Overcoming genetic heterogeneity in glioblastoma by targeting transcriptional dependencies

Tatenda Mahlokozera
Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds

Recommended Citation

This Dissertation is brought to you for free and open access by the Arts & Sciences at Washington University Open Scholarship. It has been accepted for inclusion in Arts & Sciences Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.
Overcoming Genetic Heterogeneity in Glioblastoma
by Targeting Transcriptional Dependencies
by
Tatenda Mahlokozera

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2022
St. Louis, Missouri
# Table of Contents

List of Figures ........................................ v
List of Tables ......................................... vii
Acknowledgements ..................................... viii
Abstract ................................................ ix

## Chapter 1: Introduction to Glioblastoma

1.1 Glioblastoma epidemiology ....................... 2
1.2 Therapeutic challenges ......................... 3
   1.2.1 Overview .................................. 3
   1.2.2 The blood brain barrier .................. 4
   1.2.3 Validation of target engagement and efficacy biomarkers 6
   1.2.4 Glioblastoma tumor immunology .......... 8
1.3 Genetic Heterogeneity ........................... 11
1.4 Transcriptional Heterogeneity .................... 16
1.5 Glioblastoma stem cells ......................... 21
1.6 SOX2 in Glioblastoma ............................ 26
1.7 Ubiquitin-dependent SOX2 regulation .......... 27
1.8 References ...................................... 29

## Chapter 2: Biological and therapeutic Implications of Multisector Sequencing in newly diagnosed glioblastoma

2.1 Summary .......................................... 38
2.2 Introduction ...................................... 39
2.3 Methods .......................................... 40
2.4 Compliance with ethical standards ............... 41
2.5 Results .......................................... 42
2.6 Discussion ....................................... 48
2.7 Acknowledgements ............................... 51
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.4</td>
<td>Conclusions</td>
<td>191</td>
</tr>
<tr>
<td>5.5</td>
<td>References</td>
<td>192</td>
</tr>
<tr>
<td></td>
<td><strong>Curriculum Vitae</strong></td>
<td>193</td>
</tr>
</tbody>
</table>
### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Somatic mutation landscape in treatment naïve primary glioblastoma</td>
<td>56</td>
</tr>
<tr>
<td>2.2</td>
<td>Mutational heterogeneity between tumor sectors</td>
<td>57</td>
</tr>
<tr>
<td>2.3</td>
<td>Treatment naïve hypermutated glioblastomas</td>
<td>58</td>
</tr>
<tr>
<td>2.4</td>
<td>Regional heterogeneity in subclonal architecture</td>
<td>59</td>
</tr>
<tr>
<td>2.5</td>
<td>Regional heterogeneity in potentially therapeutic gene variants</td>
<td>60</td>
</tr>
<tr>
<td>S1A</td>
<td>Recurrent copy number alterations in treatment naïve glioblastoma</td>
<td>64</td>
</tr>
<tr>
<td>S1B</td>
<td>Comprehensive representation of all copy number events</td>
<td>65</td>
</tr>
<tr>
<td>S2</td>
<td>Heterogeneity in mutational signatures between tumor sectors</td>
<td>66</td>
</tr>
<tr>
<td>S3</td>
<td>Subclonal architecture inference for all tumors</td>
<td>67</td>
</tr>
<tr>
<td>S4</td>
<td>Clonal evolution inference for all samples</td>
<td>70</td>
</tr>
<tr>
<td>S5</td>
<td>Divergence in VAFs of potentially druggable variants between sectors</td>
<td>80</td>
</tr>
<tr>
<td>3.1</td>
<td>CDC20-APC controls GSC invasion and self-renewal</td>
<td>106</td>
</tr>
<tr>
<td>3.2</td>
<td>CDC20-APC interacts with SOX2 through the WD40 domain of CDC20</td>
<td>108</td>
</tr>
<tr>
<td>3.3</td>
<td>CDC20-APC regulation of SOX2 protein and transcription controls GSC invasion and self-renewal</td>
<td>110</td>
</tr>
<tr>
<td>3.4</td>
<td>CDC20 drives GSC tumorigenicity in vivo</td>
<td>113</td>
</tr>
<tr>
<td>3.5</td>
<td>CDC20 expression is associated with decreased overall survival (OS) in Proneural subtype glioblastomas</td>
<td>115</td>
</tr>
<tr>
<td>S1</td>
<td>Characteristics of human glioblastoma stem-like cells used in this study – Related to Figure 3.1</td>
<td>129</td>
</tr>
<tr>
<td>S2</td>
<td>CDC20-APC is required for human GSC invasiveness in vitro</td>
<td>131</td>
</tr>
<tr>
<td>S3</td>
<td>The effect of CDC20-APC manipulations on cell cycle distribution, proliferation, cell death, and cell health</td>
<td>133</td>
</tr>
<tr>
<td>S4</td>
<td>CDC20-APC controls SOX2 protein stability and SOX2-dependent transcription in GSCs</td>
<td>136</td>
</tr>
<tr>
<td>S5</td>
<td>Visualization of GSC tumorigenicity in vivo by immunofluorescence</td>
<td>139</td>
</tr>
</tbody>
</table>
Figure 3.S6: The impact of CDC20 expression on overall survival in the TCGA Glioblastoma dataset

Figure 4.1: SOX2 is widely expressed among intra- and intertumoral GSC subclones and is subject to proteasomal regulation

Figure 4.2: TRIM26 maintains SOX2 expression and transcriptional activity in GSCs

Figure 4.3: TRIM26 knockdown decreases GSC self-renewal and in vivo tumorigenicity

Figure 4.4: TRIM26 protects SOX2 from polyubiquitination and subsequent proteasomal degradation

Figure 4.5: TRIM26 regulates SOX2 protein expression non-catalytically via TRIM26’s C-terminal PRYSPRY domain

Figure 4.6: WWP2 is an active E3 ubiquitin ligase in glioblastoma stem cells

Figure 4.7: TRIM26 competes with WWP2 for binding to SOX2 and blocks WWP2-dependent SOX2 polyubiquitination and proteasomal degradation
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Clinical characteristics of the study cohort</td>
<td>55</td>
</tr>
<tr>
<td>3.S1</td>
<td>Clinical characteristics of glioblastoma tumors used to derive GSCs</td>
<td>128</td>
</tr>
</tbody>
</table>
I am eternally grateful to so many people for their unwavering support and encouragement during the course of my graduate training. First, I would like to thank my mentors from before graduate school for introducing me to the wonders of scientific investigation and endowing me with a curiosity that has sustained me in my training. Thank you Das Thamattoor for believing in me, mentoring me, and advocating for me. Thank you Dr. Duncan Tate and Dr. Sallie Permar for your encouragement, support, and selfless mentorship.

I have been extremely lucky to call Albert Kim my graduate mentor. I am eternally grateful for the opportunity you gave me to embark on this incredible journey in your lab. Your mentorship, your ever-present positivity, your belief in me even when I doubted myself, and your constant encouragement have been a blessing to me. You have pushed me to be better than I thought I could be as a scientist and as a person, and for that I am deeply grateful. I would also like to thank the members of my thesis committee: Dr. Andrew Yoo, Dr. Joshua Rubin, Dr. Nima Mosammaparast, Dr. Gavin Dunn, and Dr. Azad Bonni. Thank you for all your thoughtful insights and critiques of my work and for your encouragement, I have learned a great deal from all of you. I would also like to thank Dr. Yano, not only for being on my defense committee, but for all your help and advice in my time in the lab.

Thank you also to all the members of the Kim and Yano lab that I have had the pleasure to work with and learn from over the course of my training. There are too many of you to mention, so I will not risk omitting anyone. Last but not least I want to say thank you to my family. Thank you to my wonderful sisters for always believing in me and for your constant support. Thank you most of all Mom and Dad. Thank you for your support and encouragement, and for being the perfect role models. Thank you for your unconditional love, I am incredibly lucky to be your son.

Tatenda Mahlokozera

Washington University in St. Louis

May 2022
ABSTRACT OF THE DISSERTATION

Overcoming genetic heterogeneity in glioblastoma by targeting transcriptional dependencies

by

Tatenda Mahlokozera

Doctor of Philosophy in Biology and Biomedical Sciences
Neurosciences
Washington University in St. Louis, 2022
Associate Professor Albert H. Kim, Chairperson

Glioblastoma (GBM) is the most common intrinsic central nervous system malignancy in adults, accounting for approximately 45% of such cancers. Despite advances in chemo- and radiotherapeutic approaches for various malignancies over the past decade, GBM remains an incurable disease with a dismal prognosis. Even after treatment with the current standard of care, which consists of maximal safe surgical resection, radiotherapy, and both concomitant and adjuvant chemotherapy with temozolomide, median survival is only approximately 17 months. Both treatment failure and difficulties in developing novel targeted therapies for GBM have partly been attributed to the molecular and cellular inter- and intra-tumoral heterogeneity of these neoplasms. Two major questions arise from these observations: 1) What is the extent and clinical relevance of genetic diversity in GBM? And 2) do any shared mechanisms control the malignant phenotype of GBM cells?

We first sought to profile the genetic inter- and intra-tumoral heterogeneity across ten patients harboring isocitrate dehydrogenase 1 (IDH1) wildtype GBMs, which represent 90-95% of all GBMs. Diagnostic workflows for GBM patients increasingly include DNA sequencing-based
analysis of a single tumor site following biopsy or resection. We hypothesized that sequencing of multiple sectors within a given tumor would provide a more comprehensive representation of the molecular landscape and potentially inform therapeutic strategies. We demonstrated, using image-guidance directed sampling of two to four sectors of contrast-enhancing areas of IDH1 wildtype GBM tumors, that whole-exome sequencing of individual sectors reveals a spatially divergent mutational landscape. In two extreme cases of regional heterogeneity, we described, for the first time, treatment naïve tumors with region-specific hypermutator phenotypes. In remarkable contrast to the spatial diversity of the overall mutational landscape, we demonstrated that TERT promoter mutations are unique in being recurrent in all analyzed tumors and clonal in all tumor sectors. Finally, we examined the potential therapeutic consequences of multisector sequencing data and found that multi-site analyses may be necessary to accurately characterize individual GBM tumors and identify meaningful therapeutic options.

Next, we asked whether shared epigenetic/transcriptional mechanisms might control the malignant phenotypes of genetically diverse GBM cells. GBM tumors are heterogeneous and contain a tumor-initiating pool known as glioblastoma stem-like cells (GSCs). GSCs are therapy resistant and may drive recurrence post-treatment. We reasoned that regulation of the pluripotency-related transcription factor SOX2, which is indispensable for the tumorigenicity of GSCs, may represent one such mechanism. First, we found that the mitotic E3 ubiquitin ligase, CDC20-Anaphase-Promoting Complex (APC), drives the invasiveness, self-renewal, and tumorigenic capacity of multiple, genetically heterogenous primary GSC lines. Mechanistically, we found that CDC20-APC operates through SOX2 to control GSC phenotypes by regulating SOX2 protein stability and transcriptional activity. Second, we used immunoprecipitation followed by mass spectrometry (IP-MS) to identity the E3 ubiquitin ligase TRIM26, previously reported to play a
role in immune regulation. We found that TRIM26 directly binds to SOX2 via TRIM26’s C-terminal PRY-SPRY domain. Unexpectedly, we found that TRIM26 knockdown decreased SOX2 protein stability and conversely increased SOX2 polyubiquitination in primary GSCs. Accordingly, TRIM26 knockdown reduced SOX2 transcriptional activity, self-renewal, and in vivo tumorigenicity in genetically divergent GSC lines. Mechanistically, we discovered TRIM26 stabilizes SOX2 protein by competitively reducing the interaction of SOX2 with WWP2, a bonafide SOX2 E3 ligase in GSCs. Consistent with this hypothesis, WWP2 depletion in the setting of TRIM26 knockdown rescued SOX2 protein levels, self-renewal, and in vivo tumorigenicity in GSCs. Taken together, our data provide further evidence of genetic heterogeneity among and within GBM tumors. However, epigenetic and transcriptional mechanisms controlling the GSC state appear to be a more common feature of GBM, raising the intriguing possibility that disruption of the malignant epigenetic/transcriptional landscape of GSCs represents a unifying therapeutic strategy.
CHAPTER 1:

Introduction to Glioblastoma
1.1 Glioblastoma epidemiology

Glioblastoma (GBM) is the highest grade (WHO Grade 4) and most aggressive diffuse glioma originating from the astrocytic lineage, with a 5-year overall relative survival of 6.8%. Advanced age and male sex both portend poorer clinical outcomes (Ostrom et al., 2019). Critically, this aggressive neoplasm is also the most common intrinsic brain and central nervous system malignancy in adults, accounting for approximately 45% of such malignancies (Thakkar et al., 2014). The age-adjusted overall incidence is 3.22/100,000, with increased incidence associated with male sex and advanced age at diagnosis (Ostrom et al., 2019). Depending on the genetic pathways dysregulated in tumor pathogenesis and on the age of onset, GBMs are typically segregated into two broad categories, namely primary and secondary GBMs. Primary GBMs, also known as de novo GBMs, comprise approximately 80% of disease incidence, with a mean age at diagnosis of 62 (Thakkar et al., 2014). Secondary GBMs typically arise from lower grade Astrocytomas and oligodendrogliomas and have a mean age at diagnosis of 45. Histopathologically, GBMs are characterized by hypercellularity consisting of small or giant multinucleated anaplastic cells, high mitotic activity, necrosis, and endothelial multi-layering or glomeruloid microvascular proliferation (Homma et al., 2006; Louis et al., 2007).

However, recent recommendations place less emphasis on classic pathologic features for both diagnosis and initiation of management plans for glioblastoma. The cIPMACT-NOW (the Consortium to Inform Molecular and Practical Approaches to CNS Tumor Taxonomy) update made new recommendations for criteria to diagnose IDH-wildtype diffuse astrocytic gliomas. Gliomas with histologic features of WHO grade II/III but presenting with either TERT promoter mutations, combined chromosome 7 gain and 10 loss, or EGFR amplification are now regarded as
having molecular features of WHO grade IV glioblastoma and recommended to be diagnosed and managed as such (Brat et al., 2018). Additionally, due to their distinct biology and markedly better overall prognosis, IDH-mutant tumors, which comprise 5-10% of all diagnosed glioblastomas (Louis et al., 2016; Yan et al., 2009), are no longer recommended to be classified as glioblastoma, but instead as “Astrocytoma, IDH-mutant, WHO Grade IV” (Brat et al., 2020).

1.2 Therapeutic challenges:

1.2.1 Overview

Despite advances in chemotherapeutic and radiotherapy approaches to the management of GBM and various other malignancies over the past decade, GBM remains an incurable disease with a dismal prognosis. The median survival without therapeutic intervention is around 3 months. Even after treatment with the current standard of care, which consists of maximal safe surgical resection, radiotherapy, and both adjuvant and concomitant chemotherapy with temozolomide, the median survival is only approximately 17 months (Stupp et al., 2009; Stupp et al., 2005; Thakkar et al., 2014). The majority of patients show minimal response to standard therapy, while those that initially respond universally experience disease recurrence. A number of factors contribute to the resistance to standard therapy seen in glioblastoma as well as to the lack of progress in the development of new therapeutic strategies, including 1) A severe bottleneck on deployable molecular therapies due to their inability to cross the blood-brain barrier 2) The feasibility and ethical challenges, due to anatomic location, of measuring objective target engagement and establishing reliable biomarkers of therapy response by serial biopsies 3) The phenomenon of a severe immunosuppressive environment in glioblastomas, which limits efficacy of immunotherapeutic approaches and 4) The extensive inter- and intratumoral heterogeneity of
glioblastomas and its implications for combination therapies required to attain broad coverage of actionable genetic alterations.

1.2.2 The blood brain barrier:

The blood-brain barrier (BBB) avails a selective barrier between the brain and the systemic circulation and is essential for the maintenance of normal brain physiology by preventing entry of infectious and neurotoxic agents into the central nervous system (van Tellingen et al., 2015). Anatomically, the BBB is composed of specialized endothelial cells connected by tight junctions, as well as pericytes, and astrocyte foot processes which together give rise to the unique junctional integrity of the BBB (Abbott, 2013; Abbott et al., 2006). Perivascular macrophages that line the basal lamina of the endothelial cells further regulate the rigidity of the BBB (Abbott et al., 2006; da Fonseca et al., 2014; Obermeier et al., 2013). The BBB excludes most molecules from passive transport except small, uncharged, and lipophilic molecules. Required nutrients and larger molecules gain access to the brain via transporters and receptor-mediated endocytosis (Ballabh et al., 2004). Efflux pumps, such as P-glycoprotein, impose additional restrictions by actively transporting molecules back into the blood (Harder et al., 2018). Consequently, approximately 98% of molecules are excluded from accessing the brain, presenting a significant challenge for small molecule brain cancer therapeutics. While portions of GBM tumors are known to have a leaky BBB, regions such as the infiltrative edge that is typically left behind after resection have an intact BBB (Arvanitis et al., 2020; Sarkaria et al., 2018).

A number of strategies have and are being actively developed to overcome drug delivery limitations imposed by the BBB. Intraoperative placement of a polymer impregnated with the
chemotherapeutic drug carmustine (Gliadel wafer) in the surgical resection cavity has been used to take advantage of the mechanical disruption of the BBB during surgery to allow for drug delivery(Brem et al., 1995; Westphal et al., 2003; Westphal et al., 2006). The limitation of this approach is the low therapeutic efficacy due to single dose delivery of drug over a short time frame(Fleming and Saltzman, 2002). Other groups have employed dual inhibition of two major BBB drug efflux proteins, P-glycoprotein and ABCG5, and observed increased brain penetrance of the chemotherapy drug temozolomide, which significantly enhanced the survival of tumor-bearing mice(de Gooijer et al., 2018). A more recent strategy has been to employ focused ultrasound (FUS) coupled with microbubbles to disrupt the BBB and allow for delivery of therapeutic agents. This approach successfully disrupts the blood brain barrier in animal models, allowing for both drug delivery and release of biomarkers into the bloodstream for liquid biopsy applications and is now being tested in human clinical trials(Etame et al., 2012; Harder et al., 2018; Pacia et al., 2020; Zhu et al., 2018). Finally, in a pilot study in fourteen patients with recurrent glioblastoma, hyperthermia induced by magnetic resonance imaging (MRI)-guided laser interstitial therapy (LITT) was shown to cause both tumor cytoreduction and disruption of the peritumoral BBB, with peak permeability occurring within 1-2 weeks and resolving within 4-6 weeks(Leuthardt et al., 2016). This enhanced window of BBB permeability is promising for sustained drug delivery and MRI-guided LITT is currently in human clinical trials in combination with either chemotherapeutic drug doxorubicin (NCT01851733) or immune checkpoint inhibitor pembrolizumab (NCT03277638) for treatment of recurrent GBM. Together, these recent innovations are significant and may soon increase the repertoire of efficacious small molecule drugs that can be used to treat glioblastoma patients.
1.2.3 Validation of target engagement and efficacy biomarkers:

A second therapeutic challenge in glioblastoma is imposed by the privileged anatomy of the brain and the intrinsic elevated risk of neurologic invasive procedures, including biopsies, compared to most other cancers/tissue sites. Because of the tenuous ethics and justifications for performing serial neurologic surgery/biopsy for the singular purpose of confirming drug activity by tissue assays, investigators have sought to creatively design clinical trials for both primary and recurrent GBM to address this challenge. Indeed, it is likely that many agents that have failed clinical trials did not reach therapeutic levels in the brain or did not cause the expected modulation of target protein activity. One phase II trial in recurrent GBM with PI3K pathway activation stratified patients into two cohorts. In cohort 1, patients with operable tumors were given neoadjuvant PI3K pathway inhibitor buparlisib for 1 to 2 weeks before surgery to evaluate the levels of drug penetration into the brain and assess the modulation of PI3K pathway activation in resected tumor samples using immunohistochemistry for phosphorylated AKT\textsuperscript{S473} and phosphoribosomal protein S6\textsuperscript{235/236} (Wen et al., 2019). The primary endpoint for this group was PI3K pathway inhibition in tumor tissue as well as drug pharmacokinetics. In cohort 2, patients with inoperable recurrent tumors were administered buparlisib on a 28-day schedule until progression or unacceptable toxicity. The primary endpoint for cohort 2 was 6-month progression free survival. The trial revealed minimal efficacy of single-agent buparlisib despite significant brain penetration. The lack of efficacy could largely be explained by incomplete inhibition of the PI3K pathway. In a similarly designed phase I proof of concept study in operable recurrent GBMs with loss of PTEN expression treated with mTOR inhibitor rapamycin, the magnitude of mTOR inhibition was found to associate with reduced tumor cell proliferation as measured by Ki67
staining of resected tumor specimens (Cloughesy et al., 2008). In patients where AKT activation occurred upon treatment, presumably due to loss of negative feedback, this activation resulted in a shorter time-to-progression during post-surgical rapamycin maintenance therapy. These studies suggest that neoadjuvant drug administration in the setting of recurrent GBM may be useful in measuring drug penetrance into the brain and assessing biomarkers of on-target drug activity and clinical response, while also providing biomarker-based explanations for treatment failure to inform future clinical trial design.

