PTEN), most peak and gene sets we identified were not significant. We suspect this may be due to the fact that most samples contained multiple genomic aberrations, which makes it difficult to assign contributions from a single attribute **Fig. 16a-d, Table S4a**).

By far, the most unique category of astrocytes was the IDH1mt group, which now belongs to a separate tumor classification from GBM(180, 302). These tumors are slow growing, metabolically perturbed, epigenetically distinct, and associated with improved clinical outcomes compared to IDHwt cases (303-308). We observed that astrocytes

from IDH1mt tumors have a distinctive transcriptomic signature, with some samples even bearing a closer resemblance to margin astrocytes than to astrocytes from IDHwt tumors (**Fig. 15d; Table S2f**). A differential expression analysis with IDH1mt and IDHwt tumor astrocytes showed a clear upregulation of “early”, “early/middle”, and “middle” maturation genes in IDHwt tumor astrocytes, suggesting IDHwt astrocytes may exist in a more immature state than IDH1mt astrocytes (**Fig. 15e; Fig. 17a**). To confirm that the astrocytes we purified from IDH1mt tumors were not simply bystander parenchymal cells trapped within the tumor core, we also compared their expression profiles to astrocytes from margin tissue acquired from matched samples (**Table S2g**). This analysis indicated that mature astrocyte genes are more highly expressed in margin samples (**Fig. 17b-d**), suggesting that astrocytes from IDH1mt tumors are also stalled at an intermediate developmental stage, albeit one that is more mature than IDHwt astrocytes. Altogether, our transcriptomic signatures portray a clear maturation spectrum, with IDHwt tumor astrocytes falling closest to the fetal (immature) end of the spectrum, followed by IDH1mt tumor astrocytes, and finally, margin astrocytes, which exhibit a predominantly adult quiescent signature (**Fig. 15f and 15g**).

2.3.5 *The contribution of DNA methylation to IDHwt and IDH1mt maturation differences*

Gliomas harboring IDH1 mutations produce D-2-hydroxyglutarate (D-2-HG) instead of α -ketoglutarate (α -KG) (309), which disrupts fundamental biological processes that rely on α -KG, including ten-eleven translocation (TET) enzyme activity (310, 311). TET enzymes help facilitate DNA demethylation by oxidizing 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) and downstream oxidation steps in active cytosine

demethylation (312-314). Perturbed TET activity has been shown to contribute to the altered DNA methylation landscape in IDH1mt glioma and may be linked to the CpG island methylator phenotype (CIMP), a common pathological feature of IDH1mt glioma (315, 316). Despite these studies, relatively little is known about genome-wide 5hmC and 5mC dysregulation in IDH1mt glioma. Given the implications of these important epigenetic modifications on gene regulation, we mapped genome-wide 5mC and 5hmC patterns across IDH1mt and IDHwt tumors to determine whether differences in these epigenetic features could bias astrocyte maturation in IDH1mt cells.

First, we performed methylated-DNA immunoprecipitation (MeDIP) and hydroxymethylated-DNA selective chemical labeling (hMe-Seal) using flash-frozen tissue from three IDH1mt and three IDHwt tumors, as well as three margin tissue samples as controls to compare genome-wide 5mC and 5hmC patterns (**Fig. 18a; Table S5a**). PCAs showed that the three sample groups exhibit markedly unique global 5mC and 5hmC signatures, which is consistent with existing methylation hybridization panel-based data that is routinely used in tumor diagnosis (180, 317, 318) (**Fig. 18b and 18e**). We also observed specificity within the genomic loci of differentially methylated (read counts > 20, FDR < 0.05, $|\log_{2}FC| > 0$) and hydroxymethylated (read counts > 20, FDR < 0.001, $|\log_{2}FC| > 4$) regions (DMRs and DhMRs) when comparing IDH1mt and IDHwt tumors. The majority of accumulated DMRs in IDH1mt appeared in promoter and exonic regions while DhMRs were in promoter and intronic regions (**Fig. 18c and 18f**). In comparison, the majority of DMRs and DhMRs depleted in IDH1mt were located in intergenic and intronic regions (**Fig. 18c and 18f**). Examples of differential 5hmC can be seen in EGFR, a key regulator of GBM pathogenesis, with

markedly higher 5hmC levels in IDHwt, and in *NTRK2*, a mature astrocyte gene, where 5hmC levels were higher in IDH1mt samples (**Fig. 18d**). We further validated our DhMR hits by performing 5hmC-Capture-qPCR for 10 relevant developmental genes that exhibit differential 5hmC levels between IDH1mt and IDHwt tumors. All of these capture-qPCR targets confirmed our genome-wide data (**Fig. 19a and 19b**).

We speculated that if 5mC or 5hmC was involved in the astrocyte maturation process, their genomic distributions would show biased patterns in maturation genes across IDH1mt and IDHwt tumors. We began by filtering our datasets for immature astrocyte genes upregulated in astrocytes from IDHwt tumors ($p=4.8 \times 10^{-62}$), and mature genes ($p=7.3 \times 10^{-19}$) upregulated in IDH1mt tumor astrocytes (**Fig. 18g, Table S5c**). We examined 5mC and 5hmC levels across these genes and observed a striking pattern, where gene body 5hmC levels showed a directional relationship to the expression of these maturation genes. Mature astrocyte genes upregulated in IDH1mt tumor cells exhibited significantly higher ($p < 0.001$) intragenic 5hmC levels than what we observed in IDHwt samples (**Fig. 18h and 18j**). Likewise, immature astrocyte genes upregulated in IDHwt tumor cells also demonstrated an accumulation of 5hmC that was higher in the IDHwt samples (**Fig. 18i and 18k**). In contrast, 5mC levels did not change with the expression of maturation gene sets (**Fig. 19c**). We observed a similar link between gene body 5hmC and expression when we performed these analyses across each of our astrocyte maturation genes modules, suggesting 5hmC may play a role in helping regulate these genes as well (**Fig. 19e**). Importantly, this relationship between 5hmC accumulation and gene expression was not consistent across all differentially expressed genes (**Fig. 19d-f**), suggesting that these 5hmC biases are unlikely to reflect a universal

association between transcription and intragenic hydroxymethylation. Together, these results suggest that 5hmC distribution resulting from the epigenetic consequences of IDH1mt activity may be an important factor contributing to the maturation differences between IDH1mt and IDHwt tumors.

Figure 5: A molecular trajectory of human astrocyte maturation

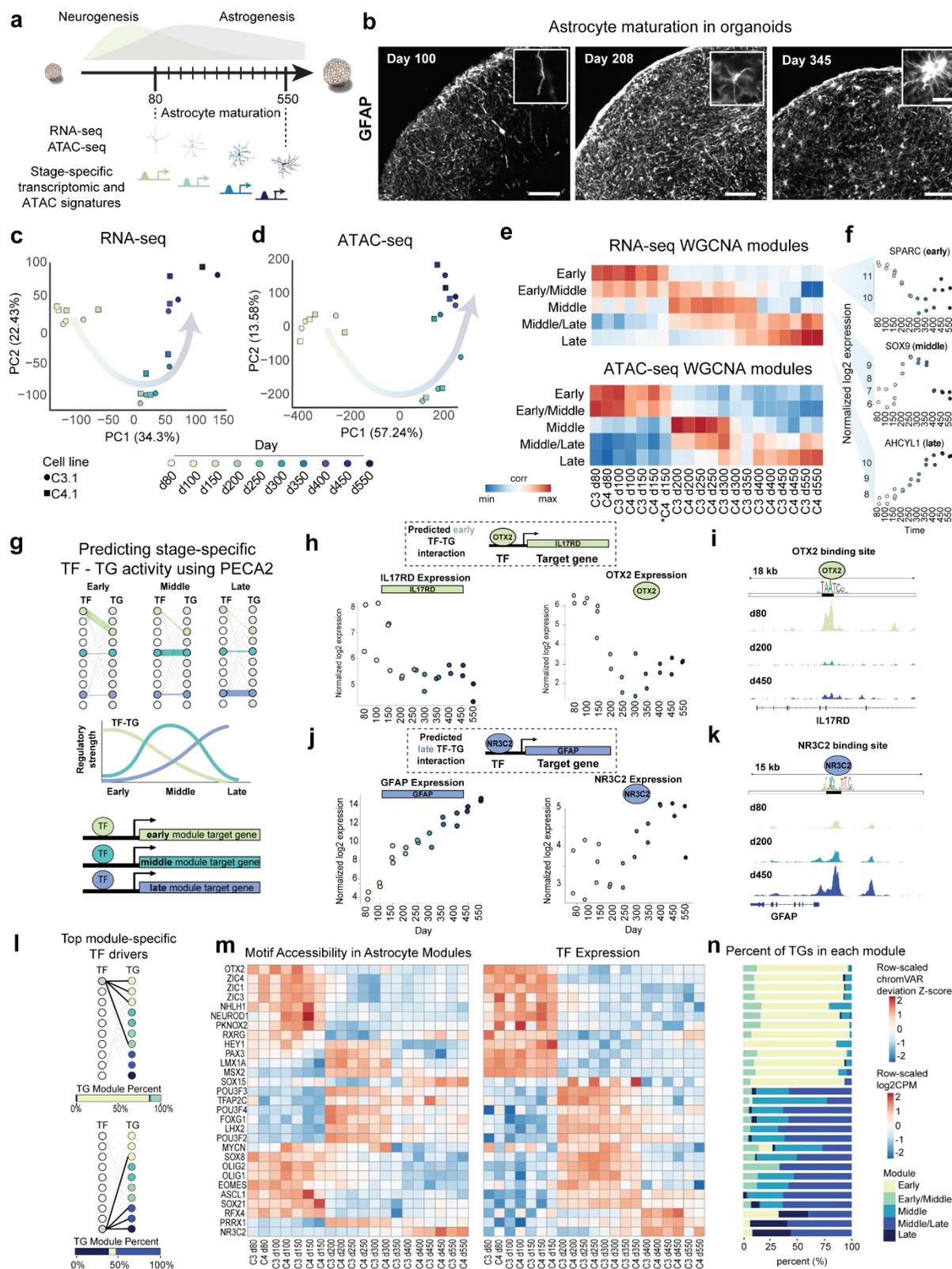


Fig. 5: A molecular trajectory of human astrocyte maturation

(a) Schematic of astrocyte maturation collection time points. (b)

Immunohistochemistry of GFAP in hCOs across three time points- d100, d208, and d345. Scale bar for large images: 100um. Scale bar for insets: 20um. (c) PCAs of RNA-

seq and (d) ATAC-seq across time points. (e) WGCNA heatmaps across maturation time points for gene expression (top) and ATAC peaks (bottom). Heatmap values are module

Eigengenes. (f) Expression of representative genes from Early, Middle, and Late gene

modules. Colors correspond to time points. (g) Schematic depicting the use of the PECA algorithm to predict transcription factor (TF) – target gene (TG) interactions during

early, middle, and late stages of astrocyte maturation. (h-k) Example validation of

PECA-predicted TF-TG networks. (h, j) Expression of TGs (IL17RD and GFAP) and TFs

(OTX2 and NR3C2) across maturation time points. (i, k) ATAC signal in regulatory

regions. TF (OTX2 and NR3C2) binding sites are located at thin black line. ATAC signal

displayed across early (C4 d80), middle (C4 d200), and late (C4 d450) time points. (l)

Approach to filter for TFs that regulate astrocyte maturation. (m) Accessibility (left) and

gene expression (right) of candidate TF drivers of astrocyte maturation. (n) Percent of

TGs in each maturation gene module for all candidate TFs.

Figure 6: Quality control of hCO astrocyte libraries

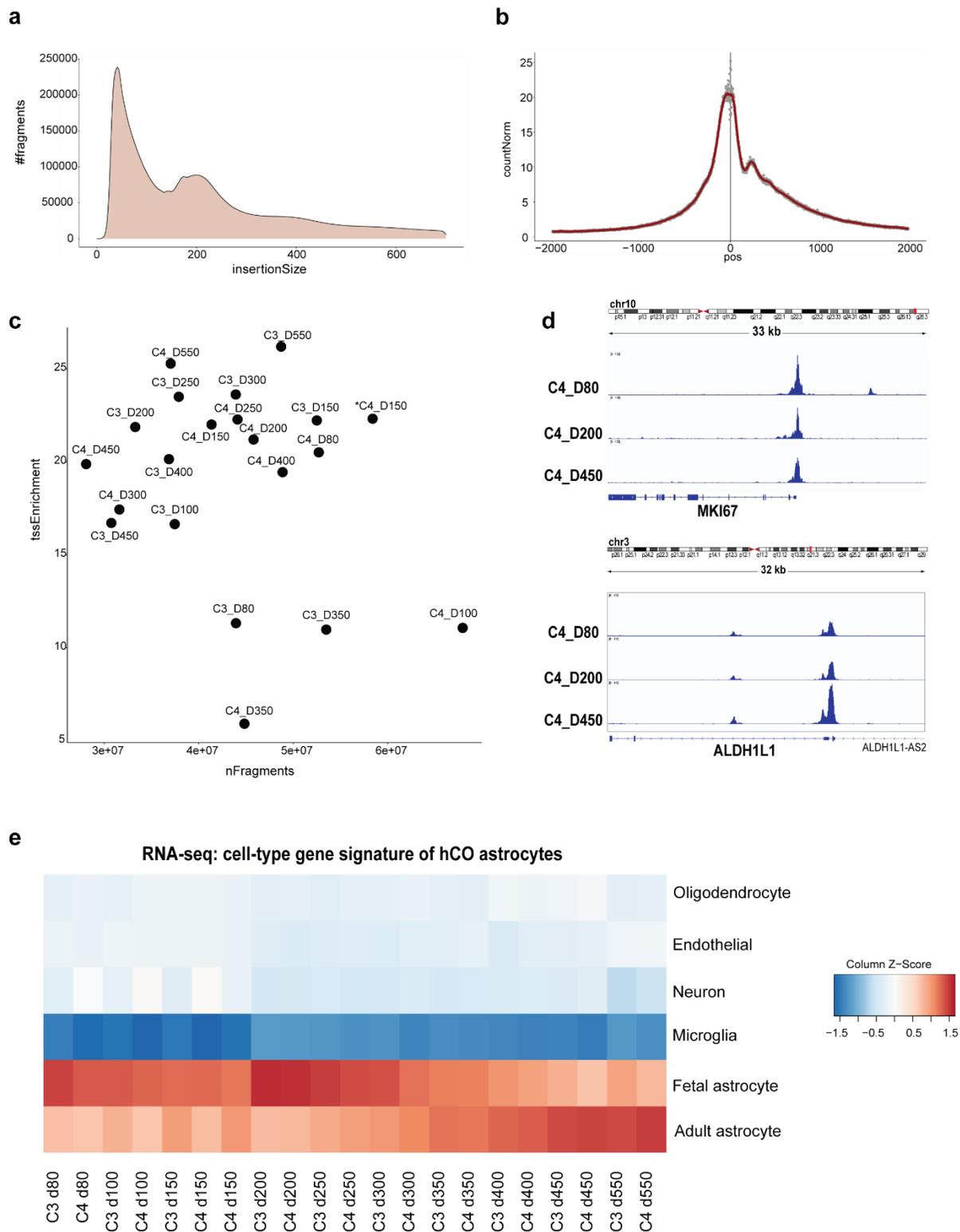


Fig. 6: Quality control of hCO astrocyte libraries

(**a, b**) Representative fragment size distribution (**a**) and line plot of centered average ATAC signal at 2 kb surrounding TSS (**b**). (**c**) Scatter plot of TSS enrichment scores for all hCO astrocyte libraries. (**d**) Examples of called ATAC peaks in biologically relevant regions. (**e**) Spearman correlation of cell type-specific gene signatures (Table S1e) using previously published data (77) with hCO astrocytes across maturation time points.

Figure 7: TF motif enrichment data in hCO-derived astrocytes

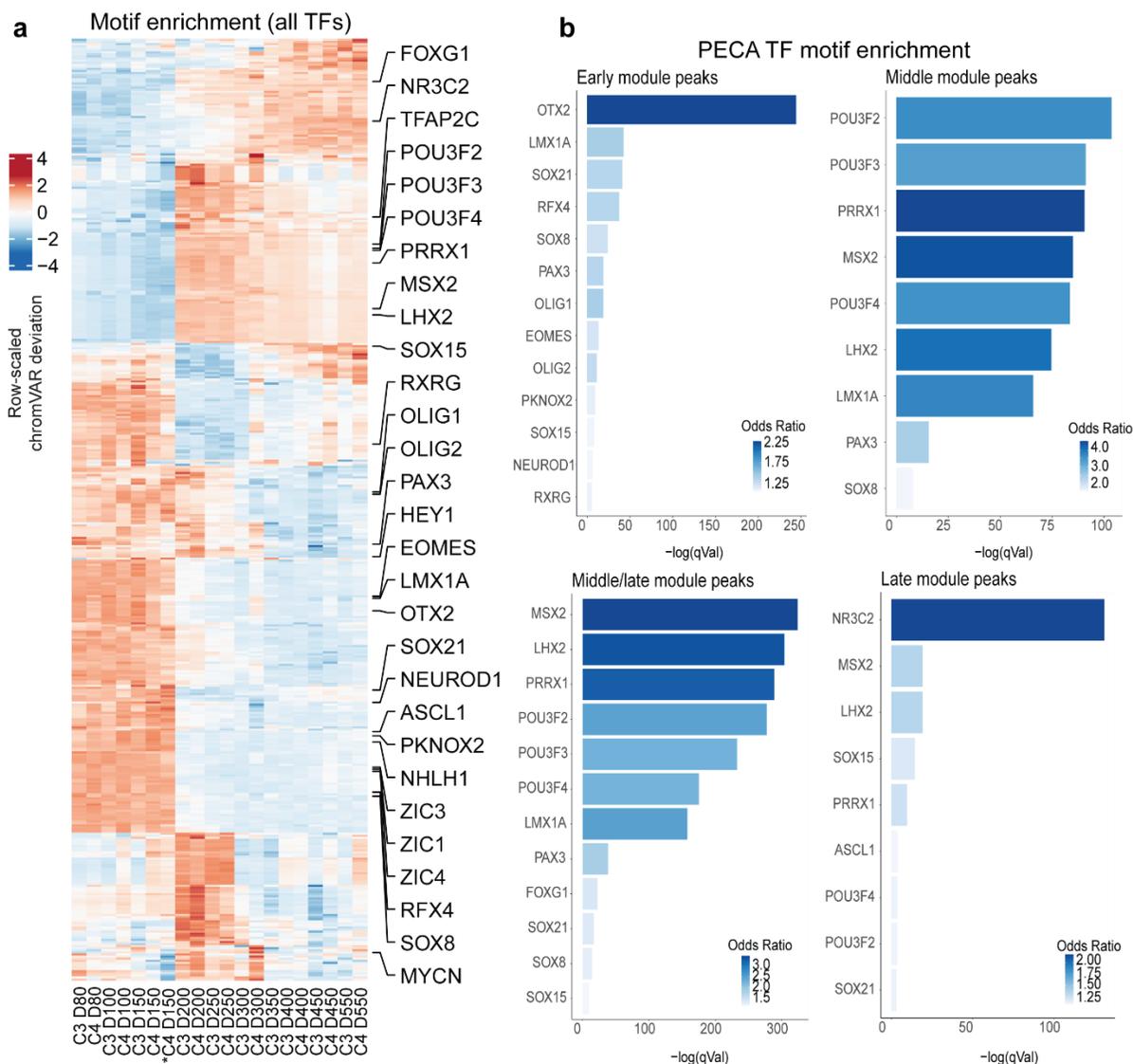


Fig. 7: TF motif enrichment data in hCO-derived astrocytes

(a) Motif enrichment in WGCNA maturation peaks (Table S1c). Annotated TFs are PECA-predicted maturation TFs. (b) Bar plots showing enrichment of maturation TFs in ATAC-seq module peaks (Table S1c). Enrichment in “early/middle” peak set was not shown because TFs were not significantly enriched. Continuous color scale indicates the odds ratio.

Figure 8: TF motif occurrence in hCO astrocytes

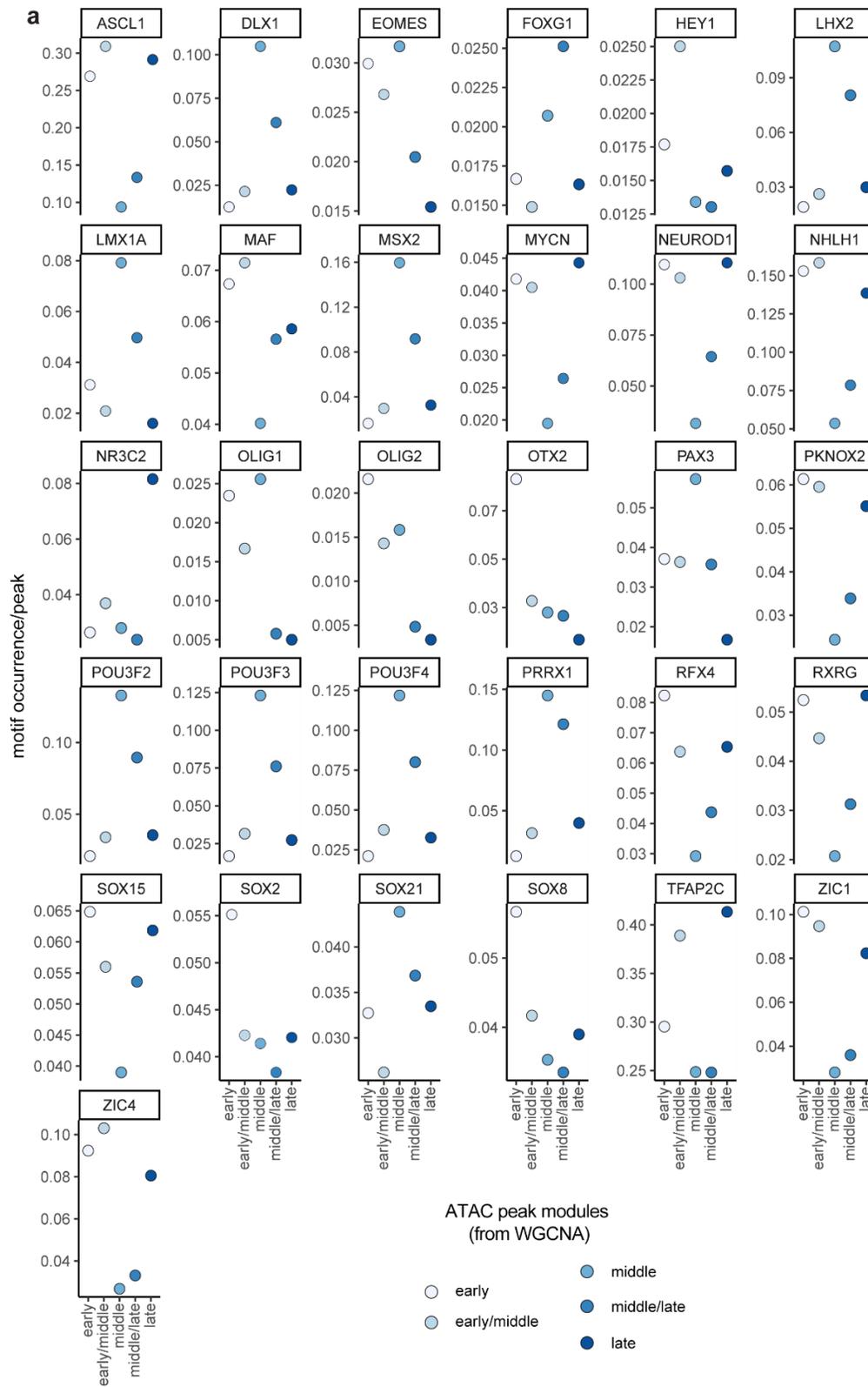


Fig. 8: TF motif occurrence in hCO astrocytes

(a) Dot plots showing motif occurrence for PECA-predicted TFs in ATAC-seq module peaks (Table S1c). Motif occurrence is normalized to number of peaks in each module. Dot color corresponds to the maturation peak module.

Figure 9: Diverging molecular profiles of GBM tumor and margin astrocyte lineage cells

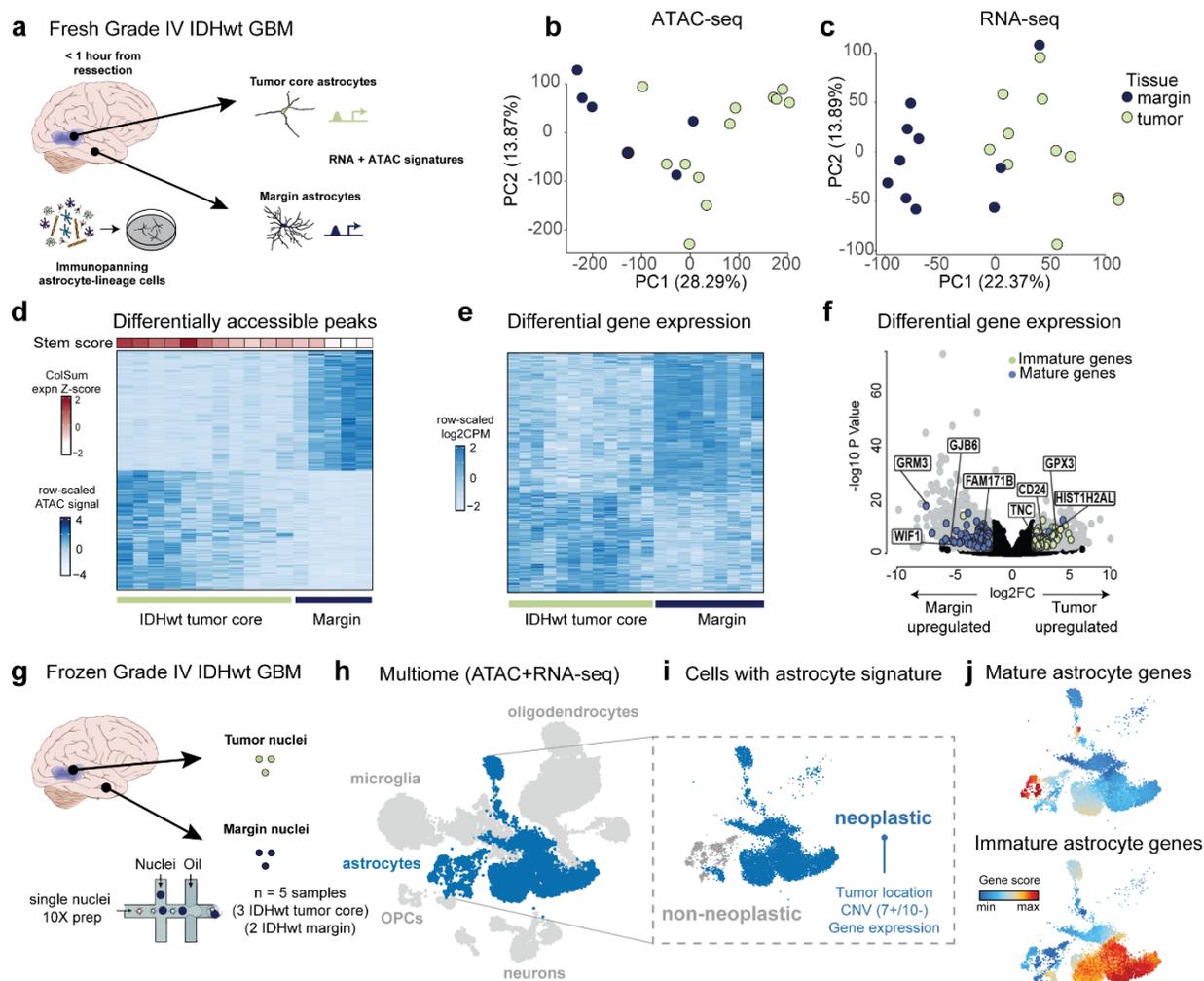


Fig. 9: Diverging molecular profiles of GBM tumor and margin astrocyte lineage cells

(a) Schematic showing collection of IDHwt tumor and margin astrocytes for joint ATAC-seq and RNA-seq profiling. (b) PCAs of ATAC-seq and (c) RNA-seq data from IDHwt tumor (green) and margin (purple) astrocyte-enriched samples. (d) Differentially accessible peaks between IDHwt tumor and margin astrocyte lineage cells ($FDR < 0.05$ and $|\log_2FC| > 1$). Columns are annotated with a stemness score derived from the scaled and summed expression of common stem markers (Table S2d). (e) Heatmap showing differential gene expression between margin ($FDR < 0.05$ and $\log_2FC < -2$) and tumor ($FDR < 0.05$ and $\log_2FC > 2$) astrocyte lineage cells. (f) Volcano plot showing DEGs between margin ($FDR < 0.05$ and $\log_2FC < -2$) and tumor ($FDR < 0.05$ and $\log_2FC > 2$) astrocyte lineage cells (gray). Mature astrocyte markers are colored purple and fetal astrocyte markers are green. (g) Schematic of collection of IDHwt tumor and margin nuclei from frozen primary tissue samples for single-nucleus multiome (ATAC-seq and RNA-seq) profiling. (h) UMAP of single nuclei (gene activity; combined RNA and ATAC). Astrocyte-lineage clusters are in blue (see Fig. 13a). (i) Non-neoplastic astrocyte-lineage cells in gray and neoplastic cells in blue, as defined by gene expression and CNV identity (see Fig. 13). (j) Combined gene activity score of mature (top) and immature (bottom) astrocyte gene signatures within astrocyte lineage clusters.