Evidence of drug penetrance and activity from trials in recurrent GBM can be leveraged in designing trials for primary GBM patients. The NCT Neuro Master Match (N²M²) trial is a phase I/IIa trial for patients with newly diagnosed IDH-wildtype GBM without MGMT promoter methylation who are unlikely to benefit from the standard chemotherapy agent temozolomide (Wick et al., 2019). In the discovery phase, molecular diagnostics in a 4 week timeframe post-surgery are used to detect predefined biomarkers that predict response to targeted treatments. Patients are then stratified into targeted treatment arms with endpoints of toxicity for phase I and 6 month progression-free-survival for phase IIa. All patients received radiotherapy, and the common control arm for all drug-trial arms received TMZ as standard of care. As with other innovative adaptive clinical trials in GBM, a major advantage of this study is the use of a single control arm to test multiple drug candidates, an approach that will accelerate the discovery of novel therapies for glioblastoma patients (Alexander et al., 2018). This trial is the first to systematically test preselected patients in patients with predictive response biomarkers in primary glioblastoma, and opens the door to designs to test additional drugs as well as other innovative combination therapies, while integrating objective measures of on-target drug activity generated
from neoadjuvant/maintenance drug regimens in the setting of recurrent glioblastoma as described above.

1.2.4 Glioblastoma tumor immunology: An “immunologically cold” tumor:

While the concept of the brain as an isolated “immune privileged” site (Medawar, 1948) has been disproved, by discovery of the afferent glial-lymphatic (glymphatic) pathway (Eide et al., 2018; Iliff et al., 2012) and meningeal lymphatic vessels (Aspelund et al., 2015; Louveau et al., 2018; Louveau et al., 2015) as well as observations of T cell immunosurveillance in the brain (Owens et al., 2008; Schlager et al., 2016; Zamvil and Steinman, 1990), glioblastomas and other brain tumors are nevertheless characterized by a significantly immunosuppressive environment. As a general phenomenon, brain tumors typically have low numbers of tumor-infiltrating lymphocytes (TILs) compared to other tumor types (Gajewski et al., 2013). This relative depletion in TILs persists even after vaccination, with the limited antigen-specific T cells that are detected in the brain largely exhibiting an exhausted phenotype (Keskin et al., 2019). One recently postulated potential mechanism for relative T cell depletion in brain tumors is sequestration of T cells in the bone marrow of patients due to tumor-driven loss of sphingosine-1-phosphate receptor 1 (S1P1) from the surface of T cells (Chongsathidkiet et al., 2018). The reduced trafficking of T cells to brain tumors has posed significant challenges for the successful implementation of immunotherapeutic regimens, such as immune checkpoint inhibitors.

In addition to the limited trafficking of T cells to the tumor, additional barriers to immunotherapy in glioblastoma are imposed by an immunosuppressive microenvironment, which results in significant TIL dysfunction and exhaustion. Glioma cells themselves contribute to an
immune suppressive environment by producing and secreting indolamine 2,3-dioxygenase (IDO). IDO in glioblastoma tumors results in stimulation and accumulation of regulatory T cells (T<sub>reg</sub>) and suppresses antigen-specific TIL function by depleting tryptophan from the microenvironment (Uyttenhove et al., 2003; Wainwright et al., 2012). The major immunomodulatory cells in the glioblastoma microenvironment are tumor associated macrophages (TAMs), which include resident microglia and bone marrow-derived macrophages (BMDM). Anti-inflammatory TAMs express high levels of immunosuppressive cytokines such as interleukin-10 (IL-10) and transforming growth factor beta (TGF-β) (Gong et al., 2012; Perng and Lim, 2015; Vitkovic et al., 2001). IL-10 downregulates MHC class II expression in monocytes and also inhibits immune cell production of interferon gamma (IFN-γ) and tumor necrosis factor alpha (TNFα). Additionally, IL10 can induce expression of PD-L1 on the surface of brain tumor macrophages, which engages the PD-1 receptor on TILs resulting in TIL anergy and apoptosis (Perng and Lim, 2015). Additionally, TAMs produce significant amounts of arginase, which results in depletion of arginine from the microenvironment and subsequent inhibition of T cell proliferation and function (Zhang et al., 2016). Because CSF amino acids, like arginine, are taken up via active transport across the blood brain barrier, resulting in approximately 90% lower levels compared to the blood, the brain is particularly susceptible to amino acid depletion (Hawkins et al., 2006). Given the significant role of TAMs in regulating immune responses in glioblastoma, approaches to target and deplete TAMs from the tumor microenvironment have been explored. The cytokine colony-stimulating factor 1 (CSF1) regulates the differentiation and survival of macrophages via CSF1 receptor (CSFR1)-dependent signaling. While, CSF1R inhibition in a preclinical studies led to significantly increased survival of tumor-bearing mice, a phase II clinical trial did not show any objective response (Butowski et al., 2016). Follow-up studies have suggested
that resistance to CSF1R inhibition may occur via aberrant PI3K signaling, further highlighting the importance of multimodality therapeutic approaches for glioblastoma (Quail et al., 2016).

Finally, another cause of immune dysfunction in glioblastoma is exposure to the current standard of care therapy. One of the well-described side effects of temozolomide (TMZ) treatment is the development of myelosuppression (Su et al., 2004). Further, combination treatment with radiation and TMZ results in an overall immunosuppressive state peripherally by altering the balance between effector and regulatory immune cells (Fadul et al., 2011; Grossman et al., 2015). However, clinical data lend a nuanced view to the issue of myelosuppression in the setting of TMZ therapy. In trials comparing standard versus intensive doses of TMZ, it was observed that intensive TMZ, while leading to a proportional increase in immunosuppressive \( T_{reg} \) cells, resulted in a substantial magnitude of humoral and delayed-type hypersensitivity responses (Asavaroengchai et al., 2002; Batich et al., 2017; Sampson et al., 2011). Indeed, vaccine studies suggest that dose dense TMZ-induced lymphopenia may result in an enhanced vaccine-elicited immune response (Migliorini et al., 2018; Padfield et al., 2015).

1.2.5 Inter- and intratumoral heterogeneity:

Glioblastoma, like other cancers, harbors extensive genetic, transcriptional, and cellular heterogeneity both among and within individual patient tumors. Intertumoral heterogeneity hampers the development of therapies that have efficacy for a large number of patients, and requires careful stratification of patients for clinical trials based on biomarkers that predict response to targeted therapies. Intratumoral heterogeneity makes the challenges of treating glioblastoma, as well as other cancers, much more significant. Indeed, a patient’s tumor can not
necessarily be regarded as one disease but rather a collection of such, given how heterogenous malignant cells in a tumor are. A deeper understanding of the biologic and therapeutic implications of intratumoral heterogeneity is critically required to move both molecular and immunologic therapies for glioblastoma forward. The levels of heterogeneity in glioblastoma are discussed in more detail in the sections below.

1.3 Genetic heterogeneity:

The revolution in cancer biology in the genomic sequencing era, which allowed multiplatform comprehensive genetic sequencing analysis of multiple tumors has greatly informed our understanding of the major alterations that drive many different cancers, the timing of these events in tumor pathogenesis, and their influence and dynamics in therapy resistance, clonal evolution, and disease recurrence. Two pioneering studies in this genomics-driven era confirmed the genetic evidence that had been accumulated in smaller series for decades, and further delineated novel molecular events in glioblastoma. Parsons et al performed whole exome sequencing and genome-wide copy number analysis of 22 glioblastoma samples (Parsons et al., 2008). They observed alterations that had been previously identified by various groups (Furnari et al., 2007): CDKN2A was altered in 50% of cases. TP53, EGFR, and PTEN were altered in 30-40% of GBMs. NF1, CDK4, and RB1mutations occurred in 12-15% of glioblastomas, while PIK3CA and PIK3R1 mutations were observed in 8-10% of GBMs. In integrative pathway analysis, the TP53 pathway (TP53, MDM2, and MDM4), the RB1 pathway (RB1, CDK4, and CDKN2A), and the PI3K/PTEN pathway (PIK3CA, PIK3R1, PTEN, and IRS1) were found to be altered in 64%, 68%, and 50% of GBMs respectively. Critically, while these findings highlighted essential pathways that are altered
in the majority of glioblastoma tumors, they revealed intertumoral heterogeneity in the driver events that led to aberrations in these pathways. Indeed, the authors observed that in all cases but one, mutations within each tumor affected only a single member of each pathway in a mutually exclusive manner. Significantly, Parsons et.al were the first to identity recurrent heterozygous somatic IDH1 mutations in glioblastoma, all of which affected the arginine residue at position 132 (R132). They further made fundamental observations that have since been confirmed by multiple subsequent studies that 1) The IDH1 mutation preferentially occurred in younger patients, 2) The mutation occurred in nearly all cases of secondary GBM, and 3) IDH1 mutation was associated with significantly improved overall survival. Clinical algorithms routinely assess IDH1 mutation status in patient care, and it is now largely accepted that IDH1 mutant glioblastoma is a distinct disease, both biologically and in terms of clinical outcomes, from IDH1-wildtype GBM(Dunn et al., 2013).

The second landmark genome-wide sequencing study that has revolutionized glioblastoma research was The Cancer Genome Atlas (TCGA) analysis of 206 glioblastomas(Cancer Genome Atlas Research, 2008). Integrating copy number analysis and somatic mutations, these authors found that the RB, TP53, and RTK pathways were altered in 87%, 78%, and 88% of GBMs respectively. Similar to Parsons et.al, the TCGA study found a statistical trend towards mutual exclusivity of alterations of components within each pathway. However, they further found that 74% of tumors harbored mutations in all 3 pathways, suggesting that dysregulation of each of these pathways was a requisite pathogenic event in glioblastoma. The TCGA study further clarified the importance of NF1 as a tumor suppressor gene in glioblastoma, finding that NF1 inactivating mutations or deletions occurred in 23% of GBMs sequenced. Finally, a major critical observation made by the TCGA group was that in approximately 10% of tumors, at least 2 of the 4 commonly
altered receptor tyrosine kinases (EGFR, ERBB2, PDGFRA, and MET) harbored amplifications or point mutations, raising the possibility of genomic alterations as a basis for co-activation of RTKs in glioblastoma. Indeed, the observation of RTK coactivation by phosphor-RTK antibody array-based methods had been made by the DePinho group using GBM tumor cell lines, xenotransplants, and primary tumor specimens (Stommel et al., 2007). RTK-coactivation rendered single agent therapies ineffective in impacting glioma cell survival, highlighting the need for combined regimens to target downstream signaling outputs of multiple RTKs. Snuderl et al further confirmed the genomic basis of RTK coactivation postulated by the TCGA study when they observed mosaic amplifications of multiple RTKs in glioblastoma specimens by fluorescence in situ hybridization (FISH) (Snuderl et al., 2011). They further found that the amplifications occurred in different intermingled cells in a tumor in a mutually exclusive manner and that these cells shared common early genetic mutations and thus represented subclones derived from a common precursor cell. These findings were corroborated in a second study that was published around the same time (Szerlip et al., 2012). In addition to highlighting the challenges for RTK-targeting therapies in GBM, these studies lent impetus to further research to fully understand the extent of intratumoral heterogeneity in glioblastoma and its therapeutic implications beyond RTK-based approaches.

An influential study that laid the groundwork for studies to dissect intratumor genetic heterogeneity was performed by Gerlinger et al in renal carcinoma (Gerlinger et al., 2012). Using multiregional whole exome sequencing of spatially separated samples from primary renal carcinomas and associated metastatic sites, they observed branched evolutionary tumor growth by phylogenetic reconstruction and found that 63-69% of somatic mutations were not detectable across every tumor region. The findings of this study suggested that single tumor-biopsy specimens potentially underestimate the mutational burden of heterogenous tumors. Further, the
authors proposed the identification and therapeutic targeting of common truncal mutations detected from phylogenetic analysis of multiregional sequencing data as a more robust treatment approach to avoid drug resistance due to Darwinian selection of preexisting drug resistant clones.

In the glioblastoma space, the first study to assess intratumoral genetic heterogeneity was by Sottoriva et al. (Sottoriva et al., 2013). They found, using genome-wide copy number analysis, that a subset of canonical driver genes in GBM (PDGFRA, MDM4, and PTEN) harbored subclonal copy number alterations (CNA) that were not shared between different regions of the same tumor.

Phylogenetic analysis based on integration if sequencing data from multiple tumor regions revealed branching tumor evolution and demonstrated that CNAs in some driver genes, like EGFR and CDKN2A, occurred early in tumor evolution, while those in PDGRA and PTEN tended to occur later in tumor evolution. Using a molecular clock based on the IRX2 genomic locus, which displays neutral methylation and can be used to count the number of cell divisions, the authors identify proliferative subclones in each region for different tumors and observe that subclones in one tumor region were not represented in other sectors from the same patient’s tumor, highlighting the genetic heterogeneity between tumor regions and the phenomenon of physical subclone regional restriction.

In a second multisector study of intratumoral heterogeneity in GBM, the Verhaak group employed whole genome and whole exome sequencing of multiple biopsy specimens from matched primary and recurrent glioblastomas (Kim et al., 2015). Across all samples sequenced, 36% of mutations were found to be private and not shared across all tumor biopsy specimens sequenced. The authors found that p53 mutations were associated with an increased number of subclonal mutations in primary tumors, and further that p53 alterations predicted a marked increase in the frequency of subclonal mutations upon recurrence. Phylogenetic analysis of matched
primary and recurrent samples for individual patients revealed two major routes of evolution. In the branched/clonal evolutionary recurrence model, the biopsies from the recurrent tumor shared a large number of clonal mutations with at least one of the regions from the initial tumor and there was a dearth of primary tumor-specific clonal mutations. Practically, such a model suggests that the recurrence developed from residual primary disease, with specific subclones resisting therapy and subsequently expanding at relapse. In the ancestral origin model of recurrence, the recurrent tumor shared a small cluster of clonal mutations with the primary tumor, however a large number of primary tumor clonal mutations disappeared in the recurrence, leading to limited clonal overlap. This model suggests that the primary and recurrent tumors evolved independently and that both tumors arose from a common ancestral cell. Practically, the ancestral model suggests effective removal of dominant primary disease subclones, with only the refractory ancestral cell making it through the therapeutic bottleneck.

While the above studies are highly informative with regards to the clonal evolutionary biology of glioblastomas in both primary disease and upon relapse, they do not necessarily give insights into how regional genetic heterogeneity in glioblastoma can be leveraged in clinical decision-making and in tailoring personalized therapeutic regimens. Parker et.al performed multiregional targeted sequencing, methylation analysis, and transcriptional profiling to begin to address these questions(Parker et al., 2016). They found heterogeneity in both the occurrence of subclonal mutations in and expression of DNA repair pathway genes, for both the mis-match repair (MMR) and base excision repair (BER) processes, across tumor sectors from the same patient. Given the association between tumor mutational load and response to immunotherapeutic agents, accurate detection of DNA repair pathway alterations, which are associated with genomic instability and higher rates of mutation, may be informative for selection of therapy. Additionally, the authors
observed that 14% of patients demonstrated heterogeneity in MGMT promoter methylation status, suggesting that a significant proportion of patients may be misclassified as MGMT unmethylated and denied potentially beneficial TMZ therapy based on single-site biopsies. In another study to assess the therapeutic implications of regional genetic heterogeneity in glioblastoma, Lee et al. assayed the response of tumor cells derived from multiple regions of GBM tumors and found that drugs targeting truncal mutations that were shared between tumor regions were more efficacious than drugs against heterogeneously altered targets (Lee et al., 2017). Taken together, these studies argue for implantation of multi-site biopsy and molecular profiling approaches to tailor personalized therapies for glioblastoma patients. A major practical question pertains to how many multisite biopsies would be needed to get enough information about heterogeneous actionable genetic alterations. It has been suggested, in the setting of medulloblastoma, that on average five biopsies would be required to cover 80% of actionable mutations, posing significant challenges with regards to workflow and cost (Morrissy et al., 2017).

1.4 Transcriptional Heterogeneity:

In parallel with the advances in understanding the genetic underpinnings of glioblastoma, significant progress has been made in elucidating the gene expression profiles of these tumors in an effort to better understand GBM pathogenesis and to find novel targets for therapeutic intervention. Two influential studies initially shaped the organization of GBM transcriptional profiles into meaningful modules that appeared to have value as prognostic indicators. Phillips et al. combined DNA microarray and comparative genomic hybridization (CGH) assays to integrate expression data with copy number alterations in high grade astrocytomas and
glioblastoma (Phillips et al., 2006). They defined 3 subtypes based on consensus clustering of microarray data. The Proneural subtype was associated with expression of OLIG2, DLL3, and BCAN and was associated with younger age at diagnosis. The Proliferative subtype was associated with expression of PCNA and TOP2A as well as gain of chromosome 7 and loss of chromosome 10. The Mesenchymal subtype was associated with expression of CHI3L1, CD44, and VEGF. The authors found that patients with Proneural tumors had longer overall survival. Further, this study posited that tumors tended to transition to Mesenchymal subtype upon disease recurrence, a phenomenon that was corroborated for Proneural tumors by both mathematical modeling and engineered mouse models years later (Ozawa et al., 2014).

A more influential study that integrated transcriptomics and genomics in GBM to define tumor subtypes was performed by Verhaak et al. based on data from The Cancer Genome Atlas Research Network (Verhaak et al., 2010). This study identified 4 subtypes of glioblastoma, 2 of which (Proneural and Mesenchymal) were consistent with those postulated by Philips et al. Indeed, both groups identified DLL3 and OLIG2 as Proneural subtype genes, as well as CHI3L1/YKL40 and CD44 as Mesenchymal subtype-associated genes. Genetically, the Proneural subtype in Verhaak et al. was enriched for PDGFRA amplifications and TP53 mutations. The Mesenchymal subtype was associated with point mutations and deletions in NF1, as well as PTEN loss. The third subtype, termed Classical, was found to typically harbor EGFR amplification, CDKN2A loss, and chromosome 7 loss paired with chromosome 10 gain. Classical tumors featured high expression of Notch (NOTCH3, JAG1, and LFNG) and Sonic hedgehog (SMO, GAS1, and GLI2) signaling pathway genes. The fourth subtype, termed Neural, was not associated with any of the canonical genetic alterations and was defined by expression of neuronal markers such as NEFL, GABRA1, and SLC12A5. In an updated analysis of the TCGA data, coupled with single cell RNA-seq of
tumor specimens and cell lines models, the Verhaak group showed that the Neural subtype was an artifact of contamination by normal brain tissue (Wang et al., 2017). From a clinical standpoint, the most significant association was the overrepresentation of younger patients in the Proneural subtype, consistent Phillips et al. In addition, the authors observed that intensive treatment, with concurrent temozolomide and radiation or greater than four cycles of temozolomide, was associated with longer survival in Classical and Mesenchymal tumors, but not in Proneural or Neural subtype cases. Taken together, the studies by Phillips et al and Verhaak et al provided a clear view of transcriptional intertumoral heterogeneity, though the subtype classifications have not proven to have much relevance to the clinical management of glioblastoma patients.

A major limitation of the initial attempts to profile transcriptional heterogeneity in glioblastoma is that they largely yielded average expression data for all the cells in a single specimen from the bulk tumor. This is particularly problematic for tumors that have extensive macrophage and microglial infiltration. Indeed, multiregion sequencing approaches suggested that the Verhaak subtypes can co-exist in an individual patient’s tumor. Sottoriva and colleagues found that in 60% of cases, fragments of tumor from the same patients could be classified into at least two different subtypes (Sottoriva et al., 2013). Similarly, Parker et al. found that 40% of their multi-regional biopsy cases demonstrated intratumoral subtype heterogeneity, leading to a non-consensus transcriptional subtype (Parker et al., 2016).

The advent of single cell RNA-sequencing (scRNA-seq) technologies has moved cancer genomics into a formidable realm where transcriptomic and genomic data can be associated with specific cell types and cell states. One of the pioneering scRNA-seq studies in glioblastoma, similar to the multisector transcriptional profiling data, provided a much more nuanced picture of tumor subtype classification, revealing that transcriptional signatures, much like genetic
alterations, exhibit appreciable intratumoral heterogeneity. Across five glioblastoma tumors, Patel et al. were able to show the admixture of cells from different subtypes in the same tumor (Patel et al., 2014). The population-level data, such as that in the bulk sequencing in the TCGA dataset, was able to capture the dominant transcriptional program (for instance the majority of cells in TCGA-Proneural samples were found to express a Proneural gene signature) but did not reveal the true diversity of transcriptional subtypes in individual tumors. Strikingly, regardless of their TCGA subtype, all tumors were found to harbor Proneural cells, perhaps in line with the suggestion that most non-GCIMP glioblastoma subtypes are derived from a common Proneural-like precursor cell (Ozawa et al., 2014). Another critical observation in the Patel study was the existence of cells in hybrid states, for example cells that could be classified as both Classical and Proneural, suggesting either an anomalous transcriptional program or plasticity of subtypes. In the same vein, there was not a clear separation between stem-like and differentiated cells in transcriptional space. Rather, cells exhibited a continuum along a stemness-differentiation axis. In terms of clinical correlates, the authors found that tumors for Proneural classified tumors, a higher Proneural score (less signatures from other subtypes) was associated with longer survival.

In addition to redefining their subtype classification system in a newer study that integrates scRNA-seq, the Verhaak group provided additional insights into the implication of intratumoral transcriptional heterogeneity (Wang et al., 2017). There was a correlation between the subtype complexity of a tumor (increased number of admixed subtypes) and both the rate and fraction of subclonal mutations. However, as has been described previously in bulk sequencing studies, high purity Mesenchymal tumors had the worst prognosis. From a microenvironmental standpoint, Mesenchymal subtype enrichment was associated with increased infiltration of macrophages and microglia, with this phenomenon being associated with NF1 loss. In line with this observation,
non-Mesenchymal tumors that transitioned to Mesenchymal on recurrence had increased infiltration of immune cells.