Figure 10: GBM astrocyte lineage ATAC-seq library quality control

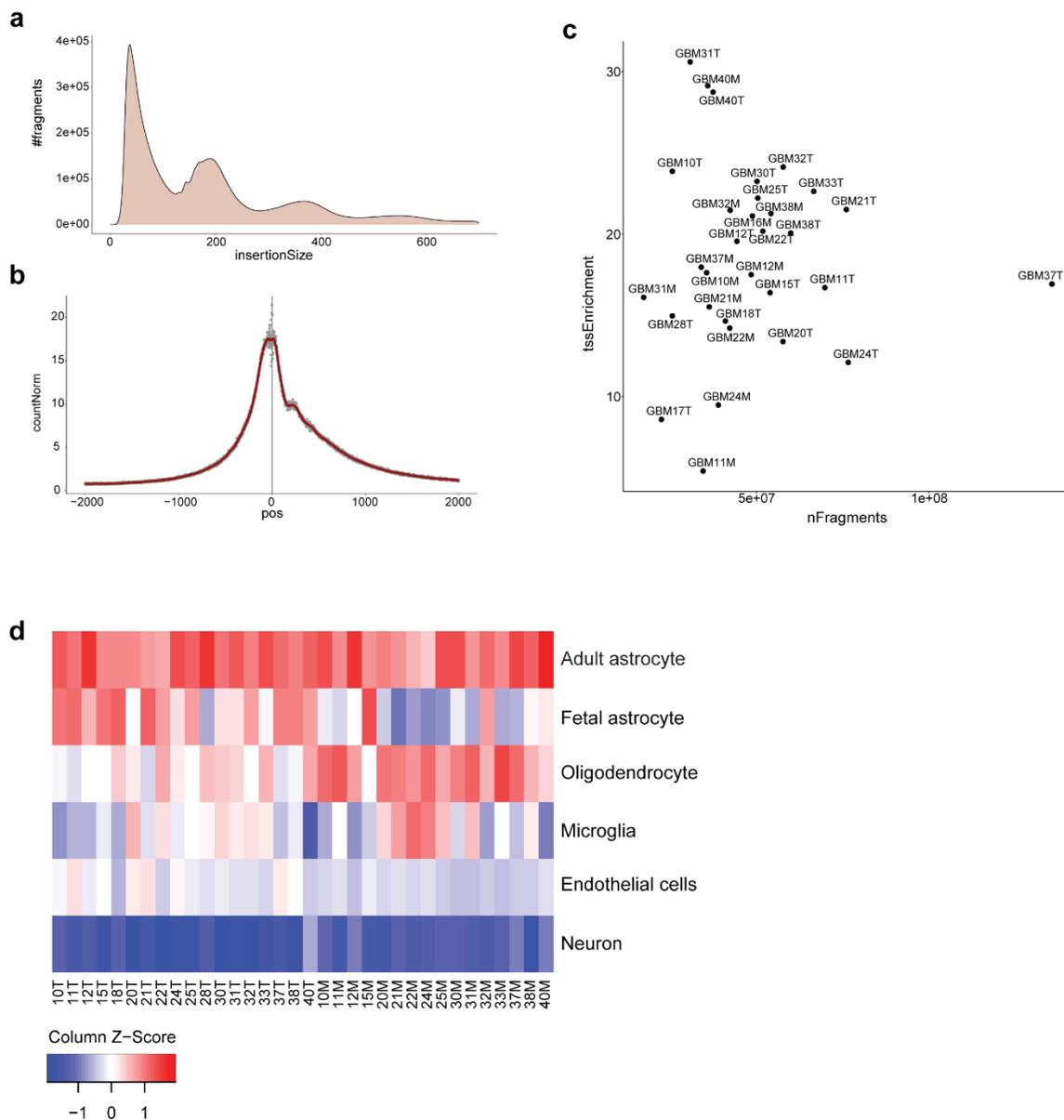


Fig. 10: GBM astrocyte lineage ATAC-seq library quality control

(a, b) Representative fragment size distribution and line plot of centered average ATAC signal at 2 kb surrounding TSS for GBM astrocyte lineage ATAC-seq libraries. **(c)** Scatter plot of TSS enrichment scores for all GBM astrocyte samples. **(d)** Spearman correlation of cell type-specific gene signatures (Table S1e) using previously published data (77) with GBM astrocyte lineage samples.

Figure 11: Molecular differences between IDHwt tumor and margin astrocytes

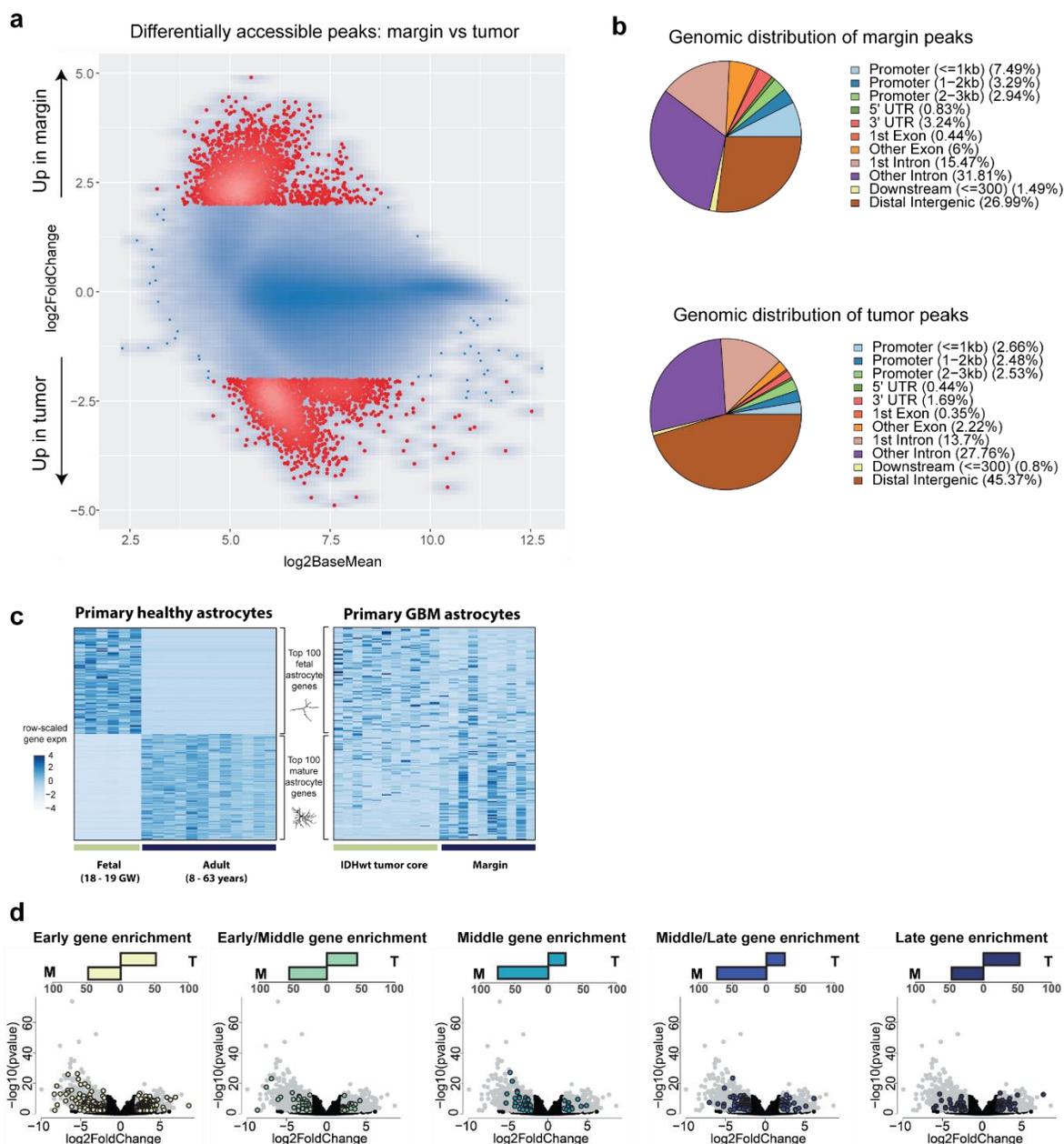


Fig. 11: Molecular differences between IDHwt tumor and margin astrocytes

(a) MA plot showing differentially accessible peaks between IDHwt margin (FDR<0.05; $\log_2FC>2$) and tumor (FDR<0.05; $\log_2FC<-2$) astrocyte lineage cells. (b) Genomic region annotations for IDHwt margin peaks (top) and IDHwt tumor peaks (bottom). (c, Left) Expression of top 100 fetal and mature astrocyte genes using data from Zhang et al. 2015 (77) (Table S2b). (c, Right) Expression of the same 200 genes in IDHwt tumor and margin astrocytes. (d) Volcano plot showing overlap of maturation module genes with DEGs between IDHwt tumor and margin astrocyte lineage cells. DEGs (FDR<0.05 and $|\log_2FC|>2$) are shown in gray and maturation genes colored by stage (see key). Percent of DEGs in respective modules quantified via diverging bar plots directly above corresponding volcanoes.

Figure 12: GBM astrocyte lineage cells deviate from the normal developmental trajectory at middle stages

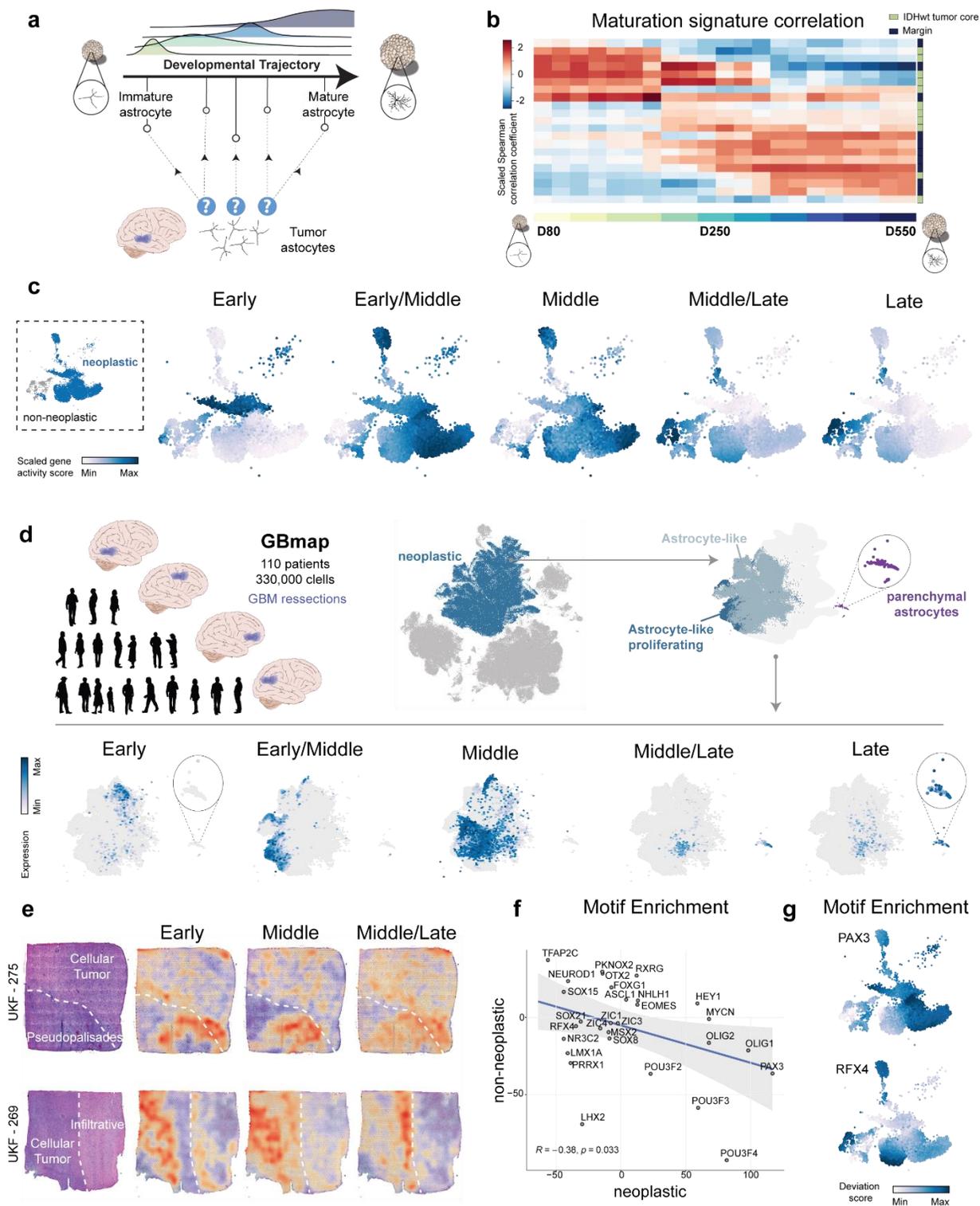


Fig. 12: GBM astrocyte lineage cells deviate from the normal developmental trajectory at middle stages

(a) Approach for projecting astrocyte maturation trajectory onto IDHwt tumor and margin astrocyte lineage cells. (b) Spearman correlation of maturation signature between IDHwt tumor and margin astrocyte lineage cells and hCO-derived astrocytes at each time point. (c) UMAPs of scaled gene activity for the top 50 genes (see Table S3f) within each maturation module across astrocyte lineage clusters. (d, Top) GBmap scRNA-seq data (293) annotated for neoplastic astrocyte-like (light blue), neoplastic astrocyte-like proliferating (dark blue), and non-neoplastic parenchymal astrocyte (purple) clusters. (d, Bottom) Scaled expression of top 100 genes (see Table S3f) in each maturation module across each GBmap astrocyte lineage cluster. (e) Spatial transcriptomic surface plots showing enrichment of “early”, “middle”, and “middle/late” maturation gene modules in (left) histologically defined regions (294). (f) Normalized motif enrichment of PECA maturation TFs in neoplastic and non-neoplastic astrocyte lineage clusters. (g) Scaled motif deviation scores for PAX3 (top) and RFX4 (bottom) across astrocyte lineage clusters.

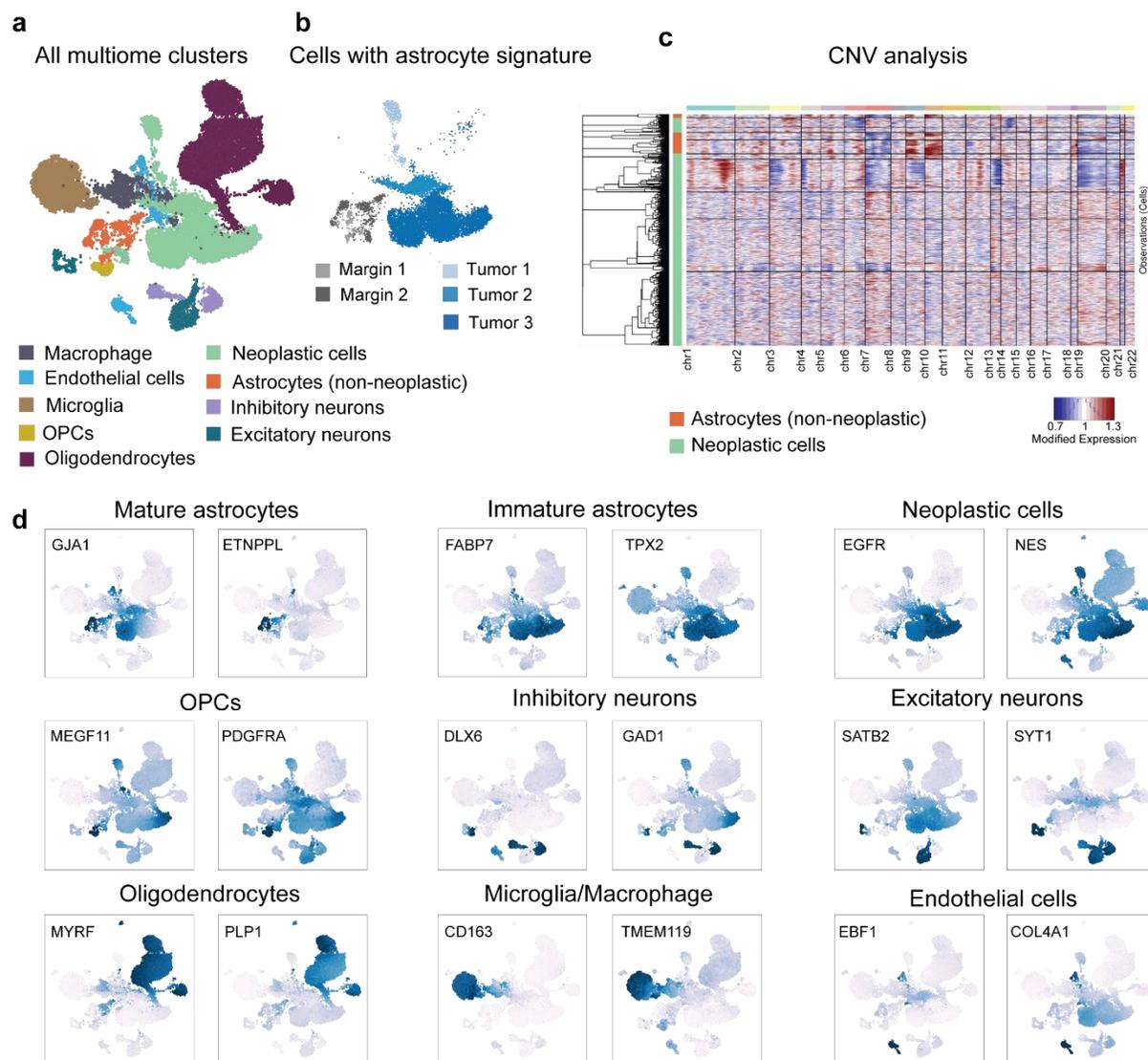
Figure 13: sn-multiome quality control and cell-type identification

Fig. 13: sn-multiome quality control and cell-type identification

(a) UMAP of GBM single nuclei with clusters colored by cell type. (b) UMAP of astrocyte lineage clusters annotated by tissue source. Cells from margin tissue are gray and cells from tumor tissue are blue. (c) CNV analysis heatmap showing expression of genes on each chromosome (column annotation) for all neoplastic (green) and non-neoplastic (orange) astrocyte lineage cells (row annotation). (d) UMAPs showing gene activity score for cell type-specific markers.

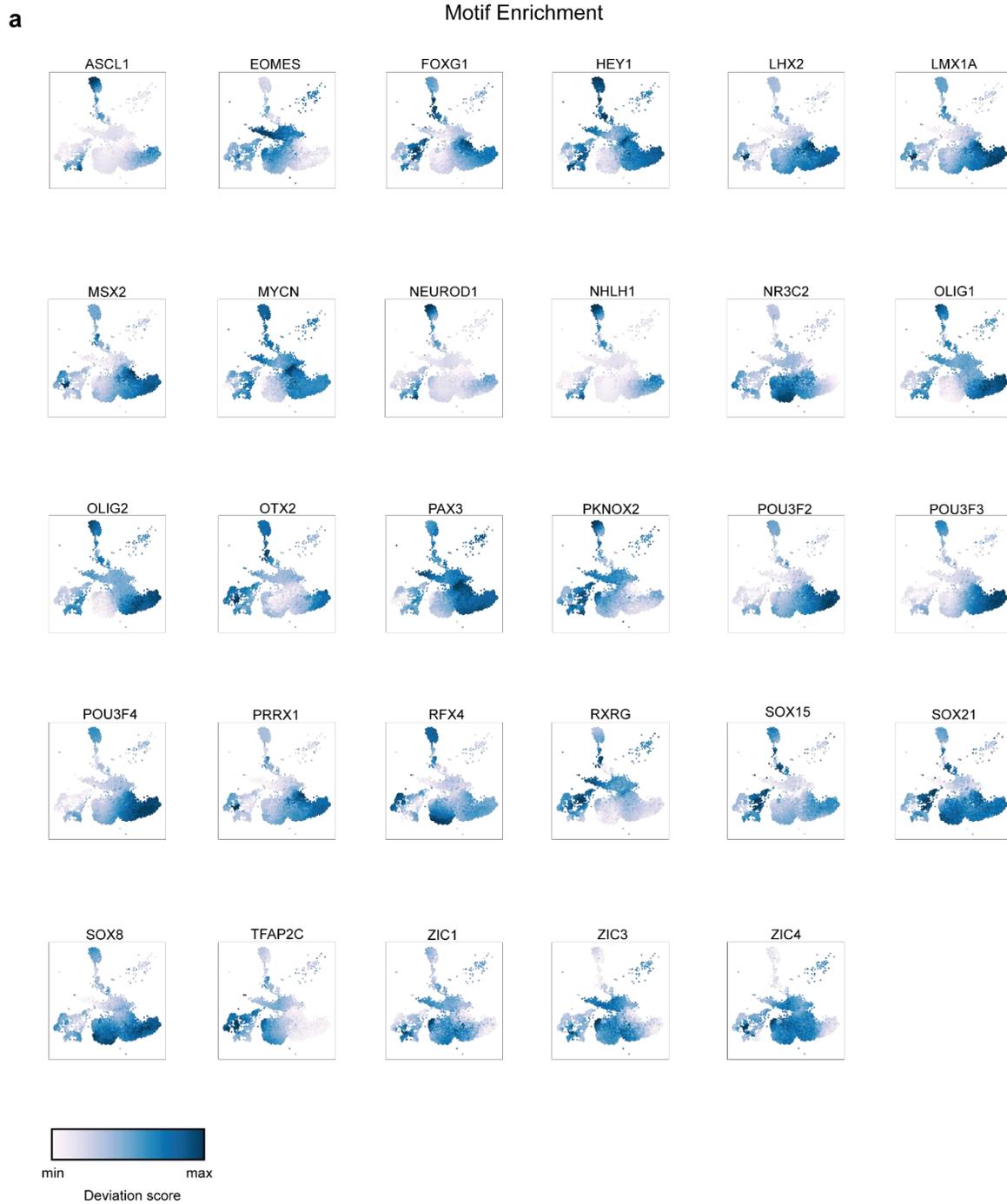
Figure 14: Measurements of maturation TFs in GBM

Fig. 14: Measurements of maturation TFs in GBM

(a) UMAPs of motif deviation scores for all PECA-predicted maturation TFs in GBM astrocyte lineage cell clusters (in addition to Fig. 5g).

Figure 15: Subtype-specific molecular signatures in tumor astrocyte lineage cells

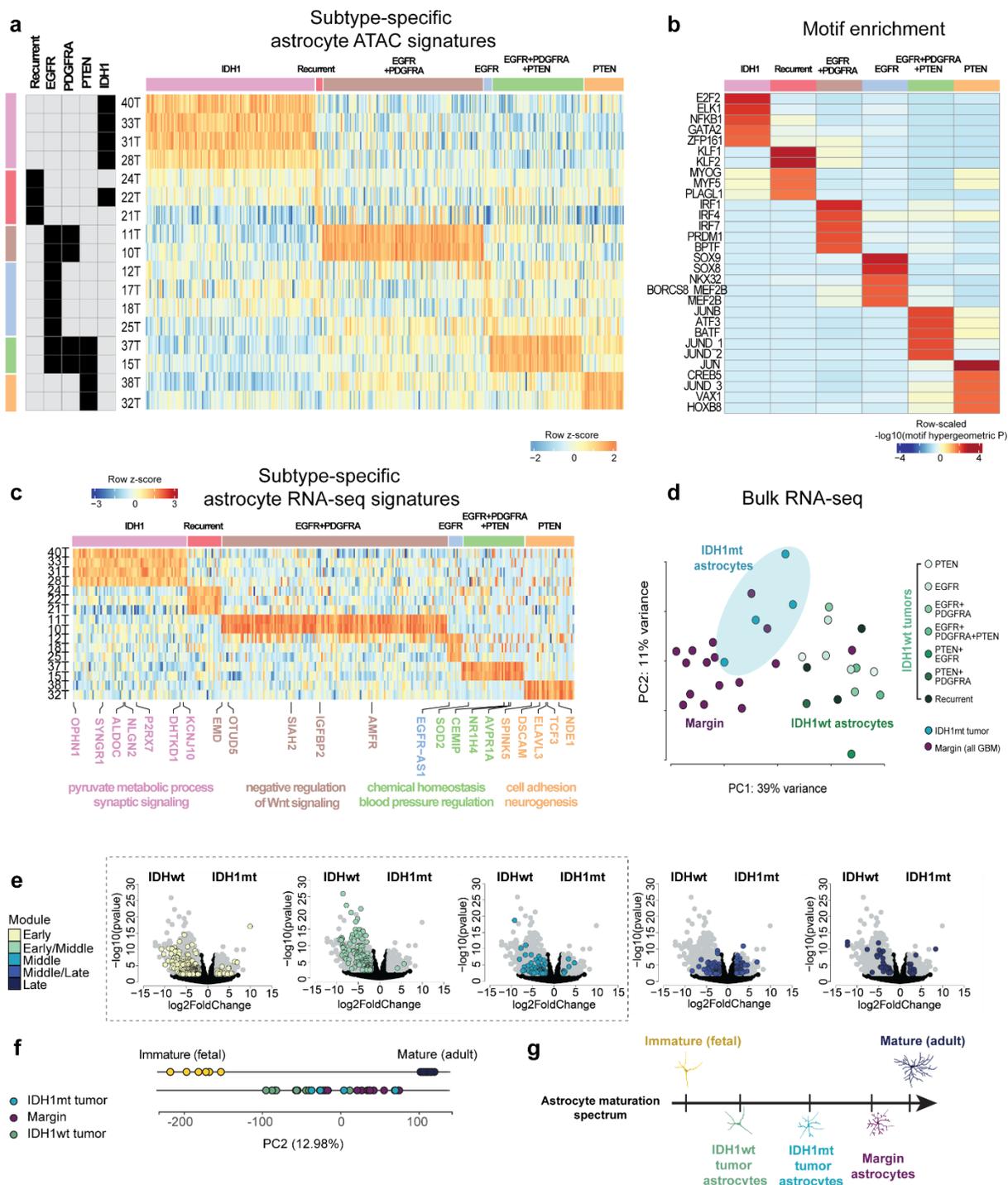


Fig. 15: Subtype-specific molecular signatures in tumor astrocyte lineage cells

(a-c) Heatmaps of ATAC and gene signatures unique to astrocytes from tumors with various genomic diagnoses. (b) Top 5 TF motifs enriched in each categorical peak set. (c) RNA-seq heatmap annotated with representative genes and enriched gene ontology terms. (d) PCA of RNA-seq data from IDHwt, IDH1mt, and margin astrocyte samples. (e) Volcano plots showing the overlay of maturation module genes with DEGs between IDHwt and IDH1mt ($FDR < 0.05$ and $|\log_2FC| > 2$) astrocyte lineage cells. DEGs in gray and colored based on maturation state. (f) (Top line) PC2 for RNA-seq data from immature (fetal, primary tissue) and mature (adult, primary tissue). (Bottom line) IDHwt tumor, IDH1mt tumor, and all GBM margin astrocyte lineage cells projected on the same PC. (g) Schematic of proposed molecular maturation spectrum across all tumor astrocyte lineage cells.

Figure 16: Additional genomic diagnosis-specific ATAC- and RNA-seq tumor astrocyte signatures

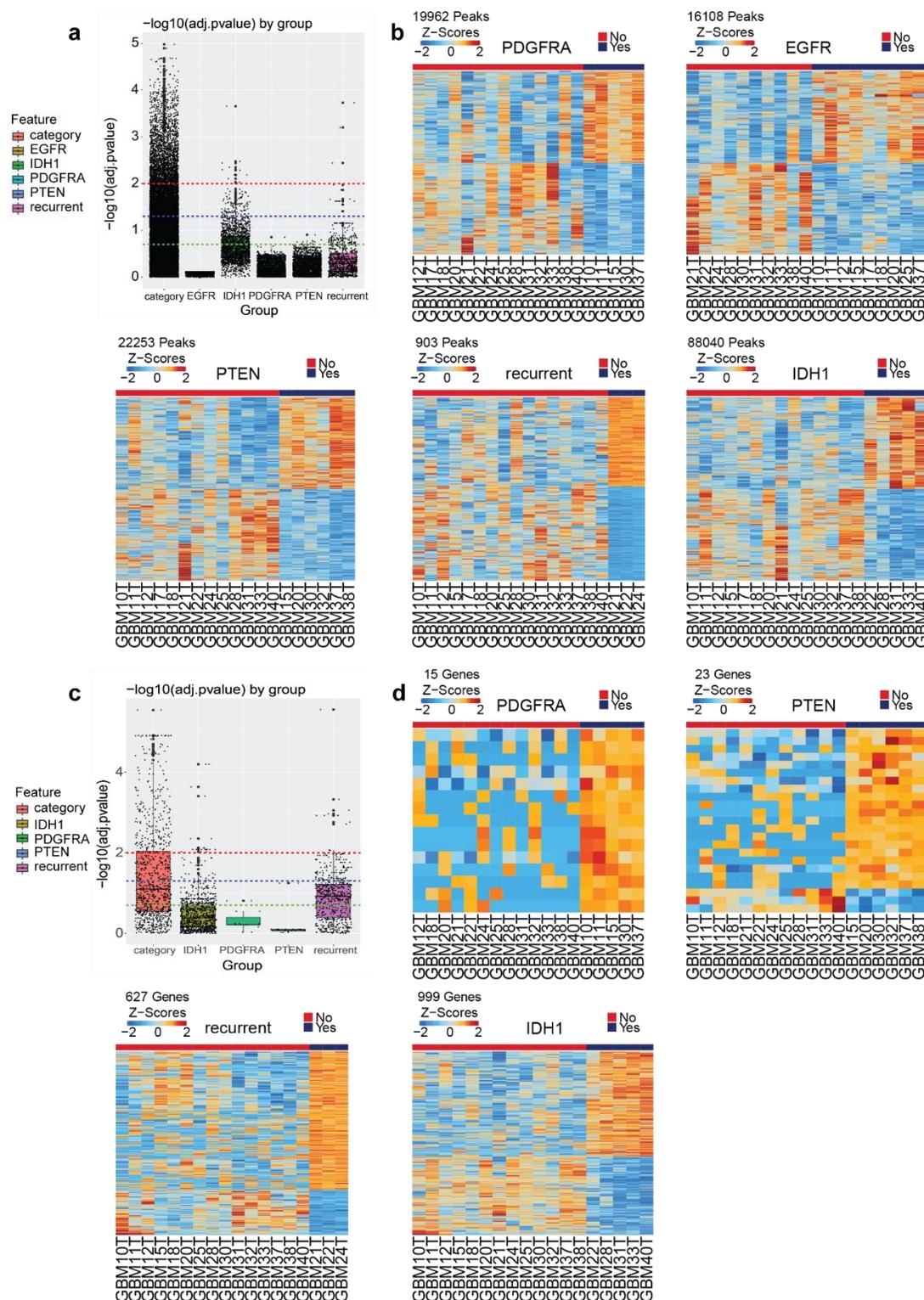


Fig. 16: Additional genomic diagnosis-specific ATAC- and RNA-seq tumor astrocyte signatures

(a, c) Box plots showing adjusted p-value distribution for peak accessibility and gene expression comparisons across individual genomic features (see Fig. 15a). Dashed green line is an adjusted p-value of 0.2, blue is 0.05, and red is 0.01. (b, d) Heatmaps of ATAC peaks and genes with variable signals across samples that do (Y) or do not (N) have each respective genomic diagnosis.