In an integrative analysis of scRNA-seq data from 28 tumors and TCGA data, the Suva group recently defined four new states that explain the transcriptional heterogeneity in glioblastoma further associated them with canonical genetic alterations in glioblastoma (Neftel et al., 2019). This study defined the AC-like state, which has expression profile similar to TCGA-Classical subtype, and is associated with EGFR amplification. The MES-like state corresponded to the TCGA-Mesenchymal subtype and was associated with NF1 mutations. OPC-like and NPC-like states both corresponded to the TCGA-Proneural subtype and were enriched for CDK4 amplification. The authors further demonstrated that intra-tumoral diversity in expression states is not driven by genetic subclones, however specific state-associated genetic alterations biased the state distribution frequency towards that particular state. Using markers derived from the transcriptomic data (CD24 for NPC-like and CD44 for MES-like) combined with barcoding, the authors show that injection of a single-subtype of cells results in the repopulation of all the original subtypes in the parent tumor and that barcoded cells can be found distributed across multiple subtypes, providing evidence for a hardwired transcriptional plasticity in glioblastoma cells.

The studies outlined above, together with many others, suggest that intratumoral transcriptional heterogeneity is the rule in glioblastoma, and further that the propensity for diversification is an inherent feature that is independent of tumor genetics or the microenvironment. In the face of such heterogeneity, it is imperative to find rules that hold within and across tumors in order to identify targetable dependencies. Indeed one common theme in the accumulating single cell data, is the existence of a developmental stratification of cancer cells, with all tumors possessing undifferentiated cells that resemble progenitors in the normal brain.
1.5 Glioblastoma stem cells:

In addition to genetic and transcriptional heterogeneity, it has been established that GBM is also characterized by appreciable cellular heterogeneity. One particular cell type that has generated significant research interest is the glioblastoma stem-like cell (GSC), also known as the glioblastoma tumor-initiating cell. The cancer stem cell hypothesis postulates the existence of a cellular hierarchy in cancer, in which only a small proportion of cells within a given bulk tumor are capable of propagating the cancer. Therefore, therapeutic eradication of GSCs is posited to be essential to achieve a complete cure in this model (Nguyen et al., 2012). By definition, GSCs are the source of all other multi-lineage, non-tumor propagating cells within the bulk malignancy. An important rationale for evoking the cancer stem cell hypothesis is that, similar to normal tissue stem cells that constitute a life-long reservoir for tissue regeneration, cancer stem cells have active self-renewal capabilities and an indefinite capacity for cell division, thus they can undergo the sufficient number of divisions necessary to accrue the requisite number of oncogenic events for malignant transformation (Nguyen et al., 2012). Though experimental evidence is currently inconclusive, the cancer stem cell model further holds that minimal residual disease and tumor recurrence post-treatment are likely due to therapy-resistant cancer stem cell subclones (Medema, 2013).

GSCs were originally characterized as a cell population that expresses CD133, a marker also expressed by normal neural stem cells, in pediatric medulloblastoma. The cells possessed in vitro stem cell characteristics, such as differentiation capacity and self-renewal potential, which were absent in the CD133− population (Singh et al., 2003). Follow-up in vivo orthotopic xenograft model studies, using both acutely sorted CD133+ cells as well as in vitro cultured GBM lines
enriched for stem cells under conditions used for normal neural stem cell enrichment, demonstrated that GSCs could initiate tumor while non-GSCs could not. Additionally, serial transplantation models resulted in phenocopying of the original tumor, thus demonstrating the in vivo self-renewal capacity of GSCs (Galli et al., 2004; Singh et al., 2004). Studies in various cancers have further validated the cancer stem cell model in other solid malignancies, including colon cancer (Barker et al., 2009), though this hierarchical model may not apply to other cancers like melanoma (Quintana et al., 2008; Valent et al., 2012).

The conceptual appeal for therapeutic targeting of GSCs is based on research that suggests that GSCs may drive both therapy resistance and post-treatment tumor recurrence. It has been demonstrated that the proportion of CD133^+ GSCs is higher in surgical resection specimens after radiotherapy compared to pre-treatment biopsy specimens (Tamura et al., 2010). In a genetically engineered mouse model of glioma incorporating transgene labeling of putative GSCs within the bulk tumors, tumor-regrowth post-treatment with temozolomide and ganciclovir was found to be due to the GSC population that was resistant to therapy (Chen et al., 2012). Various mechanisms have been proposed for the resistance of cancer stem cells to conventional therapies, including high threshold for apoptosis induction, efficient DNA repair mechanisms, quiescence, and high expression of drug efflux transporters (Colak and Medema, 2014). In human glioma xenografts and primary GBM specimens, preferential DNA-damage checkpoint activation and more efficient damage repair led to increased radiation therapy resistance of CD133^+ CSCs relative to the CD133^- tumor cells (Bao et al., 2006). GSCs are known to interact with and depend on the tumor microvascular niche. Interestingly, GSCs were shown to have the capability to differentiate into tumor microvascular endothelial cells (tMVECs) and thus establish their own supporting niche. Critically, the GSC-derived tMVECs were shown to be resistant to both chemotherapy and
radiation via induction of a senescent state in which they were still able to support GSC function (Borovski et al., 2013). Given these diverse mechanisms by which GSCs may preferentially attain therapy resistance, it is clear that approaches that target specific mechanisms may not be sufficient to eliminate GSCs (Colak and Medema, 2014). Instead, development of approaches that target the GSC cell-state itself, leading to either their differentiation or death, may be more promising.

The GSC hypothesis is not without controversy, especially due to the lack of a definitive marker that unambiguously identifies GSC in tumor specimens. In this regard, it has been reported that both CD133+ and CD133− cells from the same GBM tumor are capable of self-renewal in vitro and are able to generate serially transplantable tumors in immunocompromised mice, thus arguing against the restriction of tumor-initiating potential in GBM specifically to CD133+ GSCs that had been described in previous reports (Chen et al., 2010). In a study using freshly isolated GBM specimens, 40% of tumors were found to not contain CD133+ cells. However, cells expressing stage-specific embryonic antigen 1 (SSEA-1) were present in nearly all tumors studied and were highly tumorigenic compared to SSEA-1− cells. Additionally, most tumor cells that were CD133+ were shown to be SSEA-1+, thus arguing for SSEA-1 as a more general and superior marker for tumor initiating cells in GBM (Son et al., 2009). While the highlighted phenotypic marker inconsistencies are concerning, there are some potential explanations for the discrepancies seen in some GSC studies. Depending on the genetic background of the malignancy in question, there is a defined likelihood for dysregulation or silencing of putative stem cell markers such as CD133 (Medema, 2013). In GBM and colorectal cancer, cancer cell-specific heterogeneous CpG island DNA methylation leading to silencing of CD133 expression has been described in cultured cell lines as well as primary tumor specimens (Yi et al., 2008). A second explanation for variability
in GSC studies may be due to whether the hierarchical or stochastic model of tumor cell heterogeneity applies to a given malignancy (Nguyen et al., 2012). In the hierarchical model, a biologically distinct subpopulation of cells within the heterogeneous tumor bulk drives tumor growth, is self-propagating, and must be eliminated to cure disease definitively. The stochastic model, in contrast, posits that all cells in a tumor are equi-potent in their capacity to function as GSCs (Nguyen et al., 2012). This model suggests that temporally dynamic cell-intrinsic factors, such as oncogene activity, can interconvert cells from GSC to non-GSC behavior and may thus explain the lack of a hierarchical GSC model, as is the case in melanoma (Quintana et al., 2008). Critically, the stochastic model precludes the identification of GSCs by phenotypic or transcriptional profiles but rather by their functional behavior, such as in vitro and in vivo self-renewal as well as tumorigenicity.

Cumulative data from scRNA-seq analysis has begun to shed more light on the complexity of defining a single population as GSCs in a given tumor, as highlighted by the inconsistencies in using cell surface markers as GSC identifiers. Patel et al. found that there was a stemness-differentiation axis that is continuously occupied based on gene expression analysis in single cell data (Patel et al., 2014). This observation suggests that in vitro GSC models only represent the extremes of this axis and do not capture the full spectrum of stem-like and differentiated cells. Similarly, the Diaz group found that GSCs reside on a single axis of variation from mesenchymal to proneural in their scRNA-seq cohort (Wang et al., 2019a). The postulated that mesenchymal GSCs (mGSCs) are the progenitors of proneural GSCs (pGSCs) and that the most effective therapies were those that targeted alteration in both populations. Bhaduri et al. used scRNA-seq to demonstrate that expression of stemness markers is highly heterogenous in glioblastoma tumors, with 21 defined cell types expressing at least one marker associated with GSCs (Bhaduri et al.,
2020). They posit that stemness programs are heterogenous and that multiple cell types may recruit stemness programs, thus questioning the existence of GSCs as a unique well-defined state.

As discussed previously, the Suva group demonstrated that GBM harbors 4 subtypes that are enriched for specific genetic driver events and demonstrate plasticity in cell sorting and lineage tracing experiments (Neftel et al., 2019). In additional analysis of this data, Suva and Tirosh assessed the expression of commonly used GSC markers and found that they associate with specific states in tumors: CD24 with the NPC-like, CD44 with MES-like, CD133 with OPC-like, and NESTIN with the AC-like state (Suva and Tirosh, 2020). These observations suggest that previous studies using specific subpopulations of GSCs as defined by these marker were in fact studying a subset of GSCs. Importantly the authors outline that 3 of these states (MES-like, OPC-like, and NPC-like) demonstrate functional properties typically associated with GSCs, which are in vivo tumor initiation potential and the capacity to generate the other cellular states when presorted populations are implanted in animal models. These two criteria directly correspond to self-renewal capacity and differentiation capacity and require that all 3 states be termed “GSCs”. The authors argue for approaches that differentiate the 3 tumorigenic states into AC-like cells, which exhibit lower proliferative capacity, as potentially more efficacious strategies. Interestingly, while cell surface markers demonstrated strong enrichment to specific cell types, development-related transcription factors such as ID1, BRN2, and SOX2 were expressed in multiple states and may thus represent a potential avenue to enforce differentiation and abrogate stem cell properties of multiple states.
1.6 SOX2 in Glioblastoma:

The core pluripotency-related transcription factor SRY (sex determining region Y)-box 2 (SOX2) has a well-established oncogenic role in glioblastoma. SOX2 has been reported to be highly expressed in a majority of glioblastoma biopsy specimens, with SOX2 expression levels correlating with increased tumor grade and aggressiveness (Garros-Regulez et al., 2016). Various studies have demonstrated that SOX2 expression and transcriptional activity is indispensable for the self-renewal and tumorigenicity of GSCs (Gangemi et al., 2009; Suva et al., 2014). Remarkably, it has been reported that while oncogenic receptor tyrosine kinase (RTK) signaling induces SOX2 activity during transformation in engineered glioblastoma mouse models, the maintenance of SOX2 expression and tumor-driving transcriptional networks is independent of aberrant RTK activity (Singh et al., 2017). This suggests that once SOX2 expression is induced in glioblastoma cells, SOX2 is able to sustain its own expression independent of additional signaling input. Further, single-cell RNA sequencing analysis of IDH-wild-type glioblastomas revealed that SOX2 expression is maintained, albeit at differing levels, across all four defined cellular states that exhibit substantial plasticity and associate with divergent oncogenic drivers (Neftel et al., 2019; Suva and Tirosh, 2020). These observations of oncogene and cell state independent high level expression of SOX2 highlight its importance to glioblastoma biology.

The role of SOX2 in the induction of the stem cell state was first demonstrated by Takahashi et al. when a combination of SOX2 with the transcription factors Oct 3/4, Klf4, and c-Myc was found to be sufficient to produce induced pluripotent stem cells (iPS cells) from adult human dermal fibroblasts (Takahashi et al., 2007). In more recent work using single cell expression coupled with Bayesian network analysis, it has been demonstrated that endogenous SOX2 is
activated late in the reprogramming process and is responsible for initiating consecutive steps required for pluripotency induction, placing SOX2 at the top of the reprogramming transcription factor hierarchy (Buganim et al., 2012). Specific to GSCs, SOX2 was one of 4 transcription factors that were found to be critical in reprogramming differentiated GBM cells to the GSC phenotype, resulting in increased self-renewal and tumorigenicity (Suva et al., 2014). Indeed, loss of SOX2 alone was sufficient to abrogate GSC tumorigenicity. There has been limited success in the development of direct small molecule and nucleotide-based SOX2 inhibitors that demonstrate promise for clinical application (Garros-Regulez et al., 2016; Taniguchi et al., 2017). Part of the challenge with targeting SOX2 is that, unlike nuclear receptor transcription factors, it lacks prototypical small-molecule binding domains. Therefore, strategies that leverage canonical cellular enzymatic processes to target SOX2 transcriptional activity or protein stability may be attractive avenues for GSC differentiation therapies.

1.7 Ubiquitin-dependent SOX2 regulation:

Proteasomal inhibition has been shown to be a viable approach in targeting cancer cells. Treatment of temozolomide-resistant GSCs with bortezomib, a proteasomal inhibitor, sensitized the cells to this standard chemotherapeutic agent, suggesting that regulation of protein homeostasis is a fundamental process for GSC viability (Bota et al., 2013). Therefore, components of the ubiquitin-proteasome system (UPS) may provide critical targets for disrupting the GSC state. A growing body of evidence suggests that SOX2 protein stability is dynamically regulated by the UPS in both normal neurodevelopment and in cancer. In embryonic stem cells, it has been shown
that the E3 ubiquitin ligase WWP2 regulates SOX2 stability in a Set7/AKT-dependent fashion. AKT-driven phosphorylation of SOX2 antagonized Set7-dependent methylation, resulting in SOX2 stabilization and maintenance of ESCs (Fang et al., 2014). A role for SOX2 proteasomal regulation in neural progenitor cells has also been described, where induction of NPC differentiation resulted in upregulation of the E3 ubiquitin ligase CUL4A. Polyubiquitination and degradation of SOX2 by CUL4A^{DET1-COP1} was required for NPC differentiation (Cui et al., 2018). In the context of cancer, Wang et.al found that AKT activation promoted proliferation and stemness of esophageal cancer cell lines by phosphorylating and thus protecting SOX2 from UBR5-mediated polyubiquitination and degradation (Wang et al., 2019b). Together, these studies demonstrate a critical role of proteasomal regulation of SOX2 protein stability in normal development and in cancer stem cells.
1.8 References:


Chapter 2:
Biological and Therapeutic Implications of Multisector Sequencing in Newly Diagnosed Glioblastomas

This chapter is adapted from a manuscript published in Neuro-Oncology:

Tatenda Mahlokozera, B.S.*; Ananth K. Vellimana, M.D.*; Tiandao Li, Ph.D.*; Diane D. Mao, M.D., Zohny S. Zohny, M.D.; David H. Kim, M.S.; David D. Tran, M.D., Ph.D.; Daniel S. Marcus, Ph.D., Sarah J. Fouke, M.D.; Jian L. Campian, M.D., Ph.D.; Gavin P. Dunn, M.D., Ph.D.; Christopher A. Miller, Ph.D.; Albert H. Kim, M.D., Ph.D.
2.1 Summary:

This study provides both a biological and potential clinical rationale for pursuing multisector molecular profiling in glioblastoma patients. We demonstrate, using image-guidance directed sampling of two to four sectors of gadolinium-enhancing areas of IDH1 wildtype glioblastoma tumors, that whole-exome sequencing of individual sectors reveals a spatially divergent mutational landscape. In two extreme cases of regional heterogeneity, we describe, for the first time, treatment naïve tumors with region-specific hypermutator phenotypes. In remarkable contrast to the spatial diversity of the overall mutational landscape, we demonstrate that TERT promoter mutations are unique in being recurrent in all analyzed tumors and clonal in all tumor sectors. Finally, we examine the potential therapeutic consequences of multisector sequencing data and find that multi-site analyses may be necessary to accurately characterize individual GBM tumors and identify meaningful therapeutic options.
2.2 Introduction

Glioblastoma (GBM) is the most common malignant primary brain tumor in adults. The current standard-of-care treatment for GBM is maximal, safe, surgical resection followed by concomitant chemotherapy and radiation therapy\(^1\). Despite recent advances, the overall prognosis of the majority of GBM patients remains poor with a median survival of 15 months\(^2\) and a 5-year survival of 10%\(^1\). However, clinical outcomes vary considerably among patients. Previous studies have demonstrated marked differences among tumors at the genomic\(^3,4\) and transcriptomic\(^5,6\) levels, which may underlie differences in both the natural history of a patient’s tumor as well as responses to treatment\(^2\). For instance, patients with \textit{IDH1} mutant GBMs have a median survival at least two times longer than that of patients with \textit{IDH1} wildtype tumors\(^7\), whereas isolated \textit{TERT} promoter mutations are associated with significantly lower overall survival.\(^8\) Promoter methylation at the O\(^6\)-methylguanine-DNA-methyltransferase (MGMT) locus predicts relative sensitivity to the drug temozolomide.\(^9\)

More broadly, the genomic characterization of all cancers has led to the identification of tumor-specific alterations that may inform therapeutic options. Due to the urgent need for additional therapies in GBM, precision medicine has been introduced into the clinical space for this disease as well. The goal is to match specific tumor mutations to potentially therapeutic drugs and, more recently, to determine the neoantigen landscape as either an aggregate biomarker for immunotherapy or for vaccine approaches\(^10,11\). The current clinical workflow in most academic hospitals includes the identification of potentially actionable mutations or neoantigens from a single tumor sector. For GBM located at sites that preclude open craniotomy for resection, stereotactic needle biopsy from a single tumor region is performed and analyzed for genomic and other molecular data. However, in addition to differences among individual patients’ tumors, many solid tumors harbor substantial intratumoral genetic heterogeneity,\(^12,13\) raising the important question of whether molecular
characterization of a single sector in GBM sufficiently represents the genomic landscape of a tumor in a biologically and clinically meaningful way.

In this study, we performed whole-exome sequencing of multiple sectors of individual GBM tumors to more broadly characterize their mutational profiles, with implications for our understanding of tumor biology as it relates to clonal architecture and pre-treatment evolutionary dynamics as well as the therapeutically actionable genomic landscape.

2.3 Methods

Tumor samples
Adults with newly diagnosed, treatment naive GBM undergoing craniotomy for tumor resection were included. During tumor resection, tumor samples (n = 9) were collected from regionally distinct areas that demonstrated gadolinium enhancement on MRI, and images of the biopsy locations were captured using the neuronavigation platform (Stealth, Medtronic). In one case, four adjacent sectors (each approximately 1 cm³) of an en-bloc resected glioblastoma tumor (B103) were sampled. This study was approved by the Institutional Review Board at Washington University School of Medicine.

DNA Sequencing
Minimal tumor cellularity was 30% in all samples as determined by a pathologist. Matched tumor and blood DNA samples for each patient were subjected to whole-exome sequencing to identify somatic mutations including single nucleotide variants (SNV), insertions or deletions (INDEL), and copy number alterations (CNA). 80-100X mean target coverage of coding bases in the exome was achieved for all samples except tumor B103, for which 54-90X mean coverage was achieved. The TERT promoter was not captured well, and therefore, a set of custom capture probes from IDT
Technologies were used to target this region for additional sequencing, resulting in 4680X coverage of this locus (Supplementary Table S1).

**Somatic variant detection**

Sequences were aligned to reference build GRCh37-light using bwa mem\textsuperscript{14}. Somatic SNVs and INDELs were detected using an ensemble of six different variant callers, the calls from which were unioned and then subjected to additional filtering and manual review. Copy number alterations were detected using Varscan 2.3.6\textsuperscript{15}. See Supplemental Methods for additional details.

**Subclonal inference and clonal evolution**

Multidimensional subclonal inference was performed using copy-number neutral SNVs and SciClone\textsuperscript{16} version 1.1. (params: minimumDepth=100, maximumClusters=10). Variant clusters identified by SciClone were imported into ClonEvol ([https://github.com/hdng/clonevol](https://github.com/hdng/clonevol)) to infer each tumor’s phylogeny.

**Potentially therapeutic mutation and neoantigen prediction**

DGIdb\textsuperscript{17}, the Drug-Gene Interaction database, was used to identify potentially druggable targets based on the lists of mutations and altered genes implicated. MHC Class I neoantigen predictions were made using pVAC-seq\textsuperscript{18}, which leverages 5 algorithms from the Immune Epitope Database and Analysis resource (IEDB, iedb.org): netMHC, netmhcpan, pickpocket, smm, and smmpmbec. Predictions were retained if the best score had an IC50 < 500 nM and better binding of the mutant peptide than the wild type (fold-change > 1).

**2.4 Compliance with ethical standards**
Ethical Approval

All procedures performed were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent

Informed consent was obtained from all individual participants included in the study.

Conflict of Interest

GPD is a co-founder of Immunovalent Therapeutics. The other authors declare no conflicts of interest.

2.5 Results

Sample Characteristics

The cohort consisted of 10 patients with newly diagnosed, treatment naïve *IDH1* wildtype glioblastoma. Patient characteristics are shown in Table 1. The overall number of somatic mutations for each patient (aggregated from all samples) ranged from 112 to 1239, with a mean of 347 (Fig 1A). The number of non-synonymous mutations ranged from 47 to 556, with a mean of 132. Samples B58 and W016 harbored overall higher mutational loads than the remaining eight tumors and were classified as hypermutators (see below). Excluding these samples reduces the mean number of total mutations to 191 and that of non-synonymous mutations per tumor to 73. We observed recurrent somatic mutations in canonical GBM-associated genes, including the *TERT* promoter (10/10 cases, 100%), *EGFR* (4/10, 40%), *PIK3CA* (3/10, 33%), and *TP53* (3/10, 33%) (Fig 1B). Common copy number alterations included chromosome 7 amplifications (7/10 cases, 70%) and chromosome 10 deletions (10/10, 100%). (Supplementary Fig S1A). Paired samples from different sectors had mostly
concordant CN alterations (Supplementary Fig S1B, Supplementary Table S2), consistent with previous work suggesting that copy number (CN) events occur early in tumorigenesis19.

**Mutational heterogeneity between tumor regions**

We then compared somatic SNVs and insertions or deletions (INDEL) in different tumor sectors for each patient to understand the extent of regional intratumoral heterogeneity (Fig 2A). Mutations were classified as “private” if observed in only one sector of a tumor, and “shared” if present in two or more spatially distinct regions. Strikingly, 46% of mutations were private, while 54% were shared across sectors, indicating significant spatial heterogeneity (Fig 2B/D, Supplementary Table S3).

Using VAFs corrected for tumor purity, events were further classified as either clonal (present in all cancer cells) or subclonal (present in only a subset of cells). 51% of mutations were clonal and shared between regions, 3% were subclonal and shared, and 46% were subclonal and private to one tumor sector (Fig 2B). As both sectors appeared to share a common origin in all samples, events cannot be both clonal and private. The two hypermutator tumors influence these data due to a significantly higher load of subclonal private mutations. Excluding the hypermutator tumors resulted in 73% of mutations being classified as clonal and shared, 4% as subclonal and shared, and 23% of mutations as subclonal and private (Fig 2C). In addition, we observed that even mutations present in both sectors may differ markedly in frequency, as observed in uncorrected VAF plots (Fig 2E) for two representative cases, B72 and W059.

To better understand the impact of these differences between sectors, we focused on a subset of genes known to be recurrently mutated in cancer20-22 (Supplementary Table S4). In 8 of 10 cases, one or more mutations in cancer-related genes were private and not readily detectable in the other tumor sector(s) (Fig 2F). These private mutations include 3 of the 4 EGFR mutations (M567I in
B103, E320Q in B42, and T363I in B72) as well as both EP300 mutations (A1853T in B67; G322R in W016). This indicates that these mutations were acquired later in a tumor’s evolution. In contrast, all three samples with TP53 mutations harbored the mutation in all sectors (R248Q in B58, V173L in W016, and G245S in W059), suggesting that they are either initiating mutations (for B58 and W059, where they are high VAF and clonal) or acquired early after transformation (for W016 where the mutation has lower VAF and is subclonal). The only recurrent mutations observed in all sectors of all tumors were TERT promoter mutations (both G228A and G250A, in mutually exclusive fashion). In all cases, these TERT mutations were clonal, suggesting that they are likewise initiating events in GBM (Fig 2G).

**Treatment naïve glioblastoma with hypermutator phenotype**

Two patient tumors, W016 and B58, harbored extremely high mutational loads—1,239 and 699 mutations, respectively. We classified them as hypermutated based on a statistical outlier test for overall mutation load across all samples and on the fact that they harbored mutations greater than 3 standard deviations above the trimmed mean (>409 SNVs) of the 8 other tumors, a definition consistent with the threshold of 400 SNVs used recently. In both cases, the hypermutator phenotype was restricted to a single sector, with the non-hypermutated sectors bearing mutation loads not significantly different from the mean number of mutations among the remaining 8 cases. In both cases, there were no germline hotspot missense mutations in DNA-damage associated genes that could potentially explain the increased mutation load (Supplementary Table S5). In Patient W016, over 90% of mutations were specific to one of the two sectors (S1) (Fig 3A, B) and these consisted almost entirely of C to T transitions. (Fig 3E). The mutation spectrum of this sector did not match the oxo-8G profile (CCG > CAG) for all 1204 mutations (Supplementary Table S6), thus ruling out this artifact as a reason for increased mutation burden. We performed an exploratory analysis to
determine if mutations in known DNA damage repair genes might represent a major pathogenic event in W016 S1. W016 S1 harbored mutations in critical DNA repair-associated genes \textit{ALKBH3}, \textit{ATR}, \textit{POLB}, \textit{MLH3}, and \textit{SPO11} (Supplementary Table S7), but due to the multiple plausible drivers at sufficient VAF, the exact initiating event in the hypermutated sector could not be determined. In tumor B58, sector S1 contained approximately 70\% of the aggregate mutations from the two sequenced sectors (Fig 3C, D). Unlike other tumors, a significant proportion of the mutations in both sectors of tumor B58 exhibited COSMIC signature 3, which is associated with germline BRCA mutations\textsuperscript{25}. Importantly, both sectors harbored a common somatic BRCA missense mutation, with a higher VAF in S1 relative to S2, raising the possibility that BRCA mutation drove hypermutation in S1. These two cases highlight both the occurrence and potential for regional heterogeneity of the hypermutator phenotype in treatment naïve GBM.

\textbf{Heterogeneity in clonal architecture between tumor regions}

Given our observation of mutational heterogeneity between sectors, we next determined the clonal architecture of each tumor and reconstructed its evolution. Using the SciClone\textsuperscript{16} algorithm, the VAFs of SNVs in copy-number neutral, loss of heterozygosity (LOH)-free genomic regions were clustered in both one dimension (per sector) and two dimensions (per tumor, incorporating information from all sectors). In one representative sample, tumor B65 (Fig 4), a subset of subclones was detectable in only one of the two regions. Furthermore, subclones 2 and 3 represent “cryptic” subclones that would not be distinguishable from the founding clone without additional information from spatially distinct samples. Eight of 10 (80\%) cases contained subclonal populations that were private to one of the tumor sectors, consistent with the genetic heterogeneity observed between tumor regions (Supplementary Fig S3).
To understand how multisector sequencing can aid in determining tumor evolution, we imported mutation clusters and VAF information into ClonEvol, which reconstructs possible phylogenetic trees. For tumor B65, using individual sectors in isolation resulted in a linear clonal evolution model (Fig 4A, B), but incorporating both sectors revealed a branching structure (Fig 4C), linking subclone phylogenetic trees in the multiple sectors. In 80% of cases, multisector information added complexity to the inferred phylogenies (Supplementary Fig S4).

**Heterogeneity in putative targetable somatic variants between sectors**

We next explored the potential therapeutic implications of mutational heterogeneity in these multisector tumor samples. Using the Drug-Gene Interaction Database (DGIdb)\(^\text{17}\), we identified potentially druggable mutations in at least one of the distinct tumor sectors for all patients (Supplementary Tables S8 and S9). Eight of 10 (80%) cases had potentially targetable mutations that were not shared between sectors (Fig 5A). Even when detected in both tumor sectors, potentially druggable mutations occurred at different VAFs between sectors in a significant number (75%) of cases. This finding is exemplified by the divergent VAFs of potential druggable gene variants between two sectors (S1 and S2) for tumors B58, B65, and B42 (Fig 5B, Fig S5). In 4 of the 10 cases (40%), private, druggable mutations were more numerous than those shared between regions, suggesting GBM tumors can harbor heterogeneous, spatially restricted subclones bearing actionable mutations.

To further crystallize the clinical implications of single versus multi-site sequencing, we focused on somatic mutations in the glioblastoma driver genes, *EGFR* and *PTEN*, as demonstrative examples. Of the 4 tumors with somatic mutations in *EGFR*, only one variant, A289V, in tumor W016 was clonal and detected in both sectors. All other *EGFR* SNVs were subclonal and private to one tumor sector. Although the private *EGFR* mutation in case B42 (E320Q) has not been previously
described, the other private EGFR SNVs (T363I in B72 and M567I in B103) are listed in the COSMIC database and are predicted to be potentially pathogenic (score ≥ 0.89 for all three variants) based on the FATHMM in silico algorithmic approach. Somatic mutations in the tumor suppressor, PTEN, were detected in two patients in our cohort. For patient B42, the potentially pathogenic PTEN L112V mutation was clonal and shared between tumor sectors. In contrast, the PTEN truncation mutation E7* in tumor B65 was subclonal in one region but was absent in the second.

**Heterogeneity in the neoantigen landscape**

To determine the immunotherapeutic implications of the observed spatial mutational heterogeneity, we applied a cancer immunogenomics approach to predict the candidate neoantigen landscape in each tumor sector. In this approach, multiple computational algorithms are used to determine the affinity with which translated mutant peptides determined from identified missense variants bind to a patient’s HLA molecules. In each tumor examined, both shared and private neoantigens were identified (Fig 5C, Supplementary Table S10). This observation is exemplified by the neoantigen landscapes of two sectors for tumors B58 and B65 (Fig 5D). Within both tumors, although a subset of mostly high VAF neoantigens was shared between two spatially distinct areas, lower VAF neoantigens were identified in one of the two regions. Moreover, W016, the tumor with the highest mutational burden, harbored a commensurately high number of predicted subclonal rather than clonal neoantigens, largely specific to the hypermutated sector S1 (Figure 5C). Together, these data show that spatial mutational heterogeneity correlates with heterogeneity of the candidate neoantigen landscape and that a hypermutator state leads to an elevated number of subclonal neoepitopes.
2.6 Discussion

In this study, we examined regional genetic heterogeneity among multiple sectors of primary GBM tumors by whole-exome sequencing and targeted sequencing of the TERT promoter locus. Sottoriva et al. provided early evidence of intratumoral heterogeneity in glioblastoma using genome-wide CNA analyses, although we did not detect such CNAs using exome-derived analyses. Verhaak and colleagues performed multisector whole-exome sequencing of GBM tumors and focused on evolutionary patterns that lead to disease recurrence. Another study examined transcriptional heterogeneity and utilized targeted sequencing to demonstrate regional heterogeneity of mutations in DNA repair genes. From a precision oncology standpoint, Lee, et al. recently found that patient-derived glioblastoma cells from multisector samples were more sensitive to drugs that target truncal rather than private genetic alterations. Here we focused on the mutational burden in primary GBM tumors by whole-exome sequencing of multiple regions to examine the impact of regional heterogeneity on our understanding of pre-treatment tumor biology as well as the potential clinical implications for targeted- and immune-based therapies.

Within our cohort, 2 of 10 patients exhibited a putative hypermutator phenotype. A majority of the high mutational load in both hypermutator tumors was specific to one region. To our knowledge, the concomitant existence of non-hypermutated and hypermutated clones within distinct anatomic regions of glioblastoma has not been previously described. Also, while prior studies have reported a 20-30% incidence of a hypermutated state in recurrent GBM after treatment with temozolomide, its occurrence in primary treatment naïve adult and pediatric glioblastoma has only been reported in the setting of POLE deficiency and biallelic mismatch repair deficiency, respectively. Although POLE mutations were not observed in our patients, it is possible that mutations in base excision repair or mismatch repair genes, which were recently described in treatment naïve glioblastoma, could contribute to the development of a hypermutator phenotype.
Given the spatial heterogeneity in mutational burden, a single sector biopsy could potentially miss the hypermutated tumor region, and therefore incorrectly classify the mutational burden of the tumor. It is also possible that the high incidence of hypermutation in recurrent, post-treatment GBM\textsuperscript{31-33}, may actually represent an enrichment of pre-treatment hypermutated subclones that are potentially not observed due to undersampling. Because hypermutated clones are likely insensitive to temozolomide\textsuperscript{37}, the occurrence of these populations within the tumor may also correlate with resistance to alkylating therapy, especially in the setting of subtotal resection. Further multisector sampling studies combined with outcome data will be needed to explore this possibility further.

From an immunotherapy standpoint, elevated mutational burden—the engine for neoantigen production—correlates with response to checkpoint blockade immunotherapy in solid tumors\textsuperscript{38-40}. Although it is not the only parameter that influences responses to this drug class, it appears to be an important one that extends across cancer types. Indeed, several groups have reported compelling case studies of newly-diagnosed patients with hypermutated glioblastomas that exhibited responses to checkpoint blockade immunotherapy \textsuperscript{35,36}. Especially in the MGMT promoter unmethylated setting, in which temozolomide is not as efficacious, the finding of hypermutated subsets distributed in newly diagnosed tumors may suggest that immunotherapy could be a consideration. Finally, each tumor harbored both private and shared neoantigens in each region, suggesting that consideration of spatial diversity in neoantigen-targeting approaches, such as polyvalent vaccines, may be important.

Across our study cohort, we observed regional heterogeneity in both the occurrence and relative frequencies of several cancer driver genes, such as \textit{EGFR}, \textit{TP53}, and \textit{PIK3CA}. Additionally, clonality analysis revealed that in 90% of tumors sequenced, subclonal mutations were more likely to be private than shared between sectors, with the hypermutator cases representing extremes of regional restriction of subclonal mutations. These findings, along with convergent results on the substantial regional heterogeneity of other solid tumors, highlight the potential limitations of current
single-site diagnostic biopsy in capturing the full spectrum of mutations in an individual tumor, which may be critical to a more comprehensive understanding of a tumor’s biology and response to therapy\textsuperscript{12,13,41}. In accordance with the observed regional mutational heterogeneity, subclonal inference using Bayesian clustering methods\textsuperscript{16} demonstrated the existence of spatially-restricted subclones in 8 out of the 10 (80\%) tumors in this cohort. We recognize that percentage tumor cellularity, which differed by no more than 25\% between paired patient samples in our study, can impact the overall number of subclones detected using our model assumptions. However, we would emphasize that the evidence for pervasive subclonal heterogeneity between regions under specific invariant model assumptions is the major biologically meaningful finding in our analysis. We found that integration of information from spatially distinct tumor samples in these cases allowed us to distinguish cryptic subclones from the founding clone, which was not achievable using only one tumor sample. Our analysis may however underestimate the extent of tumor evolution and heterogeneity since we did not assess extra chromosomal DNA, which has been shown to harbor oncogenes and allow for more rapid tumor evolution than can be achieved by chromosomal amplification\textsuperscript{42}.

We also observed significant heterogeneity in potentially druggable mutations between tumor sectors. 80\% of the tumors sequenced had private targetable mutations, and these were more frequent than shared targetable mutations in 40\% of the patients. Recently, Morrissy, et al. found a predominance of mutations that were clonal in one biopsy specimen but were subclonal or undetected in additional biopsies in medulloblastomas and a limited number of high-grade gliomas. They estimated that to attain an 80\% probability of detecting at least 80\% of all potential actionable mutations, at least five biopsies would be required\textsuperscript{43}. This spatially distinct expression of druggable genes could, in part, explain the failure of prior clinical trials that employed targeted mutation approaches to treat GBM. Multisector sampling may therefore be required to select the appropriate
‘personalized’ drug regimen that would engage multiple, spatially separated, targetable mutations for broader treatment coverage of the existing subclonal landscape.

Interestingly, all patients in our cohort harbored one of the two TERT promoter mutations (G228A or G250A). This finding is in alignment with recent studies, which have reported a high incidence (54-84%) of TERT promoter mutations in GBM patients. Importantly, the TERT promoter mutation was present in the founding clone in all tumor sectors examined. This is the first report of TERT promoter mutations being clonal in IDH1 wildtype GBM, consistent with recent studies which have suggested that TERT promoter mutation is an early event in GBM pathogenesis. Intriguingly, this also suggests that TERT might represent a strategic therapeutic target in GBM patients.

Although single-site biopsies may be sufficient for the detection of known clonal mutations, including TERT (as described herein) or IDH1, in glioblastoma, our study suggests that inclusion of multi-site biopsy specimens may be critical in the design of future clinical trials that test the efficacies of small molecule or immune-based therapies due to the striking spatial divergence of most mutations. From a practical standpoint, it will be important to determine how many sites need to be sampled per tumor to capture the majority of biologically relevant, spatially restricted subclones.

2.7 Acknowledgments:
We would like to thank Dr. Timothy Ley for thoughtful discussion.

Funding Statement:
This work was supported by the Barnes-Jewish Hospital Foundation and Washington University Institute of Clinical and Translational Sciences (A.H.K.), The Christopher Davidson and Knight Family Fund (A.H.K., G.P.D.), and the NIH R01NS066905 (D.S.M., S.J.F.).
2.8 References:


### Table 2.1: Clinical characteristics of the study cohort

<table>
<thead>
<tr>
<th>Patient</th>
<th>Age at diagnosis</th>
<th>Sex</th>
<th>MGMT</th>
<th>IDH1</th>
<th>Regions sampled (n)</th>
<th>Max. tumor diameter (cm)</th>
<th>Tumor volume (cm³)</th>
<th>Cellularity (S1/S2/S3/S4)</th>
<th>Necrosis (S1/S2/S3/S4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B42</td>
<td>70.1</td>
<td>F</td>
<td>1</td>
<td>WT</td>
<td>2</td>
<td>5.3</td>
<td>13.5</td>
<td>75/75</td>
<td>10/10</td>
</tr>
<tr>
<td>B58</td>
<td>66.4</td>
<td>M</td>
<td>0</td>
<td>WT</td>
<td>2</td>
<td>5.5</td>
<td>29.3</td>
<td>83/70</td>
<td>20/5</td>
</tr>
<tr>
<td>B65</td>
<td>54</td>
<td>F</td>
<td>1</td>
<td>WT</td>
<td>2</td>
<td>5.6</td>
<td>26.5</td>
<td>60/80</td>
<td>5/0</td>
</tr>
<tr>
<td>B67</td>
<td>80.4</td>
<td>F</td>
<td>0</td>
<td>WT</td>
<td>2</td>
<td>7.2</td>
<td>48.0</td>
<td>60/60</td>
<td>50/50</td>
</tr>
<tr>
<td>B71</td>
<td>60.6</td>
<td>M</td>
<td>0</td>
<td>WT</td>
<td>2</td>
<td>4</td>
<td>13.2</td>
<td>65/70</td>
<td>50/70</td>
</tr>
<tr>
<td>B72</td>
<td>70.6</td>
<td>M</td>
<td>1</td>
<td>WT</td>
<td>2</td>
<td>5.4</td>
<td>21.1</td>
<td>80/65</td>
<td>5/15</td>
</tr>
<tr>
<td>B74</td>
<td>64.6</td>
<td>F</td>
<td>1</td>
<td>WT</td>
<td>2</td>
<td>5.2</td>
<td>23.5</td>
<td>70/70</td>
<td>10/70</td>
</tr>
<tr>
<td>W016</td>
<td>56</td>
<td>F</td>
<td></td>
<td>WT</td>
<td>2</td>
<td>5.2</td>
<td>14.1</td>
<td>60/40</td>
<td>10/0</td>
</tr>
<tr>
<td>W009</td>
<td>71</td>
<td>F</td>
<td>1</td>
<td>WT</td>
<td>2</td>
<td>5</td>
<td>14.9</td>
<td>90/85</td>
<td>5/20</td>
</tr>
<tr>
<td>B103</td>
<td>74.3</td>
<td>M</td>
<td>1</td>
<td>WT</td>
<td>4</td>
<td>5.5</td>
<td>27.5</td>
<td>70/45/70/70</td>
<td>50/65/10/5</td>
</tr>
</tbody>
</table>

- **MGMT**: 0 = methylated, 1 = unmethylated
- **IDH1**: by R132H IHC and exome sequencing
- **Age**: years
- **Years**
Figure 2.1. Somatic mutation landscape of treatment naïve primary glioblastoma. (A) Overall mutation loads for 10 patients in the study cohort. (B) Waterfall plot highlighting somatic mutations in a subset of genes that are recurrently mutated in human cancers. The mutation frequencies for each gene are depicted in the chart on the left.
Figure 2.2. Mutational heterogeneity between tumor sectors. (A) Representative T1-weighted MRIs for patients B42, W016, W059, and B71 illustrating the coordinates of two distinct tumor biopsy regions (S1 and S2) in the axial and coronal planes. (B,C) Charts depicting the proportions of clonal-shared, subclonal-shared, and subconal-private mutations for the entire study cohort (B) and eight non-hypermutant cases (C). (D) Chart showing the overall numbers of clonal-shared, subclonal-shared, and subconal-private mutations for each patient. (E) Representative VAF versus somatic mutation plots illustrating mutational heterogeneity between two distinct tumor sectors (S1 and S2) for patients B72 (left) and W059 (right). (F) Waterfall plot highlighting regional heterogeneity of somatic mutations in a subset of recurrently mutated cancer-associated genes. Mutation frequencies for each gene for all tumor regions sequenced are depicted on the left. (G) Plot of the VAFs of clonal mutations in each tumor sector normalized to the median VAF of all clonal mutations in that specific sector (black dots). The VAFs of TERT promoter mutations (G228A and G250A) are indicated in red. For patient W059, the TERT promoter region was amplified, thus VAFs after copy number correction are shown.
Figure 2.3. Treatment naïve hypermutated glioblastomas. (A, C) Charts depicting the total numbers of overall and non-synonymous mutations in the hypermutant tumors, W106 (A) and B58 (C), in each sector sequenced-S1 (top) and S2 (bottom). (B, D) VAF versus somatic mutation plots for two distinct tumor sectors, S1 (top) and S2 (bottom), for W016 (A) and B58 (B). (E) Mutational spectrum for each tumor sector for each patient in the cohort.
Figure 2.4. Regional heterogeneity in subclonal architecture. (A,B) SciClone-generated tumor coverage versus VAF plots for tumor B65 sectors 1 (A) and 2 (B) (left). Reconstructed probable phylogenies between subclones in each respective tumor sector are shown in the right panel. (C) SciClone-generated 2-dimenisonal plot of Sector 2 versus Sector 1 VAFs for patient B65 (left) and a corresponding example of the ClonEvol-generated probable branching phylogenetic trees based on VAF information from both tumor sectors (right).
Figure 2.5. Regional heterogeneity in potentially therapeutic gene variants and in the predicted neoantigen landscape. (A) Chart showing the numbers of private, clonal-shared, and subclonal-shared mutations in potentially therapeutic genes for each patient. (B) Mutation versus VAF plots highlighting heterogeneity in potentially therapeutic gene variants for patient B58 (left) and B65 (right) for two distinct tumor sectors. (C) Chart depicting the numbers of private, clonal-shared, and subclonal-shared predicted neoantigens for each tumor. (D) Mutation versus VAF plot of predicted neoantigens for patient B58 (left) and B65 (right) for two tumor sectors. Mutations along the vertical axis are grouped into 3 classes: shared, sector 1 (S1) specific, and sector 2 (S2) specific.
2.9 Supplementary methods

Somatic Variant Detection

Sequence data was aligned to reference sequence build GRCh37-lite-build37 using bwa mem[5] version 0.7.10 (params: -t 8::), then merged and deduplicated using picard version 1.113, (https://broadinstitute.github.io/picard/).