Figure 17: Gene expression comparisons between IDHwt, IDH1mt, and margin astrocyte lineage cells

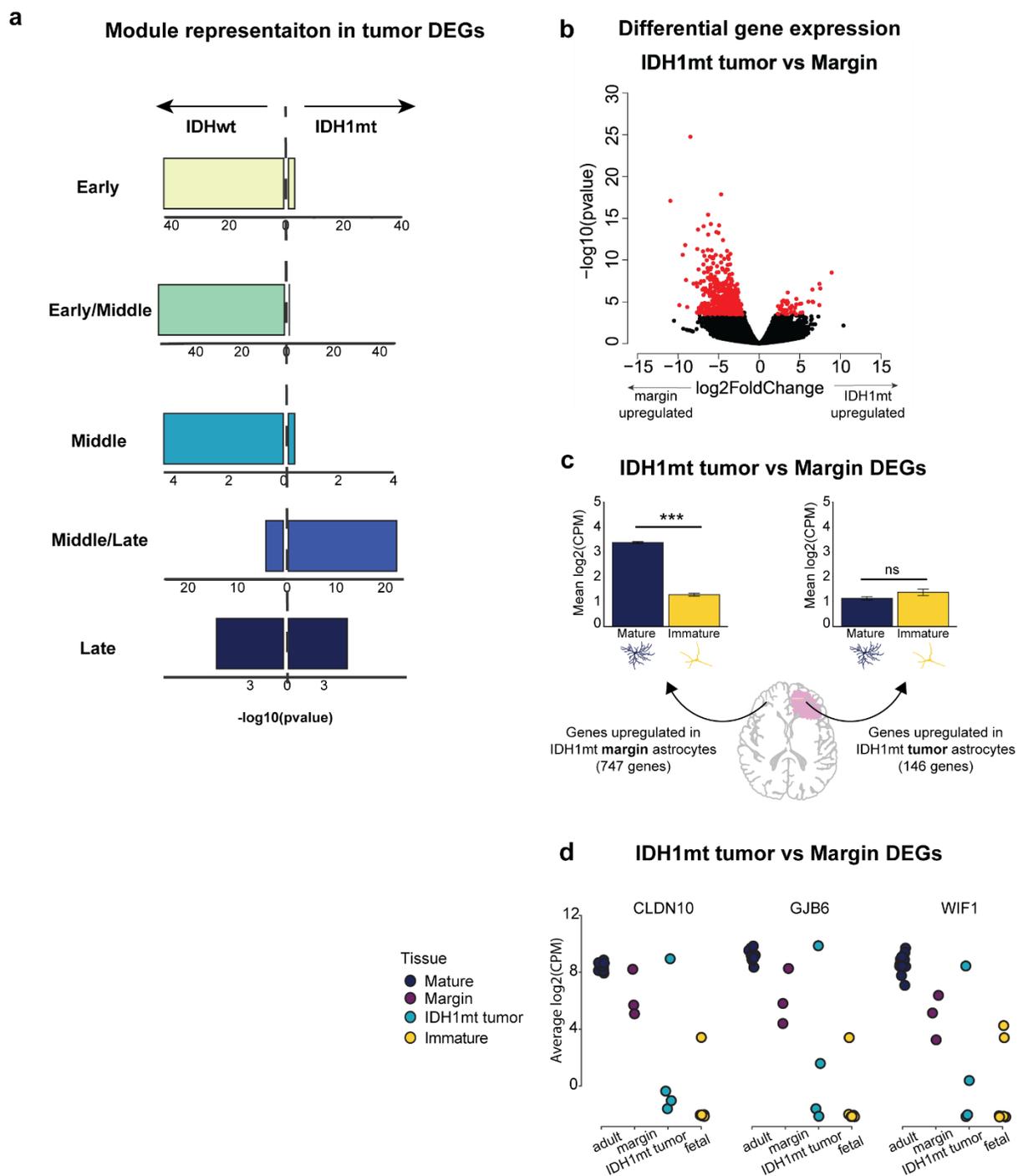


Fig. 17: Gene expression comparisons between IDHwt, IDH1mt, and margin astrocyte lineage cells

(a) Diverging bar plots of p-values for maturation module gene set enrichments in DEGs between IDHwt and IDH1mt tumor astrocytes (supplement to Fig. 14e). Bar color indicates maturation module. (b) Volcano plot showing DEGs between IDH1mt tumor (FDR<0.05; log₂FC>2) and margin (FDR<0.05; log₂FC<-2) astrocytes. DEGs are colored in red. (c) Bar plots showing DEGs from Fig. 16b in mature (adult) and immature (fetal) astrocyte transcriptomic data (77). Values are the mean log₂CPM of the 747 margin and 146 tumor DEGs. *** P < 0.0005; ns, not significant, two-tailed student's t-test. (d) Example mature DEGs upregulated in IDH1mt margin and downregulated in tumor astrocytes.

Figure 18: DNA methylation and hydroxymethylation of astrocyte lineage cells in IDH1mt and IDHwt tumors

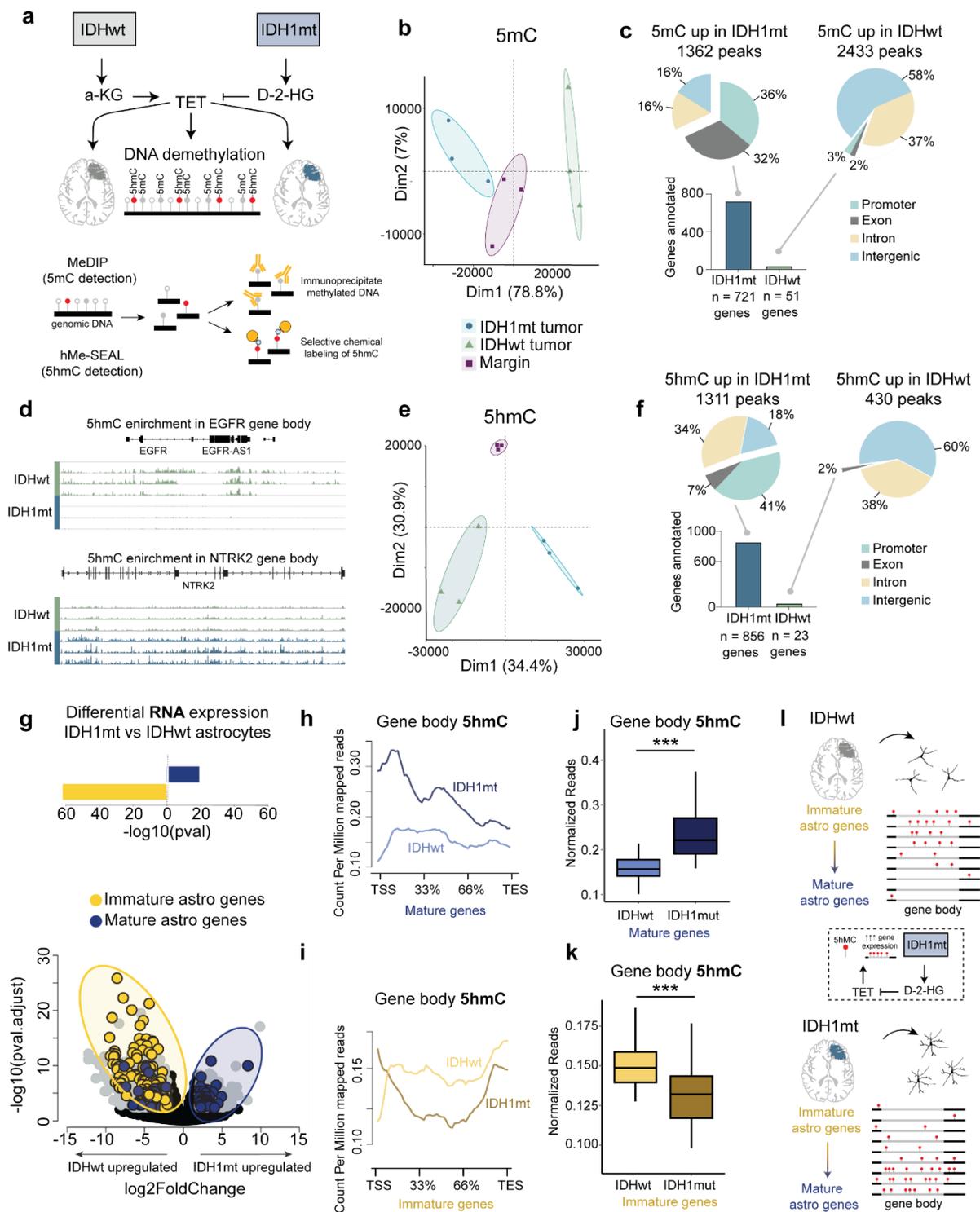


Fig. 18: DNA methylation and hydroxymethylation of astrocyte lineage cells in IDH1mt and IDHwt tumors

(a) Schematic of aberrant DNA methylation (5mC) and hydroxymethylation (5hmC) in IDH1mt cells and approach for comparing global 5mC and 5hmC levels in IDH1mt and IDHwt samples. (b, e) PCAs of 5mC (b) and 5hmC (e) data. (c, f) Pie charts of relative abundances of differentially (c) methylated regions and (f) hydroxymethylated regions across genomic features. Bar plots show the number of genes that annotate each differentially methylated/hydroxymethylated region. (d) Example plots of differential 5hmC in (top) EGFR and (bottom) NTRK2. Plots depict 5hmC signal across IDHwt (green) and IDH1mt (blue) samples. (g) DEGs ($|\log_2FC| > 2$, $FDR < 0.05$) between IDH1mt and IDHwt astrocyte lineage cells. DE immature astrocyte genes are shown in yellow and mature astrocyte genes in purple. Diverging bar plot quantifying representation of mature and immature genes in IDHwt (immature.pval= $4.8e-62$, mature.pval= 0.92) and IDH1mt (immature.pval= 0.99 , mature.pval= $7.3e-19$) DEGs. (h-k) NGS plots and corresponding box and whisker plots showing gene body 5hmC levels in IDH1mt and IDHwt tumors at (h and j) mature and (i and k) immature astrocyte genes. (l) Summary schematic for the role of gene body hydroxymethylation in maintaining IDH1mt and IDHwt maturation differences.

Figure 19: 5mC and 5hmC quantification in IDH1mt and IDHwt tumors

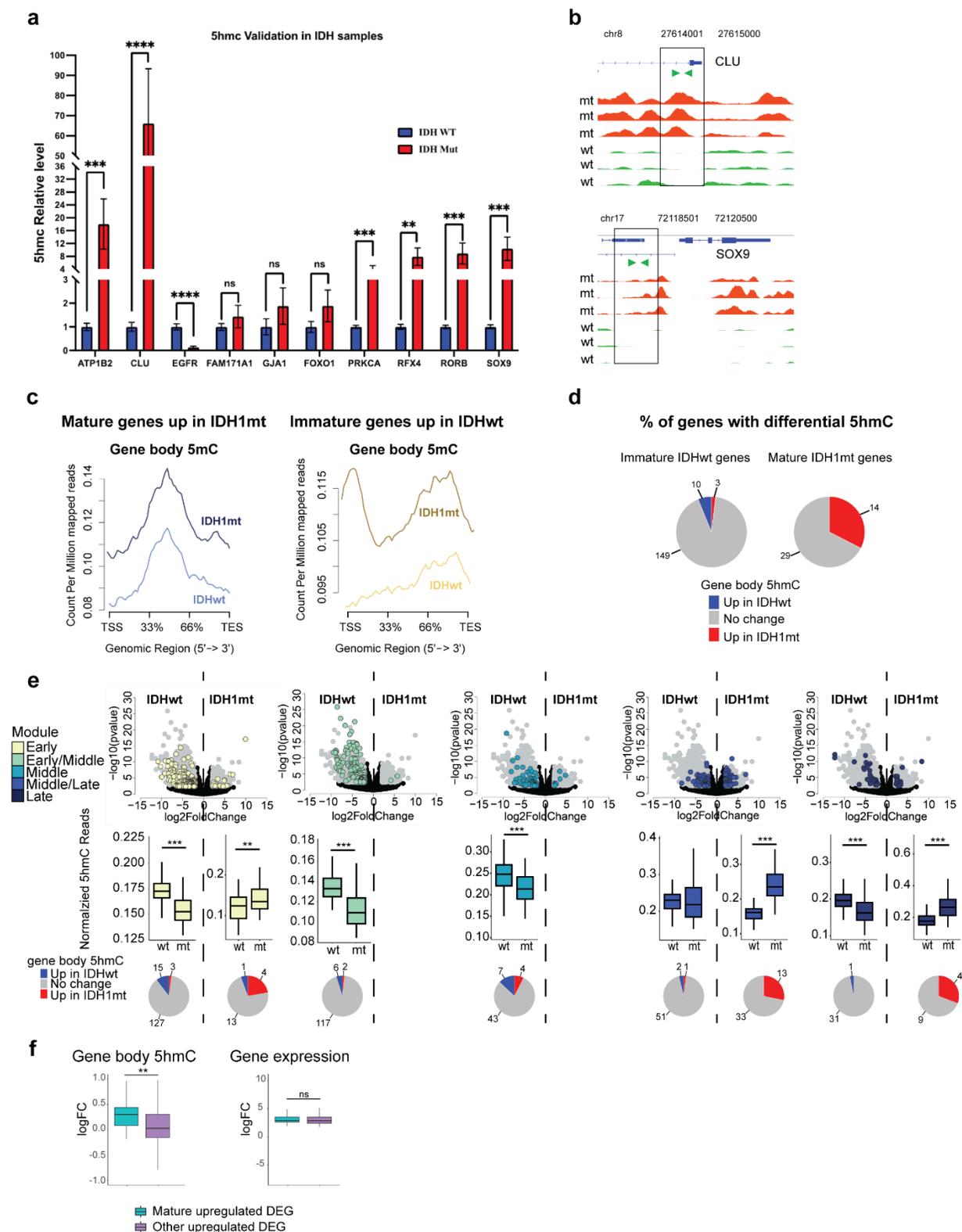


Fig. 19: 5mC and 5hmC quantification in IDH1mt and IDHwt tumors

(a) 5hmC-Capture-qPCR validation for 10 developmental and oncogenic genes. Blue bars indicate relative 5hmC levels in IDHwt samples and red bars for IDH1mt samples. ns, not significant. (b) 5hmC levels for two representative genes included in 5hmC-Capture-qPCR, CLU (top) and SOX9 (bottom). 5hmC tracks for IDHwt samples in green and IDH1mt in red. (c) NGS plots of gene body 5mC levels in IDH1mt and IDHwt for mature genes upregulated in IDH1mt astrocytes (left, purple) and immature genes upregulated in IDHwt astrocytes (right, yellow). (d) Percent of genes with differential 5hmC in (left) immature genes upregulated in IDHwt and (right) mature genes upregulated in IDH1mt. (e, Top) Volcano plots from Fig 15f. (Middle) Box and whisker plots for each respective maturation module showing normalized gene body 5hmC for genes upregulated in IDHwt (left of dashed line) and for genes upregulated in IDH1mt (right of dashed line). No 5hmC data is shown if fewer than 10 upregulated genes. (Bottom) Percent of genes with differential 5hmC. (f) Bar plots showing whole tissue gene body 5hmC levels (left) and expression (right) in differentially expressed mature genes and all other DEGs. * if $p_{\text{adj}} < 0.01$, ** if $p_{\text{adj}} < 0.001$, *** if $p_{\text{adj}} < 0.0001$

2.4 Discussion

While there is an abundance of data detailing the cell type composition of GBM tumors, we know substantially less about how these lineages deviate from normal developmental trajectories. Here, we used the hCO platform to track molecular changes throughout healthy astrocyte maturation and projected this timeline onto GBM astrocyte lineage cells to find where tumor astrocytes diverge from the typical maturation path.

Additionally, we identified unique molecular and transcriptomic signatures across tumors with differing molecular backgrounds and uncovered distinct maturation differences between IDHwt and IDH1mt tumor astrocytes.

Astrocytes undergo a three-step maturation transformation

Previous studies have described the “bookends” of astrogenesis, highlighting stark morphological, physiological, and molecular differences between fetal and adult astrocytes (30, 77, 78). These distinct astrocyte states also have functional consequences. Fetal astrocytes guide fundamental processes like synaptogenesis (77, 268, 319), neural circuit formation (268, 319), and blood-brain barrier formation (320, 321), whereas adult astrocytes are more tuned to maintaining a homeostatic environment (268). However, the dynamics of this developmental transition are more elusive. This study reveals a process whereby human astrocytes mature in three distinct stages. In particular, we define a new intermediate maturation state sandwiched between developmental landmarks of early proliferative astrocytes and quiescent mature cells. Interestingly, according to previous studies comparing hCO and fetal cortical development, our newly described intermediate maturation phase aligns most closely with near-full gestation time points (230). Additionally, this state exhibits the

highest expression of core developmental transcription factors. We speculate this may indicate a developmental window in which astrocytes are aggressively shifting their molecular focus from proliferation toward more specialized functionality.

Using paired ATAC- and RNA-seq datasets, we generate a map of stage-specific TF-target gene networks that change throughout human astrocyte maturation. This regulatory atlas highlights both known and novel elements of astrocyte development, including several TFs that are commonly associated with neuronal and oligodendrocyte precursor cell development, but that have also recently been implicated in astrogenesis (231, 272). This suggests the possibility of regulatory redundancies across neurodevelopmental lineages, which may help to explain mixed populations of neurodevelopmental cell lineages apparent in GBM (eg astrocyte/oligodendrocyte common glial progenitors). We believe this detailed longitudinal characterization of astrocyte maturation will serve as a useful resource for experimental approaches directing human astrogenesis in vitro and could refine our understanding of how and when astrocytes contribute to developmental processes and disorders.

Mapping a normal astrocyte maturation trajectory onto GBM astrocyte lineage cells

While our group and others have demonstrated that GBM astrocyte lineage cells more closely resemble fetal astrocytes (77, 291, 292), this study indicates that molecular events in the transition from “middle” to “late” stages of maturation are specifically what distinguishes neoplastic tumor astrocyte lineage cells from quiescent astrocytes in the margin. Directing attention to this narrow developmental window points to the

“middle/late” regulatory pathways that may need to be “activated” to coerce terminal differentiation in tumor astrocytes. What remains unclear is the directional progression of perturbed maturation in GBM, which could indicate how astrocyte lineage cells end up in an immature developmental stage. Our data, as well as other scRNA-seq datasets, suggest a stalling of maturation in the middle stage, resulting in an accumulation of tumor astrocytes with “early/middle” and “middle” molecular signatures. We observed only a sporadic number of neoplastic cells enriched for “middle/late” or “late” genes, which we might expect if tumor astrocytes maintained a signature of a previous quiescent state. However, we cannot rule out the possibility that some immature tumor cells are dedifferentiated mature astrocytes. There is significant evidence demonstrating that adult astrocytes can acquire immature properties when activated (78, 291, 322, 323), albeit a temporary and distinctly unique phenotype from a fetal astrocyte. Further studies, like applying lineage tracing strategies, are necessary to definitively conclude which path tumor astrocytes follow to reach their immature state and whether this route could be experimentally coerced.

Maturation differences between IDHwt and IDH1mt tumor astrocytes

We demonstrate that astrocytes from tumors harboring unique genetic aberrations exhibit specific molecular and transcriptomic features, with IDH1mt tumor astrocytes bearing a uniquely mature signature. This finding is consistent with the fact that IDH1mt tumors were recently reclassified by the World Health Organization as a separate category of diffuse glioma (180, 302) due to their distinct molecular diagnostic characteristics (308, 324), less severe pathology (303-305, 325), and prolonged patient survival rate (306-308). One unique property of IDH1mt tumors is the production of D-

2-HG, a molecule that disrupts TET enzyme activity and DNA demethylation processes (315, 326). Thus, we investigated how 5mC and 5hmC patterns differ between IDHwt and IDH1mt tumors and found a positive association between gene body 5hmC and the expression of immature and mature astrocyte genes. Other groups have demonstrated a strong correlation between gene body 5hmC and the expression of cell type-specific genes (327), insinuating a role for 5hmC in the maintenance of cell-specific identity (328-330). A similar relationship between 5hmC and gene expression has been previously documented in IDH1mt and IDHwt tumors (331), supporting our findings and suggesting that 5hmC may contribute to the regulation of key tumor transcriptional programs, including maturation gene sets. Multiple potential mechanisms have been proposed to explain how 5hmC accumulation could help modulate transcriptional activity. For example, 5hmC and TET protein aggregation on gene bodies might facilitate RNA Pol II sliding (332), maintain transcriptionally favorable euchromatic environments in the presence of H3K36 trimethylation (333), or antagonize repressors such as DNMTs and Polycomb complexes (334). Clearly, further research into the mechanisms that link 5hmC and gene expression is necessary and could help shape our understanding of how developmental trajectories are modulated in brain tumors.

The heterogeneous and promiscuous nature of GBM cells makes these tumors nearly impossible to entirely eradicate. The current treatment course—generally, resection, followed by chemotherapy and radiation—cannot target all tumor cells and has even been shown to select for certain cellular subtypes or push remaining cells to acquire new mechanisms of survival (335-337). A promising therapeutic approach could be using maps of developmental lineages, like the one laid out in this study, to coerce the

maturation of GBM cells, transforming neoplastic cells into quiescent, mature phenotypes. As differentiation therapies become a more viable option, it is crucial that we integrate principles of normal neurodevelopment to better understand which molecular networks to target to have the most effective impacts.

2.5 Materials and methods

hiPSC culture

One male (C3.1) and one female (C4.1) human induced pluripotent stem cell (hiPSC) line were used to form hCO cultures. hiPS cells were cultured on plates coated with vitronectin (ThermoFisher, Cat. A14700) and maintained in Essential 8 medium (Thermofisher, Cat. A1517001). The two iPSC lines were genotyped by SNP-array to confirm genomic integrity and screened regularly for Mycoplasma contamination.

Human cortical organoid formation and culturing

Human cortical organoids (hCOs) were formed using a previously published protocol (218). In short, once at 80-90% confluency, hiPSC colonies were detached from plates to form spheroids using the dispase method described in Sloan et al, 2018 (218). Next, spheroids were exposed to a series of small molecules to pattern toward dorsal forebrain identity. For the first 6 days in culture, spheroids were exposed to neural induction media (DMEM/F12, KSR, NEAA, Glutamax, Pen/Strep, Beta-mercaptoethanol) supplemented with Dorsomorphin (Sigma, Cat. 142 P5499-25MG, 5 μ M) and SB-431542 (Selleck Chemicals, Cat. S1067, 10 μ M) for dual-SMAD pathway inhibition. During this time, media was changed daily. Next, spheroids were treated with neural media, consisting of neurobasal A (Thermo Fisher Scientific, Cat. 10888), B-27 supplement without vitamin A (Thermo Fisher Scientific, Cat. 12587), GlutaMax (1:100, Thermo Fisher Scientific, Cat. 35050061), penicillin and streptomycin (1:100, Thermo Fisher Scientific, Cat. 15140122), supplemented with EGF (20 ng/mL, R&D Systems, Cat. 236-EG-01M) and FGF2 (20 ng/mL, R&D Systems, Cat. 233-FB01M). Neural media was changed daily through 15 days in culture, and then every other day for days 16-24. From

days 25-43 in culture, spheroids were treated every other day with neural media supplemented with BDNF (20 ng/mL, PeproTech, Cat. 450-02-147 1mg) and NT-3 (20 ng/mL, R&D Systems, Cat. 267-N3-005/CF). After day 43, maturing organoids were fed with neural media only every 3-4 days until collected for experimental time points.

GBM tissue procurement

Paired GBM “tumor” and “margin” tissue samples were obtained from surgical resections performed with informed consent under a protocol approved by the Emory University Institutional Review Board. In this study, margin tissue samples, brain tissue from around the edges of the GBM resection, serve as a patient-paired quiescent control, as it bears a closer resemblance to healthy adult astrocytes than to neoplastic fetal astrocytes (77). However, we acknowledge that this tissue contains infiltrating tumor cells, which we were able to distinguish from non-neoplastic cells via single nuclei sequencing (Fig. 9i; Fig. 13). Marginal brain tissue was obtained after safe and maximal surgical removal of the tumor as deemed appropriate by the operating neurosurgeon, using techniques including direct vision, image-guidance, and fluorescent-guidance. The surgical margin was searched for brain that is free of surgery-induced contusion and coagulation artifacts and appeared viable. These samples included white and gray matter generally taken 2-10 mm beyond the edge of what had clearly been tumor tissue. In cases where the tumor is adjacent to or in brain harboring eloquent functioning, such as speech, motor, and visual function sampling is not undertaken so as to preserve those functions. Thus, sampling is directed to cerebral tissue from more functionally quiescent regions. After tissue resection, tumor and margin samples were immediately deposited

in 4°C Hibernate-A medium (Thermo Fisher Scientific, Cat. A1247501) and prepared for tissue dissociation within 1 hr post-resection.

Single-cell dissociation and immunopanning

hCOs and primary GBM tissue samples were dissociated following a previously established protocol (77). In short, we enzymatically dissociated tissue using Papain (Worthington Biochemical, Cat. LS003126) at 20U/mL at 34°C for one-hour, quenched Papain with ovomucoid solution, mechanically dissociated tissue by triturating to obtain a single-cell suspension, and then filtered through a 40µM strainer. Once we had a single-cell suspension, we proceeded to purify individual cell types using an immunopanning protocol outlined below.

Single-cell suspensions were added to a series of 10 (for hCO) or 15 (for GBM tissue) cm Petri dishes pre-coated with cell-type specific primary antibodies (listed below) for the depletion of unwanted cell types and the final enrichment of astrocytes. Cell suspensions were incubated on plates for 5 or 15 minutes, depending on the coating antibody, and then unbound cells were transferred to the next plate in the sequence until the final positive selection plate.

hCO plate sequence: one anti-THY1 for 5 mins (BD Biosciences, Cat. 550402) to deplete neurons, followed by either an anti-CD49f (immature astrocytes; BioLegend, Cat. 313602) plate for 15 mins or an anti-HepaCAM (astrocytes; R&D Systems, Cat. MAB4108) for 15 mins. At early time points (<100 days), only CD49f+ cells are abundant with very few HepaCAM+ cells. By day 150, the ratio of CD49f+ cells to

HepaCAM+ cells started to shift, such that we had roughly equivalent numbers of each, and then continues to shift towards a HepaCAM+ enrichment over time. For this reason, we collected CD49f+ cells for the 80- and 100-day time points, both CD49+ and HepaCAM+ cells for the 150-day time points and for all later time points we used HepaCAM+ cells. In our analyses, we found that CD49f+ and HepaCAM+ cells from d150 hCOs had nearly identical transcriptomic and genomic accessibility profiles, suggesting that these antibodies capture a similar population as these cells mature.

GBM tissue plate sequence: one anti-CD45 (R&D Systems, Cat. MAB1430) plate for 5 mins to deplete macrophages, two sequential anti-O4 (in-house hybridoma) plates for 5 mins each to deplete pre- and mature/myelinating oligodendrocytes, and an anti-HepaCAM (R&D Systems, Cat. MAB4108) plate for 15 mins each to positively-select for astrocytes.

For both hCO and GBM tissue samples, we ended plate sequences with two positive-selection astrocyte plates- one for RNA-seq libraries and the other for ATAC-seq libraries. For RNA-seq, we scraped cells using 700uL Qiazol (Qiagen, Cat. 79306) and froze at -80°C for downstream processing. For ATAC-seq library preparation, we incubated cells in a trypsin solution at 37°C for 5 mins and gently dislodged cells from the plates for further processing.

Bulk ATAC-seq library prep

Bulk ATAC-seq libraries were prepared following the previously established Omni-ATAC protocol (338) with minor modifications. Briefly, astrocytes that were trypsinized

during immunopanning were counted and 10,000-50,000 cells were washed with cold ATAC-seq resuspension buffer (RSB; 10 mM Tris-HCl pH 7.4, 10 mM NaCl, and 3 mM MgCl₂ in water), and permeabilized with ATAC-seq lysis buffer (RSB supplemented with 0.1% NP40, 0.1% Tween-20, and 0.01% digitonin). For the transposition step, nuclei were resuspended in 50 µl of transposition mix (25 µl 2× TD buffer, 2.5 µl transposase, 16.5 µl PBS, 0.5 µl 1% digitonin, 0.5 µl 10% Tween-20, and 5 µl water) and incubated at 37 °C for 30 min in a thermomixer with shaking at 1,000 r.p.m. The DNA fragments were cleaned up with the Qiagen MinElute PCR Purification kit (Qiagen, Cat. 28004) and PCR amplified for 5 cycles with Illumina Nextera adaptors using NEBNext High Fidelity 2x Master Mix (New England Biolabs, Cat. M0541S). Next, a qPCR reaction was performed using 10% of the pre-amplified PCR product. We looked for the Ct value at ¼ maximum fluorescence to determine the number of additional cycles to amplify the ATAC-seq libraries. After a second round of PCR amplification (all together, a total of 8-10 cycles), we performed a second cleanup with 1.8x Ampure XP beads (Aline biosciences, Cat. C-1003-50). Fragment size distribution and concentration were evaluated via Bioanalyzer and libraries were sequenced using 2x150-bp reads on an Illumina HiSeq 2500 instrument at a targeted depth of 50 million paired-end reads per sample.

Bulk RNA-seq library prep

For bulk RNA-seq libraries, we extracted RNA from purified astrocytes using the miRNeasy kit (QIAGEN, Cat. 217004) according to the manufacturer's protocol. We assessed the quality of RNA via Bioanalyzer (Agilent, Eukaryotic Total RNA Pico kit, Cat. 5067-1513) and RNA samples with an RNA integrity number (RIN) less than 8 were

discarded. We prepared bulk RNA-seq libraries with the NEBNext Ultra II kit (New England Biolabs, Cat. E7805S) using poly-A selection, and cDNA library quality was assessed via Bioanalyzer (Agilent, High Sensitivity DNA kit, Cat. 5067-4626). Libraries were sequenced using 2x150-bp reads on an Illumina HiSeq 2500 instrument at a targeted depth of 20 million paired-end reads per sample.

Single-nucleus multiome library preparation and sequencing from GBM tissue

Nuclei were extracted from frozen GBM tissues following a previously established protocol (338, 339). For all samples, ~20 mg of tissue was dissociated with a 2-ml Dounce homogenizer in homogenization buffer (0.26 M sucrose, 0.03 M KCl, 0.01 M MgCl₂, 0.02 M Tricine-KOH pH 7.8, 0.001 M DTT, 0.5 mM Spermidine, 0.15 mM Spermine, 0.3% NP40, and cOmplete Protease inhibitor). This was followed by filtering through a 40um Flowmi cell strainer, then a 20um bucket-style cell strainer, and centrifugation for 10 min at 600 r.c.f. After the majority of the supernatant was carefully removed, the pellet was resuspended in homogenization buffer and mixed with an equal volume of 50% iodixanol solution to make a final concentration of 25% iodixanol. Next, a 30% iodixanol solution, followed by a 40% iodixanol solution, was layered under the 25% mixture and centrifuged for 20 min at 3000 r.c.f without the centrifuge brake. Post-centrifugation, a thin white nuclei band was carefully collected from the interface of the 30% and 40% iodixanol solutions. Nuclei underwent 1-2 wash steps to remove any additional debris by gently mixing nuclei in ATAC- RSB-Tween buffer (0.01 M Tris-HCl pH 7.5, 0.01 M NaCl, 0.003 M MgCl₂, 0.1% Tween-20) and centrifuging for 10 min at 600 r.c.f. Nuclei were then counted and a total of 16,100 nuclei were collected for

library preparation with the 10x Genomics Single Cell Multiome ATAC + Gene Expression kit (10x Genomics, Cat. 1000285) following the manufacturer's protocol for transposition, 10x capture, and library preparation. Libraries were assessed via Bioanalyzer (Agilent, High Sensitivity DNA kit, Cat. 5067-4626) and sequenced at a target depth of at least 50,000 read-pairs per nucleus using 2x150-bp reads on an Illumina Novaseq 6000 instrument through Admera Health. Libraries were processed using the 10x Genomics Cell Ranger ARC (cellranger-arc-2.0.0) pipeline with default parameters and aligned to the hg38 reference genome (refdata-cellranger-arc-GRCh38-2020-A-2.0.0). The quality control matrix containing raw and filtered read numbers is provided in Table S3b.

DNA isolation for 5mC/5hmC experiments

Margin, IDH1mt, and IDHwt tumor samples were harvested and immediately frozen on dry ice and stored at -80°C. Tissue was digested in a lysis buffer (10 mM Tris pH 8.0, 5 mM EDTA, 200 mM NaCl, 0.2% SDS) with 30 µl proteinase K (20 mg/ml) and incubated at 55°C overnight. After the overnight digestion, the lysates were brought to room temperature and incubated with 5 µL of RNase A solution (20 mg/ml) for 1 hour at 37 °C. This was followed by adding an equal volume of buffered phenol:chloroform:isoamyl alcohol (25:24:1 ratio) and centrifuging at 14,000 RPM at room temperature for 20 minutes. The supernatant was transferred to clean tubes and 5 µl of 5 M NaCl, 2 µl Glycogen, and an equal volume of 100% ethanol were added. After overnight incubation at -20°C, samples were centrifuged at 10,000 g for 10 minutes at 4°C, and the DNA pellet was washed in 70% ethanol twice. After all ethanol was

removed, the DNA pellet was dissolved in nuclease-free water and incubated overnight at 4°C, before being quantified by Nanodrop.

Methylated DNA Immunoprecipitation (MeDIP)

3 µg of genomic DNA was fragmented to 300-400 base pairs using a Covaris focused-ultrasonicator Me220. DNA fragments were then subjected to end repair, A-tailing, adaptor ligation, and USER digestion using the NEBNext Ultra II DNA Library Prep kit for Illumina (New England BioLabs, Cat. E7645S), according to the manufacturer's protocol. Following USER digestion and purification, DNA was denatured for 10 minutes at 95°C and immunoprecipitated overnight at 4°C with 4 µL of either 5mC antibody (Active Motif, Cat. 39649) or IgG antibody (Sigma, Cat. 12-371) in IP buffer (500 mM Tris-HCl, pH 7.4, 750 mM NaCl and 0.25% TritonX). The mixture was then incubated with Protein G coated Dynabeads (ThermoFisher, Cat. 10004D) for at least 2 hours at 4°C, washed with ice cold IP buffer, and finally washed in ice cold high salt IP buffer (IP buffer plus extra 300 mM NaCl). After the final washing, the beads were treated with 200 µL of digestion buffer (1X TE Buffer, pH 7.4, 0.25% SDS, 0.25% Proteinase K (2.5 mg/mL)) and shaken at 1000 rpm for 2 hours at 55°C. The methylated DNA was recovered by phenol:chloroform:isoamyl alcohol (25:24:1) extraction, followed by precipitation in 3X volume of 100% ethanol, supplemented with 3 µL glycogen (5mg/mL) and 15 µL NaAC (pH 5.2), and then incubated overnight at -20°C. The next day, DNA was pelleted and washed with 75% ethanol and dissolved in nuclease-free water. Libraries were sequenced 2x150-bp on an Illumina HiSeq 2500 platform at a targeted depth of 30 million paired-end reads per sample.

5hmC capture

5hmC capture was performed according to the methods described in Kuehner et al, 2021 (340). In brief, 5 µg of genomic DNA was fragmented to 300-400 base pairs through sonication, and 5hmCs were glycosylated by T4 phage β-glucosyltransferase enzyme and UDP-6-N₃-glucose. The glycosylated fragments were purified with Ampure XP beads (BeckmanCoulter, Cat. A63881), biotinylated by click reaction with disulfide biotin, and pulled down with Dynabeads MyOne Streptavidin C1 beads (ThermoFisher, Cat. 65001). The 5hmC fragments were released from the beads using dithiothreitol and purified with Ampure XP beads for a final time. DNA fragments were eluted in nuclease-free water and quantified by Qubit. Libraries were sequenced 2x150-bp on an Illumina HiSeq 2500 platform at a targeted depth of 30 million paired-end reads per sample.

qPCR validation of 5hmC–enriched regions

5hmC-captured DNA products were used as templates in triplicate 20 µL qPCR reactions, which included 1x PerfeCTa SYBR Green FastMix, low ROX (QuantaBio, Cat. 95074), 0.25 µM forward and reverse primers, and nuclease-free water. The PCR cycling was performed on QuantStudio™ 3 System using Fast mode and the conditions were as follows: 95 °C for 10 minutes, 40 cycles at 95 °C for 15 seconds, 60 °C for 1 minute. Fold enrichment was calculated as 2^{-dCt} , where $dCt = Ct$ (5-hmC enriched in IDH1-mutant group) – Ct (5-hmC enriched in IDHwt group). Primer sequences are provided in Table S5b. Primers were designed using the Primer3 online tool (<https://bioinfo.ut.ee/primer3-0.4.0/>).

Whole tissue RNA isolation and library preparation

Total RNA was isolated from the tumor and margin brain tissue with TRIzol. Tissues were homogenized in TRIzol and incubated for at least 5 minutes at room temperature. Chloroform was added to the homogenate in a 1:5 ratio, the tubes were shaken, incubated at room temperature for another 15 minutes, and samples were centrifuged at 20,000 g for 15 minutes at 4°C. The top aqueous layer was transferred to a clean tube, a tenth volume of 3 M NaAc (pH 5.2), 4 µl of glycogen (5 mg/ml), and 100% isopropanol (1:1 ratio) was added and then incubated at -80°C overnight. The next day, samples were centrifuged at 15,000 RPM for 20 minutes at 4°C. The resulting RNA pellet was washed twice in 75% ethanol and centrifuged at 7,500 g for 10 minutes at 4°C. The washed RNA pellet was dissolved in nuclease-free water. RNA was quantified by Nanodrop and the quality was confirmed by a gel. Library construction was performed by Admera Health, and libraries were sequenced 2x150-bp on an Illumina HiSeq 2500 platform at a targeted depth of 30 million paired-end reads per sample.

Bulk RNA-seq data processing (hCO and GBM astrocyte samples)

Fastq files were trimmed using the Trimmomatic software (341), reads were mapped to the GRCh38/hg38 reference genome using STAR aligner (342), and read summarization was carried out with the featureCounts software (343).

We assessed transcriptomic signatures of hCO astrocytes at each time point and GBM tissue astrocytes to confirm cell identity by correlating the expression of cell type-

specific signatures (Table S1) between our samples and a published primary human RNA-seq dataset (77) that includes data for purified forebrain fetal astrocytes, postnatal astrocytes, neurons, oligodendrocytes, endothelial cells, and microglia/macrophages.

Bulk ATAC-seq data processing (hCO and GBM astrocyte samples)

ATAC-seq file processing included the following steps: (1) Nextera adapters were trimmed using TrimGalore (344), (2) reads were mapped to the hg19 reference genome using Bowtie2 (345), (3) mitochondrial and non-unique alignment reads were removed using Samtools (346), (4) PCR duplicates were removed with Picard tools (347), (5) black-listed reads were removed using Bedtools (348), and (6) peaks were called using MACS2 (349) with the following parameters: shift -75 –extsize 150 –nomodel –call-summits –nolambda -p 0.01 -B –SPMR

Bam files from the data processing pipeline were loaded into R for further analysis. We first used the ChrAccR (v0.9.17) R package's setConfigElement function to remove low-coverage peaks by filtering to only include those that had a minimum insertion count of 4 and removed peaks on the X and Y chromosomes. Next, a quality check report was generated using the run_atac_qc function in the ChrAccR R package. hCO and GBM astrocyte ATAC-seq libraries were evaluated for nucleosomal-read periodicity in the fragment size distribution, high signal enrichment at transcription start sites (TSSs; TSS enrichment score > 6), and high peak signal with minimal noise around biologically-relevant regions in the genome.

Principal component analyses

Principal component analysis was performed using the `getDimRedPlot` function in the `muRtools` (v0.9.5) R package, using the first and second principal components.

Differential expression and differential accessibility

Differential expression (RNA-seq) and accessibility (ATAC-seq) analyses were performed in R using the `DESeq2` package, accounting for sample pairing between tumor and margin samples (\sim Patient + Tissue_type) (350). Prior to finding differentially accessible peaks, consensus peak sets were generated using the `getConsensusPeakSet` function in the `ChrAccR` R package where peaks had to be consistently present across 75% of biological replicates within a group (IDHwt tumor or margin) to be retained. Differentially accessible peaks were visualized using the `ggmaplot` R package. Peaks and genes that had an $FDR < 0.05$ and absolute $\log_2FC > 2$ were deemed differentially expressed or accessible, respectively.

To see if specific genes were overrepresented in our differentially expressed gene sets, we implemented the `testGeneOverlap` function in the `GeneOverlap` R package, which applies a Fisher's Exact Test to check for significant association between sets of genes. The primary gene sets that we looked at were the top 1000 and top 100 fetal astrocyte-enriched genes and the top 1000 and top 100 adult astrocyte-enriched genes published in Zhang et al, 2016 (77) (Tables S1e and S2b). Additionally, we also looked for the overrepresentation of maturation WGCNA module genes (Table S1b) in several analyses.

Weighted gene co-expression network analysis (WGCNA)

Prior to running a WGCNA, we first performed a variance stabilizing transformation to normalize data across astrocyte maturation time points using the DESeq2 R package. We then subsetted the data to include the top 15% most variably expressed genes and peaks to focus on genes and regions of the genome that exhibit substantial changes. To identify sets of genes and peaks that change across astrocyte maturation, we performed a WGCNA using the WGCNA R package (351), with the guidance of the following tutorial: [WGCNA Gene Correlation Network Analysis - Bioinformatics Workbook](#). Top module genes (Table S3f) were determined using the adjacency function in the WGCNA package, which we used to calculate gene network adjacency values and rank genes with the highest module connectivity.

Bulk ATAC-seq motif analyses

Motif enrichment and occurrence values for the hCO astrocyte maturation timeline were generated using WGCNA ATAC peak sets (Table S1c).

For motif enrichment, we evaluated several parameters, including motif deviation scores, q-values, and odds ratios. For calculating deviation scores, peak count matrices were RPKM normalized, log₂ transformed, and quantile normalized. ChromVAR deviation scores were then calculated with the getChromVarDev function in the ChrAccR R package, using the JASPAR TF motif database, and motif deviation z-scores were plotted using the ComplexHeatmap R package (352). We calculated q-values and odds ratios with the getMotifEnrichment function in the ChrAccR R package using the JASPAR TF motif database. Motif enrichment bar plots depict -log(q value) and odds ratios for PECA-predicted maturation TF candidates where there was significant

enrichment ($p_{\text{adj}} < 0.05$) in the respective WGCNA ATAC peak module (early, early/middle, middle, middle/late, late).

Motif occurrence data was generated using the FIMO (353) program which is part of the MEME Suite of motif analysis tools. Dot plots depict the occurrence of PECA-predicted maturation TF motifs in each WGCNA ATAC peak module, normalized to the total number of peaks in each module.

Paired Expression and Chromatin Accessibility (PECA) analysis

We employed the PECA2 (282, 283) computational algorithm to integrate our hCO astrocyte ATAC- and RNA-seq data to predict TFs that might regulate the expression of our WGCNA maturation genes. To start, we first binned our maturation time point data into three groups based on the results of the ATAC- and RNA-seq PCAs- early (d80-150), middle (d200-350), and late (d400-550). We did this by generating average gene expression values (TPM) for each developmental group and used Samtools to merge ATAC .bam files into a single file for each bin.

Using PECA2, we generated matrices for early, middle, and late groups which include a score for the likelihood of a TF to regulate a TG (for all known human TFs and protein-coding genes). Since we were most interested in the TF-TG networks that are different across maturation groups, we next ran the PECA network comparison algorithm, comparing between all possible maturation group combinations (early x middle, early x late, and middle x late). For each comparison, PECA produced a table of TF-TG

networks that were specific to each group in the comparison, as well as a table of networks that are shared between the compared groups.

For each group-specific TF-TG network matrix, we (1) filtered for TF-TG networks with a fold change > 1.5 , (2) filtered again for TF-TG pairs in which the TG belonged to a WGCNA maturation module gene, and (3) for each group, compiled a list of remaining TFs that were predicted to regulate at least 5 maturation TGs. This yielded a final list of 29 TF candidates (Table S1d) that we predict to regulate astrocyte maturation genes. For each of these TF candidates, we looked at motif enrichment and occurrence, as well as TF expression, across hCO astrocyte maturation. Additionally, we quantified the relative percentage of genes in each maturation module that these TF candidates are predicted to regulate.

Initial processing and analysis of single-nucleus multiome data

We first removed reads mapping to the mitochondrial genome, chromosome Y, and common blacklisted regions. We then used ArchR (v1.0.2) (354) to process fragment data and perform quality control analyses, dimensionality reduction, and clustering. Harmony batch correction was performed prior to unsupervised dimensionality reduction and clustering (355). We filtered out nuclei that had a TSS enrichment score < 5 , less than 3500 unique nuclear fragments, and those without matched RNA reads. Using the ArchR doublet detection tool (default parameters), we removed 2,683 doublets, leaving a total of 27,036 nuclei for further analysis. The remaining nuclei had a median TSS enrichment score of 10.96 and a median of 19,872 fragments per nucleus.

Identifying clusters and cell-type assignments

We first used ArchR to generate gene activity scores, which are correlated with gene expression and calculated based on chromatin accessibility at the gene body, promoter, and distal regulatory regions (354, 356). We identified marker genes for each cluster using the ArchR `getMarkerFeatures()` function (filtering threshold: $FDR \leq 0.01$ & $\log_2FC \geq 1$) and assigned cluster names based on the enrichment of well-known CNS cell type markers. After excluding clusters that had less than 100 nuclei, we identified clusters representing 9 major CNS cell types (Table S3c and S3d).

Inference of astrocyte maturation modules onto single-nucleus multiomic clusters

Maturation module enrichment was computed using the ArchR `addModuleScore()` function with `geneScoreMatrix` using the top 50 genes in each module (Table S3f).

Copy Number Variation (CNV) estimation from single-nuclei RNA-seq data

Genome-wide large-scale chromosomal CNV score was estimated using the ‘inferCNV’ R package (v1.10.1) with default parameters (357). The raw gene expression matrix was extracted from the ArchR project (v1.0.2) (354) project using the function `getMatrixFromProject()` with default parameters. The CNV score per nuclei is defined as the mean squares of CNV values across the genome.

snATAC-seq peak calling and TF motif deviations analysis

ArchR was used to call peaks with default parameters. Briefly, a pseudo-bulk dataset was created for the 2 main cell types- neoplastic and non-neoplastic astrocyte lineage

cells. The reproducible peak sets were analyzed using `addReproduciblePeakSet()` with MACS2 (358) with a fixed-width peak size of 501 bp and iterative overlap peak merging. The resulting `PeakMatrix`, with a total of 194,115 peaks, was used for TF deviation analysis with the default parameter with ArchR's `addDeviationsMatrix()` function. TFs with a high correlation of motif accessibility can be identified based on the correlation of the inferred gene activation score to the motif deviation, thus identifying the known TFs in driving differences among neoplastic vs non-neoplastic lineages.

Identifying unique astrocyte chromatin and gene signatures based on molecular diagnoses for tumors

The ATAC-seq data consisting of 485,908 peaks, and RNA-seq counts data was normalized using the DESeq2 package. Features of interest (peaks for ATAC-seq and genes for RNA-seq data) that are unique to each group were identified using the 'feature binarization' method as outlined in Corces et al. (359). The group details for the 19 samples are shown in Table S4a.

Spatial transcriptomics

Using our maturation gene modules, we evaluated spatial gene expression in two IDHwt samples that were included in Ravi et al. 2022, following their analysis pipeline (294).

5mC/5hmC analysis

Paired-end sequencing reads of MeDIP-seq libraries and 5hmC-capture-seq libraries were first mapped to the human genome (hg38) using Bowtie 2 (2.4.4) (360) with "--no-

mixed --no-discordant” parameters. To identify differential 5mC and 5hmC signal between IDH1mt and IDHwt samples, the human genome sequence was binned to 500 bp bins and reads of each sample were counted in each bin. Read counts were normalized to total mapped reads and bins with averages $RC > 20$ in either IDH1mt or IDHwt groups were used for differential analysis by DESeq2 (361). Significant differential 5mC regions were defined by $FDR < 0.05$ whereas significant differential 5mC Regions were defined by $FDR < 0.001$ and $|\log_{2}FC| > 4$. Differential 5mC and 5hmC regions were annotated to adjacent genes using HOMER (362). To compare gene body 5hmC levels, we used the human gene body coordinates (hg38) downloaded from UCSC Table, and counted gene body 5hmC reads for differential expression analysis using DESeq2. PCA analyses were conducted using the “factoextra” R package. We used ngs.plot and igvtools to compare global and loci-specific 5hmC signals between IDH1mt and IDHwt samples, respectively (363, 364).

Whole tissue bulk RNA-seq

Paired-end RNA-seq reads were first mapped to the human genome (hg38) by TopHat2 (2.1.0) with default parameter (365). Aligned Bam files were used to perform differential expression analysis between IDH1mt and IDHwt triplicates by Cuffdiff (2.2.1) (366). $FDR < 0.05$ were used to define significant differential expressed genes.

Contributions

C.S. and S.A.S conceptualized the scientific ideas and designed the experiments. J.O., K.H., and E.N. provided the surgical GBM sample resections used in this study. C.S. performed hCO formation, hCO and tissue dissociations and immunopanning, nuclei isolations, bulk ATAC- and RNA-seq library preparations and computational analyses. H.V.W performed analysis for single nucleus multiome data. Y.L., F.W., and B.Y. performed the 5mC and 5hmC captures and analyses. P.C. helped with analysis of genotype specific astrocyte signatures. A.S., A.K., and A.V. assisted with single nucleus captures, hCO formation and maintenance, and analytical pipelines, respectively. K.J. and V.R. analyzed the spatial transcriptomic data. C.S. and S.A.S. wrote the initial manuscript, with input from H.V.W., V.C., and B.Y. All co-authors contributed to the final manuscript.

CHAPTER 3: DISCUSSION

3.1 Summary

Glioblastoma is a highly aggressive diffusely infiltrating brain neoplasm that consists of malignant tumor cells reflecting multiple immature cell states, including NPC-like, AC-like, and OPC-like cells. While it is well-accepted that GBM cells hijack normal developmental regulatory programs to drive lineages similar to those in embryonic stages, it remains unclear which aspects of the normal trajectory of glial development are aberrant within these tumors. To uncover the precise phases of glial development that are mirrored within GBM, we first need comprehensive maps of normal glial maturation, an obstacle that has historically been hindered by a lack of access to primary human tissue samples. To overcome this hurdle, we used human iPSC-derived cortical organoids, an in vitro model system that recapitulates many features of human cortical development, including astrocyte and neuronal maturation. The work presented in this dissertation leveraged this model system to precisely track transcriptomic and chromatin accessibility changes across normal human astrocyte maturation and project this developmental trajectory onto the molecular landscape of GBM astrocytes.

Using cortical organoids maintained in culture for nearly two years, we identified temporally graded shifts in gene expression and TF activity across astrocyte maturation. This time course revealed a novel intermediate phase of astrocyte maturation that is enriched for glial lineage-commitment TFs. When projecting these maturation signatures onto bulk and single-nucleus transcriptomic and chromatin accessibility data derived from primary IDHwt GBM tissue samples, we found a distinct activation of the intermediate phase of astrocyte maturation in malignant tumor AS-like cells, whereas

non-malignant astrocytes in the margin reflect a more mature signature. Interestingly, AS-like cells from IDH1mt glioma exhibit a strikingly unique signature, generally reflecting a more mature astrocyte maturation stage, albeit still not as mature as non-neoplastic margin astrocytes. Lastly, we investigated the potential role of IDH1mt-associated DNA hydroxymethylation patterns around maturation genes as a possible protective mechanism for preserving a mature state in IDH1mt tumors.

3.2 Key findings and future directions

3.2.1 Human astrocyte maturation: filling in the gaps

While multiple groups have demonstrated vast transcriptomic, morphological, and functional differences between fetal and adult human astrocytes, the intermediate changes that occur between these timeline extremes were unknown. By maintaining organoid cultures for almost two years, we were able to precisely track maturation events during this elusive developmental window. In this work, we present RNA-seq and ATAC-seq data that reveal a developmental process benchmarked by three distinct phases of maturation— “early”, “middle”, and “late” (Fig. 5). While the “early” and “late” phases best reflect embryonic (17-20 GW) and postnatal/adult maturation signatures (Fig. 6), respectively, the “middle” phase captured in our time course most closely aligns to full gestational and very early postnatal stages of development (230, 235).

Notably, we observed significantly high expression of glial fate-determining TFs during the “middle” phase of maturation, including OLIG1/2, ASCL1, SOX8, and SOX9 (Fig. 5), all of which play important roles in driving both astrocyte and oligodendrocyte lineages.

This suggests that the “middle” stage of maturation may represent a critical developmental step where precursor populations begin to diverge down separate glial lineage trajectories before maturing into quiescent cells. Cells approaching the “middle” stage of maturation may in fact represent a shared glial precursor with bipotent glial commitment capabilities to both AS and OL lineages. Single-cell transcriptomic studies of primary fetal tissue samples from the developing cortex and spinal cord indicate the presence of a shared multi-potent glial intermediate progenitor cell (gIPC) that is EGFR+/OLIG2+/OLIG1+/ASCL1+ (117-119), markers that are highly expressed during the “middle” phase of our maturation timeline. While it is conceivable that we captured a glial progenitor stage in our timeline, most cortical organoid protocols do not consistently generate OL lineage cells, so this might be a challenging hypothesis to confirm in this brain region. However, both astrocytes and oligodendrocytes are present in spinal cord organoids. One future direction could include assessing the presence of the “middle” maturation signature in hindbrain patterned organoids (brainstem or spinal cord) to see if it is enriched in a bipotent progenitor population. Perhaps comparing the “middle” maturation signature from cortical organoids to that in hindbrain organoids or the published transcriptomic signature of glial IPCs could elucidate which factors restrict bipotent glial progenitor population to an astrocytic fate in cortical organoids, which could have important implications for glioma biology.

3.2.2 *What can the principles of astrocyte maturation tell us about GBM biology?*

Our group and others have demonstrated a resemblance between GBM tumor AS-like cells and fetal astrocytes (77); however, exactly where along the astrocyte maturation

spectrum tumor cells reside has yet to be elucidated. For the first time, we projected a complete map of astrocyte maturation onto astrocyte lineage cells from primary IDHwt GBM tumor and margin samples. We found that tumor AS-like cells reflect early to middle maturation stages, where the majority of cells show an enrichment of the “middle” signature (Fig. 12). This is in contrast to margin astrocytes, which more closely reflect “late” molecular profiles (Fig 12). These findings are substantial given that the “middle” signature may represent a multi-potent shared glial progenitor capable of progressing down AS or OL lineages. GBMs present a high degree of cellular plasticity and demonstrate the capacity to transdifferentiate between the two states when xenografted into the mouse brain (see “*Glioblastoma*” section). Tumor cells emulating a phase of maturation where either glial fate is still flexible could explain the presence of both glial signatures in the tumor and the capacity of tumor cells to convert between the two lineages.

Understanding the neurodevelopmental events that coincide with the “middle” maturation stage of astrocyte development could inform us about GBM tumor biology and why it would be advantageous for tumor cells to revert to or get stuck in this particular maturation stage. Perhaps the most prominent developmental change that overlaps with the timing of the “middle” stage of maturation is a peak in synaptogenesis (367). In normal brain development, glial cells play an important role in facilitating synaptogenesis by promoting synapse formation and modulating synaptic signaling (368-370). Likewise, recent work by the Monje group and collaborators demonstrated that glial-facilitated neuronal activity may also promote glioma development (371). Their initial studies found that neuronal activity-dependent secretion of neuroligin-3

(NLGN3) induces PI3K-mTOR activity and leads to increased proliferation and tumor growth in patient-derived pediatric and adult glioma xenograft models (372, 373). More recently, their team discovered that neuronal activity-mediated glioma growth occurs via synaptic and non-synaptic mechanisms that echo normal developmental neuron-glia signaling processes (374). For example, OPCs form synapses with neurons during normal development (375, 376), which subsequently induces progenitor proliferation, differentiation, and survival (377-380). A similar synaptic signaling mechanism is apparent in glioma, where AMPA receptor-dependent neuron-glioma synapses are formed with OPC-like tumor cells (374). The team also found potassium-induced prolonged currents in glioma that were amplified by gap junction-mediated connections (381), a mechanism that is reminiscent of astrocytic glutamate transporter currents and inward potassium currents that result from a flood of extracellular potassium from neurons (382-384). Additionally, neuronal activity-driven glioma growth does not appear to be a unidirectional process, but rather a positive feedback loop where gliomas induce neuronal hyperexcitability through secretion of glutamate and other synaptogenic factors (385-387). In return, neuronal excitation can increase extracellular potassium levels (388), inducing prolonged glioma potassium currents and neuron-to-glioma synaptic signaling. Given this information, the “middle” maturation stage we identified may help promote synaptogenesis in a manner that glioma cells could exploit to promote tumorigenesis.

Altogether, the “middle” stage of maturation that we identified could be a particularly favorable developmental phase for tumor cells to exist in as it may afford them the

ability to transition between glial states and could provide a glial-induced neuronal signaling mechanism to accommodate tumor growth.

3.2.3 *Can maturation TFs be harnessed for therapeutic intervention?*

We demonstrated that several key glial development TFs are highly active in malignant GBM AS-like cells, particularly OLIG1 and OLIG2 (Fig. 12). Both of these TFs play a role in the development of astrocyte and oligodendrocyte lineages, with particular emphasis on OLIG2, which plays a role in AS and OL fate determination (123). OLIG2 is highly expressed in the gIPC cell population from human fetal datasets (118, 119) and has been shown to promote interactions between NFIA and SOX10, which antagonize each other to promote a singular glial fate (123). The high activity of OLIG2 that we observe in both our “middle” stage and in GBM cells is consistent with our hypothesis that this stage may represent a shared glial progenitor that is enriched in a malignant tumor population. Another TF that is active in GBM tumor AS-like cells is RFX4 (Fig. 12). Our dataset of maturing astrocytes within organoids indicates that this TF, which has previously been implicated in establishing AS fate (273), may drive both “early” and “late” stages of astrocyte maturation. RFX4 is also highly expressed in GBM cells and significantly correlates with a poor patient prognosis, indicating it may play a role in tumor pathogenesis (207). Together, this suggests that RFX4 may activate both immature and mature astrocytic gene sets and when expressed in GBM might favor a more immature signature. There is some evidence suggesting that different RFX4 isoforms are expressed in glioma as compared to normal brain development (389). One way to functionally test if each of these different isoforms activate separate maturation

gene sets would involve overexpressing each of the two isoforms separately during astrogenesis within organoids to assess their maturation capabilities. Additionally, long-read sequencing could be employed to determine if unique RFX4 isoforms are expressed at early and late developmental time points. One might predict in such a setting that the early developmental RFX4 isoform might be differentially expressed in GBM tumors compared to margin.

Understanding which TF programs need to be activated and deactivated to promote permanent quiescence is a critical hurdle for the differentiation therapy field. Previous approaches targeting developmental signaling pathways, such as BMP4 treatment, have yielded inconsistent results, where in some cases GSCs are able to revert back to a stem-like state (243, 254). In this instance, the observed dedifferentiation was the result of incomplete chromatin accessibility changes that allowed for the maintenance of an immature gene program (254). To overcome this hurdle, one approach is to target a specific class of TFs called pioneer TFs, which play a substantial role in neurodevelopment and cell reprogramming. Pioneer TFs can bind at condensed nucleosomal DNA regions, recruiting chromatin remodelers, cofactors, and additional TFs, effectively changing the local epigenomic landscape (390). Pioneer TFs are frequently perturbed in glioma, including OCT3/4, SOX2, KLF4, P53, and ASCL1, several of which are active in our GBM astrocyte dataset (Fig. 12 and 14) (247). OCT3/4, SOX2, and KLF4, also known as the OSK trinity, is a powerful trio that promotes the maintenance of a pluripotent state and self-renewal (391-393). Higher expression of these three TFs corresponds to greater tumorigenesis and maintenance of GSCs in vitro (170, 394-397). ASCL1 is a pioneer TF that is also important in early neurodevelopment

and is associated with Wnt and Notch signaling, promoting neurogenesis and a proneural transcriptional state in GSCs (169, 248). ASCL1 levels in GBM positively correlate with SOX2 and OLIG2 levels, two factors that are important in driving the classical (more AC-like) or proneural (OPC-like) GSC states, respectively (248). Additionally, ASCL1 appears to regulate the expression of OLIG1/2, SOX2, NFIA/B/ X, and POU3F2/3, which are upregulated in GBM and play vital roles in driving early glial programs (248).

However, targeting pioneer TFs is not trivial. One approach is the delivery of miRNAs that target aberrantly expressed pioneer TFs. One particularly successful example of this is treatment with miR145 via a polyurethane-short branch polyethylenimine (PU-PEI) delivery vehicle, which suppressed Oct4 and Sox2 levels in GSCs, suppressed tumorigenesis in vitro and in vivo, and increased GSC sensitivity to radiation and TMZ treatment (398). Several chemical compounds have also been identified that can relieve the effects of impaired pioneer TF activity, including COTI-2 {E)-N'-(6,7-dihydroquinolin-8(5H)-ylidene)-4-(pyridin-2-yl)piperazine-1-carbothiohydrazide and APR-246 {(2-(hydroxymethyl)-2 (methoxymethyl)quinuclidin-3-one, which helped to reshape the p53 mutant protein form back to its wildtype form to partially rescue its tumor suppressor function (399). Both of these compounds are currently undergoing testing in phase I and phase II clinical trials, respectively (399).

While differentiation therapy is an attractive idea because it targets the developmental roots of tumorigenesis, there are several concerns regarding the efficacy of this therapeutic approach. One complication is the cellular heterogeneity in GBM and other

gliomas. This raises questions about whether it is possible to identify targets capable of pushing all malignant cells to differentiate and the consequences of missing even just a small percentage of these cells. Given the aggressive nature of GSCs, it is likely that leaving remnant malignant cells could result in future tumor growth. While it seems unlikely that there is a pan-differentiation factor that could be exploited to target all cells, perhaps focusing on lineages that have the highest proliferative and migratory capacity could at least slow tumor reoccurrence and prolong patient survival. Another question to consider is the biological consequences of essentially generating new populations of differentiated cell types in the adult brain. Assuming malignant tumor cells could be terminally differentiated, how would those cells integrate into and impact healthy tissue? It is unclear whether these newly differentiated cells could influence neural circuitry and activity levels or induce neuroinflammation; however, these second-hand effects must be accounted for in differentiation therapy pursuits.

3.2.4 *Revisiting the GBM cell of origin*

The work presented in this dissertation informs us about GBM tumor biology (namely, which astrocyte developmental stage is reflected in malignant tumor cells), the possible biological consequences of tumor cells existing in this developmental stage, and which molecular cues could be driving tumor cells to accumulate in this particular phase of development. However, this work does not seek to address the cellular origins of GBM—whether tumors are driven by a dedifferentiated quiescent cell or a remnant proliferative population—a question that will likely require elegant lineage tracing experiments and analytical tools to investigate. Prospective lineage tracing approaches

involve the labeling of a founder cell and tracking its descendants (400). One method for doing so in rodent models, mosaic analysis with double markers (MADM), utilizes a Cre-loxP genetic recombination system with dual fluorescent reporters to trace recombination events and decipher lineage relationships (401). This method was employed to investigate glioma cell of origin by inducing concurrent Tp53/Nf1 mutations in NSCs in mice, which led to substantial expansion of an OPC-like population, but not NSCs or other typical NSC descendants, suggesting that OPCs could serve as the cell of origin for glioma with Tp53 and Nf1 mutations (205).