SNVs were detected using the union of four callers: 1) samtools[6] version r982 (params: mpileup -BuDs) intersected with Somatic Sniper[4] version 1.0.4 (params: -F vcf -G -L -q 1 -Q 15) and processed through false-positive filter v1 (params: --bam-readcount-version 0.4 -bam-readcount-min-base-quality 15 --min- mapping-quality 40 --min-somatic-score 40), 2) VarScan[3] version 2.3.6 filtered by varscan-high-confidence filter version v1 and processed through false-positive filter v1 (params: --bam-readcount-version 0.4 --bam-readcount-min-base-quality 15), 3) Strelka[10] version 1.0.11 (params: isSkipDepthFilters = 1), and 4) mutect[2] v1.1.4 (params: number-of-chunks=50).

Indels were detected using the union of 4 callers: 1) GATK[7] somatic-indel version 5336 2) pindel[11] version 0.5 filtered with pindel somatic calls and VAF filters (params: --variant-freq-cutoff=0.08), and pindel read support, 3) VarScan[3] version 2.3.6 filtered by varscan-high-confidence-indel version v1 and 4) Strelka[10] version 1.0.11 (params: isSkipDepthFilters = 1).

SNVs and Indels were further filtered by removing artifacts found in a panel of 905 normal exomes, removing sites that exceeded 0.1% frequency in the 1000 genomes or NHLBI exome
sequencing projects, and then using a Bayesian classifier (https://github.com/genome/genome/blob/master/lib/perl/Genome/Model/Tools/Validation/IdentifyOutliers.pm) and retaining variants classified as somatic with a binomial log-likelihood of at least 10.

Copy number analysis was performed using Varscan 2.3.6 [24], and segmented with the DNAcopy package[8]. Data was recentered when necessary, then segments of less than 50 probes were filtered to remove noise, followed by merging of adjacent segments with absolute CN difference of less than 0.2.

**Mutation signature**

The mutational spectra of mutations of all samples were analyzed using deconstructSigs[9] to extract signatures based on the Wellcome Trust Sanger Institute Mutational Signature Framework[1] and statistically quantify the contribution of each signature for each tumor.
2.10 Supplementary Methods References:


**Figure 2.S1A.** Recurrent copy number aberrations in treatment naïve glioblastoma. Common chromosomal gains and losses observed in the glioblastoma cohort are highlighted (n =10).
Figure 2.S1B. Comprehensive representation of all copy number events in all tumor regions sequenced for the entire study cohort (n = 10). Copy number gains and losses are highlighted in red and blue respectively.
Figure 2.S2. Heterogeneity in mutational signatures between tumor sectors. This chart demonstrates the relative proportions of a total of 17 mutational signatures associated with each tumor sector across the 10 glioblastomas sequenced.
Figure 2.S3. Subclonal architecture inference for all tumors. The top and middle panels show SciClone-generated tumor coverage versus VAF plots based on SNVs in copy-number neutral, loss of heterozygosity (LOH)-free genomic regions for two distinct tumor sectors. The bottom panels depict the SciClone-generated 2-dimensional plot integrating VAF information from two sectors for improved resolution of subclones.
Figure 2.S3. Subclonal architecture inference for all tumors. The top and middle panels show SciClone-generated tumor coverage versus VAF plots based on SNVs in copy-number neutral, loss of heterozygosity (LOH)-free genomic regions for two distinct tumor sectors. The bottom panels depict the SciClone-generated 2-dimensional plot integrating VAF information from two sectors for improved resolution of subclones.
Figure 2.S3. Subclonal architecture inference for all tumors. The top and middle panels show SciClone-generated tumor coverage versus VAF plots based on SNVs in copy-number neutral, loss of heterozygosity (LOH)-free genomic regions for two distinct tumor sectors. The bottom panels depict the SciClone-generated 2-dimensional plot integrating VAF information from two sectors for improved resolution of subclones.
Figure 2.S4. Clonal evolution inference for sample B103. The left panels show box plots for the VAFs of mutations clusters detected using variants from copy number neutral regions in two distinct sectors (top and bottom). The middle panels show the ClonEvol-generated probable evolutionary relationships between identified subclones in each sector (top and bottom). The right hand panels illustrate the integrated phylogenetic relationships between all subclones in both tumor sectors. All possible predictions generated are shown.
Figure 2.S4. Clonal evolution inference for sample B42. The left panels show box plots for the VAFs of mutations clusters detected using variants from copy number neutral regions in two distinct sectors (top and bottom). The middle panels show the ClonEvol-generated probable evolutionary relationships between identified subclones in each sector (top and bottom). The right hand panels illustrate the integrated phylogenetic relationships between all subclones in both tumor sectors. All possible predictions generated are shown.
Figure 2.S4. Clonal evolution inference for sample B58. The left panels show box plots for the VAFs of mutations clusters detected using variants from copy number neutral regions in two distinct sectors (top and bottom). The middle panels show the ClonEvol-generated probable evolutionary relationships between identified subclones in each sector (top and bottom). The right hand panels illustrate the integrated phylogenetic relationships between all subclones in both tumor sectors. All possible predictions generated are shown.
Figure 2.S4. Clonal evolution inference for sample B65. The left panels show box plots for the VAFs of mutations clusters detected using variants from copy number neutral regions in two distinct sectors (top and bottom). The middle panels show the ClonEvol-generated probable evolutionary relationships between identified subclones in each sector (top and bottom). The right hand panels illustrate the integrated phylogenetic relationships between all subclones in both tumor sectors. All possible predictions generated are shown.
Figure 2.S4. Clonal evolution inference for sample B67. The left panels show box plots for the VAFs of mutations clusters detected using variants from copy number neutral regions in two distinct sectors (top and bottom). The middle panels show the ClonEvol-generated probable evolutionary relationships between identified subclones in each sector (top and bottom). The right hand panels illustrate the integrated phylogenetic relationships between all subclones in both tumor sectors. All possible predictions generated are shown.
Figure 2.S4. Clonal evolution inference for sample B71. The left panels show box plots for the VAFs of mutations clusters detected using variants from copy number neutral regions in two distinct sectors (top and bottom). The middle panels show the ClonEvol-generated probable evolutionary relationships between identified subclones in each sector (top and bottom). The right hand panels illustrate the integrated phylogenetic relationships between all subclones in both tumor sectors. All possible predictions generated are shown.
Figure 2.S4. Clonal evolution inference for sample B72. The left panels show box plots for the VAFs of mutations clusters detected using variants from copy number neutral regions in two distinct sectors (top and bottom). The middle panels show the ClonEvol-generated probable evolutionary relationships between identified subclones in each sector (top and bottom). The right hand panels illustrate the integrated phylogenetic relationships between all subclones in both tumor sectors. All possible predictions generated are shown.
Figure 2.S4. Clonal evolution inference for sample B74. The left panels show box plots for the VAFs of mutations clusters detected using variants from copy number neutral regions in two distinct sectors (top and bottom). The middle panels show the ClonEvol-generated probable evolutionary relationships between identified subclones in each sector (top and bottom). The right hand panels illustrate the integrated phylogenetic relationships between all subclones in both tumor sectors. All possible predictions generated are shown.
Figure 2.S4. Clonal evolution inference for sample W016. The left panels show box plots for the VAFs of mutations clusters detected using variants from copy number neutral regions in two distinct sectors (top and bottom). The middle panels show the ClonEvol-generated probable evolutionary relationships between identified subclones in each sector (top and bottom). The right hand panels illustrate the integrated phylogenetic relationships between all subclones in both tumor sectors. All possible predictions generated are shown.
Figure 2.S4. Clonal evolution inference for sample W059. The left panels show box plots for the VAFs of mutations clusters detected using variants from copy number neutral regions in two distinct sectors (top and bottom). The middle panels show the ClonEvol-generated probable evolutionary relationships between identified subclones in each sector (top and bottom). The right hand panels illustrate the integrated phylogenetic relationships between all subclones in both tumor sectors. All possible predictions generated are shown.
Figure S5. Divergence in VAFs of potentially druggable gene variants between tumor sectors. Representative therapeutic gene variant versus VAF plot highlighting divergent VAFs for the depicted gene variants between sector 1 (S1) and sector 2 (S2) for patient B42.
Chapter 3:

A CDC20-APC/SOX2 Signaling Axis Regulates Human Glioblastoma Stem-Like Cells

This chapter is adapted from a manuscript published in Cell Reports:

3.1 Summary:

Glioblastoma harbors a dynamic subpopulation of glioblastoma stem-like cells (GSCs) that can propagate tumors \textit{in vivo} and is resistant to standard chemoradiation. Identification of the cell-intrinsic mechanisms governing this clinically important cell state may lead to the discovery of novel therapeutic strategies for this challenging malignancy. Here, we demonstrate that the mitotic E3 ubiquitin ligase CDC20-Anaphase-Promoting Complex (CDC20-APC) drives invasiveness and self-renewal in patient tumor-derived GSCs. Moreover, \textit{CDC20} knockdown inhibited and CDC20 overexpression increased the ability of human GSCs to generate brain tumors in an orthotopic xenograft model \textit{in vivo}. CDC20-APC control of GSC invasion and self-renewal operates through pluripotency-related transcription factor \textit{SOX2}. Our results identify a CDC20-APC/SOX2 signaling axis that controls key biological properties of GSCs, with implications for CDC20-APC-targeted strategies in the treatment of glioblastoma.
3.2 Introduction:

Glioblastoma, the most common malignant primary brain tumor in adults remains a challenging disease with a poor prognosis (Wen and Kesari, 2008). Increasing appreciation of the cancer cell heterogeneity within glioblastomas has focused attention on a subpopulation of cells called tumor-initiating cells or glioblastoma stem-like cells (GSCs) (Singh et al., 2004). GSCs contribute to overall tumor growth as well as tumor recurrence following chemoradiation and exhibit elevated invasive potential compared to their non-stem cell counterparts (Bao et al., 2006; Chen et al., 2012; Cheng et al., 2011). GSCs also retain the genetic features of parental tumors, suggesting they are a faithful model system for human glioblastoma (Lee et al., 2006; Pollard et al., 2009).

The Anaphase-Promoting Complex (APC) E3 ubiquitin ligase functions with co-activator CDC20 to drive mitosis (Peters, 2006). CDC20-APC has been viewed as a potential strategic target in several human cancers (Wang et al., 2015). CDC20 mRNA is elevated in glioblastoma compared to low-grade gliomas, and CDC20 immunoreactivity in gliomas correlates with pathological grade, but little is known about the biological roles of CDC20-APC in glioblastoma (Bie et al., 2011; Marucci et al., 2008). Recent studies have revealed unexpected non-mitotic roles for CDC20-APC in the developing mammalian brain, indicating CDC20-APC executes functions beyond the cell cycle (Kim et al., 2009; Puram et al., 2011; Yang et al., 2009). These observations have important ramifications not only for brain development but also raise the possibility that CDC20-APC may function in the aberrant developmental state of GSCs.

Here we report CDC20-APC is required for GSC invasiveness and self-renewal in a manner distinct from its role in cell cycle control. We identify pluripotency-related transcription factor SOX2 as a CDC20-interacting protein and show CDC20-APC operates through SOX2 to regulate human
GSC invasion and self-renewal. Finally, we demonstrate CDC20-APC is essential for GSC tumorigenicity in orthotopic xenografts and that CDC20 expression has prognostic value in a subset of glioblastoma patients. These results highlight a critical role for CDC20-APC in the maintenance of human GSC function and suggest that targeting this pathway in glioblastoma may disrupt the GSC state.

3.3 Results:

We have generated low-passage patient-derived glioblastoma stem-like cell lines (GSCs) (Table S1), which express neural stem cell markers (Figure 1A, S1A-C), exhibit self-renewal in vitro (Figure S1D), and form infiltrative brain tumors in immunocompromised mice (Figure 1B, S1E) (Pollard et al., 2009). We examined CDC20 expression by immunoblotting in multiple GSC lines and found increased protein levels in GSCs compared to primary human astrocytes (Figure 1C). To test the role of CDC20 in GSCs, we used RNA interference (RNAi) lentiviruses to target human CDC20 (CDC20i.1 and CDC20i.2), which resulted in efficient CDC20 knockdown (Figure 1D). We focused first on invasiveness, a defining clinical feature of gliomas. GSCs transduced with CDC20 RNAi were subjected to an in vitro Matrigel invasion assay, which quantitatively assesses invasion through an extracellular matrix-coated filter (Figure 1E). CDC20 knockdown using by two distinct RNAi viruses inhibited GSC invasiveness by 55% and 95%, respectively (Figure 1E).

To demonstrate the specificity of the CDC20 RNAi phenotype, we performed a rescue experiment using rat Cdc20 (herein CDC20-Res), which shares 94.8% amino acid identity with human CDC20 but harbors 4 base mismatches within the sequence targeted by CDC20i.2, rendering it insensitive to CDC20i.2 (Figure S2A). The inhibition of GSC invasiveness by CDC20 knockdown was reversed by co-expression of CDC20-Res, demonstrating the specificity of the CDC20 RNAi phenotype (Figure 1F). To test the generalizability of CDC20’s role in GSC invasion, we subjected
two additional patient tumor-derived GSC lines to CDC20 knockdown and similarly found that CDC20 RNAi decreased invasiveness (Figure S2B,C). CDC20 overexpression also increased the invasive capacity of three human GSC lines (Figure 1G,H, S2D,E). Thus, through both loss-of-function and gain-of-function approaches, CDC20 is necessary and sufficient for GSC invasion in vitro.

We next determined if CDC20 operates with the APC to control GSC invasiveness. We knocked down Anaphase-Promoting Complex 2 (ANAPC2), the essential catalytic subunit of the APC, and found that ANAPC2 RNAi inhibited GSC invasiveness in three human GSC lines (Figure 1I, J, S2B,C). We also tested if the interaction between CDC20 and the APC is essential for GSC invasiveness by using a pharmacological inhibitor of the APC, ProTAME, which interferes with the binding of the CDC20 IR tail with the APC (Figure 1K, Figure S2F) (Zeng et al., 2010). We confirmed exposure to ProTAME disrupts the interaction between CDC20 and APC subunit CDC27 in GSCs (Figure S2F). ProTAME treatment inhibited invasiveness in three human GSC lines, suggesting CDC20 acts with the APC to control GSC invasion (Figure 1K, Figure S2G,H).

We next examined the role of CDC20 in GSC self-renewal, a property which often parallels tumorigenic potential (Suva et al., 2014). We performed the extreme limiting dilution assay to measure the frequency of self-renewing cells and found that CDC20 knockdown decreased the percentage of self-renewing GSCs by 45%. (Figure 1L) (Singh et al., 2004). In complementary experiments, CDC20 overexpression increased the frequency of self-renewing cells by 56% and 89% in two GSC lines, respectively (Figure 1M,N). Exposure to APC inhibitor ProTAME also inhibited GSC self-renewal (Figure 1O). Together, these experiments indicate CDC20 operates with the APC to promote GSC invasion and self-renewal.

We next asked if cell cycle perturbations triggered by CDC20-APC manipulations might be responsible for the observed effects on GSC invasion and self-renewal. Examination of cell cycle
profile revealed little to no change in the distribution of cell cycle phases in CDC20 knockdown GSCs compared to that of control infected cells (Figure S3A). Additionally, the degree of CDC20 knockdown achieved in these experiments did not significantly alter cellular proliferation by the MTS assay, although ANAPC2 knockdown modestly decreased proliferation (Figure S3B). These data are consistent with the previously reported observation that CDC20 knockdown does not significantly alter mitotic transition until CDC20 levels drop below a critical threshold (Wolthuis et al., 2008). In other experiments, CDC20 overexpression had little to no effect on the cell cycle distribution or proliferation of GSCs (Figure S3C-F). These results support the hypothesis that CDC20 control of GSC invasiveness and self-renewal can be separated from CDC20 regulation of the cell cycle.

We next asked if CDC20-APC control of GSC function might be a consequence of decreased cellular survival. Importantly, we found that the degree of CDC20 and ANAPC2 knockdown achieved did not significantly alter cell survival in GSCs (Figure S3G). Additionally, CDC20 RNAi did not significantly increase caspase-3 activity in GSCs, and ANAPC2 RNAi caused a mild increase in caspase-3 activity in only one of two GSC lines (Figure S3H,I). In other experiments, short-term treatment with APC inhibitor ProTAME revealed minimal to no cell death in two GSC lines (Figure S3J and data not shown). These results suggest alterations in cell survival were not significantly contributing to the invasion and self-renewal phenotypes observed with CDC20-APC manipulations.

To understand the mechanism of CDC20-APC in GSC invasiveness, we turned to the question of where in the cell CDC20 operates to mediate invasiveness. Previous reports demonstrated that specific subcellular pools of CDC20 dictate distinct biological responses in neural development (Kim et al., 2009; Puram et al., 2011). CDC20 localizes to both cytoplasmic and nuclear compartments in GSCs (Figure 1G) (Kallio et al., 1998). To localize CDC20 to distinct subcellular locations, we generated viruses that express mutant CDC20 fusion proteins carrying either a nuclear
localization sequence (GFP-NLS-CDC20) or nuclear export sequence (GFP-NES-CDC20), the latter localizing CDC20 to the cytoplasm (Figure 2A). Expression of nuclear CDC20 enhanced GSC invasiveness, whereas expression of cytoplasmic CDC20 did not significantly alter invasive capacity, suggesting CDC20-APC stimulates a nuclear program to drive invasion (Figure 2A).

To elucidate the signal transduction pathway downstream of CDC20-APC, we considered nuclear proteins implicated in GSC invasiveness and self-renewal. The stem cell regulatory gene SOX2 has received recent attention in the glioblastoma field due to its critical roles in glioblastoma self-renewal, invasion, and tumor propagation (Alonso et al., 2011; Gangemi et al., 2009). We first tested if a physical interaction exists between CDC20 and SOX2. Remarkably, epitope-tagged CDC20 and SOX2 were found in a complex in transfected 293 cells (Figure 2B). Moreover, we found CDC20 endogenously interacts with SOX2 in two distinct GSC lines (Figure 2C,D). APC subunit CDC27 was also found in an endogenous complex with SOX2, suggesting CDC20-APC interacts with SOX2 (Figure 2C and data not shown). To determine if CDC20 binds directly to SOX2, we performed GST-pull down assays using recombinant GST-SOX2 fusion proteins and in vitro translated CDC20, which revealed a robust direct interaction (Figure 2E). Deletion mapping indicated the WD40 repeat domain of CDC20 interacts directly with SOX2 in vitro (Figure 2E,F). Reciprocal GST pull-down assays using GST-fusion proteins carrying the WD40 repeat domain of CDC20 (GST-CDC20(WD40)) and in vitro translated deletion mutants of SOX2 revealed CDC20(WD40) binds to SOX2 aa1-200 and aa124-317, suggesting SOX2 aa124-200 are required for CDC20 binding (Figure 2G,H). Indeed, the SOX2 deletion mutant lacking aa110-200 failed to bind CDC20(WD40) (Figure 2G,H). These data indicate CDC20-APC endogenously interacts with SOX2 in GSCs likely via direct binding between SOX2 amino acids 124-200 and the WD40 repeat domain of CDC20, suggesting a mechanistic link between CDC20-APC and SOX2.
Differentiation of human GSCs in culture led to a dramatic decrease in CDC20 protein levels, suggesting that as with SOX2, CDC20 is enriched in the GSC state (Figure 3A). To test if CDC20-APC regulates SOX2 in GSCs, we subjected GSCs to \textit{CDC20} knockdown (Figure 3B, S4A). Intriguingly, \textit{CDC20} RNAi decreased SOX2 protein levels in GSCs, and co-expression of RNAi-resistant CDC20-Res with \textit{CDC20} RNAi reversed this decrease, suggesting CDC20 specifically promotes SOX2 protein expression (Figure 3B,C, S4A). Conversely, CDC20 overexpression in two GSC lines increased SOX2 protein (Figure 3D, Figure S4B). In other experiments, both \textit{ANAPC2} knockdown and APC inhibitor ProTAME decreased SOX2 protein in two GSC lines, suggesting collectively that CDC20 collaborates with the APC to maintain SOX2 levels (Figure 3E-G, Figure S4C).

We next turned to the question of how CDC20-APC regulates SOX2 protein levels and examined the effect of APC inhibitor ProTAME on SOX2 protein over time in GSCs (Figure 3G and data not shown). SOX2 protein levels began to decrease about 4 hours after ProTAME exposure (Figure 3G and data not shown), but SOX2 mRNA demonstrated little to no change after ProTAME treatment over a similar timeframe (Figure S4D,E). We therefore examined the possibility that CDC20-APC controls SOX2 protein stability. Consistent with this hypothesis, treatment with proteasome inhibitor MG132 reversed the decrease in SOX2 protein triggered by both ProTAME and \textit{CDC20} RNAi in two GSC lines (Figure 3H,I, Figure S4F). Similar results were seen using a different proteasome inhibitor, bortezomib, in the setting of ProTAME, suggesting CDC20-APC stabilizes SOX2 protein (data not shown).

To determine the biochemical consequences of CDC20-APC control of SOX2, we established a SOX2 transcriptional activity reporter using a lentiviral GFP T2A luciferase expression vector driven by the SOX2-responsive human \textit{SOX2} regulatory region 2 enhancer (hSRR2) (Figure 3J) (Sikorska et al., 2008). We confirmed GSCs infected with this SOX2 reporter virus exhibited a
hSRR2-specific GFP and luciferase signal compared to control reporter-infected cells; COS-1 cells, which do not express SOX2, did not exhibit a hSRR2-dependent signal (Figure S4G). CDC20 knockdown in GSCs substantially decreased the hSRR2-driven luciferase signal compared to control RNAi, suggesting CDC20 promotes SOX2-mediated transcription (Figure 3J). Accordingly, CDC20 knockdown and APC inhibitor ProTAME decreased the mRNA levels of SOX2 target gene NES (Nestin) in GSCs (Figure 3K, Figure S4H) (Berezovsky et al., 2014). Together, these data indicate CDC20-APC positively regulates SOX2 transcriptional activity in GSCs.

To determine the biological consequences of CDC20 regulation of SOX2, we performed epistasis experiments using the Matrigel invasion assay. We first confirmed that SOX2 knockdown decreases GSC invasiveness in three GSC lines (Figure 3L, S4I,J) (Alonso et al., 2011). SOX2 RNAi did not significantly affect cellular survival or health by the propidium iodide exclusion, MTS, and caspase-3 activity assays, consistent with a prior report (Figure S4L,M and data not shown) (Gangemi et al., 2009). Whereas CDC20 overexpression increased GSC invasiveness, the combination of CDC20 overexpression and SOX2 RNAi decreased invasiveness to a level similar to that of SOX2 RNAi alone (Figure 3L). In a second GSC line, SOX2 RNAi also inhibited the ability of CDC20 overexpression to enhance invasiveness (Figure S4K). Conversely, SOX2 overexpression partially but significantly reversed the CDC20 RNAi-induced invasion phenotype (Figure 3M), together suggesting that SOX2 acts downstream of CDC20 to drive invasiveness. The increase in GSC self-renewal triggered by CDC20 overexpression was also inhibited by SOX2 knockdown in two GSC lines (Figure 3N and data not shown), indicating SOX2 functions downstream of CDC20 to control self-renewal. To test if the binding of CDC20 to SOX2 is critical for GSC invasion, structure-function experiments were performed in the setting of SOX2 RNAi (Figure 3O, S4N). Using a SOX2 cDNA carrying 7 base mismatches in the sequence targeted by SOX2 RNAi (SOX2-Res), we generated lentiviruses that express full-length SOX2-Res and mutant SOX2-Res1110-200,
the latter of which does not bind CDC20 *in vitro* (Figure 2G,H). Whereas expression of full-length SOX2-Res rescued the SOX2 RNAi-triggered deficit in invasion, SOX2-Res1110-200 did not, suggesting the binding of SOX2 to CDC20 is important for GSC invasion (Figure 3O).

To examine the relevance of CDC20-APC in GSC tumorigenicity *in vivo*, we used two GSC lines stably expressing GFP T2A luciferase, enabling GFP immunofluorescence as well as bioluminescence imaging (BLI) in live animals to monitor tumor burden (Figure 4A,B). GSCs infected with *CDC20* RNAi or control virus were injected into the brains of NOD-SCID1 mice. BLI performed over several months revealed *CDC20* knockdown inhibited brain tumor formation (Figure 4A,B). GFP immunofluorescence in brain sections of injected mice demonstrated infiltrative tumors corresponding to the BLI signal (Figure 4B and data not shown). In other experiments, we infected a third GSC line with *CDC20* RNAi or control virus, injected these cells into the brains of NOD-SCID mice, and sacrificed mice 3 months later to assess tumorigenicity by immunofluorescence (Figure 4C,S5). Control-infected GSCs formed brain tumors in all animals (8/8), while *CDC20* RNAi-infected GSCs formed tumors in only 2 of 6 animals, suggesting again that CDC20 is critical for the tumor-initiating potential of GSCs (Figure 4C). In complementary experiments, GSCs stably expressing luciferase were infected with CDC20-expressing or control virus, injected into NOD-SCID1 mice, and assessed for brain tumor growth by BLI, which showed that CDC20 overexpression enhances tumor growth *in vivo* (Figure 4D). Together, these experiments indicate CDC20 drives the *in vivo* tumorigenicity of human GSCs.

We interrogated The Cancer Genome Atlas (TCGA) to investigate if *CDC20* expression correlates with clinical outcomes in glioblastoma patients. Consistent with prior reports, we found CDC20 mRNA is significantly elevated in glioblastomas compared to normal brain (Figure 5A) (Bie et al., 2011; Marucci et al., 2008). We then assessed CDC20 expression in the four TCGA-based molecular subtypes—Proneural, Mesenchymal, Classical, and Neural—and found the Proneural
subtype demonstrated significantly higher CDC20 expression compared to the other subtypes (Figure 5B) (Verhaak et al., 2010). We stratified the glioblastoma patients with valid survival data into high (two-fold-change or greater compared to normal brain) and low CDC20 mRNA groups and observed that CDC20 expression in the entire population was not significantly associated with overall survival (OS) (Figure 4C). We then performed Kaplan-Meier survival analyses on patients with high or low CDC20 mRNA expression within each subtype (Figure 4D). Although CDC20 expression did not correlate with OS within the Mesenchymal, Classical, or Neural subtypes, patients with high CDC20-expressing Proneural tumors exhibited a substantially shorter OS (median 53.9 weeks) compared to that of patients with low CDC20-expressing tumors (median 219.6 weeks) (Figure 4D). We confirmed this association using a Cox proportional hazard model to identify an optimal cutoff for CDC20 expression in relation to OS, which also indicated a significant correlation between high CDC20 expression and shorter OS specifically in the Proneural subtype (Figure S6A, B).

Somatic mutations in the isocitrate dehydrogenase 1 gene (IDH1) are found in a subset of Proneural patients with longer OS than patients with IDH1-wildtype tumors (Hartmann et al., 2010; Parsons et al., 2008; Yan et al., 2009). We asked if CDC20 expression might interact with IDH1 mutation status or represent an independent prognostic marker in Proneural glioblastomas (Figure S6C,D). When IDH1 MUT tumors were included, Proneural tumor patients with high CDC20 expression again had a poorer prognosis (Figure S6C). When IDH1 MUT tumors were excluded, the number of Proneural tumor patients with low CDC20 expression was small (6 patients with 4 censored), but the OS of patients with high and low CDC20 tumors was not appreciably different, suggesting an interaction between IDH1 mutation and CDC20 expression (Figure S6C). We then examined gene expression data for Proneural tumors only and found IDH1 MUT tumors exhibit significantly lower CDC20 expression compared to that of IDH1 WT tumors (Figure S6D). Together,
these data indicate CDC20 expression is prognostic of OS in Proneural glioblastomas and, in a limited subset analysis, appears to interact with IDH1 mutation status.

3.4 Discussion:

In this study, we have demonstrated CDC20-APC operates through SOX2 to control human GSC invasion and self-renewal. Additionally, we have found CDC20 is critical for human GSC tumorigenicity in vivo. Interrogation of the TCGA revealed high CDC20 expression was associated with decreased overall survival in Proneural subtype glioblastomas.

CDC20-APC has been intensively studied in the cell cycle field and is viewed as a promising target in several human cancers (Wang et al., 2015). As proof of concept, conditional Cdc20 knockout in mouse models of skin cancer and fibrosarcoma caused mitotic arrest and apoptotic tumor regression (Manchado et al., 2010). Intriguingly, the essential role of CDC20 in GSC invasiveness and self-renewal appears to be separable from CDC20’s known role in cell cycle regulation; the CDC20 manipulations used herein did not obviously affect proliferation or cell cycle parameters, consistent with the previous finding that only a minimal level of CDC20 is needed for mitotic transition (Wolthuis et al., 2008). More recently, CDC20 knockdown was shown to sensitize cancer cells to chemotherapy and radiation therapy (Wan et al., 2014). Our results reinforce the rationale for the development of CDC20-APC inhibitors in glioblastoma not only to reduce tumor burden through cell cycle and cell death mechanisms but also to disrupt key functional properties of GSCs.

As with SOX2, CDC20 protein is enriched in human GSCs compared to glioblastoma cells differentiated in vitro. This finding, which remains to be validated in human tumor samples ex vivo, raises interesting questions about how CDC20 is regulated in the GSC state. Downstream of CDC20, regulation of SOX2 appears to occur at two—not necessarily mutually exclusive—levels: CDC20
binding to SOX2 and CDC20-APC control of SOX2 protein stability. SOX2 binding to CDC20 appears to be important for SOX2 control of GSC invasiveness (Figure 2H, 3O). It is possible that CDC20 binding enhances SOX2 function, perhaps through the CDC20-APC-dependent recruitment of transcriptional activators (Turnell et al., 2005). The exact mechanistic link between CDC20-APC and SOX2 protein stability remains an important open question. SOX2 regulation by the ubiquitin-proteasome system has only recently begun to be examined in the context of non-cancerous cells, such as embryonic stem (ES) cells. Whereas SOX2 acetylation and methylation increase SOX2 degradation, phosphorylation of murine SOX2 at Thr118 (Thr116 in human SOX2) by AKT stabilizes SOX2 protein, raising the possibility that CDC20-APC might affect SOX2 stability by altering SOX2 post-translational modifications (Baltus et al., 2009; Fang et al., 2014; Jeong et al., 2010). Alternatively, CDC20-APC may act indirectly on SOX2 by ubiquitinating and destroying a critical E3 ligase, which targets SOX2. Only two E3 ligases that target and degrade SOX2 have been reported so far. One is FZR1 (also CDH1), an alternative co-activator of the APC, which is responsible for G1 maintenance (Fukushima et al., 2013). However, we found little to no change in SOX2 protein levels in the setting of CDH1 RNAi in GSCs, and ANAPC2 RNAi and APC inhibitor ProTAME, which inhibit both CDH1-APC and CDC20-APC, decreased SOX2 protein, suggesting a dominant role for CDC20-APC in SOX2 protein regulation in GSCs (Figure 3E-G and data not shown). More recently, WWP2 was identified as a SOX2 ubiquitin ligase in ES cells (Fang et al., 2014). Whether WWP2 or other E3 ligases contribute to SOX2 stability in glioblastoma remains to be determined.

Our results have several intriguing implications for CDC20-APC’s role in the transcriptional networks governing glioblastoma molecular subtypes as well as non-cancerous stem/progenitor cells. Interestingly, we have found in the TCGA dataset that CDC20 expression is particularly elevated in the Proneural subtype. Since SOX2 is a known Proneural signature gene, the finding that CDC20-
APC promotes SOX2-dependent transcription raises the intriguing hypothesis that CDC20-APC stimulates Proneural signature gene transcription (Verhaak et al., 2010). The CDC20-APC/SOX2 mechanism might therefore be particularly relevant for the biology underlying this molecular subtype. As a prognostic marker, CDC20 expression appears to interact with IDH1 mutation, suggesting a potential mechanistic link between IDH1 mutant status and low CDC20 expression. But the exact relationship between CDC20 expression and survival in the bulk tumor data of the TCGA and the CDC20-APC/SOX2 mechanism in GSCs requires further investigation. For instance, in contrast to bulk tumor, human GSCs cluster into predominantly two molecular subtypes—Proneural and Mesenchymal (Bhat et al., 2013). The molecular subtyping of the human GSC lines utilized in this study suggests the control of core GSC functions by the CDC20-APC/SOX2 signaling axis is generalizable and independent of GSC subtype (Figure S1A). Additionally, current mRNA and genome-based bulk tumor datasets may not reflect the SOX2 protein regulatory mechanisms reported herein, which will require interrogation of proteomic datasets. More speculatively, the CDC20-APC/SOX2 pathway may play a role in the transcriptional program in other cellular contexts, including the regulation of neural stem cells and potentially, the maintenance of pluripotency in embryonic or induced pluripotent stem cells (Lewitzky and Yamanaka, 2007; Pevny and Nicolis, 2010).

Although the mechanisms of SOX2’s critical role in self-renewal have been extensively investigated in the context of stem cell biology and cancer (He et al., 2009), the downstream mechanisms that specifically drive SOX2-dependent invasion in glioblastoma remain to be identified. SOX2 has been implicated in promoting the invasive potential of other cancers, raising the possibility that CDC20-APC control of SOX2 might regulate invasion in diverse cancers (Forghanifard et al., 2014; Girouard et al., 2012; Han et al., 2012; Lou et al., 2013; Xia et al., 2014). Future analyses of SOX2 transcriptional targets will be important to elucidate the precise mechanisms of SOX2-mediated
Invasiveness specifically in glioblastoma. Moreover, given the multitude of identified CDC20-APC substrates, it is likely that additional, SOX2-independent mechanisms contribute to CDC20-APC regulation of GSC invasiveness and self-renewal.

3.5 Experimental procedures:

**Cell Culture.** The generation of adherent human GSC cultures has been described (Pollard et al., 2009). Briefly, tumor samples obtained directly from surgery were dissociated by mincing and incubation in Accutase (SIGMA) for 20-60 minutes at 37°C. Cell suspensions were passed through a 70 micron cell strainer (Falcon) and plated using Ndiff RHB-A media (Stem Cell, UK) with EGF and FGF-2 (Peprotech) (hereafter “GSC media”), each at 20 ng/ml, on polyornithine and laminin (SIGMA)-coated Primaria dishes/flasks (BD Bioscience). Media was replaced with half fresh GSC media every 2-3 days. Cells were routinely used between passages 5 and 20. Informed consent was obtained from patients for use of human tissue and cells, and all human tissue-related protocols used in this study were approved by the Institutional Review Board (Washington University). Primary human astrocytes (Lonza) were cultured in astrocyte growth media (Lonza). Human embryonic kidney 293 cells were cultured in Dulbecco's modified Eagle's medium with 10% fetal bovine serum (FBS) and penicillin/streptomycin (Life Technologies). All cell lines were incubated at 37°C with 5% CO₂. Lentiviral transduction was performed by adding virus with 4 1g/mL of polybrene for 4 hours to cells. For rescue or epistasis experiments, GSCs were transduced with RNAi or control lentivirus one day after plating and then transduced with CDC20-Res expression virus or control virus the following day. Cells were selected in 2 1g/mL of puromycin 1-2 days after infection. For
self-renewal and in vivo tumorigenicity experiments, GSCs were utilized 4 days following indicated viral infections.

**Cell Invasion Assay.** The in vitro cell invasion assay was performed using Matrigel-coated invasion chambers (BD Bioscience) (Valster et al., 2005). In 24-well plates, 5 x 10^4 GSCs in GSC media was added to the upper chamber of a rehydrated, Matrigel-coated polycarbonate membrane filter. The bottom chamber of the well was pre-filled with RHB-A media containing 10% FBS as chemoattractant. After 24 hrs, non-invasive cells from the upper side of the filter were removed using a moist cotton swab. The invasive cells on the reverse side of the filter were then fixed and stained with DAPI nuclear dye, and images of the cells were captured in a blinded fashion in 3 different low-power fields (5X objective) per condition using a fluorescence microscope (Leica Microsystems, DMI4000 B). Quantitation of invasion was also performed in a blinded fashion using Image J software (NIH).

**Extreme limiting dilution analysis.** Cells were plated at five-fold dilutions (3000, 600, 120, 24, 5 or 1 cell/well) in Corning ultra-low attachment 96-well plates. 7-10 days later, the number of wells containing spheres was counted and used to calculate the frequency of self-renewing GSCs by online software (http://bioinf.wehi.edu.au/software/elda/) (Hu and Smyth, 2009; Singh et al., 2004).

**Xenotransplantation.** Animals were used in accordance with a protocol approved by the Animal Studies Committee of the Washington University School of Medicine per the recommendations of the Guide for the Care and Use of Laboratory Animals (NIH). 250,000 cells per animal (unless otherwise noted) were injected stereotactically into the right putamen of approximately 6-week-old male NOD-SCID1 mice (for B36, B49) or female NOD-SCID mice (for B18) (Hope Center Animal
Surgery Core, Washington University). The coordinates used were: 1 mm rostral to bregma, 2 mm lateral, and 2.5 mm deep.

**Statistics.** All images are representative of results from 3 independent experiments unless otherwise stated. Statistical analyses were performed with XLSTAT (Addinsoft), Excel (Microsoft), or R Version 3.1.1 software. The unpaired Student’s t-test was used for comparisons in experiments with only two groups. In experiments with more than two comparison groups, analysis of variance (ANOVA) was performed followed by Fisher’s least significant difference or the Bonferroni test for pairwise comparisons among three and greater than three groups, respectively.
3.6 Author contributions:

D.D.M., A.D.G., T.M., and A.H.K. performed and analyzed experiments; H.Y., Y.P., T.B., E.A.T., S.A., and A.T. contributed to experiments; J.L. and I.C. performed bioinformatic analyses; A.H.K., E.C.L., M.R.C., K.M.R., J.L.D., G.J.Z., and R.G.D. provided clinical material; H.Y., G.P.D., and D.D.T. edited the manuscript; and A.H.K. conceived the research project, analyzed the data, and wrote the manuscript.
3.7 Acknowledgments:

This work was supported by National Institutes of Health Grant K08NS081105, American Cancer Society-Institutional Research Grant, Voices Against Brain Cancer, the Elsa U. Pardee Foundation, the Concern Foundation, the Duesenberg Research Fund, (to A.H.K.), National Institutes of Health Grant K01AG033724 (to H.Y.), and National Institutes of Health Grant P50 CA094056 (to S.A.). We thank members of the Kim and Yano laboratories for helpful discussions and critical reading of the manuscript.
3.8 References:


Lee, J., Kotliarova, S., Kotliarov, Y., Li, A., Su, Q., Donin, N.M., Pastorino, S., Purow, B.W., Christopher, N., Zhang, W., et al. (2006). Tumor stem cells derived from glioblastomas cultured in bFGF and EGF more closely mirror the phenotype and genotype of primary tumors than do serum-cultured cell lines. Cancer cell 9, 391-403.


Figure 3.1 CDC20-APC controls glioblastoma stem-like cell invasion and self-renewal.
(A) GSC lines were subjected to immunofluorescence with indicated antibodies and Hoechst nuclear stain. Bar = 50µm.

(B) B18 GSCs were injected into the right putamen of NOD-SCID mice, and animals sacrificed after 3 months. Sectioned brains were subjected to immunohistochemistry with indicated antibodies. Nuclei were stained with DAPI. White box highlights tumor. Bar = 100µm.

(C) Lysates from GSC lines and normal human astrocytes (NHA) were processed for immunoblotting using indicated antibodies.

(D) B18 GSCs were transduced with CDC20 RNAi (CDC20i.1 and CDC20i.2) or control scrambled (Scr) lentivirus. 7 days later, cell lysates were subjected to immunoblotting using indicated antibodies. Similar results were seen with control viruses SHC002 and LacZ RNAi (data not shown).

(E) GSCs treated as in (D) were subjected to the in vitro Matrigel transwell assay 6 days after infection. Data represent mean+SEM. CDC20 knockdown inhibited GSC invasiveness compared to control (ANOVA, P = 0.001 and P < 0.0001 for CDC20i.1 and CDC20i.2, respectively, n = 4).

(F) GSCs transduced with the indicated lentiviruses were subjected to the in vitro Matrigel transwell assay as in (E). Data represent mean+SEM. CDC20 RNAi decreased GSC invasiveness compared to control (ANOVA; P < 0.003, n = 3). Expression of CDC20-Res rescued the CDC20 RNAi-triggered invasion phenotype (ANOVA; P = 0.003). Vec = control vector virus.

(G) B18 GSCs transduced with GFP-CDC20-Res-expressing lentivirus were subjected to live fluorescence microscopy. Bar = 10µm.

(H) GSCs transduced with GFP-CDC20-Res-expressing or control vector lentiviruses (Vec) were assessed for invasion 5 days later. Data represent mean+SEM. CDC20 overexpression increased GSC invasion compared to control (unpaired t-test, P = 0.01, n = 5).

(I) B18 GSCs were transduced with ANAPC2 RNAi (ANAPC2i) or control scrambled (Scr) lentivirus. 7 days later, cell lysates were subjected to immunoblotting using indicated antibodies. Similar results were seen using control virus SHC002 (data not shown).

(J) GSCs treated as in (I) were subjected to the in vitro Matrigel transwell assay 6 days later. Data represent mean+SEM. ANAPC2 knockdown inhibited GSC invasiveness compared to control (unpaired t-test, P = 0.001, n = 4).

(K) B18 GSCs were subjected to the in vitro Matrigel transwell assay in the presence of ProTAME or DMSO (Veh). Data represent mean+SEM. ProTAME inhibited invasion in a dose-dependent manner (ANOVA, P < 0.003, n = 3).

(L) GSCs infected with CDC20i.2 or control (Scr) virus were subjected to the extreme limiting dilution assay. 7 days later, the number of wells with spheres was counted and analyzed. Data represent mean+SEM. CDC20 RNAi decreased the percentage of self-renewing GSCs compared to control (unpaired t-test, P = 0.0002, n = 3).