The ideal scenario for lineage tracing experiments is to pair a prospective method with a retrospective approach, which leverages naturally occurring somatic mutations, such as long interspersed nuclear element-1 (LINE-1) retrotransposition events, copy-number variants (CNV), single-nucleotide variants (SNV), and microsatellite length variants, to delineate cellular relationships (400). Retrospective lineage tracing methods are particularly useful for investigating tumor evolution by tracking cell subclones that are generated by varying patterns in mutational stasis and expansion. Several single-cell transcriptomic studies of GBM have attempted to use these methods to unravel tumor cell phylogenies by looking at CNVs and reconstructing the most parsimonious lineage between CNV events (262). Others have tried RNA velocity, a method to extrapolate temporal state changes by looking at splicing and mRNA degradation across cells (402), and pseudotime analyses, which aim to order cells along a continuous trajectory based on similarities and differences in gene expression (403) to infer GBM differentiation hierarchies (263). While all of these methods have thus far implicated an early glial progenitor cell type at the apex of the GBM developmental hierarchy (262, 263), there

do not seem to be clear differentiation paths that are consistent across varying patient samples. Epigenetic profiling may be more informative for deciphering the cellular roots of GBM because purely transcriptomic information captures only a snapshot of cell state, whereas epigenetic signatures are stable, heritable, and often foreshadow changes in cell state (404-406). Coupling epigenetic and transcriptomic profiling has been particularly effective for delineating the developmental hierarchy within diffuse pediatric gliomas and provided striking evidence for different OPC-like cells of origin for H3.1 and H3.3 K27-gliomas (199). However, further exploration into the GBM cell of origin should integrate both retrospective and prospective experiments to both account for the complexity of patient tumors, while simultaneously validating findings through functional studies.

While we cannot definitively draw conclusions about the GBM cell of origin from the data presented in this dissertation, it does seem to support one particular hypothesis. OPCs are a highly likely candidate for driving gliomagenesis, as OPCs are the most abundant proliferative cell type in the adult brain. This bestows a greater endogenous potential than quiescent cell types for acquiring driver oncogene and/or tumor suppressor mutations that could facilitate tumor initiation. Additionally, there is now mounting evidence suggesting that OPCs, or a subpopulation of NG2+ OPCs, can give rise to both OL and AS lineages, both of which are present in the tumor bulk (see section “Evidence of a shared glial precursor”). Given that our “middle” astrocyte maturation signature appears to share many features with published multipotent glial progenitor populations, it seems reasonable that we would see an enrichment of this signature in GBM malignant cells if they were in fact an OPC-like population of malignant cells that

could drive multiple glial lineages in the tumor. That said, while far less abundant, there is some evidence suggesting that quiescent mature astrocytes can dedifferentiate to a NSC-like state if they acquire potent mutations (TP53), followed by a second neuroinflammatory hit (162). In this cell of origin scenario, it is peculiar that we would see an accumulation of malignant cells in a “middle” maturation phase and not an earlier developmental stage. This might suggest that the “middle” phase represents a preferential attractor cell state with a molecular program that allows malignant cells to thrive in a harsh TME.

3.2.5 *Developmental differences across glioma subtypes*

While investigating subtleties in astrocyte maturation signatures across tumors with varying molecular diagnostic backgrounds, we identified a striking difference between astrocytes from IDH-wildtype (IDHwt) and IDH1-mutant (IDH1mt) tumors (Fig. 15). This finding was consistent with the recent reclassification of IDHmt gliomas by the World Health Organization, which now categorizes IDHmt glioma as a separate class from IDHwt GBM based on distinct histology, molecular diagnostic features, and a slightly better prognosis (180, 181).

When comparing astrocytes from IDHwt and IDH1mt tumors, we found that astrocytes from the former class were enriched for immature markers, more specifically, “early” – “middle” astrocyte gene signatures. This finding is particularly interesting when considering the differences in cell developmental hierarchies between the two types of tumors. Venteicher et al recently demonstrated that IDHmt glioma possess a hierarchy

that is quite distinct from IDHwt tumors. Using single-cell transcriptomic profiling they suggested that IDHmt gliomas contain a relatively small pool of proliferative cells resembling NSCs that bifurcate into both AS-like and OL-like lineages (184). This is distinct from IDHwt GBMs where lineage trajectories are more mixed and disordered. Chaligne et al confirmed this same hierarchical structure in IDHmt tumors by employing single-cell joint profiling of transcriptional, genetic, and epigenetic signatures across a cohort of diffuse gliomas (407). Interestingly, AS and OL lineages in the IDHmt tumors are differentially enriched for distinct CNVs and marked by relatively low dedifferentiation rates (407). This is in contrast to IDHwt GBM, where phylogenies are dictated by heritable passenger DNA methylation events; however, in their datasets individual clades did not demonstrate enrichment for specific cell states (OPC-like, AC-like, NPC-like, MES-like) and dedifferentiation rates were much greater (407). Together, these data suggest that in IDHmt tumors, AS and OL lineage commitment is more stable, unlike in IDHwt tumors, where plasticity and dedifferentiation prevail and promote changes in cell state. Our data corroborate these findings as the enrichment of immature astrocyte maturation signatures, and particularly the “middle” signature, in astrocytes from IDHwt tumors might be indicative of a multipotent AS-lineage that is more plastic than IDH1mt cells that are farther along the differentiation spectrum.

When considering why astrocytes from IDH1mt tumors may be protected from unpredictable changes in cell state, we turned to a key feature of these tumors that results from the IDH mutation—aberrant DNA methylation patterns. Most IDH mutations in cancer cause the IDH protein to lose its affinity for isocitrate, resulting in an accumulation of the oncometabolite D-2-hydroxyglutarate (D-2-HG) instead of α -

ketoglutarate (α -KG) (309, 326). The switch from producing α -KG to D-2-HG has many biological consequences, including the disruption of DNA demethylation through impaired ten-eleven translocation (TET) enzyme activity (311). In the typical demethylation process, TET enzymes oxidize 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) followed by several additional oxidation steps as part of active cytosine demethylation (310). While it is well known that altered DNA methylation is a pathological feature of glioma harboring the IDH mutation (315, 326), it is unclear how this may play a role in glial cell differentiation and ultimately, drive a developmental hierarchy that is different from what is observed in IDHwt glioma. In the work presented here, we investigated differences in 5mC and 5hmC accumulation in immature and mature gene sets and found a positive association between 5hmC gene body levels and expression of maturation gene sets (Fig. 18). Notably, regardless of gene expression levels, gene body 5mC was consistently higher in IDH-mutant samples, even for mature genes, which were more highly expressed in IDHmut samples (Fig 19).

Other groups have demonstrated a strong positive correlation between gene body 5hmC and the expression of cell type-specific gene programs (330). However, it is not yet clear if this association between expression and gene body 5hmC is a byproduct of transcription or if 5hmC deposition may play an active role in facilitating transcription. It is possible that gene body hydroxymethylation helps facilitate transcriptional elongation, as it has been shown to co-localize with RNA Pol II and may help facilitate Pol sliding (332, 408). However, a study by Colquitt et al proposes an alternative role for gene body 5hmC in establishing cell identity and state changes. The authors investigated the consequences of activating DNA hydroxymethylation by inducing Tet3 expression,

which oxidizes 5mC (333). After overexpressing Tet3 in mature olfactory sensory neurons, the authors found that exonic and intronic regions that initially had moderate 5hmC levels exhibited only a slight gain of the mark, whereas regions that had high levels of 5hmC exhibited a significant reduction (328). However, the loss of 5hmC did not result in the enrichment of additionally oxidized forms, but rather a return to an unmodified state (328). Interestingly, genes that were downregulated as a result of Tet3 overexpression also showed a significant depletion of gene body 5hmC, whereas genes that were upregulated only displayed a slight increase in gene body 5hmC (328). These results suggest that gene body 5hmC may play a more substantial role in maintaining constant levels of transcriptional expression within mature postmitotic populations, rather than initiating the activation of de novo gene sets.

One possible scenario where this hypothesis converges on our observations with IDHmut and IDHwt glioma is that a loss of 5hmC at immature genes in the IDH1mut could result in decreased expression of those gene sets and a transcriptional state that is not as immature as in IDHwt tumors. Likewise, the maintenance of 5hmC at mature genes may help sustain the expression of those genes in IDHmut cells, promoting a more mature phenotype. To test this hypothesis, one future experiment could be to parallel the approach of Colquitt et al and see if depleting 5hmC in immature astrocytes results in a subsequent decrease in immature gene expression. A follow-up experiment might be to then expose immature astrocytes to D-2-HG or introduce the IDH mutation to determine if this mechanism is downstream of the IDH driver mutation.

Alternatively, it is possible that the loss of 5hmC is driven by another factor unique to IDHmt brain tumors beyond the IDH mutation itself, such as additional passenger

mutations, TME niches, or developmental origins. Additional epigenetic mechanisms may also be at play, including histone methylation, which is also impacted by D2HG accumulation. The role of histone modifications in regulating a more mature signature in IDH-mutant glioma should be further explored as it has for other diffuse gliomas (409).

3.3 Conclusion

Comprehensive timelines of glial lineage development could greatly inform our understanding of glioma resilience and highlight new angles for therapeutic exploration. However, the field is still young, and we are continuing to generate human-based systems that will allow us to map precise changes in molecular and functional programs throughout neurodevelopment. In this dissertation, we leveraged hiPSC-derived cortical organoids to chronicle transcriptome- and chromatin-level changes across normal human astrocyte maturation. We identified a novel intermediate developmental phase of astrocyte maturation that may reflect an elusive stage of glial fate determination. Projecting this maturation map onto astrocyte-like cells from primary GBM tumor samples indicated that this intermediate signature is highly enriched in malignant cells, which may play a key role in promoting a plastic and adaptable state. Lastly, we demonstrated that this intermediate astrocyte maturation signature was noticeably absent from IDH1mt tumors, which may at least in some part arise due to differences in DNA hydroxymethylation at key maturation gene sets. Together, these findings inform how glial maturation state may contribute to glioma differentiation and heterogeneity

and reveal abundant new targets to explore for coercing a mature and quiescent phenotype.

APPENDIX: ONGOING AND FUTURE FUNCTIONAL EXPERIMENTS

Abstract

This section of my dissertation builds on our findings in Chapter 2 and includes preliminary results for experiments that investigate mechanisms of astrocyte maturation. First, we outline our approach for generating stable iPSC lines to overexpress candidate maturation TFs in hCSs at specific developmental time points to assess their role(s) in regulating astrocyte maturation. Next, we map our astrocyte maturation modules onto published human prenatal neurodevelopment datasets, highlighting specific glial progenitor populations that are captured in our hCS trajectory. Lastly, we investigate how an oncometabolite produced by IDH-mutant gliomas, D2HG, affects maturation and quiescence when added to astrocytes purified from primary fetal cortical tissue. While these experiments and analyses are still ongoing, the evidence presented in this portion of my dissertation supports our findings in Chapter 2 and will guide further functional testing.

Introduction

In Chapter 2 of this dissertation, we identified maturation gene programs in hCS astrocytes and demonstrated that specific stages of this trajectory are enriched in GBM astrocytes. Here, we present preliminary findings from new and ongoing investigations into the mechanisms that drive astrocyte maturation and the biological relevance of these maturation stages in human cortical development. Our major focus has been (1) testing the functional role of predicted driver TFs in astrocyte development and maturation, (2) benchmarking hCS astrocyte maturation stages against single cell datasets of the developing prenatal cortex, and (3) determining if oncogenic features directly impact astrocyte maturation programs.

In this section, we first test the functional consequences of overexpressing predicted maturation TF candidates (identified in Chapter 2) in hCSs. We also explore how our hCS astrocyte maturation trajectory is reflected in an existing human cortical development dataset, demonstrating that “middle” and “late” maturation programs are enriched in unique astrocyte and glial progenitor prenatal cell populations. Lastly, we expand upon our previous analysis that showed that astrocytes from IDH1mt tumors are more mature than astrocytes from IDHwt tumors. We generated new evidence that D2HG, the oncometabolite produced by IDH-mutant tumors, represses immature genes that are upregulated in IDHwt tumors and inhibits cell proliferation.

While these experiments and analyses are not yet complete, the preliminary results in this section support our findings in Chapter 2 and will help provide mechanistic insight into regulators of astrocyte maturation in healthy and oncogenic contexts.

Results

Overexpressing TF candidates in hCSs

To test how our predicted maturation TF candidates contribute to astrocyte maturation in vitro, we sought to create stable iPSC lines that would allow us to overexpress TF candidates at specific developmental time points. We generated constructs that (1) express a TF of interest—RFX4, ASCL1, SOX21, and NR3C2—and GFP under a doxycycline-inducible promoter and (2) puromycin resistance under the constitutive hPGK promoter to act as a selectable marker (**Fig. 20**). We started by testing a construct for the TF RFX4. We generated lentivirus containing this transfer plasmid, infected C4.1 iPSCs, and treated with puromycin to select for successfully integrated cells (**Fig. 21a**). After expanding surviving iPSC colonies, we performed a serial dilution and identified five isolated clones for further expansion and evaluation (**Fig. 21a**). By DNA PCR, four of the five clones contained the GFP sequence; however, one clone—RBG—demonstrated substantially more GFP-positive cells and greater RFX4 expression upon doxycycline treatment (**Fig. 21b-d**).

We used the RBG RFX4 iPSC clone (hereto referred as RBG) to form hCSs, which we plan to treat with doxycycline at timepoints either before or after endogenous hCS astrogenesis. Inducing early RFX4 expression provides an opportunity to test whether RFX4 is sufficient to induce precocious astrogenesis, whereas later timepoints will test whether RFX4 impacts astrocyte maturation once lineage commitment has already occurred. (**Fig. 22a and 22b**). We treated RBG hCS cultures with media supplemented

with doxycycline for 10 days and observed GFP-positive cells only in cultures treated with doxycycline (**Fig. 22c**). Additionally, in preliminary testing, doxycycline-treated hCS cultures exhibit greater numbers of GFAP- and RFX4-positive cells than in hCSs that did not receive doxycycline (**Fig. 22d**). These RBG hCSs are currently being cultured until the appropriate age for our experimental paradigm.

Mapping molecular maturation programs onto published fetal datasets

To determine how hCS maturation modules are reflected across in vivo human cortical development, we assessed the expression of maturation gene modules and predicted candidate maturation TFs in a published human fetal cortical snRNA-seq dataset (117). Specifically, we focused on data generated from nuclei extracted from micro-dissected prenatal cortical plate tissue procured during the second and third trimesters (17-41 GW). Following the described methods in Ramos et al. 2022, we identified nuclei clusters reflecting major neurodevelopmental cell types, including: transit-amplifying cells (TACs), oligodendrocyte precursor cells (OPCs), astrocytes, neural intermediate progenitor cells (nIPCs), additional differentiated neuronal subtypes, and a glial intermediate progenitor cell (gIPC) population (**Fig. 23a**).

We next subsetted this dataset to include glial populations and relevant precursor cell types for further analysis (**Fig. 23b**). We evaluated the aggregated expression of maturation gene modules across clusters and pseudotime and observed an enrichment of the “middle” maturation gene signature in a small population of gIPCs, as well as one of two populations of astrocytes (**Fig. 23c and 23d**). In contrast, the “middle/late” and

“late” signatures are more highly expressed in a second astrocyte population that appears to have a slightly higher pseudotime score, as well as in gIPCs and OPCs (**Fig. 23c and 23d**).

Additionally, we examined the expression of middle and late TF candidates identified in Chapter 2 in the primary human fetal dataset. Our analysis indicates that a group of “middle” TFs, including ASCL1, OLIG1, and OLIG2, are more specifically enriched in the gIPC population, whereas other TFs—LHX2 and FOXP1—are more broadly expressed across glial precursor populations (**Fig. 23e**). EOMES, a TF previously thought to be specific to neural progenitor cells, also appears to be expressed in subpopulations of gIPCs and astrocytes, where the “middle” gene module is enriched (**Fig. 23e**). Uniquely, RFX4 is very specific to one of the two astrocyte populations that highly expresses “middle/late” and “late” maturation modules (**Fig. 23e**). Lastly, NR3C2, one of the more mature TF candidates is highly expressed by a population of astrocytes, gIPCs, and OPCs, where “middle/late” and “late” maturation modules are highly expressed (**Fig. 23e**). Together, these findings support our previous classification of maturation TFs, with gIPCs demonstrating high expression of “middle” TFs and later astrocyte populations showing higher expression of “late” maturation TFs.

The effects of D2HG on astrocyte maturation

In Chapter 2 we demonstrate that astrocytes from IDH1mt tumors reflect a more mature transcriptomic signature compared to astrocytes from IDHwt tumors. This finding was

particularly interesting given that patients with IDH-mutant gliomas tend to have a more favorable prognosis and a higher percentage of differentiated astrocytes (184, 185). To assess how the IDH1 mutation affects astrocyte maturation, we exposed human fetal astrocytes to D2HG, the oncometabolite that is produced in IDH-mutant tumors. We purified astrocytes from primary human fetal cortical tissue and cultured cells in either (1) un-supplemented astrocyte growth media (AGM), (2) AGM supplemented with DMSO, or (3) AGM supplemented with D2GH (**Fig. 24a**). Cells exposed to 0.5 mM and 1.0 mM D2HG died after only 2 - 3 days, so we performed experiments with 0.1 mM D2HG and the corresponding DMSO concentration (**Fig. 24b**).

We identified hundreds of DEGs between astrocytes exposed to D2HG and DMSO, with DMSO-exposed cells demonstrating few gene expression differences compared to untreated cells (**Fig. 24c-e**). Astrocytes exposed to D2HG demonstrate a downregulation of immature astrocyte genes and a slight enrichment of mature genes compared to the vehicle control group (**Fig. 24d**). Additionally, we asked whether genes that are differentially expressed between primary IDHwt and IDH1mt tumors are impacted by D2HG treatment. The presence of D2HG significantly inhibited IDHwt gene expression, including genes involved with proliferation and stemness (**Fig. 24e and 24f**). These findings were further supported by immunocytochemistry experiments where fetal astrocytes were given a 24-hr EdU pulse on the final day of D2HG and DMSO treatment. Notably, the D2HG condition demonstrates substantially fewer actively dividing cells compared to the vehicle control group and untreated cells (**Fig. 24g and 24h**). Across all three conditions, there is an abundance of GFAP-positive cells

that do not appear to have substantial differences in morphology (**Fig. 24g**); however, a more detailed analysis is prudent for confirmation.

Figure 20: Vector maps for overexpressing maturation TFs

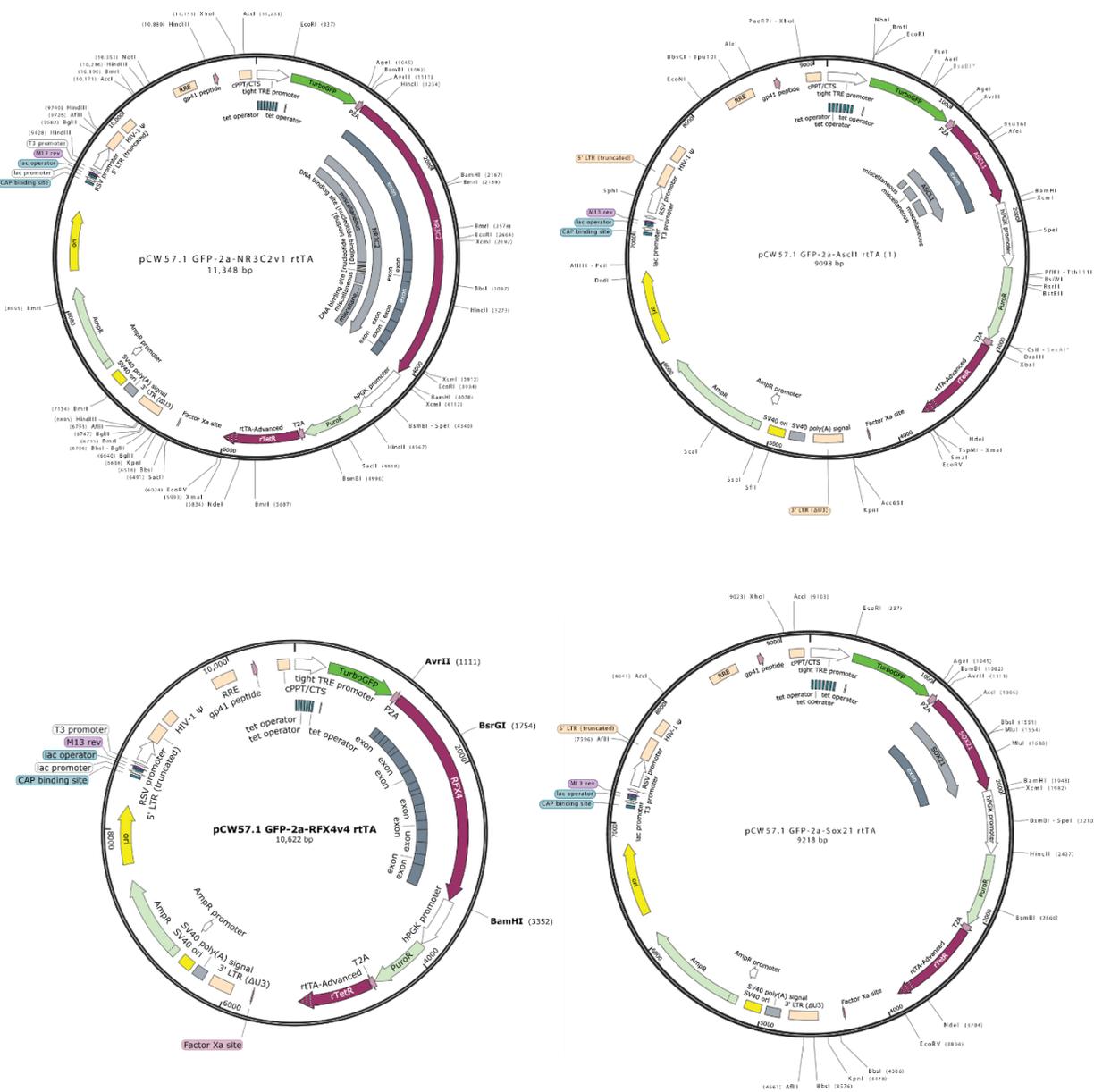


Fig. 20: Vector maps for overexpressing maturation TFs

Vector maps for the doxycycline-inducible lentiviruses to overexpress NR3C2 (top, left), ASCL1 (top, right), RFX4 (bottom, left), and SOX21 (bottom, right).

Figure 21: Approach for overexpressing maturation TFs

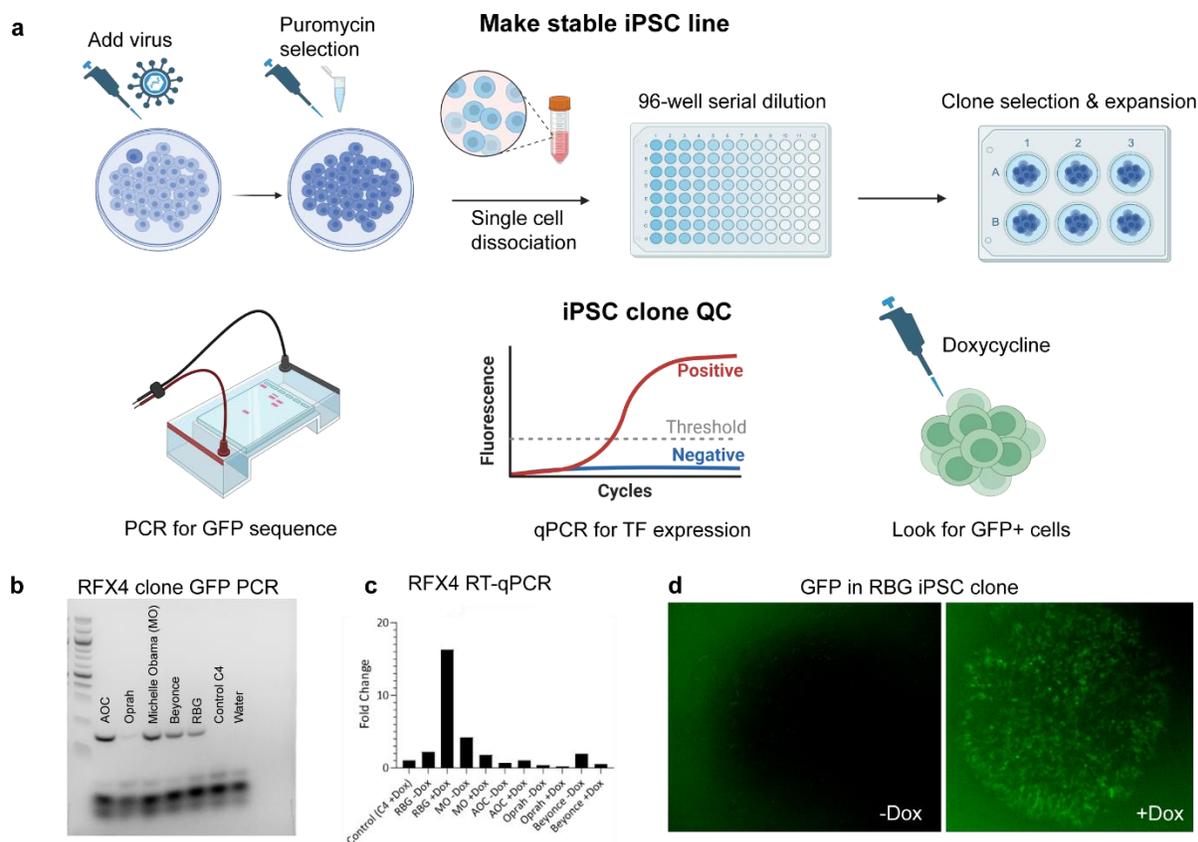


Fig. 21: Approach for overexpressing maturation TFs

(a) Schematic outlining method for generating stable iPSC lines and assessing clone quality. (b-d) QC for inducible RFX4 iPSC clones. (b) PCR gel depicting presence or absence of GFP sequence across selected clones. (c) RT-qPCR of RFX4 across iPSC clones treated with (+) and without (-) doxycycline. (d) GFP signal in selected “RBG” iPSC clone treated with (right) and without (left) doxycycline.

Figure 22: Test induction of RFX4 in hCSs

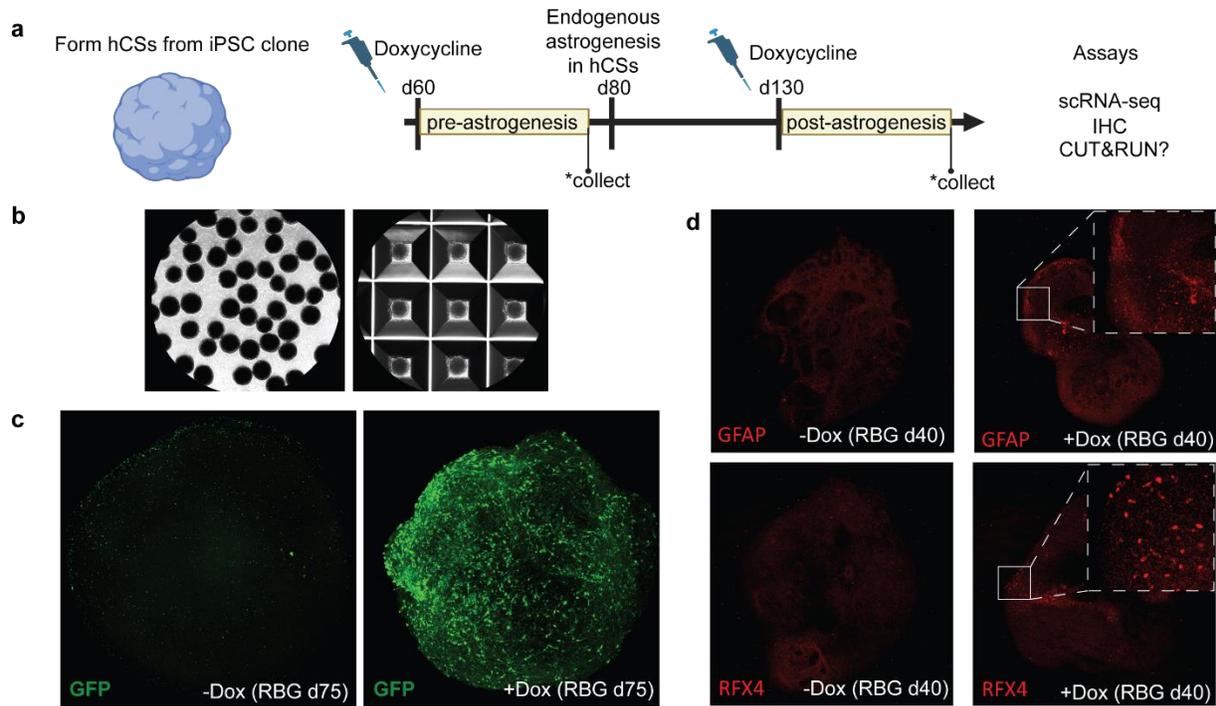


Fig. 22: Test induction of RFX4 in hCSs

(a) Schematic of experimental paradigm using (b) hCSs generated from selected iPSC clones. (c) GFP signal in hCSs formed from RBG RFX4 iPSC clone with hCSs treated with (right) and without (left) doxycycline for 10 days. (d) GFAP (top) and RFX4 (bottom) expression in RBG hCSs treated with (right) and without (left) doxycycline.

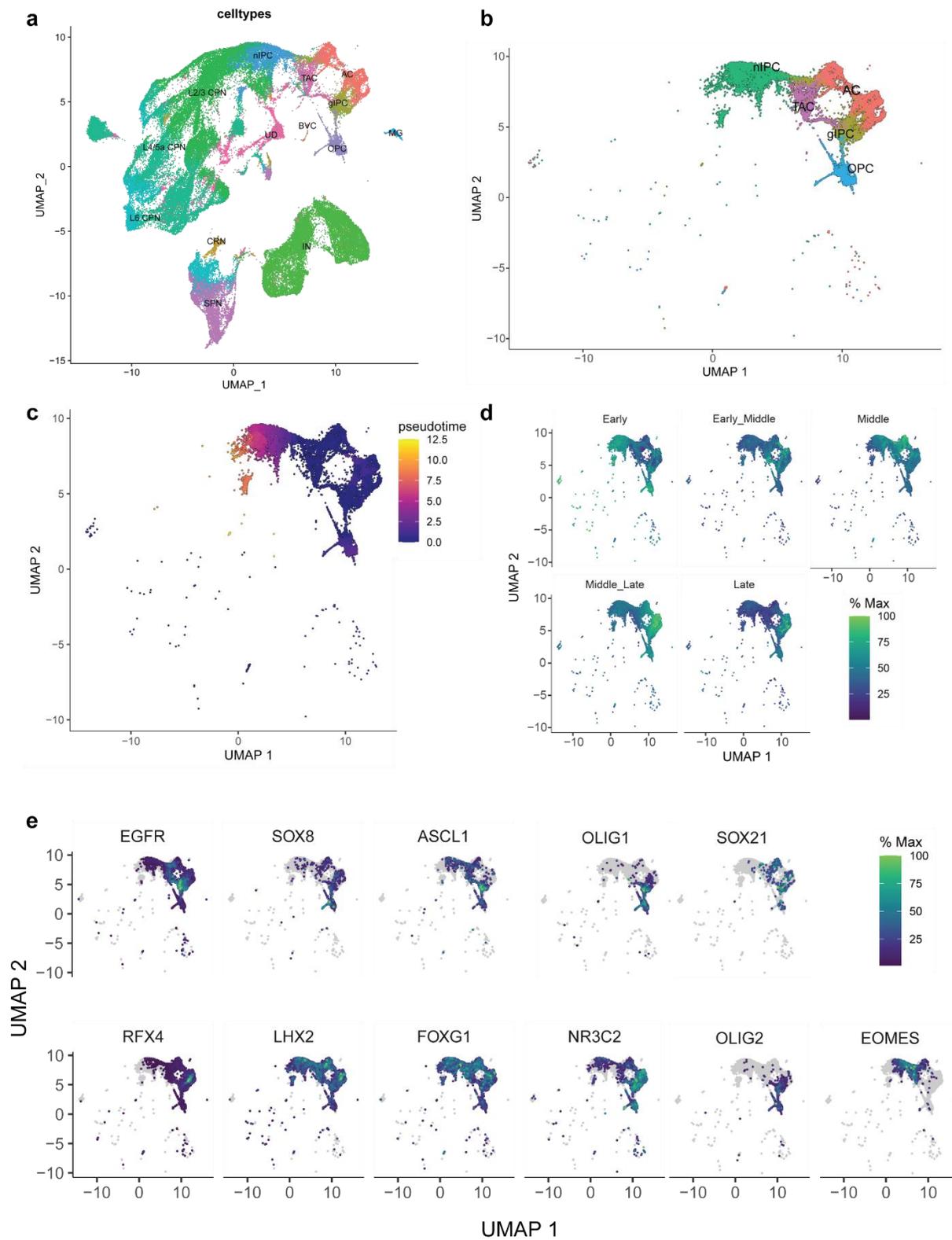
Figure 23: Mapping maturation modules onto fetal datasets

Fig. 23: Mapping maturation modules onto fetal datasets

(a) Cell clusters from Ramos et al. 2022 human fetal cortical plate dataset. (b) Subsetted nuclei for those in early glial and neural progenitor clusters. (c) Pseudotime across progenitor clusters. (d) Expression of “early”, “early/middle”, “middle”, “middle/late”, and “late” maturation gene modules in progenitor clusters. (e) Expression of PECA-identified maturation TFs across progenitor clusters. Astrocyte (AC), oligodendrocyte precursor cell (OPC), glial intermediate progenitor cell (gIPC), neural intermediate progenitor cell (nIPC), transit-amplifying cell (TAC).