(M) B18 GSCs infected with CDC20-expressing or control virus were treated as in (L). Data represent mean+SEM. CDC20 overexpression increased the percentage of self-renewing GSCs compared to control (unpaired t-test, P = 0.0005, n = 3).

(N) A1 GSCs treated as in (M) were subjected to the extreme limiting dilution assay. Data represent mean+SEM. CDC20 overexpression increased the percentage of self-renewing GSCs compared to control infection (unpaired t-test, P = 0.009, n = 3).

(O) GSCs were subjected to the extreme limiting dilution assay with 10µM ProTAME or DMSO (Veh) and analyzed as in (L). Data represent mean+SEM. ProTAME decreased the percentage of self-renewing GSCs compared to vehicle (unpaired t-test, P < 0.0001, n = 3).

See also Table S1 and Figure S1-3.
Figure 3.2 CDC20-APC interacts with SOX2 through the WD40 repeat domain of CDC20.
(A) Top: B18 GSCs transduced with lentiviruses expressing indicated GFP-tagged mutant CDC20 proteins were subjected to live fluorescence microscopy. Bar = 25 µm. Bottom: GSCs treated as above were assessed for invasion 6 days later. Data represent mean±SEM. Expression of nuclear localized CDC20 (NLS-CDC20) but not cytoplasmic CDC20 (NES-CDC20) increased GSC invasiveness compared to control (ANOVA, P = 0.005, n = 5).

(B) Lysates of 293 cells transfected with GFP-CDC20 together with the myc-SOX2 expression plasmid or control vector were immunoprecipitated using myc antibody and immunoblotted with the indicated antibodies.

(C) GSC line B18 lysates were immunoprecipitated with the CDC20 or control IgG antibody and immunoblotted with the indicated antibodies.

(D) GSC line B36 lysates were immunoprecipitated with the SOX2 or control IgG antibody and immunoblotted with the indicated antibodies.

(E) In vitro translated, 35S-methionine-labeled CDC20 mutant proteins were used in GST pull-down assays using recombinant GST-SOX2 and GST proteins (left panel). The middle panel (Input) confirms comparable levels of CDC20 mutants. CDC20 proteins were visualized by fluorography. Similar amounts of GST and GST-SOX2 were used for pull-downs (Coomassie brilliant blue staining (CBB), right panel).

(F) Schematic depicting CDC20 domain structure (top) and summary of in vitro binding experiments (bottom).

(G) In vitro translated, 35S-methionine-labeled SOX2 mutant proteins were used in GST pull-down assays using recombinant GST-CDC20 WD40 repeat domain (aa 168-477) (referred to as GST-CDC20 in this panel) and GST proteins. SOX2 proteins were visualized by fluorography. The first panel shows a low exposure and the second panel a high exposure of a representative experiment. Input panels confirm SOX2 mutants were produced at comparable levels. Similar amounts of GST and GST-CDC20 were used for pull-downs (CBB, far right panel). Images were background corrected.

(H) Schematic depicting SOX2 domain structure (top) and summary of in vitro binding experiments (bottom).
Figure 3.3 CDC20-APC regulation of SOX2 protein and transcription controls GSC invasion and self-renewal.
(A) GSCs (B18, A1) were maintained in GSC or differentiating medium (Diff) (containing fetal bovine serum and no growth factors) for 14 days. Cell lysates were subjected to immunoblotting using indicated antibodies.

(B) B18 GSCs were transduced with CDC20 RNAi (CDC20i.1 and CDC20i.2) or control LacZ RNAi (C) lentivirus. 7 days later, cell lysates were subjected to immunoblotting using indicated antibodies.

(C) GSCs were transduced with the indicated lentiviruses. 7 days later, cell lysates were subjected to immunoblotting using indicated antibodies. Expression of CDC20-Res rescued the CDC20 RNAi-triggered decrease in SOX2 protein. C = SHC002 virus. Vec = control vector virus.

(D) GSCs transduced with CDC20-expressing lentivirus or control vector virus (Vec) were maintained in RHB-A media for 5 days. Cell lysates were subjected to immunoblotting using indicated antibodies.

(E) GSCs were transduced with ANAPC2 RNAi or control LacZ RNAi (C) lentivirus. 7 days later, cell lysates were subjected to immunoblotting using indicated antibodies.

(F) GSCs were treated with ProTAME or DMSO (Veh) for 12 hours. Cell lysates were subjected to immunoblotting using indicated antibodies.

(G) GSCs were treated with 20 µM of ProTAME or DMSO (Veh) as indicated. Cell lysates were subjected to immunoblotting using indicated antibodies.

(H) GSCs were treated with 20 µM of ProTAME, 5 µM of proteasome inhibitor MG132, or a combination of both for 8 hours. Cell lysates were subjected to immunoblotting using indicated antibodies. Veh = DMSO.

(I) GSCs were transduced with CDC20 RNAi (CDC20i.2) or control SHC002 (C) lentivirus for 7 days and treated with 10 µM of MG132 or DMSO (Veh) for 6 hours. Cell lysates were subjected to immunoblotting using indicated antibodies.

(J) GSCs stably infected with the SOX2 transcriptional reporter (hSRR2) or control reporter (mCMV) were transduced with CDC20 RNAi (CDC20i.2) or control scrambled (Scr) lentivirus. 7 days later, luciferase assays were performed. Luciferase values were normalized by total protein, and fold-change calculated by scaling to Scr + mCMV values (=1). Data represent mean+SEM. CDC20 RNAi decreased SOX2 reporter activity compared to control (ANOVA, P < 0.0001, n = 3).

(K) GSCs were infected with CDC20 RNAi (CDC20i.1 and CDC20i.2) or control scrambled (Scr) lentivirus. RNA was harvested 7 days later and reverse transcribed into cDNA. qPCR was performed on samples using specific primers for human NES, GAPDH and ACTB were used as reference genes. Data represent mean+SEM. CDC20 RNAi decreased NES mRNA in GSCs compared to control. (ANOVA, P < 0.0001, n = 3).

(L) GSCs infected with the CDC20-expressing or control vector (Vec) lentivirus together with the SOX2 RNAi (SOX2i) or control SHC002 RNAi (Scr) virus were subjected to the in vitro Matrigel transwell assay 7 days later. Data represent mean+SEM. Expression of CDC20 increased invasion compared to control infection (ANOVA, P = 0.007, n = 6). Expression of CDC20 plus SOX2 RNAi reduced invasion compared to infection with Scr plus either the CDC20-expressing or control vector virus. (ANOVA; P < 0.0001 and P = 0.001, respectively).

(M) GSCs infected with SOX2-expressing or control vector (Vec) lentivirus together with CDC20 RNAi or control SHC002 RNAi (Scr) were treated as in (L). Data represent mean+SEM. CDC20 knockdown decreased invasiveness compared to control (ANOVA, P < 0.0001, n = 4). Expression of SOX2 plus CDC20 RNAi increased invasion compared to infection with Vec plus CDC20 RNAi viruses (ANOVA; P = 0.011).

(N) GSCs infected as in (L) were subjected to the extreme limiting dilution assay as in Figure 1L. Data represent mean+SEM. Expression of CDC20 plus Scr increased self-renewal compared to control (ANOVA, P = 0.004, n = 3). Expression of CDC20 plus SOX2 RNAi reduced self-renewal
compared to infection with Scr plus either the CDC20-expressing or control virus. (ANOVA; $P < 0.0001$ and $P = 0.001$, respectively).

(0) GSCs were transduced with the indicated lentiviruses and subjected to the in vitro Matrigel transwell assay as in (L). Data represent mean+SEM. SOX2 RNAi plus control vector virus (Vec) decreased GSC invasiveness compared to control (ANOVA; $P < 0.001$, $n = 6$). Expression of full-length SOX2-Res (FL) rescued the SOX2 RNAi-triggered defect in invasiveness (ANOVA; $P < 0.0001$) whereas SOX2-Res1'110-200 did not (ANOVA; $P = 0.9$). See also Figure S4.
Figure 3.4 CDC20 drives GSC tumorigenicity in vivo
(A) B36 GSCs stably infected with CMV-driven GFP T2A luciferase lentivirus were transduced with CDC20 RNAi or control LacZ RNAi (LacZi) lentivirus and injected into the right putamen of NOD-SCIDy mice. Injected mice were subjected to live bioluminescence imaging (BLI). Top: Data represent mean±SEM (n = 5 animals per condition). CDC20 knockdown decreased GSC tumorigenicity compared to control (unpaired t-test, *P < 0.01, #P < 0.05). Bottom: Representative animals subjected to BLI are shown.

(B) B49 GSCs stably infected with CMV-driven GFP T2A luciferase lentivirus were treated as in (A) and injected into the brains of NOD-SCIDy mice. Injected mice were subjected to live BLI. Top: Data represent mean±SEM (n = 5 animals per condition). CDC20 knockdown decreased GSC tumorigenicity compared to control (unpaired t-test, *P < 0.05). Bottom: Injected mice were sacrificed at 5 months. Representative coronal brain sections subjected to GFP immunofluorescence to visualize tumor are shown. Nuclei were stained with DAPI. Bar = 100 µM

(C) B18 GSCs infected with CDC20i.2 or control SHC002 (Scr) virus were injected into the brains of NOD-SCID as in (A). 3 months after injection, animals were sacrificed, and brains were processed for immunohistochemistry using antibodies against NES and GFAP. Nuclei were stained with Hoechst 33342. The number of animals harboring a brain tumor in each treatment group is indicated. CDC20 knockdown decreased GSC tumorigenicity compared to control (Fisher’s exact test, P = 0.015).

(D) B36 GSCs stably infected with CMV-driven GFP T2A luciferase lentivirus were transduced with GFP-CDC20-expressing or control vector lentivirus and injected into the brains of NOD-SCIDy mice as in (A). Injected mice were subjected to live BLI. Left: Data presented are mean±SEM (n = 4 animals per condition). CDC20 overexpression increased tumor formation compared to control (unpaired t-test, P < 0.04). Right: Representative animals subjected to BLI are shown.

See also Figure S5.
Figure 3.5 High CDC20 expression is associated with decreased overall survival (OS) in Proneural subtype glioblastomas.
(A) Box plot (median and middle 50% of data represented in each box) for CDC20 mRNA expression in TCGA glioblastoma (n = 473) and normal brain tissue samples (n = 10). CDC20 expression is higher in glioblastoma samples compared to normal tissue (unpaired t-test, \( P = 9.62 \times 10^{-14} \)).

(B) Box plot for normalized CDC20 gene expression (compared to normal samples) demonstrates the highest level of CDC20 expression in the Proneural subtype compared to other subtypes (Holm's adjustment for multiplicity, \( P < 0.0001 \)).

(C) Kaplan-Meier curves showing OS of 466 newly diagnosed glioblastoma patients from the TCGA based on \( CDC20 \) expression. High \( CDC20 \) represents two-fold or greater expression and low \( CDC20 \) less than two-fold expression compared to mean \( CDC20 \) expression in normal brain samples (log-rank test, \( P = 0.390 \)).

(D) Kaplan-Meier curves showing OS of TCGA patients separated by molecular subtype based on \( CDC20 \) expression. Data were analyzed as in (C). High CDC20 expression was associated with decreased OS only in patients with Proneural tumors (log-rank test, \( P = 0.002 \)).
3.9 Supplemental Experimental Procedures

**Antibodies and drugs.**

Antibodies used for this study include polyclonal rabbit anti-CDC20 (H-175) (Santa Cruz Biotechnology) and mouse anti-CDC27 (C-4) (Santa Cruz Biotechnology), polyclonal rabbit anti-ANAPC2 (APC2 (H-295)) (Santa Cruz Biotechnology), mouse anti-β-Actin(C4) (Santa Cruz Biotechnology), polyclonal rabbit anti-Nestin (Millipore), mouse anti-Nestin MAB1259 (R&D systems), rabbit anti-Nestin ABD69 (Millipore), monoclonal rabbit anti-SOX2 86D9 (Cell Signaling Technology), rabbit polyclonal anti GFAP Z0334 (Dako), mouse anti-α-tubulin(Sigma), and polyclonal rabbit anti-GFP serum (A6455) (Molecular Probes, Life Technologies), monoclonal rabbit anti-EZH2 D2C9 (Cell Signaling Technology), mouse anti-BMI1 05 637 (Millipore), polyclonal rabbit anti-myc (A-14) (Santa Cruz Biotechnology), mouse anti-myc (9E10) (SIGMA), monoclonal rabbit anti-c-MYC D84C12 (Cell Signaling Technology), rabbit anti-MSI1 ab21628 (Abcam), and mouse anti-FLAG (M2) (SIGMA). CD133/1(AC133) antibody conjugated to PE was purchased from Miltenyi Biotec Inc. Drug used include cell permeable APC inhibitor ProTAME (Boston Biochem) proteasome inhibitor MG-132 (EMD Millipore Corporation), and bortezomib. Hoechst 33342 was purchased from Molecular Probes, Life Technologies, and propidium iodide and Fluoroshield with DAPI purchased from SIGMA.

**Plasmids.**

The following human gene target-directed shRNA plasmids (in pLKO.1) from the RNAi Consortium were used: CDC20 RNAi (CDC20i.1 and 2), targeting 5’-AGACCAACCCATCACCCTCAGT-3’ and 5’-TGTTGGTAATGATAACTTGGT-3’, respectively (Washington University RNAi Core); ANAPC2 RNAi, targeting 5’-CCCGGCACCTTCTCTGTCTT-3’; SOX2 RNAi, targeting 5’-AGCGTGACTATCCCTTCTTTC-3’ (SIGMA); LacZ RNAi, targeting 5’-TGTTGCATTATCCGAACCAT-3’; control RNAi plasmids SHC002 and Scrambled containing
sequences
CCGGCAACAAGATGAAGAGCACAACACTCGAGTTGCTCTCATCTTGTGTTTTT
and CCTAAGGTTAAGTCGCCCTCGAGCGAGCGGCGACTTAACCTTAGG (Addgene), respectively. Plasmids encoding FLAG-tagged rat Cdc20, the N-terminal GFP and rat Cdc20 fusion protein (GFP-CDC20-Res) as well as GFP-NLS-CDC20-Res and GFP-NES-CDC20-Res have been described (Kim et al., 2009). GFP-CDC20-Res and mutant GFP-CDC20-Res cDNAs were subcloned into N103 lentiviral vector (kindly provided by Dr. Andrew Yoo (Washington University)). Full-length human SOX2 was cloned from a B18 GSC cDNA library. Myc epitope tagged rat Cdc20 deletion and human SOX2 deletion expression plasmids were generated in pcDNA3 by standard PCR and subcloning techniques. GST-human SOX2 and GSTCDC20(WD40) were generated by using the bacterial expression plasmid pGEX-4T1. The HAtagged full-length human SOX2 rescue construct (HA-SOX2-Res), which harbors 7 base mismatches in the region targeted by SOX2 RNAi (GAAAAAGACAAATATAACCTT, mismatches highlighted), and SOX2 deletion mutants were generated by gene synthesis (Integrated DNA Technologies) and subcloned into pcDNA3 and N103. The human SOX2 regulatory region 2 (hSRR2)/minimal CMV promoter (mCMV)-driven GFP T2A luciferase lentiviral plasmid was generated by first PCR cloning SRR2 from 293 cell genomic DNA using the following primers: forward: CCATCGATTTTTAGGATAAACATTGTACTGGGAAG and reverse: CCACTAGTATCAAAAAGCTTTATATTGCAAAAC. The PCR product was then subcloned into the ClaI and SpeI sites in pGreenfire1 mCMV-GFP T2A luciferase plasmid (Systems Biosciences). CMV-driven GFP T2A luciferase lentiviral plasmid was generated by standard PCR subcloning. All plasmids were confirmed by sequencing.

Real-Time Quantitative PCR.
RNA was isolated using the RNeasy Plus Mini Kit (QIAGEN). Reverse transcription was performed with High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Power SYBR Green PCR Master Mix (Applied Biosystems) was used for amplification on the CFX Connect Real-Time System (Bio-Rad). All samples were run in triplicate with a corresponding β-Actin and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) control for each. Relative transcript copy number for each transcript was normalized to β-Actin and GAPDH and was calculated using the \( \Delta \Delta C_q \) method. The primers sets used for each gene are as follows: Human SOX2 forward: 5’-GGGGAAAGTAGTTTGCTGCC-3’; Human SOX2 reverse: 5’-CGCCGCCGATGATTGTATT-3’; Human NES forward: 5’-GAGAGCCCTGAGCCCCAAAGA-3’; Human NES reverse: 5’-CTCCCGCAGCAGACTCACC; Human β-Actin forward: 5’-ATGATATCGCCGCCTCGTCGTC-3’; Human β-Actin reverse: 5’-TGACCCATGCCCACCATCACG-3’; Human GAPDH forward: 5’-ATGGGGAAGGTGAAGGTCG-3’; Human GAPDH reverse: 5’-GGGGTCATTGATGGCAACAATA-3’.

**Transient transfection.**

Polyethyleneimine (PEI, Polysciences #24765-2) was dissolved in ddH₂O to a concentration of 1 mg/mL. The solution was then adjusted to pH 7.0 and filter sterilized. A mixture of plasmid DNA and PEI solution (1:4::μg DNA: μL PEI) was made in OPTiMEM (Life Technologies) and incubated at room temperature for 15 minutes. Total plasmid DNA amount was equalized by addition of vector pcDNA3. DNA/PEI complexes were applied to cells, and media was changed after 12-16 hours. Experiments were performed 24-36 hours following transfection.
Lentiviral production.

293LE cells were plated with a goal density of 70-80% after 1 day. The next day, transfection was performed using the PEI transfection method to introduce the plasmid of interest along with packaging plasmid psPAX2 and envelope plasmid pCMV-VSVG to Opti-MEM (Life Technologies). On day 6, medium from the plates was collected and spun down at 1200 x g for 5 min at 4°C. Supernatant was filtered through 0.45 micron filters. Lenti-X Concentrator (Clontech) was then added to the filtrate and mixed, and the tubes were incubated at 4°C for 6-7 hours. Lentiviruses were then centrifuged at 1500 x g for 45 minutes at 4°C. Supernatant was aspirated; pellets were re-suspended in one tenth of the original medium volume of cold PBS, and stored at -80°C in aliquots. Viral copy number was adjusted for transduction of GSCs on the basis of titer measured using the Lenti-X qRT-PCR titration kit (Clontech).

Co-immunoprecipitation and immunoblot analysis.

Cells were lysed in 1% NP-40 lysis buffer containing 20 mM Tris [pH 8], 200 mM NaCl, 10% glycerol, 1 mM EDTA, 10 mM NaF, 1 mM sodium orthovanadate, and a protease inhibitor cocktail (Calbiochem). Clarified lysates were precleared with protein A/G-sepharose (Life Technologies), and immunoprecipitations were performed overnight at 4°C with indicated antibodies followed by protein A/G-sepharose. Pellets were washed 6 times with lysis buffer and boiled in Laemmli sample buffer. Samples with were separated by SDS-PAGE and transferred to 0.45 μm Immobilon-P PVDF membrane (EMD Millipore). Membranes were blocked in 5% Milk in Tris-buffered saline with Tween-20 (TBST) at room temperature and
incubated with primary antibodies at 4°C overnight or at room temperature for 2-4 hours. Membranes were washed with TBST and incubated with appropriate horseradish peroxidase-conjugated secondary antibodies for 1 hour at room temperature. Membranes were then washed with TBST and developed using Pierce ECL western blotting substrate (Thermo Scientific).

**In vitro binding.**

GST-human SOX2 and the corresponding GST control proteins were bacterially expressed and isolated using Glutathione Sepharose 4 Fast Flow (GE Healthcare) according to manufacturer's instructions. GST-CDC20(WD40) and the corresponding GST proteins were produced in transiently transfected 293 cells following a similar protocol as for bacterially expressed proteins. Using the pcDNA3-myc-rat Cdc20 and pcDNA3-myc or HA- human SOX2 deletion mutant plasmids, 35S-labeled products were *in vitro* translated with the TNT-coupled reticulocyte lysate system (Promega) and incubated with indicated GST-fusion proteins bound to glutathione-sepharose 4 Fast Flow beads in 1% NP-40 lysis buffer at 4°C for 16 hours (GE Healthcare). The beads were washed 6 times with lysis buffer and boiled in LaemmlI sample buffer. Proteins were resolved by SDS-PAGE, and 35S-labeled proteins were visualized by autoradiography with signal enhancement (Amplify, GE Healthcare). GST-fusion proteins were assessed by Coomassie Brilliant Blue R-250 (Biorad).

**Flow cytometry for CD133.**

5 x 10^5 live GSCs were dissociated and labeled with CD133/1(AC133) antibodies conjugated to PE or mouse IgG1 isotype control antibodies conjugated to PE (Miltenyi Biotec Inc). Normal mouse serum (SIGMA) was used as blocking reagent. Cells were subjected to flow
cytometry (Siteman Flow Cytometry Core) using the Blue fluorescence-2 channel to detect PE and analyzed using Flowjo 10.

**Cell survival and proliferation.**

2 x 10^4 GSCs were plated per well in a 24-well plate. Fresh medium was added every 2 or 3 days. One week after plating, cells were assessed for the number of viable cells as determined by Trypan blue exclusion (Life Technologies) was quantified using the Countess Automated Cell Counter (Life Technologies).

**MTS assay.**

3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS) assays were performed per manufacturer’s instructions (Promega), and MTS- reducing activity was normalized for each condition to control scrambled RNAi or vector lentiviruses as appropriate (=1).