Figure 24: The effects of D2HG on astrocyte maturity

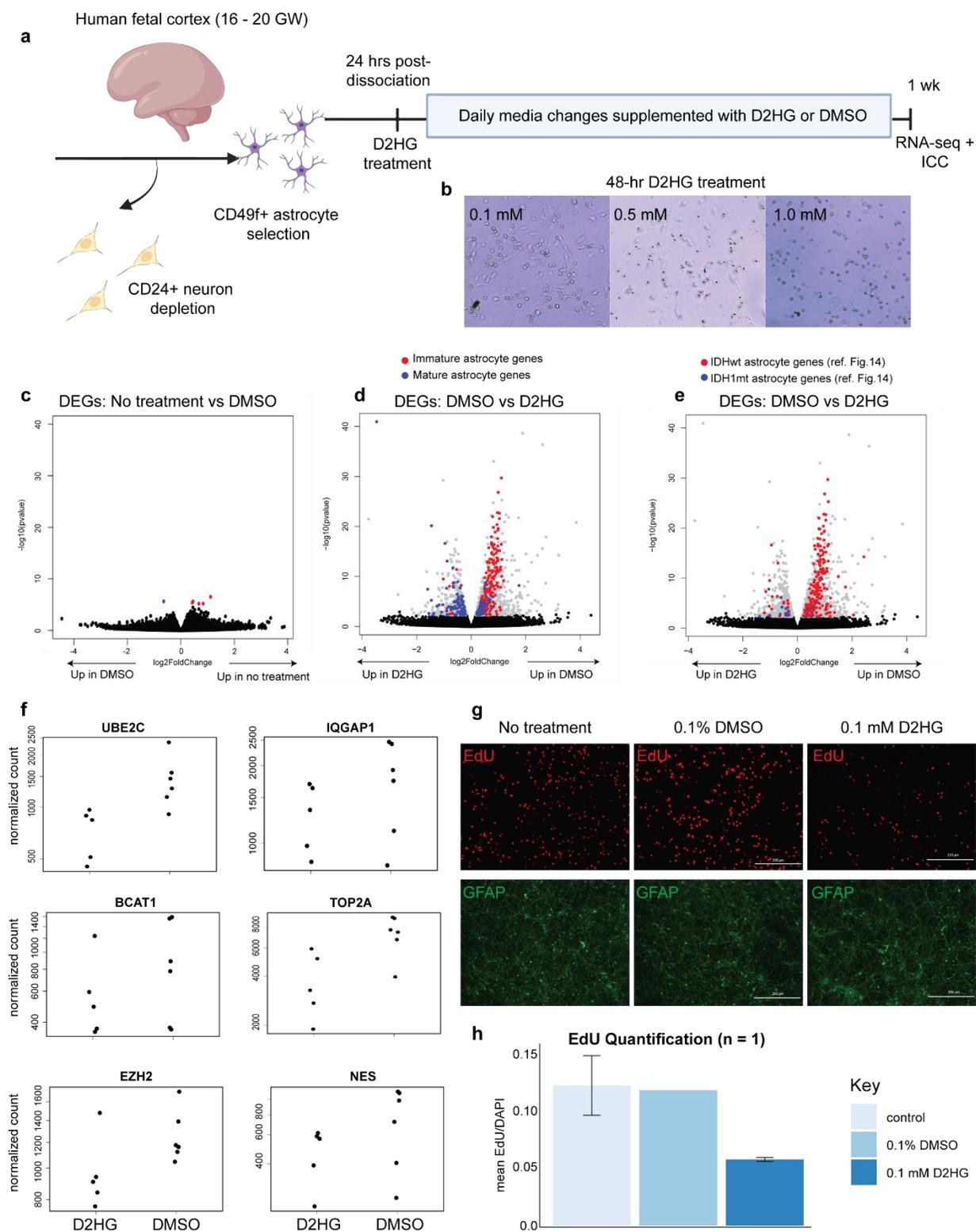


Fig. 24: The effects of D2HG on astrocyte maturity

(a) Schematic of experimental paradigm where astrocytes were purified from human fetal cortex and treated with D2HG or DMSO. (b) Brightfield images of fetal astrocytes treated with (left) 0.1 mM, (middle) 0.5 mM, and (right) 1.0 mM D2HG for 48 hrs. (c) DEGs between astrocytes that received normal media (no treatment) and media supplemented with 0.1% DMSO for 1 wk. (d and e) DEGs between astrocytes that received media supplemented with 0.1% DMSO and 0.1 mM D2HG for 1 wk. (d) Mature and immature DEGs are colored in blue and red, respectively. (e) DEGs that are upregulated in IDHwt and IDH1mt tumors (ref. Fig. 15) are colored in red and blue, respectively. (f) Dot plots of normalized expression of representative DEGs, where each dot represents one biological replicate. (g) EdU (red) and GFAP (green) in fetal astrocytes that received (left) normal media, (middle) 0.1% DMSO, and (right) 0.1 mM D2HG for 1 wk. (h) Bar graph showing mean EdU+ cells normalized to total DAPI. Quantification is for one of the biological replicates and error bars correspond to standard deviation across technical replicates.

Discussion

The experiments and analyses presented in this section aim to provide more mechanistic and biological insight into our findings regarding astrocyte maturation in Chapter 2. Specifically, we sought to validate several of our predicted maturation TF candidates in the hCS model system and identify how our maturation trajectory aligns with in vivo glial development in the prenatal cortex. Additionally, we build upon our previous finding that astrocyte maturation state greatly differs between IDHwt and IDH1mt tumors, and present evidence directly implicating the D2HG oncometabolite in coercing a quiescent state.

The preliminary experiments with our RFX4 viral construct indicate that we successfully developed an approach for directly testing the contribution of individual TFs to astrocyte development and maturation. We are in the process of aging our RBG hCS cultures and will administer doxycycline at pre- and post-astrogenesis (which occurs endogenously around d80) time points to determine if RFX4 expression plays a role in initiating astrogenesis and/or driving maturation. Now that we have established a protocol for making the iPSC stable lines, we are currently working with Emory's Stem Cell Core to generate additional lines and have two (RFX4 and ACSL1) out of the four already validated.

We also tested whether our new astrocyte maturation trajectory is supported by existing datasets of in vivo human cortical development. In particular, we wanted to see if our maturation programs are enriched in specific glial progenitor subpopulations. We used one of the only published human single cell datasets that captures cells from the elusive

third trimester of prenatal development, published in Ramos et al. 2022. Interestingly, our novel intermediate stage of maturation appears to be expressed by a specific prenatal astrocyte population and small subpopulation of gIPCs, whereas the “middle/late” and “late” gene programs are enriched in separate astrocyte and gIPC populations. Two of our predicted “late” maturation TFs—NR3C2 and RFX4—are also enriched in the same astrocyte population as the “middle/late” and “late” gene programs. While some of our predicted “middle” stage TFs (ASCL1 and OLIG1/2) are enriched in the larger of the two gIPC populations, others (EOMES and FOXG1) are more highly expressed in the smaller subpopulation of gIPCs and astrocytes. Together, this suggests that our “middle” stage of maturation captured in hCSs may correspond to early gIPC and astrocyte populations, while the “late” stages correspond to an older population of astrocytes, defined by their slightly higher pseudotime score. Partitioning these subpopulations of glial cells by the age of tissue they originate from would help confirm if the subpopulations of astrocytes and gIPCs with differential maturation profiles are specific to different developmental windows (ex. second vs third trimester).

In addition to identifying TF and gene programs that play key roles in the normal astrocyte maturation trajectory, we have also demonstrated that a specific glioma feature, mutations in the gene IDH1, may contribute to astrocyte maturation status in high grade gliomas. In Chapter 2 we showed that astrocytes from primary IDHwt and IDH1mt tumors have stark transcriptomic differences, with IDHwt tumors harboring a significantly more immature profile and cells from IDH1mt tumors appearing more mature. We next tested whether a key feature of IDH1mt tumors, the production of D2HG, contributes to the preservation of a more mature astrocyte state. Our data

suggests that in only one week, D2HG inhibits the expression of immature genes that are highly upregulated in IDHwt tumors and reduced overall fetal cell proliferative capacity. Thus, while IDH1/2 mutations are thought to be driver tumor mutations, the consequence of D2HG production does appear to promote a more quiescent cell state, which could explain why the IDH mutation is a favorable marker in gliomas.

While it is not yet clear through which mechanism(s) D2HG acts to induce cell quiescence, there are a few possibilities supported by our data. D2HG acts as a competitive inhibitor of α -KG, which regulates over 60 dioxygenases, including JmjC domain-containing histone demethylases (KDMs) and TET enzymes, responsible for DNA and histone demethylation (311, 410, 411). While multiple groups have explored how aberrant DNA methylation patterns may contribute to a more differentiated cell state in IDHmt tumors (including in Chapter 2 of this dissertation), there is far less research on differences in histone methylation patterns between IDHwt and IDHmt tumors and how this may contribute to cell maturation in IDH-mutant gliomas.

Interestingly, a recent study demonstrated that ONC201, a drug that improves outcomes for H3K27M-mutant diffuse midline glioma (DMG) patients, appears to increase 2-hydroxyglutarate levels and the H3K27me3 repressive mark, which was accompanied by a downregulation of cell cycle regulation and neuro-glial differentiation genes (412).

This raises the question—could D2HG inhibit stemness through a similar mechanism in IDH-mutant gliomas? While there is one study demonstrating that H3K9 and H3K27 methylation increases in response to D2HG and the IDH mutation, it remains unclear how and if this mechanism impacts key regulators of differentiation in astrocytes (413).

While the work discussed in this section is still ongoing, our preliminary data will provide new insight into the mechanisms that regulate astrocyte maturation and will reveal new areas for further exploration.

Methods

Generating and validating stable iPSC lines

ORF sequences for candidate TFs *Ascl1* (NM_004316), *NR3C2* (NM_000901.5), *RFX4* (NM_001206691.2), and *SOX21* (NM_007084) were amplified (from either human reference cDNA library (Takara 639654) or cDNA libraries made from cortical organoids) and cloned into the multiple cloning sites of pCW57-GFP-2A-MCS (Addgene #71783) using *AvrII* and *BamHI* restriction sites. These transfer plasmids were then transfected into HEK298FT cells with Lipofectamine 3000 (Invitrogen - L3000-015) to produce lentivirus following manufacturer's protocols.

C4.1 iPSCs were cultured as described in section 2.5 and incubated with virus for 48 hrs. Cells were then cultured for 1 week in E8 media supplemented with 1 ug/ml of puromycin (VWR; Cat #: 0210055225) to select for infected cells. Remaining colonies were used for clonal isolation using Corning's "Cell Cloning by Serial Dilution in 96 Well Plates" protocol (8/08 Rev2). Individual colonies were scraped with a p200 pipette tip, transferred to vitronectin-coated 6-well plates, and expanded for making freeze-downs and quality control testing. iPSC clones that passed quality control testing were used for generating hCSs (using methods described in section 2.5).

iPSC clones were evaluated based on three metrics: (1) presence of GFP band after PCR and gel electrophoresis, (2) GFP fluorescence after treatment with doxycycline for 72 hrs, and (3) increased expression of the TF of interest using RT-qPCR in clones after treatment with doxycycline for 72 hrs. To assess presence/absence of GFP sequence, DNA extraction was performed using the Qiagen DNeasy Blood & Tissue kit (Qiagen; Cat #: 69504) following the manufacturer's instructions. PCR was carried out using Taq DNA Polymerase (Thermo Scientific; Cat #: EPO402) following manufacturer's protocol and DNA was amplified for 35 cycles. Samples were run on a 2% gel at 100V for 90 mins.

RT-qPCR of iPSC clones and hCSs treated with doxycycline

For RT-qPCR, RNA was extracted from cells/hCSs using the RNeasy kit (Qiagen; Cat. #: 74104) according to the manufacturer's instructions. RNA concentration was checked via NanoDrop and cDNA synthesis was performed using SuperScript™ IV Reverse Transcriptase (Invitrogen, cat. 18090050) with a mix of random hexamers (Invitrogen, cat. N8080127) and oligo d(T) (Invitrogen, cat. 18418020), according to the manufacturer's instructions. Samples were prepared for qPCR using the SYBR Green PowerUp Master Mix (Applied Biosystems, cat. A25741), according to the manufacturer's instructions. We ordered primers through Invitrogen's custom DNA oligos. Primers include:

RFX4_forward: caccaattatatcaggagtttgacct

RFX4_reverse: acctcacaacacagcggtc

Primer pairs were validated for specificity and efficiency using qPCR with serial dilutions. Samples were run in triplicates to determine the proper cycle threshold (CT)

of each gene. According to the manufacturer's protocol, we performed 40 cycles of amplification. When determining fold changes in gene expression across samples, the CT of each gene was normalized according to the CT of the housekeeping gene in the same sample:

GAPDH_forward: catgagaagtatgacaacagcct

GAPDH_reverse: agtccttccacgataccaaagt

Immunohistochemistry (IHC) with hCSs

hCSs received 0 or 2 ug/mL doxycycline every other day for 10 days, starting at day 30. At day 40, organoids were fixed with 4% paraformaldehyde for 3-4 hours at 4 degrees and then incubated with 30% sucrose for 24-48 hours. Next, hCSs were washed in O.C.T. compound and immersed in cryomolds containing O.C.T. Cryomolds were stored at -80 until the blocks were sectioned on a cryostat at 70 um thickness. Cryosections were washed three times (10 mins each) in PBS containing 0.01% triton X-100 and blocked in PBS containing 10% normal donkey serum and 0.3% triton X-100 for 60 mins at room temperature. Sections were incubated with primary antibodies (Rb anti-GFAP; Agilent Dako ZO334; 1:1500 or Rb anti-RFX4; Atlas Antibodies HPA05052; 1:100) in a PBS buffer containing 10% normal donkey serum and 0.3% triton X-100, overnight at 4 degrees. The next day, sections were washed three times (10 mins each) in PBS containing 0.01% triton X-100 to remove unbound primary antibodies. Sections were incubated with secondary antibodies (Donkey anti-rabbit 594; 1:1000) and the Hoechst nuclear reagent in PBS containing 10% normal donkey serum and 0.3% triton X-100 for 60 mins at room temperature, protected from light. Unbound secondary antibodies were washed three times (10 mins each) with PBS containing 0.01% triton X-

100 and coverslipped using Fluoromount-G mounting medium (Thermo Fisher Scientific; Cat. #: 00-4958-02). Slides were air dried overnight and then imaged on a confocal microscope (Leica).

Fetal tissue dissociation and astrocyte culture with D2HG

Fresh fetal cortical tissue (16-20 GW) was dissociated as described in section 2.5. Once a single-cell suspension was reached, neurons were first depleted via immunopanning with an anti-CD24 antibody (Miltenyi Biotec; Cat. #: 130-108-037). Unbound cells were then immunopanned for CD49f+ cells, which were trypsonized from the immunopanning plate for culturing. Trypsinized cells were centrifuged at 300 g for 10 mins, counted using a hemocytometer, and 20,000 cells were plated on poly-D-lysine-coated plastic coverslips in a Neurobasal-DMEM-based serum-free medium supplemented with HBEGF (R&D Systems; Cat. #: 259-HE).

Purified astrocytes were treated with either just media, or media supplemented with 0.1% DMSO or 0.1 mM Octyl-D-2HG (Sigma-Aldrich; Cat #: SML2200), a membrane-permeant precursor form of the oncometabolite D-2-hydroxyglutarate (D-2HG) produced by tumor cells. Media was changed every day for 1 week, with respective supplements. Cells used for EdU assay were treated with respective media solution, supplemented with EdU, approximately 24 hrs prior to collection for ICC.

Cells used for bulk RNA-seq were scraped in Qiazol and RNA was extracted using Qiagen's RNeasy kit (Qiagen; Cat. #: 74104). Libraries were prepared using Takara Bio's SMART-Seq HT kit (Takara Bio; Cat. #: R400749). Computational analysis, including

read trimming mapping, count matrix generation, and differential gene expression were carried out as described in section 2.5.

ICC with fetal astrocytes

Cells collected for ICC were fixed in 4% PFA for 15 mins at room temperature. Cells were then washed in 3% BSA and incubated for 20 mins in 0.5% Triton-X at room temperature. First, the EdU protocol was followed according to the manufacturer's directions (Thermo Scientific; Cat. #: C10339). After incubating cells in the EdU reaction cocktail for 30 mins in the dark, cells were rinsed once in 3% BSA, and incubated in blocking solution (0.3% Triton-X with 10% donkey serum). Cells were incubated in primary antibody (Chicken anti-GFAP; 1:1000; BioLegend; Cat. #: 829401) overnight at 4 degrees. The following day, cells were washed three times with PBS and incubated in secondary antibody (Donkey anti-chicken 488; 1:1000; Jackson ImmunoResearch; Cat. #: 703-545-155), supplemented with DAPI (1:1000; Thermo Fisher Scientific; Cat. #: D3571), for 2 hrs at room temperature. Cells were washed again, 3 times with PBS and coverslips were mounted using Fluoromount-G mounting medium. Slides were air dried overnight and then imaged on Keyence BZ-X810. DAPI and EdU+ cell counts were quantified using the hybrid cell count function in the BZ-X800 software.

Pseudotime analysis

Data used in pseudotime analysis was from Ramos et al. 2022

(<https://doi.org/10.1038/s41467-022-34975-2>). Specifically, we show data from nuclei isolated from micro-dissected fetal cortical plate tissue (17-41 GW). Initial processing,

including filtering and count normalization were performed according to the “Data integration and clustering” section of Ramos et al. 2022, using the R package Seurat. The processed Seurat object was converted using the function `as.cell_data_set()` in the SeuratWrappers package to a Monocle 3 cell data set (cde) object for further analysis with the Monocle 3 R package.

We used Monocle 3 to perform unsupervised clustering with the `cluster_cells()` function with a UMAP reduction method, Leiden community detection clustering method, and a q-value cutoff of 0.05. We next implemented the `learn_graph()` function with default parameters to “learn” how cells transition through gene expression changes, forming the predicted trajectory that cells follow. The Monocle 3 function `order_cells()` was used to calculate where each cell falls in pseudotime using “UMAP” as a reduction method and setting transit-amplifying cells (TACs) as the root cell population. Finally, the `plot_cells()` function in the Monocle 3 package was used to visualize UMAPs with nuclei colored by cell type and pseudotime, and to show individual expression of maturation TFs, as well as aggregated expression of maturation gene modules.

REFERENCES

1. J. Takouda, S. Katada, K. Nakashima, Emerging mechanisms underlying astrogenesis in the developing mammalian brain. *Proc Jpn Acad Ser B Phys Biol Sci* **93**, 386-398 (2017).
2. A. R. Desai, S. K. McConnell, Progressive restriction in fate potential by neural progenitors during cerebral cortical development. *Development* **127**, 2863-2872 (2000).
3. R. Kanski, M. E. van Strien, P. van Tijn, E. M. Hol, A star is born: new insights into the mechanism of astrogenesis. *Cell Mol Life Sci* **71**, 433-447 (2014).
4. D. Derouet *et al.*, Neuropoietin, a new IL-6-related cytokine signaling through the ciliary neurotrophic factor receptor. *Proc Natl Acad Sci U S A* **101**, 4827-4832 (2004).
5. A. Uemura *et al.*, Cardiotrophin-like cytokine induces astrocyte differentiation of fetal neuroepithelial cells via activation of STAT3. *Cytokine* **18**, 1-7 (2002).
6. P. Y. Cheng *et al.*, Interplay between SIN3A and STAT3 mediates chromatin conformational changes and GFAP expression during cellular differentiation. *PLoS One* **6**, e22018 (2011).
7. F. He *et al.*, A positive autoregulatory loop of Jak-STAT signaling controls the onset of astroglialogenesis. *Nat Neurosci* **8**, 616-625 (2005).
8. C. Menard *et al.*, An essential role for a MEK-C/EBP pathway during growth factor-regulated cortical neurogenesis. *Neuron* **36**, 597-610 (2002).
9. A. S. Gauthier *et al.*, Control of CNS cell-fate decisions by SHP-2 and its dysregulation in Noonan syndrome. *Neuron* **54**, 245-262 (2007).
10. M. Ernst, B. J. Jenkins, Acquiring signalling specificity from the cytokine receptor gp130. *Trends Genet* **20**, 23-32 (2004).
11. Y. Hirabayashi *et al.*, Polycomb limits the neurogenic competence of neural precursor cells to promote astrogenic fate transition. *Neuron* **63**, 600-613 (2009).
12. Z. L. Yuan, Y. J. Guan, D. Chatterjee, Y. E. Chin, Stat3 dimerization regulated by reversible acetylation of a single lysine residue. *Science* **307**, 269-273 (2005).
13. K. Nakashima *et al.*, Synergistic signaling in fetal brain by STAT3-Smad1 complex bridged by p300. *Science* **284**, 479-482 (1999).
14. S. Kamakura *et al.*, Hes binding to STAT3 mediates crosstalk between Notch and JAK-STAT signalling. *Nat Cell Biol* **6**, 547-554 (2004).
15. K. M. Wilczynska *et al.*, Nuclear factor I isoforms regulate gene expression during the differentiation of human neural progenitors to astrocytes. *Stem Cells* **27**, 1173-1181 (2009).
16. B. Cebolla, M. Vallejo, Nuclear factor-I regulates glial fibrillary acidic protein gene expression in astrocytes differentiated from cortical precursor cells. *J Neurochem* **97**, 1057-1070 (2006).
17. M. Namihira *et al.*, Committed neuronal precursors confer astrocytic potential on residual neural precursor cells. *Dev Cell* **16**, 245-255 (2009).
18. L. das Neves *et al.*, Disruption of the murine nuclear factor I-A gene (Nfia) results in perinatal lethality, hydrocephalus, and agenesis of the corpus callosum. *Proc Natl Acad Sci U S A* **96**, 11946-11951 (1999).
19. X. Yu, J. Nagai, B. S. Khakh, Improved tools to study astrocytes. *Nat Rev Neurosci* **21**, 121-138 (2020).
20. S. N. Lanjewar, S. A. Sloan, Growing Glia: Cultivating Human Stem Cell Models of Gliogenesis in Health and Disease. *Front Cell Dev Biol* **9**, 649538 (2021).

21. J. Tchieu *et al.*, NFIA is a gliogenic switch enabling rapid derivation of functional human astrocytes from pluripotent stem cells. *Nat Biotechnol* **37**, 267-275 (2019).
22. N. Tiwari *et al.*, Stage-Specific Transcription Factors Drive Astroglialogenesis by Remodeling Gene Regulatory Landscapes. *Cell Stem Cell* **23**, 557-571 e558 (2018).
23. B. Deneen *et al.*, The transcription factor NFIA controls the onset of gliogenesis in the developing spinal cord. *Neuron* **52**, 953-968 (2006).
24. C. C. Stolt *et al.*, The Sox9 transcription factor determines glial fate choice in the developing spinal cord. *Genes Dev* **17**, 1677-1689 (2003).
25. P. Kang *et al.*, Sox9 and NFIA coordinate a transcriptional regulatory cascade during the initiation of gliogenesis. *Neuron* **74**, 79-94 (2012).
26. M. Nagao, T. Ogata, Y. Sawada, Y. Gotoh, Zbtb20 promotes astrocytogenesis during neocortical development. *Nat Commun* **7**, 11102 (2016).
27. G. B. Yeon *et al.*, NFIB induces functional astrocytes from human pluripotent stem cell-derived neural precursor cells mimicking in vivo astroglialogenesis. *J Cell Physiol* **236**, 7625-7641 (2021).
28. J. S. Byun *et al.*, The transcription factor PITX1 drives astrocyte differentiation by regulating the SOX9 gene. *J Biol Chem* **295**, 13677-13690 (2020).
29. S. M. Glasgow *et al.*, Glia-specific enhancers and chromatin structure regulate NFIA expression and glioma tumorigenesis. *Nat Neurosci* **20**, 1520-1528 (2017).
30. M. Lattke *et al.*, Extensive transcriptional and chromatin changes underlie astrocyte maturation in vivo and in culture. *Nat Commun* **12**, 4335 (2021).
31. Z. Liu *et al.*, Induction of oligodendrocyte differentiation by Olig2 and Sox10: evidence for reciprocal interactions and dosage-dependent mechanisms. *Dev Biol* **302**, 683-693 (2007).
32. M. Finzsch, C. C. Stolt, P. Lommes, M. Wegner, Sox9 and Sox10 influence survival and migration of oligodendrocyte precursors in the spinal cord by regulating PDGF receptor alpha expression. *Development* **135**, 637-646 (2008).
33. J. O. Watzlawik, A. E. Warrington, M. Rodriguez, PDGF is required for remyelination-promoting IgM stimulation of oligodendrocyte progenitor cell proliferation. *PLoS One* **8**, e55149 (2013).
34. M. Weider *et al.*, Elevated in vivo levels of a single transcription factor directly convert satellite glia into oligodendrocyte-like cells. *PLoS Genet* **11**, e1005008 (2015).
35. C. C. Stolt *et al.*, SoxD proteins influence multiple stages of oligodendrocyte development and modulate SoxE protein function. *Dev Cell* **11**, 697-709 (2006).
36. J. Samanta, J. A. Kessler, Interactions between ID and OLIG proteins mediate the inhibitory effects of BMP4 on oligodendroglial differentiation. *Development* **131**, 4131-4142 (2004).
37. A. Liu *et al.*, A molecular insight of Hes5-dependent inhibition of myelin gene expression: old partners and new players. *EMBO J* **25**, 4833-4842 (2006).
38. T. Morrow, M. R. Song, A. Ghosh, Sequential specification of neurons and glia by developmentally regulated extracellular factors. *Development* **128**, 3585-3594 (2001).
39. F. D. Miller, A. S. Gauthier, Timing is everything: making neurons versus glia in the developing cortex. *Neuron* **54**, 357-369 (2007).
40. F. Barnabe-Heider *et al.*, Evidence that embryonic neurons regulate the onset of cortical gliogenesis via cardiotrophin-1. *Neuron* **48**, 253-265 (2005).

41. A. Bonni *et al.*, Regulation of gliogenesis in the central nervous system by the JAK-STAT signaling pathway. *Science* **278**, 477-483 (1997).
42. L. Bugga, R. A. Gadiant, K. Kwan, C. L. Stewart, P. H. Patterson, Analysis of neuronal and glial phenotypes in brains of mice deficient in leukemia inhibitory factor. *J Neurobiol* **36**, 509-524 (1998).
43. K. Nakashima *et al.*, BMP2-mediated alteration in the developmental pathway of fetal mouse brain cells from neurogenesis to astrocytogenesis. *Proc Natl Acad Sci U S A* **98**, 5868-5873 (2001).
44. K. Nakashima, M. Yanagisawa, H. Arakawa, T. Taga, Astrocyte differentiation mediated by LIF in cooperation with BMP2. *FEBS Lett* **457**, 43-46 (1999).
45. M. A. Bonaguidi *et al.*, LIF and BMP signaling generate separate and discrete types of GFAP-expressing cells. *Development* **132**, 5503-5514 (2005).
46. R. E. Gross *et al.*, Bone morphogenetic proteins promote astroglial lineage commitment by mammalian subventricular zone progenitor cells. *Neuron* **17**, 595-606 (1996).
47. M. W. Miller, Expression of transforming growth factor-beta in developing rat cerebral cortex: effects of prenatal exposure to ethanol. *J Comp Neurol* **460**, 410-424 (2003).
48. M. R. Song, A. Ghosh, FGF2-induced chromatin remodeling regulates CNTF-mediated gene expression and astrocyte differentiation. *Nat Neurosci* **7**, 229-235 (2004).
49. K. Irmady, S. Zechel, K. Unsicker, Fibroblast growth factor 2 regulates astrocyte differentiation in a region-specific manner in the hindbrain. *Glia* **59**, 708-719 (2011).
50. H. Asano *et al.*, Astrocyte differentiation of neural precursor cells is enhanced by retinoic acid through a change in epigenetic modification. *Stem Cells* **27**, 2744-2752 (2009).
51. A. Voss *et al.*, Computational Identification of Ligand-Receptor Pairs that Drive Human Astrocyte Development. *bioRxiv*, 2022.2005.2031.491513 (2022).
52. M. Noble, K. Murray, P. Stroobant, M. D. Waterfield, P. Riddle, Platelet-derived growth factor promotes division and motility and inhibits premature differentiation of the oligodendrocyte/type-2 astrocyte progenitor cell. *Nature* **333**, 560-562 (1988).
53. M. C. Raff, L. E. Lillien, W. D. Richardson, J. F. Burne, M. D. Noble, Platelet-derived growth factor from astrocytes drives the clock that times oligodendrocyte development in culture. *Nature* **333**, 562-565 (1988).
54. C. H. Heldin, B. Westermark, Mechanism of action and in vivo role of platelet-derived growth factor. *Physiol Rev* **79**, 1283-1316 (1999).
55. D. R. Goddard, M. Berry, S. L. Kirvell, A. M. Butt, Fibroblast growth factor-2 inhibits myelin production by oligodendrocytes in vivo. *Mol Cell Neurosci* **18**, 557-569 (2001).
56. C. Fressinaud, J. M. Vallat, G. Labourdette, Basic fibroblast growth factor down-regulates myelin basic protein gene expression and alters myelin compaction of mature oligodendrocytes in vitro. *J Neurosci Res* **40**, 285-293 (1995).
57. R. Bansal, S. E. Pfeiffer, Inhibition of protein and lipid sulfation in oligodendrocytes blocks biological responses to FGF-2 and retards cytoarchitectural maturation, but not developmental lineage progression. *Dev Biol* **162**, 511-524 (1994).
58. F. Jiang, T. J. Frederick, T. L. Wood, IGF-I synergizes with FGF-2 to stimulate oligodendrocyte progenitor entry into the cell cycle. *Dev Biol* **232**, 414-423 (2001).
59. F. A. McMorris, R. W. Furlanetto, R. L. Mozell, M. J. Carson, D. W. Raible, Regulation of oligodendrocyte development by insulin-like growth factors and cyclic nucleotides. *Ann N Y Acad Sci* **605**, 101-109 (1990).

60. F. Bandeira, R. Lent, S. Herculano-Houzel, Changing numbers of neuronal and non-neuronal cells underlie postnatal brain growth in the rat. *Proc Natl Acad Sci U S A* **106**, 14108-14113 (2009).
61. A. V. Molofsky, B. Deneen, Astrocyte development: A Guide for the Perplexed. *Glia* **63**, 1320-1329 (2015).
62. W. P. Ge, A. Miyawaki, F. H. Gage, Y. N. Jan, L. Y. Jan, Local generation of glia is a major astrocyte source in postnatal cortex. *Nature* **484**, 376-380 (2012).
63. B. Emery, Q. R. Lu, Transcriptional and Epigenetic Regulation of Oligodendrocyte Development and Myelination in the Central Nervous System. *Cold Spring Harb Perspect Biol* **7**, a020461 (2015).
64. H. Li, J. P. de Faria, P. Andrew, J. Nitarska, W. D. Richardson, Phosphorylation regulates OLIG2 cofactor choice and the motor neuron-oligodendrocyte fate switch. *Neuron* **69**, 918-929 (2011).
65. B. C. Warf, J. Fok-Seang, R. H. Miller, Evidence for the ventral origin of oligodendrocyte precursors in the rat spinal cord. *J Neurosci* **11**, 2477-2488 (1991).
66. J. Cai *et al.*, Generation of oligodendrocyte precursor cells from mouse dorsal spinal cord independent of Nkx6 regulation and Shh signaling. *Neuron* **45**, 41-53 (2005).
67. M. Fogarty, W. D. Richardson, N. Kessaris, A subset of oligodendrocytes generated from radial glia in the dorsal spinal cord. *Development* **132**, 1951-1959 (2005).
68. N. Spassky *et al.*, The early steps of oligodendrogenesis: insights from the study of the plp lineage in the brain of chicks and rodents. *Dev Neurosci* **23**, 318-326 (2001).
69. N. Tekki-Kessaris *et al.*, Hedgehog-dependent oligodendrocyte lineage specification in the telencephalon. *Development* **128**, 2545-2554 (2001).
70. N. Kessaris *et al.*, Competing waves of oligodendrocytes in the forebrain and postnatal elimination of an embryonic lineage. *Nat Neurosci* **9**, 173-179 (2006).
71. I. Jakovcevski, R. Filipovic, Z. Mo, S. Rakic, N. Zecevic, Oligodendrocyte development and the onset of myelination in the human fetal brain. *Front Neuroanat* **3**, 5 (2009).
72. S. A. Back *et al.*, Late oligodendrocyte progenitors coincide with the developmental window of vulnerability for human perinatal white matter injury. *J Neurosci* **21**, 1302-1312 (2001).
73. E. A. Bushong, M. E. Martone, M. H. Ellisman, Maturation of astrocyte morphology and the establishment of astrocyte domains during postnatal hippocampal development. *Int J Dev Neurosci* **22**, 73-86 (2004).
74. E. A. Bushong, M. E. Martone, Y. Z. Jones, M. H. Ellisman, Protoplasmic astrocytes in CA1 stratum radiatum occupy separate anatomical domains. *J Neurosci* **22**, 183-192 (2002).
75. M. M. Halassa, T. Fellin, H. Takano, J. H. Dong, P. G. Haydon, Synaptic islands defined by the territory of a single astrocyte. *J Neurosci* **27**, 6473-6477 (2007).
76. K. Ogata, T. Kosaka, Structural and quantitative analysis of astrocytes in the mouse hippocampus. *Neuroscience* **113**, 221-233 (2002).
77. Y. Zhang *et al.*, Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse. *Neuron* **89**, 37-53 (2016).
78. M. C. Krawczyk *et al.*, Human Astrocytes Exhibit Tumor Microenvironment-, Age-, and Sex-Related Transcriptomic Signatures. *J Neurosci* **42**, 1587-1603 (2022).

79. P. D. Le Roux, T. A. Reh, Independent regulation of primary dendritic and axonal growth by maturing astrocytes in vitro. *Neurosci Lett* **198**, 5-8 (1995).
80. R. Le, S. Esquenazi, Astrocytes mediate cerebral cortical neuronal axon and dendrite growth, in part, by release of fibroblast growth factor. *Neurol Res* **24**, 81-92 (2002).
81. G. Z. Tau, B. S. Peterson, Normal development of brain circuits. *Neuropsychopharmacology* **35**, 147-168 (2010).
82. P. R. Huttenlocher, A. S. Dabholkar, Regional differences in synaptogenesis in human cerebral cortex. *J Comp Neurol* **387**, 167-178 (1997).
83. C. Garel *et al.*, Fetal cerebral cortex: normal gestational landmarks identified using prenatal MR imaging. *AJNR Am J Neuroradiol* **22**, 184-189 (2001).
84. N. J. Allen *et al.*, Astrocyte glypicans 4 and 6 promote formation of excitatory synapses via GluA1 AMPA receptors. *Nature* **486**, 410-414 (2012).
85. K. S. Christopherson *et al.*, Thrombospondins are astrocyte-secreted proteins that promote CNS synaptogenesis. *Cell* **120**, 421-433 (2005).
86. C. Eroglu *et al.*, Gabapentin receptor alpha2delta-1 is a neuronal thrombospondin receptor responsible for excitatory CNS synaptogenesis. *Cell* **139**, 380-392 (2009).
87. W. S. Chung *et al.*, Astrocytes mediate synapse elimination through MEGF10 and MERTK pathways. *Nature* **504**, 394-400 (2013).
88. B. Stevens *et al.*, The classical complement cascade mediates CNS synapse elimination. *Cell* **131**, 1164-1178 (2007).
89. R. Daneman, A. Prat, The blood-brain barrier. *Cold Spring Harb Perspect Biol* **7**, a020412 (2015).
90. R. Araya *et al.*, BMP signaling through BMPRIA in astrocytes is essential for proper cerebral angiogenesis and formation of the blood-brain-barrier. *Mol Cell Neurosci* **38**, 417-430 (2008).
91. R. C. Janzer, M. C. Raff, Astrocytes induce blood-brain barrier properties in endothelial cells. *Nature* **325**, 253-257 (1987).
92. L. Hosli *et al.*, Direct vascular contact is a hallmark of cerebral astrocytes. *Cell Rep* **39**, 110599 (2022).
93. P. J. Magistretti, I. Allaman, Lactate in the brain: from metabolic end-product to signalling molecule. *Nat Rev Neurosci* **19**, 235-249 (2018).
94. D. Attwell *et al.*, Glial and neuronal control of brain blood flow. *Nature* **468**, 232-243 (2010).
95. C. Howarth, P. Gleeson, D. Attwell, Updated energy budgets for neural computation in the neocortex and cerebellum. *J Cereb Blood Flow Metab* **32**, 1222-1232 (2012).
96. G. Dallerac, J. Zapata, N. Rouach, Versatile control of synaptic circuits by astrocytes: where, when and how? *Nat Rev Neurosci* **19**, 729-743 (2018).
97. R. Chittajallu, A. Aguirre, V. Gallo, NG2-positive cells in the mouse white and grey matter display distinct physiological properties. *J Physiol* **561**, 109-122 (2004).
98. J. P. Michalski, R. Kothary, Oligodendrocytes in a Nutshell. *Front Cell Neurosci* **9**, 340 (2015).
99. S. Gokhan *et al.*, Combinatorial profiles of oligodendrocyte-selective classes of transcriptional regulators differentially modulate myelin basic protein gene expression. *J Neurosci* **25**, 8311-8321 (2005).
100. J. Hornig *et al.*, The transcription factors Sox10 and Myrf define an essential regulatory network module in differentiating oligodendrocytes. *PLoS Genet* **9**, e1003907 (2013).

101. Y. Yu *et al.*, Olig2 targets chromatin remodelers to enhancers to initiate oligodendrocyte differentiation. *Cell* **152**, 248-261 (2013).
102. H. Bujalka *et al.*, MYRF is a membrane-associated transcription factor that autoproteolytically cleaves to directly activate myelin genes. *PLoS Biol* **11**, e1001625 (2013).
103. T. Thurnherr *et al.*, Cdc42 and Rac1 signaling are both required for and act synergistically in the correct formation of myelin sheaths in the CNS. *J Neurosci* **26**, 10110-10119 (2006).
104. S. C. Zhang, Defining glial cells during CNS development. *Nat Rev Neurosci* **2**, 840-843 (2001).
105. N. J. Scolding *et al.*, Myelin-oligodendrocyte glycoprotein (MOG) is a surface marker of oligodendrocyte maturation. *J Neuroimmunol* **22**, 169-176 (1989).
106. A. Nishiyama, M. Komitova, R. Suzuki, X. Zhu, Polydendrocytes (NG2 cells): multifunctional cells with lineage plasticity. *Nat Rev Neurosci* **10**, 9-22 (2009).
107. A. Nishiyama, L. Boshans, C. M. Goncalves, J. Wegrzyn, K. D. Patel, Lineage, fate, and fate potential of NG2-glia. *Brain Res* **1638**, 116-128 (2016).
108. M. C. Raff, R. H. Miller, M. Noble, A glial progenitor cell that develops in vitro into an astrocyte or an oligodendrocyte depending on culture medium. *Nature* **303**, 390-396 (1983).
109. M. S. Rao, M. Mayer-Proschel, Glial-restricted precursors are derived from multipotent neuroepithelial stem cells. *Dev Biol* **188**, 48-63 (1997).
110. R. J. Franklin, S. A. Bayley, R. Milner, C. Ffrench-Constant, W. F. Blakemore, Differentiation of the O-2A progenitor cell line CG-4 into oligodendrocytes and astrocytes following transplantation into glia-deficient areas of CNS white matter. *Glia* **13**, 39-44 (1995).
111. M. S. Windrem *et al.*, Fetal and adult human oligodendrocyte progenitor cell isolates myelinate the congenitally dysmyelinated brain. *Nat Med* **10**, 93-97 (2004).
112. X. Zhu, D. E. Bergles, A. Nishiyama, NG2 cells generate both oligodendrocytes and gray matter astrocytes. *Development* **135**, 145-157 (2008).
113. X. Zhu *et al.*, Age-dependent fate and lineage restriction of single NG2 cells. *Development* **138**, 745-753 (2011).
114. P. Bedner, R. Jabs, C. Steinhauser, Properties of human astrocytes and NG2 glia. *Glia* **68**, 756-767 (2020).
115. D. V. Hansen, J. H. Lui, P. R. Parker, A. R. Kriegstein, Neurogenic radial glia in the outer subventricular zone of human neocortex. *Nature* **464**, 554-561 (2010).
116. U. C. Eze, A. Bhaduri, M. Haeussler, T. J. Nowakowski, A. R. Kriegstein, Single-cell atlas of early human brain development highlights heterogeneity of human neuroepithelial cells and early radial glia. *Nat Neurosci* **24**, 584-594 (2021).
117. S. I. Ramos *et al.*, An atlas of late prenatal human neurodevelopment resolved by single-nucleus transcriptomics. *Nat Commun* **13**, 7671 (2022).
118. L. Yang, Z. Li, G. Liu, X. Li, Z. Yang, Developmental Origins of Human Cortical Oligodendrocytes and Astrocytes. *Neurosci Bull* **38**, 47-68 (2022).
119. J. Andersen *et al.*, Single-cell transcriptomic landscape of the developing human spinal cord. *Nat Neurosci* **26**, 902-914 (2023).

120. C. D. Pozniak *et al.*, Sox10 directs neural stem cells toward the oligodendrocyte lineage by decreasing Suppressor of Fused expression. *Proc Natl Acad Sci U S A* **107**, 21795-21800 (2010).
121. S. Klum *et al.*, Sequentially acting SOX proteins orchestrate astrocyte- and oligodendrocyte-specific gene expression. *EMBO Rep* **19**, (2018).
122. L. A. Hassel *et al.*, Differential activity of transcription factor Sox9 in early and adult oligodendroglial progenitor cells. *Glia* **71**, 1890-1905 (2023).
123. S. M. Glasgow *et al.*, Mutual antagonism between Sox10 and NFIA regulates diversification of glial lineages and glioma subtypes. *Nat Neurosci* **17**, 1322-1329 (2014).
124. J. D. Lathia, S. C. Mack, E. E. Mulkearns-Hubert, C. L. Valentim, J. N. Rich, Cancer stem cells in glioblastoma. *Genes Dev* **29**, 1203-1217 (2015).
125. R. C. Gimple, S. Bhargava, D. Dixit, J. N. Rich, Glioblastoma stem cells: lessons from the tumor hierarchy in a lethal cancer. *Genes Dev* **33**, 591-609 (2019).
126. A. Kreso, J. E. Dick, Evolution of the cancer stem cell model. *Cell Stem Cell* **14**, 275-291 (2014).
127. J. Chen *et al.*, A restricted cell population propagates glioblastoma growth after chemotherapy. *Nature* **488**, 522-526 (2012).
128. S. Bao *et al.*, Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature* **444**, 756-760 (2006).
129. B. Bao, A. Ahmad, A. S. Azmi, S. Ali, F. H. Sarkar, Overview of cancer stem cells (CSCs) and mechanisms of their regulation: implications for cancer therapy. *Curr Protoc Pharmacol* **Chapter 14**, Unit 14 25 (2013).
130. S. Li *et al.*, Signaling pathways in brain tumors and therapeutic interventions. *Signal Transduct Target Ther* **8**, 8 (2023).
131. T. Reya, H. Clevers, Wnt signalling in stem cells and cancer. *Nature* **434**, 843-850 (2005).
132. M. Teodorczyk, M. H. H. Schmidt, Notching on Cancer's Door: Notch Signaling in Brain Tumors. *Front Oncol* **4**, 341 (2014).
133. Z. Yu, T. G. Pestell, M. P. Lisanti, R. G. Pestell, Cancer stem cells. *Int J Biochem Cell Biol* **44**, 2144-2151 (2012).
134. R. N. Curry, S. M. Glasgow, The Role of Neurodevelopmental Pathways in Brain Tumors. *Front Cell Dev Biol* **9**, 659055 (2021).
135. Q. T. Ostrom *et al.*, CBTRUS statistical report: Primary brain and central nervous system tumors diagnosed in the United States in 2006-2010. *Neuro Oncol* **15 Suppl 2**, ii1-56 (2013).
136. J. P. Thakkar *et al.*, Epidemiologic and molecular prognostic review of glioblastoma. *Cancer Epidemiol Biomarkers Prev* **23**, 1985-1996 (2014).
137. A. Mansouri, J. Karamchandani, S. Das, in *Glioblastoma*, S. De Vleeschouwer, Ed. (Brisbane (AU), 2017).
138. R. Stupp *et al.*, Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med* **352**, 987-996 (2005).
139. B. Tran, M. A. Rosenthal, Survival comparison between glioblastoma multiforme and other incurable cancers. *J Clin Neurosci* **17**, 417-421 (2010).
140. M. Koshy *et al.*, Improved survival time trends for glioblastoma using the SEER 17 population-based registries. *J Neurooncol* **107**, 207-212 (2012).

141. N. Cancer Genome Atlas Research, Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061-1068 (2008).
142. R. G. Verhaak *et al.*, Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell* **17**, 98-110 (2010).
143. A. Sottoriva *et al.*, Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc Natl Acad Sci U S A* **110**, 4009-4014 (2013).
144. Q. Wang *et al.*, Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. *Cancer Cell* **32**, 42-56 e46 (2017).
145. C. Neftel *et al.*, An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell* **178**, 835-849 e821 (2019).
146. A. Dirkse *et al.*, Stem cell-associated heterogeneity in Glioblastoma results from intrinsic tumor plasticity shaped by the microenvironment. *Nat Commun* **10**, 1787 (2019).
147. A. R. Pine *et al.*, Tumor Microenvironment Is Critical for the Maintenance of Cellular States Found in Primary Glioblastomas. *Cancer Discov* **10**, 964-979 (2020).
148. L. Cheng *et al.*, Glioblastoma stem cells generate vascular pericytes to support vessel function and tumor growth. *Cell* **153**, 139-152 (2013).
149. N. Charles *et al.*, Perivascular nitric oxide activates notch signaling and promotes stem-like character in PDGF-induced glioma cells. *Cell Stem Cell* **6**, 141-152 (2010).
150. M. D. Brooks, R. Sengupta, S. C. Snyder, J. B. Rubin, Hitting Them Where They Live: Targeting the Glioblastoma Perivascular Stem Cell Niche. *Curr Pathobiol Rep* **1**, 101-110 (2013).
151. A. Sharma, A. Shiras, Cancer stem cell-vascular endothelial cell interactions in glioblastoma. *Biochem Biophys Res Commun* **473**, 688-692 (2016).
152. K. L. Covelto *et al.*, HIF-2 α regulates Oct-4: effects of hypoxia on stem cell function, embryonic development, and tumor growth. *Genes Dev* **20**, 557-570 (2006).
153. J. M. Heddleston, Z. Li, R. E. McLendon, A. B. Hjelmeland, J. N. Rich, The hypoxic microenvironment maintains glioblastoma stem cells and promotes reprogramming towards a cancer stem cell phenotype. *Cell Cycle* **8**, 3274-3284 (2009).
154. Z. Li *et al.*, Hypoxia-inducible factors regulate tumorigenic capacity of glioma stem cells. *Cancer Cell* **15**, 501-513 (2009).
155. S. Seidel *et al.*, A hypoxic niche regulates glioblastoma stem cells through hypoxia inducible factor 2 α . *Brain* **133**, 983-995 (2010).
156. B. Ortensi, M. Setti, D. Osti, G. Pelicci, Cancer stem cell contribution to glioblastoma invasiveness. *Stem Cell Res Ther* **4**, 18 (2013).
157. Z. Wang, H. Zhang, S. Xu, Z. Liu, Q. Cheng, The adaptive transition of glioblastoma stem cells and its implications on treatments. *Signal Transduct Target Ther* **6**, 124 (2021).
158. B. B. Liao *et al.*, Adaptive Chromatin Remodeling Drives Glioblastoma Stem Cell Plasticity and Drug Tolerance. *Cell Stem Cell* **20**, 233-246 e237 (2017).
159. I. Simeonova, E. Huillard, In vivo models of brain tumors: roles of genetically engineered mouse models in understanding tumor biology and use in preclinical studies. *Cell Mol Life Sci* **71**, 4007-4026 (2014).
160. D. Hambardzumyan, L. F. Parada, E. C. Holland, A. Charest, Genetic modeling of gliomas in mice: new tools to tackle old problems. *Glia* **59**, 1155-1168 (2011).

161. D. Friedmann-Morvinski *et al.*, Dedifferentiation of neurons and astrocytes by oncogenes can induce gliomas in mice. *Science* **338**, 1080-1084 (2012).
162. H. Simpson Ragdale *et al.*, Injury primes mutation-bearing astrocytes for dedifferentiation in later life. *Curr Biol* **33**, 1082-1098 e1088 (2023).
163. R. M. Bachoo *et al.*, Epidermal growth factor receptor and Ink4a/Arf: convergent mechanisms governing terminal differentiation and transformation along the neural stem cell to astrocyte axis. *Cancer Cell* **1**, 269-277 (2002).
164. D. Friedmann-Morvinski, I. M. Verma, Dedifferentiation and reprogramming: origins of cancer stem cells. *EMBO Rep* **15**, 244-253 (2014).
165. B. Liu, A. H. Neufeld, Activation of epidermal growth factor receptors in astrocytes: from development to neural injury. *J Neurosci Res* **85**, 3523-3529 (2007).
166. Q. Zhu *et al.*, Genetic evidence that Nkx2.2 and Pdgfra are major determinants of the timing of oligodendrocyte differentiation in the developing CNS. *Development* **141**, 548-555 (2014).
167. A. R. Calver *et al.*, Oligodendrocyte population dynamics and the role of PDGF in vivo. *Neuron* **20**, 869-882 (1998).
168. X. Liu *et al.*, EGF signaling promotes the lineage conversion of astrocytes into oligodendrocytes. *Mol Med* **28**, 50 (2022).
169. E. Rheinbay *et al.*, An aberrant transcription factor network essential for Wnt signaling and stem cell maintenance in glioblastoma. *Cell Rep* **3**, 1567-1579 (2013).
170. M. L. Suva *et al.*, Reconstructing and reprogramming the tumor-propagating potential of glioblastoma stem-like cells. *Cell* **157**, 580-594 (2014).
171. M. L. Suva, N. Riggi, B. E. Bernstein, Epigenetic reprogramming in cancer. *Science* **339**, 1567-1570 (2013).
172. F. Lu *et al.*, Olig2-Dependent Reciprocal Shift in PDGF and EGF Receptor Signaling Regulates Tumor Phenotype and Mitotic Growth in Malignant Glioma. *Cancer Cell* **29**, 669-683 (2016).
173. A. Y. Galvez-Contreras, A. Quinones-Hinojosa, O. Gonzalez-Perez, The role of EGFR and ErbB family related proteins in the oligodendrocyte specification in germinal niches of the adult mammalian brain. *Front Cell Neurosci* **7**, 258 (2013).
174. Y. Fu *et al.*, Heterogeneity of glial progenitor cells during the neurogenesis-to-gliogenesis switch in the developing human cerebral cortex. *Cell Rep* **34**, 108788 (2021).
175. X. Zhang *et al.*, Bulk and mosaic deletions of Egfr reveal regionally defined gliogenesis in the developing mouse forebrain. *iScience* **26**, 106242 (2023).
176. F. B. Furnari, T. F. Cloughesy, W. K. Cavenee, P. S. Mischel, Heterogeneity of epidermal growth factor receptor signalling networks in glioblastoma. *Nat Rev Cancer* **15**, 302-310 (2015).
177. A. B. Lassman *et al.*, Epidermal growth factor receptor (EGFR) amplification rates observed in screening patients for randomized trials in glioblastoma. *J Neurooncol* **144**, 205-210 (2019).
178. F. Liu *et al.*, EGFR Mutation Promotes Glioblastoma through Epigenome and Transcription Factor Network Remodeling. *Mol Cell* **60**, 307-318 (2015).
179. D. Sardar *et al.*, Sox9 directs divergent epigenomic states in brain tumor subtypes. *Proc Natl Acad Sci U S A* **119**, e2202015119 (2022).
180. D. N. Louis *et al.*, The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro Oncol* **23**, 1231-1251 (2021).

181. B. T. Whitfield, J. T. Huse, Classification of adult-type diffuse gliomas: Impact of the World Health Organization 2021 update. *Brain Pathol* **32**, e13062 (2022).
182. N. Lindberg, M. Kastemar, T. Olofsson, A. Smits, L. Uhrbom, Oligodendrocyte progenitor cells can act as cell of origin for experimental glioma. *Oncogene* **28**, 2266-2275 (2009).
183. C. Dai *et al.*, PDGF autocrine stimulation dedifferentiates cultured astrocytes and induces oligodendrogliomas and oligoastrocytomas from neural progenitors and astrocytes in vivo. *Genes Dev* **15**, 1913-1925 (2001).
184. A. S. Venteicher *et al.*, Decoupling genetics, lineages, and microenvironment in IDH-mutant gliomas by single-cell RNA-seq. *Science* **355**, (2017).
185. I. Tirosh *et al.*, Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. *Nature* **539**, 309-313 (2016).
186. Y. H. Chen, D. H. Gutmann, The molecular and cell biology of pediatric low-grade gliomas. *Oncogene* **33**, 2019-2026 (2014).
187. D. Y. Lee, T. H. Yeh, R. J. Emmett, C. R. White, D. H. Gutmann, Neurofibromatosis-1 regulates neuroglial progenitor proliferation and glial differentiation in a brain region-specific manner. *Genes Dev* **24**, 2317-2329 (2010).
188. B. Dasgupta, Y. Yi, D. Y. Chen, J. D. Weber, D. H. Gutmann, Proteomic analysis reveals hyperactivation of the mammalian target of rapamycin pathway in neurofibromatosis 1-associated human and mouse brain tumors. *Cancer Res* **65**, 2755-2760 (2005).
189. A. Kaul, Y. H. Chen, R. J. Emmett, S. Dahiya, D. H. Gutmann, Pediatric glioma-associated KIAA1549:BRAF expression regulates neuroglial cell growth in a cell type-specific and mTOR-dependent manner. *Genes Dev* **26**, 2561-2566 (2012).
190. A. Kaul, Y. H. Chen, R. J. Emmett, S. M. Gianino, D. H. Gutmann, Conditional KIAA1549:BRAF mice reveal brain region- and cell type-specific effects. *Genesis* **51**, 708-716 (2013).
191. J. A. S. Lammers *et al.*, LGG-15. SINGLE-NUCLEUS AND BULK RNA SEQUENCING OF PEDIATRIC PILOCYTIC ASTROCYTOMAS REVEALS CNS LOCATION-ASSOCIATED TUMOR CELL AND MICROENVIRONMENTAL HETEROGENEITY. *Neuro-Oncology* **25**, i58-i59 (2023).
192. Z. J. Reitman *et al.*, Mitogenic and progenitor gene programmes in single pilocytic astrocytoma cells. *Nat Commun* **10**, 3731 (2019).
193. C. J. Pachocki, E. M. Hol, Current perspectives on diffuse midline glioma and a different role for the immune microenvironment compared to glioblastoma. *J Neuroinflammation* **19**, 276 (2022).
194. V. Di Ruscio *et al.*, Pediatric Diffuse Midline Gliomas: An Unfinished Puzzle. *Diagnostics (Basel)* **12**, (2022).
195. R. Duan, W. Du, W. Guo, EZH2: a novel target for cancer treatment. *J Hematol Oncol* **13**, 104 (2020).
196. P. W. Lewis *et al.*, Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. *Science* **340**, 857-861 (2013).
197. K. Funato, T. Major, P. W. Lewis, C. D. Allis, V. Tabar, Use of human embryonic stem cells to model pediatric gliomas with H3.3K27M histone mutation. *Science* **346**, 1529-1533 (2014).
198. M. G. Filbin *et al.*, Developmental and oncogenic programs in H3K27M gliomas dissected by single-cell RNA-seq. *Science* **360**, 331-335 (2018).

199. S. Jessa *et al.*, K27M in canonical and noncanonical H3 variants occurs in distinct oligodendroglial cell lineages in brain midline gliomas. *Nat Genet* **54**, 1865-1880 (2022).
200. I. Liu *et al.*, The landscape of tumor cell states and spatial organization in H3-K27M mutant diffuse midline glioma across age and location. *Nat Genet* **54**, 1881-1894 (2022).
201. J. Neradil, R. Veselska, Nestin as a marker of cancer stem cells. *Cancer Sci* **106**, 803-811 (2015).
202. M. Hassn Mesrati, A. B. Behrooz, Y. A. A, A. Syahir, Understanding Glioblastoma Biomarkers: Knocking a Mountain with a Hammer. *Cells* **9**, (2020).
203. T. N. Ignatova *et al.*, Human cortical glial tumors contain neural stem-like cells expressing astroglial and neuronal markers in vitro. *Glia* **39**, 193-206 (2002).
204. E. O. Vik-Mo *et al.*, A comparative study of the structural organization of spheres derived from the adult human subventricular zone and glioblastoma biopsies. *Exp Cell Res* **317**, 1049-1059 (2011).
205. C. Liu *et al.*, Mosaic analysis with double markers reveals tumor cell of origin in glioma. *Cell* **146**, 209-221 (2011).
206. C. Bouchart, A. L. Trepant, M. Hein, D. Van Gestel, P. Demetter, Prognostic impact of glioblastoma stem cell markers OLIG2 and CCND2. *Cancer Med* **9**, 1069-1078 (2020).
207. H. Y. Jeong *et al.*, High expression of RFX4 is associated with tumor progression and poor prognosis in patients with glioblastoma. *Int J Neurosci* **131**, 7-14 (2021).
208. R. Uceda-Castro *et al.*, GFAP splice variants fine-tune glioma cell invasion and tumour dynamics by modulating migration persistence. *Sci Rep* **12**, 424 (2022).
209. S. O. Ghazi *et al.*, Cell of origin determines tumor phenotype in an oncogenic Ras/p53 knockout transgenic model of high-grade glioma. *J Neuropathol Exp Neurol* **71**, 729-740 (2012).
210. J. Radke, G. Bortolussi, A. Pagenstecher, Akt and c-Myc induce stem-cell markers in mature primary p53(-)/(-) astrocytes and render these cells gliomagenic in the brain of immunocompetent mice. *PLoS One* **8**, e56691 (2013).
211. T. Kondo, M. Raff, Oligodendrocyte precursor cells reprogrammed to become multipotential CNS stem cells. *Science* **289**, 1754-1757 (2000).
212. M. R. Dawson, A. Polito, J. M. Levine, R. Reynolds, NG2-expressing glial progenitor cells: an abundant and widespread population of cycling cells in the adult rat CNS. *Mol Cell Neurosci* **24**, 476-488 (2003).
213. S. Bardehle *et al.*, Live imaging of astrocyte responses to acute injury reveals selective juxtavascular proliferation. *Nat Neurosci* **16**, 580-586 (2013).
214. A. Beauchamp *et al.*, Whole-brain comparison of rodent and human brains using spatial transcriptomics. *Elife* **11**, (2022).
215. X. Zhao, A. Bhattacharyya, Human Models Are Needed for Studying Human Neurodevelopmental Disorders. *Am J Hum Genet* **103**, 829-857 (2018).
216. B. D. Semple, K. Blomgren, K. Gimlin, D. M. Ferriero, L. J. Noble-Haeusslein, Brain development in rodents and humans: Identifying benchmarks of maturation and vulnerability to injury across species. *Prog Neurobiol* **106-107**, 1-16 (2013).
217. C. J. Zeiss, Comparative Milestones in Rodent and Human Postnatal Central Nervous System Development. *Toxicol Pathol* **49**, 1368-1373 (2021).
218. S. A. Sloan, J. Andersen, A. M. Pasca, F. Birey, S. P. Pasca, Generation and assembly of human brain region-specific three-dimensional cultures. *Nat Protoc* **13**, 2062-2085 (2018).

219. F. Birey *et al.*, Assembly of functionally integrated human forebrain spheroids. *Nature* **545**, 54-59 (2017).
220. J. Jo *et al.*, Midbrain-like Organoids from Human Pluripotent Stem Cells Contain Functional Dopaminergic and Neuromelanin-Producing Neurons. *Cell Stem Cell* **19**, 248-257 (2016).
221. A. Zagare, M. Gobin, A. S. Monzel, J. C. Schwamborn, A robust protocol for the generation of human midbrain organoids. *STAR Protoc* **2**, 100524 (2021).
222. K. Muguruma, A. Nishiyama, H. Kawakami, K. Hashimoto, Y. Sasai, Self-organization of polarized cerebellar tissue in 3D culture of human pluripotent stem cells. *Cell Rep* **10**, 537-550 (2015).
223. P. Valiulahi *et al.*, Generation of caudal-type serotonin neurons and hindbrain-fate organoids from hPSCs. *Stem Cell Reports* **16**, 1938-1952 (2021).
224. T. Ogura, H. Sakaguchi, S. Miyamoto, J. Takahashi, Three-dimensional induction of dorsal, intermediate and ventral spinal cord tissues from human pluripotent stem cells. *Development* **145**, (2018).
225. N. Duval *et al.*, BMP4 patterns Smad activity and generates stereotyped cell fate organization in spinal organoids. *Development* **146**, (2019).
226. S. P. Pasca, Assembling human brain organoids. *Science* **363**, 126-127 (2019).
227. I. Chiaradia, M. A. Lancaster, Brain organoids for the study of human neurobiology at the interface of in vitro and in vivo. *Nat Neurosci* **23**, 1496-1508 (2020).
228. S. P. Pasca, The rise of three-dimensional human brain cultures. *Nature* **553**, 437-445 (2018).
229. A. Fiorenzano *et al.*, Single-cell transcriptomics captures features of human midbrain development and dopamine neuron diversity in brain organoids. *Nat Commun* **12**, 7302 (2021).
230. S. A. Sloan *et al.*, Human Astrocyte Maturation Captured in 3D Cerebral Cortical Spheroids Derived from Pluripotent Stem Cells. *Neuron* **95**, 779-790 e776 (2017).
231. A. E. Trevino *et al.*, Chromatin accessibility dynamics in a model of human forebrain development. *Science* **367**, (2020).
232. R. M. Marton *et al.*, Differentiation and maturation of oligodendrocytes in human three-dimensional neural cultures. *Nat Neurosci* **22**, 484-491 (2019).
233. M. Madhavan *et al.*, Induction of myelinating oligodendrocytes in human cortical spheroids. *Nat Methods* **15**, 700-706 (2018).
234. J. A. Garcia-Leon *et al.*, Generation of oligodendrocytes and establishment of an all-human myelinating platform from human pluripotent stem cells. *Nat Protoc* **15**, 3716-3744 (2020).
235. A. Gordon *et al.*, Long-term maturation of human cortical organoids matches key early postnatal transitions. *Nat Neurosci* **24**, 331-342 (2021).
236. R. J. Stubbins, A. Karsan, Differentiation therapy for myeloid malignancies: beyond cytotoxicity. *Blood Cancer J* **11**, 193 (2021).
237. J. Jin, F. Grigore, C. C. Chen, M. Li, Self-renewal signaling pathways and differentiation therapies of glioblastoma stem cells (Review). *Int J Oncol* **59**, (2021).
238. F. J. Swartling, M. Cancer, A. Frantz, H. Weishaupt, A. I. Persson, Dereglated proliferation and differentiation in brain tumors. *Cell Tissue Res* **359**, 225-254 (2015).
239. R. Wang, C. Liu, All-trans retinoic acid therapy induces asymmetric division of glioma stem cells from the U87MG cell line. *Oncol Lett* **18**, 3646-3654 (2019).

240. M. Ying *et al.*, Regulation of glioblastoma stem cells by retinoic acid: role for Notch pathway inhibition. *Oncogene* **30**, 3454-3467 (2011).
241. M. Koguchi *et al.*, BMP4 induces asymmetric cell division in human glioma stem-like cells. *Oncol Lett* **19**, 1247-1254 (2020).
242. S. Nayak *et al.*, Bone Morphogenetic Protein 4 Targeting Glioma Stem-Like Cells for Malignant Glioma Treatment: Latest Advances and Implications for Clinical Application. *Cancers (Basel)* **12**, (2020).
243. G. A. Videla Richardson *et al.*, Specific Preferences in Lineage Choice and Phenotypic Plasticity of Glioma Stem Cells Under BMP4 and Noggin Influence. *Brain Pathol* **26**, 43-61 (2016).
244. H. Caren, S. Beck, S. M. Pollard, Differentiation therapy for glioblastoma - too many obstacles? *Mol Cell Oncol* **3**, e1124174 (2016).
245. M. T. Stockhausen, K. Kristoffersen, L. Stobbe, H. S. Poulsen, Differentiation of glioblastoma multiforme stem-like cells leads to downregulation of EGFR and EGFRvIII and decreased tumorigenic and stem-like cell potential. *Cancer Biol Ther* **15**, 216-224 (2014).
246. A. I. Giannopoulou, D. S. Kanakoglou, C. Piperi, Transcription Factors with Targeting Potential in Gliomas. *Int J Mol Sci* **23**, (2022).
247. A. I. Giannopoulou, D. S. Kanakoglou, A. G. Papavassiliou, C. Piperi, Insights into the multi-faceted role of Pioneer transcription factors in glioma formation and progression with targeting options. *Biochim Biophys Acta Rev Cancer* **1877**, 188801 (2022).
248. T. Y. Vue *et al.*, ASCL1 regulates neurodevelopmental transcription factors and cell cycle genes in brain tumors of glioma mouse models. *Glia* **68**, 2613-2630 (2020).
249. X. Fang *et al.*, Inhibiting DNA-PK induces glioma stem cell differentiation and sensitizes glioblastoma to radiation in mice. *Sci Transl Med* **13**, (2021).
250. H. W. Lo, X. Cao, H. Zhu, F. Ali-Osman, Constitutively activated STAT3 frequently coexpresses with epidermal growth factor receptor in high-grade gliomas and targeting STAT3 sensitizes them to Iressa and alkylators. *Clin Cancer Res* **14**, 6042-6054 (2008).
251. Y. Shi *et al.*, Ibrutinib inactivates BMX-STAT3 in glioma stem cells to impair malignant growth and radioresistance. *Sci Transl Med* **10**, (2018).
252. B. Fuh *et al.*, LLL-3 inhibits STAT3 activity, suppresses glioblastoma cell growth and prolongs survival in a mouse glioblastoma model. *Br J Cancer* **100**, 106-112 (2009).
253. F. Pistollato *et al.*, Hypoxia and HIF1alpha repress the differentiative effects of BMPs in high-grade glioma. *Stem Cells* **27**, 7-17 (2009).
254. H. Caren *et al.*, Glioblastoma Stem Cells Respond to Differentiation Cues but Fail to Undergo Commitment and Terminal Cell-Cycle Arrest. *Stem Cell Reports* **5**, 829-842 (2015).
255. N. M. Aiello, B. Z. Stanger, Echoes of the embryo: using the developmental biology toolkit to study cancer. *Dis Model Mech* **9**, 105-114 (2016).
256. L. A. Garraway, W. R. Sellers, Lineage dependency and lineage-survival oncogenes in human cancer. *Nat Rev Cancer* **6**, 593-602 (2006).
257. M. Filbin, M. Monje, Developmental origins and emerging therapeutic opportunities for childhood cancer. *Nat Med* **25**, 367-376 (2019).
258. P. Dalerba *et al.*, Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat Biotechnol* **29**, 1120-1127 (2011).

259. D. A. Lawson *et al.*, Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells. *Nature* **526**, 131-135 (2015).
260. I. Ben-Porath *et al.*, An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet* **40**, 499-507 (2008).
261. Q. Weng *et al.*, Single-Cell Transcriptomics Uncovers Glial Progenitor Diversity and Cell Fate Determinants during Development and Gliomagenesis. *Cell Stem Cell* **24**, 707-723 e708 (2019).
262. A. Bhaduri *et al.*, Outer Radial Glia-like Cancer Stem Cells Contribute to Heterogeneity of Glioblastoma. *Cell Stem Cell* **26**, 48-63 e46 (2020).
263. C. P. Couturier *et al.*, Single-cell RNA-seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. *Nat Commun* **11**, 3406 (2020).
264. S. Muller *et al.*, Single-cell sequencing maps gene expression to mutational phylogenies in PDGF- and EGF-driven gliomas. *Mol Syst Biol* **12**, 889 (2016).
265. A. P. Patel *et al.*, Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396-1401 (2014).
266. V. Venkataramani *et al.*, Glioblastoma hijacks neuronal mechanisms for brain invasion. *Cell* **185**, 2899-2917 e2831 (2022).
267. K. S. Smith *et al.*, Unified rhombic lip origins of group 3 and group 4 medulloblastoma. *Nature* **609**, 1012-1020 (2022).
268. A. V. Molofsky *et al.*, Astrocytes and disease: a neurodevelopmental perspective. *Genes Dev* **26**, 891-907 (2012).
269. E. S. Akdemir, A. Y. Huang, B. Deneen, Astrocytogenesis: where, when, and how. *F1000Res* **9**, (2020).
270. M. Lattke, F. Guillemot, Understanding astrocyte differentiation: Clinical relevance, technical challenges, and new opportunities in the omics era. *WIREs Mech Dis* **14**, e1557 (2022).
271. C. Falcone *et al.*, Cortical Interlaminar Astrocytes Are Generated Prenatally, Mature Postnatally, and Express Unique Markers in Human and Nonhuman Primates. *Cereb Cortex* **31**, 379-395 (2021).
272. A. E. Trevino *et al.*, Chromatin and gene-regulatory dynamics of the developing human cerebral cortex at single-cell resolution. *Cell* **184**, 5053-5069 e5023 (2021).
273. J. D. Cahoy *et al.*, A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J Neurosci* **28**, 264-278 (2008).
274. S. Clavreul, L. Dumas, K. Loulier, Astrocyte development in the cerebral cortex: Complexity of their origin, genesis, and maturation. *Front Neurosci* **16**, 916055 (2022).
275. D. E. Allen *et al.*, Fate mapping of neural stem cell niches reveals distinct origins of human cortical astrocytes. *Science* **376**, 1441-1446 (2022).
276. Y. Yang, H. Higashimori, L. Morel, Developmental maturation of astrocytes and pathogenesis of neurodevelopmental disorders. *J Neurodev Disord* **5**, 22 (2013).
277. J. A. Stogsdill *et al.*, Astrocytic neuroligins control astrocyte morphogenesis and synaptogenesis. *Nature* **551**, 192-197 (2017).
278. H. Liu *et al.*, SOX9 Overexpression Promotes Glioma Metastasis via Wnt/beta-Catenin Signaling. *Cell Biochem Biophys* **73**, 205-212 (2015).
279. S. M. Glasgow *et al.*, The miR-223/nuclear factor I-A axis regulates glial precursor proliferation and tumorigenesis in the CNS. *J Neurosci* **33**, 13560-13568 (2013).

280. A. M. Pasca *et al.*, Functional cortical neurons and astrocytes from human pluripotent stem cells in 3D culture. *Nat Methods* **12**, 671-678 (2015).
281. J. G. Camp *et al.*, Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. *Proc Natl Acad Sci U S A* **112**, 15672-15677 (2015).
282. Z. Duren, X. Chen, R. Jiang, Y. Wang, W. H. Wong, Modeling gene regulation from paired expression and chromatin accessibility data. *Proc Natl Acad Sci U S A* **114**, E4914-E4923 (2017).
283. Z. Duren, X. Chen, J. Xin, Y. Wang, W. H. Wong, Time course regulatory analysis based on paired expression and chromatin accessibility data. *Genome Res* **30**, 622-634 (2020).
284. A. F. Chen *et al.*, NEAT-seq: simultaneous profiling of intra-nuclear proteins, chromatin accessibility and gene expression in single cells. *Nat Methods* **19**, 547-553 (2022).
285. D. Lamparter, D. Marbach, R. Rueedi, S. Bergmann, Z. Kutalik, Genome-Wide Association between Transcription Factor Expression and Chromatin Accessibility Reveals Regulators of Chromatin Accessibility. *PLoS Comput Biol* **13**, e1005311 (2017).
286. Y. Liu *et al.*, Downregulation of Pax3 expression correlates with acquired GFAP expression during NSC differentiation towards astrocytes. *FEBS Lett* **585**, 1014-1020 (2011).
287. L. Xia *et al.*, PAX3 is overexpressed in human glioblastomas and critically regulates the tumorigenicity of glioma cells. *Brain Res* **1521**, 68-78 (2013).
288. R. Tanabe *et al.*, PRRX1 induced by BMP signaling decreases tumorigenesis by epigenetically regulating glioma-initiating cell properties via DNA methyltransferase 3A. *Mol Oncol* **16**, 269-288 (2022).
289. N. Malik *et al.*, Comparison of the gene expression profiles of human fetal cortical astrocytes with pluripotent stem cell derived neural stem cells identifies human astrocyte markers and signaling pathways and transcription factors active in human astrocytes. *PLoS One* **9**, e96139 (2014).
290. T. Ando, R. Kato, H. Honda, Identification of an early cell fate regulator by detecting dynamics in transcriptional heterogeneity and co-regulation during astrocyte differentiation. *NPJ Syst Biol Appl* **5**, 18 (2019).
291. D. Henrik Heiland *et al.*, Tumor-associated reactive astrocytes aid the evolution of immunosuppressive environment in glioblastoma. *Nat Commun* **10**, 2541 (2019).
292. S. Darmanis *et al.*, Single-Cell RNA-Seq Analysis of Infiltrating Neoplastic Cells at the Migrating Front of Human Glioblastoma. *Cell Rep* **21**, 1399-1410 (2017).
293. C. Ruiz-Moreno *et al.*, Harmonized single-cell landscape, intercellular crosstalk and tumor architecture of glioblastoma. *bioRxiv*, 2022.2008.2027.505439 (2022).
294. V. M. Ravi *et al.*, Spatially resolved multi-omics deciphers bidirectional tumor-host interdependence in glioblastoma. *Cancer Cell* **40**, 639-655 e613 (2022).
295. E. Hulleman *et al.*, A role for the transcription factor HEY1 in glioblastoma. *J Cell Mol Med* **13**, 136-146 (2009).
296. A. B. Hui, K. W. Lo, X. L. Yin, W. S. Poon, H. K. Ng, Detection of multiple gene amplifications in glioblastoma multiforme using array-based comparative genomic hybridization. *Lab Invest* **81**, 717-723 (2001).
297. A. Borgenvik, M. Cancer, S. Hutter, F. J. Swartling, Targeting MYCN in Molecularly Defined Malignant Brain Tumors. *Front Oncol* **10**, 626751 (2020).
298. X. Su *et al.*, GFAP expression is regulated by Pax3 in brain glioma stem cells. *Oncol Rep* **36**, 1277-1284 (2016).

299. H. Zhu *et al.*, Transcriptional Repression of p53 by PAX3 Contributes to Gliomagenesis and Differentiation of Glioma Stem Cells. *Front Mol Neurosci* **11**, 187 (2018).
300. B. Azzarelli, L. Miravalle, R. Vidal, Immunolocalization of the oligodendrocyte transcription factor 1 (Olig1) in brain tumors. *J Neuropathol Exp Neurol* **63**, 170-179 (2004).
301. K. L. Ligon *et al.*, The oligodendroglial lineage marker OLIG2 is universally expressed in diffuse gliomas. *J Neuropathol Exp Neurol* **63**, 499-509 (2004).
302. C. Horbinski, T. Berger, R. J. Packer, P. Y. Wen, Clinical implications of the 2021 edition of the WHO classification of central nervous system tumours. *Nat Rev Neurol* **18**, 515-529 (2022).
303. P. Kickingeder *et al.*, IDH mutation status is associated with a distinct hypoxia/angiogenesis transcriptome signature which is non-invasively predictable with rCBV imaging in human glioma. *Sci Rep* **5**, 16238 (2015).
304. J. Bi *et al.*, Altered cellular metabolism in gliomas - an emerging landscape of actionable co-dependency targets. *Nat Rev Cancer* **20**, 57-70 (2020).
305. M. E. Berens *et al.*, Multiscale, multimodal analysis of tumor heterogeneity in IDH1 mutant vs wild-type diffuse gliomas. *PLoS One* **14**, e0219724 (2019).
306. K. Tateishi, H. Wakimoto, D. P. Cahill, IDH1 Mutation and World Health Organization 2016 Diagnostic Criteria for Adult Diffuse Gliomas: Advances in Surgical Strategy. *Neurosurgery* **64**, 134-138 (2017).
307. M. Sanson *et al.*, Isocitrate dehydrogenase 1 codon 132 mutation is an important prognostic biomarker in gliomas. *J Clin Oncol* **27**, 4150-4154 (2009).
308. H. Yan *et al.*, IDH1 and IDH2 mutations in gliomas. *N Engl J Med* **360**, 765-773 (2009).
309. L. Dang *et al.*, Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* **462**, 739-744 (2009).
310. W. A. Pastor, L. Aravind, A. Rao, TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat Rev Mol Cell Biol* **14**, 341-356 (2013).
311. W. Xu *et al.*, Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. *Cancer Cell* **19**, 17-30 (2011).
312. S. Ito *et al.*, Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129-1133 (2010).
313. M. Tahiliani *et al.*, Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930-935 (2009).
314. S. Kriaucionis, N. Heintz, The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324**, 929-930 (2009).
315. S. Turcan *et al.*, IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature* **483**, 479-483 (2012).
316. T. M. Malta *et al.*, Glioma CpG island methylator phenotype (G-CIMP): biological and clinical implications. *Neuro Oncol* **20**, 608-620 (2018).
317. L. P. Priesterbach-Ackley *et al.*, Brain tumour diagnostics using a DNA methylation-based classifier as a diagnostic support tool. *Neuropathol Appl Neurobiol* **46**, 478-492 (2020).
318. Z. Jaunmuktane *et al.*, Methylation array profiling of adult brain tumours: diagnostic outcomes in a large, single centre. *Acta Neuropathol Commun* **7**, 24 (2019).
319. N. J. Allen, D. A. Lyons, Glia as architects of central nervous system formation and function. *Science* **362**, 181-185 (2018).

320. P. Ballabh, A. Braun, M. Nedergaard, The blood-brain barrier: an overview: structure, regulation, and clinical implications. *Neurobiol Dis* **16**, 1-13 (2004).
321. N. J. Abbott, L. Ronnback, E. Hansson, Astrocyte-endothelial interactions at the blood-brain barrier. *Nat Rev Neurosci* **7**, 41-53 (2006).
322. S. A. Liddelow, B. A. Barres, Reactive Astrocytes: Production, Function, and Therapeutic Potential. *Immunity* **46**, 957-967 (2017).
323. M. V. Sofroniew, Astrocyte Reactivity: Subtypes, States, and Functions in CNS Innate Immunity. *Trends Immunol* **41**, 758-770 (2020).
324. D. W. Parsons *et al.*, An integrated genomic analysis of human glioblastoma multiforme. *Science* **321**, 1807-1812 (2008).
325. L. E. Huang, Friend or foe-IDH1 mutations in glioma 10 years on. *Carcinogenesis* **40**, 1299-1307 (2019).
326. S. Han *et al.*, IDH mutation in glioma: molecular mechanisms and potential therapeutic targets. *Br J Cancer* **122**, 1580-1589 (2020).
327. I. H. Lin, Y. F. Chen, M. T. Hsu, Correlated 5-Hydroxymethylcytosine (5hmC) and Gene Expression Profiles Underpin Gene and Organ-Specific Epigenetic Regulation in Adult Mouse Brain and Liver. *PLoS One* **12**, e0170779 (2017).
328. B. M. Colquitt, W. E. Allen, G. Barnea, S. Lomvardas, Alteration of genic 5-hydroxymethylcytosine patterning in olfactory neurons correlates with changes in gene expression and cell identity. *Proc Natl Acad Sci U S A* **110**, 14682-14687 (2013).
329. X. L. Cui *et al.*, A human tissue map of 5-hydroxymethylcytosines exhibits tissue specificity through gene and enhancer modulation. *Nat Commun* **11**, 6161 (2020).
330. B. He *et al.*, Tissue-specific 5-hydroxymethylcytosine landscape of the human genome. *Nat Commun* **12**, 4249 (2021).
331. W. K. Glowacka *et al.*, 5-Hydroxymethylcytosine preferentially targets genes upregulated in isocitrate dehydrogenase 1 mutant high-grade glioma. *Acta Neuropathol* **135**, 617-634 (2018).
332. J. P. Thomson, R. R. Meehan, The application of genome-wide 5-hydroxymethylcytosine studies in cancer research. *Epigenomics* **9**, 77-91 (2017).
333. A. Perera *et al.*, TET3 is recruited by REST for context-specific hydroxymethylation and induction of gene expression. *Cell Rep* **11**, 283-294 (2015).
334. Y. Xu *et al.*, Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol Cell* **42**, 451-464 (2011).
335. R. Sachdeva *et al.*, BMP signaling mediates glioma stem cell quiescence and confers treatment resistance in glioblastoma. *Sci Rep* **9**, 14569 (2019).
336. K. P. L. Bhat *et al.*, Mesenchymal differentiation mediated by NF-kappaB promotes radiation resistance in glioblastoma. *Cancer Cell* **24**, 331-346 (2013).
337. B. C. Prager, S. Bhargava, V. Mahadev, C. G. Hubert, J. N. Rich, Glioblastoma Stem Cells: Driving Resilience through Chaos. *Trends Cancer* **6**, 223-235 (2020).
338. M. R. Corces *et al.*, An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods* **14**, 959-962 (2017).
339. W. J. G. Ryan Corces, Howard Y. Chang, Isolation of nuclei from frozen tissue for ATAC-seq and other epigenomic assays. (2019).
340. J. N. Kuehner *et al.*, 5-hydroxymethylcytosine is dynamically regulated during forebrain organoid development and aberrantly altered in Alzheimer's disease. *Cell Reports* **35**, 109042 (2021).

341. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).
342. A. Dobin *et al.*, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
343. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).
344. F. J. Felix Krueger, Phil Ewels, Ebrahim Afyounian, Michael Weinstein, Benjamin Schuster-Boeckler, Gert Hulselmans, sclamons.
345. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359 (2012).
346. H. Li *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
347. B. Institute. (Broad Institute, GitHub repository).
348. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842 (2010).
349. Y. Zhang *et al.*, Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).
350. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
351. P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
352. Z. Gu, R. Eils, M. Schlesner, Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847-2849 (2016).
353. C. E. Grant, T. L. Bailey, W. S. Noble, FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017-1018 (2011).
354. J. M. Granja *et al.*, ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat Genet* **53**, 403-411 (2021).
355. I. Korsunsky *et al.*, Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* **16**, 1289-1296 (2019).
356. J. M. Granja *et al.*, Single-cell multiomic analysis identifies regulatory programs in mixed-phenotype acute leukemia. *Nat Biotechnol* **37**, 1458-1465 (2019).
357. I. T. Timothy Tickle, Christophe Georgescu, Maxwell Brown, Brian Haas. (Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA, 2019).
358. T. Liu, Use model-based Analysis of ChIP-Seq (MACS) to analyze short reads generated by sequencing protein-DNA interactions in embryonic stem cells. *Methods in molecular biology (Clifton, N.J.)* **1150**, 81-95 (2014).
359. M. R. Corces *et al.*, Single-cell epigenomic analyses implicate candidate causal variants at inherited risk loci for Alzheimer's and Parkinson's diseases. *Nat Genet* **52**, 1158-1168 (2020).
360. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357-359 (2012).
361. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, (2014).
362. S. Heinz *et al.*, Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Molecular Cell* **38**, 576-589 (2010).

363. L. Shen, N. Shao, X. Liu, E. Nestler, ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics* **15**, 284 (2014).
364. J. T. Robinson *et al.*, Integrative genomics viewer. *Nature Biotechnology* **29**, 24-26 (2011).
365. D. Kim *et al.*, TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* **14**, R36 (2013).
366. C. Trapnell *et al.*, Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* **7**, 562-578 (2012).
367. S. Budday, P. Steinmann, E. Kuhl, Physical biology of human brain development. *Front Cell Neurosci* **9**, 257 (2015).
368. G. Fossati, M. Matteoli, E. Menna, Astrocytic Factors Controlling Synaptogenesis: A Team Play. *Cells* **9**, (2020).
369. N. J. Allen, Role of glia in developmental synapse formation. *Curr Opin Neurobiol* **23**, 1027-1033 (2013).
370. L. A. Akay, A. H. Effenberger, L. H. Tsai, Cell of all trades: oligodendrocyte precursor cells in synaptic, vascular, and immune function. *Genes Dev* **35**, 180-198 (2021).
371. A. Barria, Dangerous liaisons as tumour cells form synapses with neurons. *Nature* **573**, 499-501 (2019).
372. H. S. Venkatesh *et al.*, Neuronal Activity Promotes Glioma Growth through Neuroligin-3 Secretion. *Cell* **161**, 803-816 (2015).
373. H. S. Venkatesh *et al.*, Targeting neuronal activity-regulated neuroligin-3 dependency in high-grade glioma. *Nature* **549**, 533-537 (2017).
374. H. S. Venkatesh *et al.*, Electrical and synaptic integration of glioma into neural circuits. *Nature*, (2019).
375. D. E. Bergles, J. D. Roberts, P. Somogyi, C. E. Jahr, Glutamatergic synapses on oligodendrocyte precursor cells in the hippocampus. *Nature* **405**, 187-191 (2000).
376. R. Karadottir, P. Cavelier, L. H. Bergersen, D. Attwell, NMDA receptors are expressed in oligodendrocytes and activated in ischaemia. *Nature* **438**, 1162-1166 (2005).
377. X. Liu, Q. Wang, T. F. Haydar, A. Bordey, Nonsynaptic GABA signaling in postnatal subventricular zone controls proliferation of GFAP-expressing progenitors. *Nat Neurosci* **8**, 1179-1187 (2005).
378. E. Kougioumtzidou *et al.*, Signalling through AMPA receptors on oligodendrocyte precursors promotes myelination by enhancing oligodendrocyte survival. *Elife* **6**, (2017).
379. E. M. Gibson *et al.*, Neuronal activity promotes oligodendrogenesis and adaptive myelination in the mammalian brain. *Science* **344**, 1252304 (2014).
380. K. Deisseroth *et al.*, Excitation-neurogenesis coupling in adult neural stem/progenitor cells. *Neuron* **42**, 535-552 (2004).
381. H. S. Venkatesh *et al.*, Electrical and synaptic integration of glioma into neural circuits. *Nature* **573**, 539-545 (2019).
382. D. E. Bergles, C. E. Jahr, Synaptic activation of glutamate transporters in hippocampal astrocytes. *Neuron* **19**, 1297-1308 (1997).
383. C. Luscher, R. C. Malenka, R. A. Nicoll, Monitoring glutamate release during LTP with glial transporter currents. *Neuron* **21**, 435-441 (1998).

384. J. Sibille, U. Pannasch, N. Rouach, Astroglial potassium clearance contributes to short-term plasticity of synaptically evoked currents at the tripartite synapse. *J Physiol* **592**, 87-102 (2014).
385. S. L. Campbell, S. C. Buckingham, H. Sontheimer, Human glioma cells induce hyperexcitability in cortical networks. *Epilepsia* **53**, 1360-1370 (2012).
386. C. C. John Lin *et al.*, Identification of diverse astrocyte populations and their malignant analogs. *Nat Neurosci* **20**, 396-405 (2017).
387. S. C. Buckingham *et al.*, Glutamate release by primary brain tumors induces epileptic activity. *Nat Med* **17**, 1269-1274 (2011).
388. B. R. Ransom, S. Goldring, Ionic determinants of membrane potential of cells presumed to be glia in cerebral cortex of cat. *J Neurophysiol* **36**, 855-868 (1973).
389. H. Matsushita *et al.*, Identification of glioma-specific RFX4-E and -F isoforms and humoral immune response in patients. *Cancer Sci* **96**, 801-809 (2005).
390. M. Iwafuchi-Doi, K. S. Zaret, Pioneer transcription factors in cell reprogramming. *Genes Dev* **28**, 2679-2692 (2014).
391. S. Jerabek, F. Merino, H. R. Scholer, V. Cojocaru, OCT4: dynamic DNA binding pioneers stem cell pluripotency. *Biochim Biophys Acta* **1839**, 138-154 (2014).
392. U. Basu-Roy *et al.*, Sox2 antagonizes the Hippo pathway to maintain stemness in cancer cells. *Nat Commun* **6**, 6411 (2015).
393. D. C. Di Giammartino *et al.*, KLF4 is involved in the organization and regulation of pluripotency-associated three-dimensional enhancer networks. *Nat Cell Biol* **21**, 1179-1190 (2019).
394. W. S. Song *et al.*, Sox2, a stemness gene, regulates tumor-initiating and drug-resistant properties in CD133-positive glioblastoma stem cells. *J Chin Med Assoc* **79**, 538-545 (2016).
395. K. Kobayashi *et al.*, Oct-3/4 promotes migration and invasion of glioblastoma cells. *J Cell Biochem* **113**, 508-517 (2012).
396. T. N. Ignatova *et al.*, Gliomagenesis is orchestrated by the Oct3/4 regulatory network. *J Neurosurg Sci*, (2021).
397. S. Wang *et al.*, Kruppel-like factor 4 (KLF4) induces mitochondrial fusion and increases spare respiratory capacity of human glioblastoma cells. *J Biol Chem* **293**, 6544-6555 (2018).
398. Y. P. Yang *et al.*, Inhibition of cancer stem cell-like properties and reduced chemoradioresistance of glioblastoma using microRNA145 with cationic polyurethane-short branch PEI. *Biomaterials* **33**, 1462-1476 (2012).
399. M. J. Duffy, N. C. Synnott, S. O'Grady, J. Crown, Targeting p53 for the treatment of cancer. *Semin Cancer Biol* **79**, 58-67 (2022).
400. M. B. Woodworth, K. M. Girsakis, C. A. Walsh, Building a lineage from single cells: genetic techniques for cell lineage tracking. *Nat Rev Genet* **18**, 230-244 (2017).
401. H. Zong, J. S. Espinosa, H. H. Su, M. D. Muzumdar, L. Luo, Mosaic analysis with double markers in mice. *Cell* **121**, 479-492 (2005).
402. G. La Manno *et al.*, RNA velocity of single cells. *Nature* **560**, 494-498 (2018).
403. L. Haghverdi, M. Buttner, F. A. Wolf, F. Buettner, F. J. Theis, Diffusion pseudotime robustly reconstructs lineage branching. *Nat Methods* **13**, 845-848 (2016).

404. Z. Meir, Z. Mukamel, E. Chomsky, A. Lifshitz, A. Tanay, Single-cell analysis of clonal maintenance of transcriptional and epigenetic states in cancer cells. *Nat Genet* **52**, 709-718 (2020).
405. W. A. Flavahan, E. Gaskell, B. E. Bernstein, Epigenetic plasticity and the hallmarks of cancer. *Science* **357**, (2017).
406. D. Shibata, Mutation and epigenetic molecular clocks in cancer. *Carcinogenesis* **32**, 123-128 (2011).
407. R. Chaligne *et al.*, Epigenetic encoding, heritability and plasticity of glioma transcriptional cell states. *Nat Genet* **53**, 1469-1479 (2021).
408. H. Zeng *et al.*, Bisulfite-Free, Nanoscale Analysis of 5-Hydroxymethylcytosine at Single Base Resolution. *J Am Chem Soc* **140**, 13190-13194 (2018).
409. Y. Z. Kim, Altered histone modifications in gliomas. *Brain Tumor Res Treat* **2**, 7-21 (2014).
410. N. R. Rose, M. A. McDonough, O. N. King, A. Kawamura, C. J. Schofield, Inhibition of 2-oxoglutarate dependent oxygenases. *Chem Soc Rev* **40**, 4364-4397 (2011).
411. R. Chowdhury *et al.*, The oncometabolite 2-hydroxyglutarate inhibits histone lysine demethylases. *EMBO Rep* **12**, 463-469 (2011).
412. S. Venneti *et al.*, Clinical efficacy of ONC201 in H3K27M-mutant diffuse midline gliomas is driven by disruption of integrated metabolic and epigenetic pathways. *Cancer Discov*, (2023).
413. C. Lu *et al.*, IDH mutation impairs histone demethylation and results in a block to cell differentiation. *Nature* **483**, 474-478 (2012).