**Cell cycle analysis.**

GSCs were plated at a density of 2x10⁵/60mm dish 18 hrs prior to transduction with CDC20 RNAi or CDC20 overexpression virus. Cells were fixed with 50% Ethanol/ phosphate-buffered saline (PBS)/0.01% IGEPAL CA-630 overnight at 4°, treated with Trypsin, RNAse, and then spermine. DNA nuclei were stained with propidium iodide. Cells were then assessed using FACScan (BD Bioscience) (Siteman Flow Cytometry Core) with fluorescence-2 (FL2) detector to detect light emitted between 564 and 604 nm. The data were analyzed using Flow Jo 10 software.
**Cell death assays.** The lactate dehydrogenase (LDH) release assay has been described previously (Kim et al., 2002). Briefly, LDH values were normalized by subtracting the background LDH released by control cells from treated cells and scaling to the signal triggered by complete cell death induced by 24 hours exposure to 30 mM of A23187.

**Caspase-3 activity assay.**

4 X 10^4 cells were plated in 24-well plates in triplicate per experiment. One day after plating, cells were transduced with indicated lentiviruses. Puromycin was added to the culture medium 24 hours after transduction and the cells were incubated for an additional 6 days. On day 7 after viral transduction, 200 µl of Caspase-Glo 3/7 Reagent (Promega G8091) was added to 200 µl of culture media per well. Plates were mixed on a plate shaker for 30 seconds then allowed to incubate at room temperature for 1 hour, and luminescence readings were acquired (TECAN Infinite M200 Pro).

**Immunocytochemistry.**

Polyornithine and laminin-coated German glass coverslips (Bello Glass, Inc) were seeded with 2 x 10^4 GSCs in a 24-well plate. Cells were fixed with 4% paraformaldehyde for 15 minutes at room temperature and subjected to immunofluorescence analysis with indicated antibodies according to standard protocols.

**Immunohistochemistry.**

At the indicated time post-injection, animals were sacrificed, and brains harvested for analysis. Brains were fixed in 4% paraformaldehyde and placed in 30% sucrose for
cryoprotection. 10 micron-thick frozen sections were generated using a cryostat. Frozen sections were fixed in ice-cold acetone and then rehydrated in PBS. Sections were blocked in PBS containing 4% normal goat serum (Vector laboratories) for 1 hour at room temperature. Double immunostaining for Nestin and GFAP or single immunostaining for GFP was then carried out overnight at 4°C, and sections were washed with PBS and incubated with appropriate secondary antibodies conjugated to fluorescent dyes (Alexa Fluor 568 Goat Anti-Mouse, Alexa Fluor 488 Goat Anti-Rabbit (Life Technologies) for 60 minutes at room temperature. After additional PBS washes, Hoechst 33342 or DAPI was added to sections to stain nuclei. Mounting medium was then added to the sections, and slides were cover-slipped. Fluorescent images of the sections were taken using an automated inverted microscope (Leica Microsystems, Cat. No. DMI4000 B). For whole coronal brain images, processed sections were imaged using the Zeiss ApoTome microscope (Histology and Microscopy Core, Washington University). Multichannel images of section parts were initially acquired using the blue, FITC, and rhodamine (Cy3) channels in multidimensional acquisition. The acquired images were then merged to obtain the image of the whole brain section.

**In vitro luciferase assay.**

Cells seeded on PLO and laminin-coated 24-well Primaria plates stably infected with pGreenfire1 hSRR2/mCMV-GFP T2A luciferase or control pGreenfire1 mCMV-GFP T2A luciferase lentivirus was transduced with indicated RNAi viruses and selected with puromycin. Seven days later, cells were visualized for GFP expression by live fluorescence microscopy and then assayed for luciferase activity using the One-Glo system (Promega). Luciferase signal was quantified by luminometry using a Tecan Infinite M200 Pro Microplate
Reader and divided by total protein levels. Data were normalized to the control condition (Scr + mCMV-GFP T2A luciferase infection = 1).

**Live bioluminescence imaging.**

GSCs stably expressing GFP T2A luciferase by lentiviral transduction were then infected with indicated lentiviruses and injected into the brains of NOD- SCIDγ mice as above. For bioluminescence imaging, animals were given 150 mg/mL D-luciferin (Gold Biotech) in PBS and imaged with a charge-coupled device (CCD) camera-based bioluminescence imaging system (IVIS Lumina; Caliper, Hopkinton, MA; exposure time 3-5 minutes, binning 8, field of view 12, f/stop 1, open filter) (Molecular Imaging Center, Washington University). Signal was displayed as photons/sec/cm²/sr (Gross and Piwnica-Worms, 2005).

**TCGA analysis.**

Glioblastoma clinical and Agilent G4502A gene expression microarray data was obtained for 483 glioblastoma tumors (including 10 normal brain tissue samples) from The Cancer Genome Atlas (TCGA) ([http://cancergenome.nih.gov/](http://cancergenome.nih.gov/)). The subtype classification as defined by Verhaak and colleagues for the TCGA data into four groups—Neural, Proneural, Mesenchymal and Classical—was obtained from [https://genome-cancer.ucsc.edu](https://genome-cancer.ucsc.edu) (Verhaak et al., 2010). 466 patients had valid survival data. CDC20 expression in glioblastomas was normalized to the mean of the normal brain samples. The overall survival of patients with *CDC20* expression level ≥2 times and <2 times that of normal brain samples was compared using Kaplan Meier survival curves, and the log-rank test was performed. This analysis was
performed on the data as a whole, as well as for each of the four molecular subtypes (IBM SPSS statistics (Version 21) software package). Maximally selected statistics on the derived linear predictor from a Cox proportional hazard model was used to identify an optimal OS cut-point for CDC20 expression while maintaining a minimum of 20% and a maximum of 80% of all patients in each group to avoid assigning too few patients to one group (Lausen et al., 2004). This analysis was performed on the data as a whole, as well as for individual molecular subtypes where indicated (IBM SPSS statistics (Version 21) software package). The TCGA Glioblastoma dataset has mutation data on 9658 genes in 291 patients. 15 IDH1 gene mutations (SNPs and missense) were identified in 15 independent patients. We restricted analysis to the 235 patients with available gene expression (Agilent G4502A gene expression microarray) and mutation data in the TCGA. Correlating subtype and IDH1 mutation data among these 235 patients, the Proneural subtype encompasses 49 patients in total with all of the IDH1 mutations (11 of 15) represented in the Proneural subtype as expected.
3.10 Supplemental References


Table 3.S1. Clinical characteristics of glioblastoma tumors used to derive GSCs, Related to Figure 3.1.

<table>
<thead>
<tr>
<th>GSC line</th>
<th>Pathology of tumor</th>
<th>Clinical status</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>GBM</td>
<td>Primary, right temporal lobe</td>
</tr>
<tr>
<td>B3</td>
<td>GBM</td>
<td>Primary, right temporo-parietal lobes</td>
</tr>
<tr>
<td>B18</td>
<td>GBM</td>
<td>Recurrent, right frontal lobe, (distant from primary left frontal lobe site)</td>
</tr>
<tr>
<td>B36</td>
<td>GBM with oligodendroglioma component</td>
<td>Primary, left temporal lobe</td>
</tr>
<tr>
<td>B49</td>
<td>GBM</td>
<td>Primary, right temporal</td>
</tr>
</tbody>
</table>
Figure 3. S1 Characteristics of human glioblastoma stem-like lines used in this study, Related to Figure 3.1.
(A) Human GSC lines were subjected to Nestin and SOX2 immunofluorescence. Percent immunopositive cells was calculated by dividing by total number nuclei by Hoechst 33342. Data represent mean±SEM (n = 3).

(B) Cell lysates from human GSC lines were subjected to immunoblot analysis using indicated antibodies.

(C) Human GSC lines were dissociated and subjected to CD133-PE labeling followed by flow cytometry. A representative plot is shown from two independent experiments for each line. The mean CD133+ percentage from two experiments is shown in the top right corner of each plot.

(D) Human GSC lines were subjected to the extreme limiting dilution assay to determine the frequency self-renewing cells. Data represent mean±SEM (n = 3).

(E) GSC lines were injected into the right putamen of NOD-SCID (B18) or NOD-SCIDγ (A1, B3, B36, B49) mice, and mice were sacrificed at 2-3 months. Brains were processed for Nestin and GFAP immunofluorescence to assess for tumor formation (Figure 1B and data not shown). All lines tested were capable of generating orthotopic xenograft tumors.

(F) Left: Mean normalized Illumina HT12v4 gene expression data for four Proneural and Mesenchymal signature GSC genes as described by Bhat et al. (2013) are shown for each GSC line from two independent experiments. Right: Expression for each gene was Z-corrected, and then the Proneural gene composite score was subtracted from the Mesenchymal gene composite score for each line and plotted as a heatmap to determine molecular subtype (Bhat et al., 2013).
Figure 3.S2. CDC20-APC is required for human GSC invasiveness in vitro, Related to Figure 3.1.
(A) GSCs were transduced with the indicated lentiviruses, and seven days later, cell lysates were subjected to immunoblot analysis using indicated antibodies. Whereas endogenous CDC20 is efficiently knocked down by CDC20i.2, GFP-CDC20-Res is resistant to CDC20 RNAi.

(B) A1 GSCs treated with the indicated RNAi lentiviruses (Scr = control scrambled SH002 virus) were subjected to the in vitro Matrigel transwell assay six days after infection to monitor invasiveness. Data represent mean+SEM (n = 3). CDC20 RNAi and ANAPC2 RNAi inhibited GSC invasiveness compared to control infection (ANOVA, P < 0.0001).

(C) B36 GSCs treated with the indicated RNAi lentiviruses were subjected to the in vitro Matrigel transwell assay six days after infection. Data represent mean+SEM (n = 4). CDC20 RNAi and ANAPC2 RNAi inhibited GSC invasiveness compared to control infection (ANOVA, P ≤ 0.001).

(D) A1 GSCs treated with GFP-tagged CDC20-expressing lentivirus or control vector virus (Vec) were subjected to the in vitro Matrigel transwell assay as in (B). Data represent mean+SEM (n = 4). CDC20 overexpression increased the invasive capacity of GSCs compared to control infection (unpaired t-test, P < 0.02).

(E) B36 GSCs treated with GFP-tagged CDC20-expressing lentivirus or control vector virus were subjected to the in vitro Matrigel transwell assay as in (B). Data represent mean+SEM (n = 3). CDC20 overexpression increased the invasive capacity of GSCs compared to control infection (unpaired t-test, P < 0.05).

(F) B18 GSCs were treated with APC inhibitor ProTAME (20 µM) or DMSO (Veh) for indicated times, and cell lysates were subjected to co-immunoprecipitation with anti-CDC20 or control IgG. Immunoprecipitates were resolved by SDS-PAGE and subjected to immunoblot analysis using indicated antibodies.

(G) A1 GSCs were subjected to the in vitro Matrigel transwell assay over 24 hours in the presence of indicated concentrations of ProTAME or DMSO (Veh). Data represent mean+SEM (n = 3). APC inhibitor ProTAME inhibited A1 GSC invasiveness in a dose-dependent manner (ANOVA, P < 0.0001).

(H) B3 GSCs were subjected to the in vitro Matrigel transwell assay over 24 hours in the presence of indicated concentrations of ProTAME or DMSO (Veh). Data represent mean+SEM (n = 3). APC inhibitor ProTAME inhibited B3 GSC invasiveness in a dose-dependent manner (ANOVA, P < 0.0001).
Figure 3.S3 The effect of molecular and pharmacological manipulations of CDC20-APC on cell cycle distribution, proliferation, cell death, and cell health, Related to Figure 3.1.
(A) GSCs transduced with CDC20 RNAi (CDC20i.2) or control scrambled (Scr) lentivirus were subjected to propidium iodide DNA labeling seven days later followed by flow cytometry to assess the distribution of cell cycle phases. Scr = control scrambled lentivirus. Data represent mean+SEM (n = 2). CDC20 knockdown did not significantly alter the cell cycle profile of GSCs compared to control infection. (ANOVA, P ≥ 0.222).

(B) GSCs infected with the indicated RNAi lentiviruses or control scrambled virus were subjected to the MTS cellular viability/proliferation assay seven days later. Data represent mean+SEM (n = 3). CDC20 RNAi did not significantly alter cellular proliferation compared to control infection (ANOVA, P = 0.717). ANAPC2 RNAi significantly decreased cellular proliferation compared to control infection (ANOVA, P = 0.004).

(C) GSCs transduced with GFP-tagged CDC20-expressing (CDC20) or control vector (Vec) lentivirus for five days were assayed as in (A). Data represent mean+SEM (n = 2). CDC20 overexpression did not significantly alter the cell cycle profile of GSCs compared to that of control infection. (ANOVA, P ≥ 0.284).

(D) GSCs transduced with GFP-tagged CDC20-expressing (CDC20) or control vector (Vec) lentivirus for five days were assayed as in (B). Data represent mean+SEM (n = 3). CDC20 overexpression did not significantly alter cellular viability/proliferation as monitored by the MTS assay compared to control infection. (ANOVA, P = 0.69).

(E) B18 GSCs treated as in (C) were plated in 24-well plates, and seven days later, the number of live cells was counted by Trypan blue exclusion. Data represent mean+SEM (n = 3). CDC20 overexpression did not significantly alter cell number compared to control infection (unpaired t-test, P = 0.74).

(F) A1 GSCs treated as in (C) were processed as in (E). Data represent mean+SEM (n = 3). CDC20 overexpression did not significantly alter cell proliferation compared to control infection (unpaired t-test, P = 1.0).

(G) GSCs were transduced with indicated RNAi lentiviruses or control scrambled (Scr) lentivirus, and live cell counts were quantified by Trypan blue exclusion six days later. Data represent mean+SEM (n = 3). CDC20 knockdown (CDC20i.1 and CDC20i.2) did not significantly alter cell survival in GSCs compared to control infection. (ANOVA, P ≥ 0.09). ANAPC2 knockdown did not significantly alter cell survival in GSCs compared to control infection. (ANOVA, P = 0.064).

(H) B18 GSCs were transduced with indicated RNAi lentiviruses or control scrambled (Scr) lentivirus, and caspase-3 activity was quantified by a fluorometric assay. The caspase-3 signal was divided by the total nuclear count (Hoechst 33342) and normalized to control infection (=1). Data represent mean+SEM (n = 3). CDC20 knockdown did not significantly alter caspase-3 activity in GSCs compared to control infection with a trend towards decreased caspase-3 activity. (ANOVA, P = 0.04). ANAPC2 knockdown did not significantly alter caspase-3 activity in GSCs compared to control infection. (ANOVA, P = 0.06).

(I) A1 GSCs were transduced with indicated RNAi lentiviruses or control scrambled (Scr) lentivirus, and caspase-3 activity was quantified by a fluorometric assay as in (H). Data represent mean+SEM (n = 3). CDC20 knockdown did not significantly alter caspase-3 activity in GSCs compared to control infection (ANOVA, P = 0.166). ANAPC2 knockdown significantly increased caspase-3 activity in GSCs compared to control infection. (ANOVA, P = 0.001).

(J) A1 GSCs were treated with the indicated concentrations of ProTAME for 48 hours and assessed for cell death by the LDH release assay. MG132 served as a positive control for cell death. Data represent mean+SEM (n = 3). 10 µM ProTAME, which was sufficient to inhibit invasion (Figure S2G), did not significantly increase cell death compared to DMSO (Veh) control
(ANOVA, P = 0.068). 20 μM ProTAME triggered modest levels of cell death at 48 hours compared to Veh (ANOVA, P = 0.0003).
Figure 3.S4. CDC20-APC controls SOX2 protein stability and SOX2-dependent transcription in GSCs, Related to Figure 3.3.
(A) A1 GSCs were transduced with CDC20 RNAi (CDC20i.1 and CDC20i.2) or control LacZ RNAi (Scr) lentivirus, and seven days later, cell lysates were subjected to immunoblotting using indicated antibodies.

(B) A1 GSCs were transduced with GFP-tagged CDC20-expressing lentiviruses or control GFP expressing virus (GFP), maintained in RHB-A media for five days, and cell lysates were subjected to immunoblotting using indicated antibodies. Similar results were obtained using control N103 vector virus (data not shown).

(C) A1 GSCs were transduced with ANAPC2 RNAi or control scrambled RNAi (Scr) lentivirus, and seven days later, cell lysates were subjected to immunoblotting using indicated antibodies.

(D) B18 GSCs were treated with 20 µM of ProTAME or DMSO (Veh) for indicated time points, and RNA was harvested from treated cells and reverse transcribed into cDNA. qPCR was performed on samples using specific primers for human SOX2. GAPDH and ACTB were used as reference genes. Data represent mean±SEM (n = 3). ProTAME treatment did not significantly alter SOX2 RNA levels (ANOVA, P = 0.128).

(E) A1 GSCs were treated with 10 µM of ProTAME or DMSO (Veh), and RNA was harvested from treated cells and reverse transcribed into cDNA. qPCR was performed as in (D). Data represent mean±SEM (n = 5). ProTAME treatment did not significantly alter SOX2 RNA levels (unpaired t-test, P = 0.97).

(F) B36 GSCs were transduced with CDC20 RNAi (CDC20i.2) or control SHC002 (Scr) lentiviruses for seven days and then treated with 10 µM of MG132 or DMSO (Veh) for 6 hours. Cell lysates were subjected to immunoblotting using indicated antibodies.

(G) Top: B18 GSCs stably infected with either control reporter (mCMV) or the SOX2 reporter (hSRR2) were monitored for reporter activity by visualizing GFP using live fluorescence microscopy. Scale bar = 50 µm. Bottom: B18 GSCs stably infected with either control mCMV or hSRR2 reporter were assessed for luciferase activity by luminometry, and luciferase values were normalized by total protein (left). COS-1 cells stably infected with either control mCMV or hSRR2 reporter were assessed for luciferase activity by luminometry, and luciferase values were normalized by total protein (right). Data represent mean±SEM (n = 3). GSCs exhibited significant SOX2 reporter luciferase activity above control mCMV reporter (unpaired t-test, P = 0.0006).

(H) B18 GSCs were treated with 20 µM of ProTAME or DMSO (Veh) for 4 hours, and RNA was processed as in (D). qPCR was performed by using primers specific for human Nestin. GAPDH and ACTB were used as reference genes. Data represent mean±SEM (n = 3). ProTAME decreased Nestin mRNA in GSCs compared to vehicle. (unpaired t-test, P = 0.002).

(I) B36 GSCs were transduced with SOX2 RNAi or control scrambled (Scr) lentiviruses and subjected to in vitro Matrigel invasion assays seven days later. Data represent mean±SEM (n = 5). SOX2 RNAi decreased GSC invasiveness compared to control infection (unpaired t-test, P < 0.0001).

(J) A1 GSCs were transduced with SOX2 RNAi or control scrambled (Scr) lentiviruses and subjected to in vitro Matrigel invasion assays seven days later. Data represent mean±SEM (n = 3). SOX2 RNAi decreased GSC invasiveness compared to control infection (unpaired t-test, P < 0.0009).

(K) A1 GSCs were transduced with SOX2 RNAi or control scrambled (Scr) lentiviruses along with CDC20-expressing or control vector (Vec) lentiviruses and subjected to in vitro Matrigel invasion assays seven days later. Data represent mean±SEM (n = 3). CDC20 overexpression +Scr increased GSC invasiveness compared to that of control infected cells, and SOX2 RNAi in the setting of CDC20 overexpression decreased invasiveness compared to that of CDC20 overexpression + Scr (ANOVA, both P = 0.013).
(L) B18 GSCs were transduced with SOX2 RNAi or control scrambled (Scr) lentivirus, and caspase-3 activity was quantified by a fluorometric assay as in Figure S3H. Data represent mean±SEM (n = 3). SOX2 knockdown did not significantly alter caspase-3 activity in GSCs compared to that of control infection (unpaired t-test, P = 0.07).

(M) A1 GSCs were transduced with SOX2 RNAi or control scrambled (Scr) lentivirus and caspase-3 activity was quantified by a fluorometric assay as in Figure S3H. Data represent mean±SEM (n = 3). SOX2 knockdown did not significantly alter caspase-3 activity in GSCs compared to that of control infection (unpaired t-test, P = 0.73).

(N) B18 GSCs transduced with SOX2 RNAi or control scrambled (Scr) lentiviruses were cotransduced with HA-tagged full-length SOX2-Res, SOX2-Res Δ110-200, or control vector viruses, cell lysates were subjected to immunoblot analysis using the indicated antibodies seven days later.
Figure 3.S5. Visualization of GSC tumorigenicity in vivo by immunofluorescence, Related to Figure 3.4.

B18 GSCs were infected with CDC20 RNAi (CDC20i.2) or control SHC002 (Scr) virus and selected with puromycin for 72 hours. 250,000 cells were injected per animal into the right putamen of NOD-SCID mice using a stereotactic apparatus. Three months after injection, animals were sacrificed. Brains were harvested, processed to generate 10 micron-thick frozen sections, and then subjected to immunohistochemistry using antibodies against Nestin and GFAP. Nuclei were stained with Hoechst 33342. Images of processed sections were obtained using fluorescence microscopy. LV = lateral ventricle. Scale bar = 100 µm.
Figure 3.6 The impact of CDC20 expression on overall survival in the TCGA Glioblastoma dataset, Related to Figure 3.5.
A) Kaplan-Meier curves showing overall survival (OS) of 466 newly diagnosed glioblastoma patients from the TCGA divided by normalized CDC20 expression from the Agilent microarray (compared to normal samples). Stratification of patients using an optimized CDC20 expression cut-point did not predict longer or short OS in the entire patient set (log-rank test, P = 0.202).

B) Kaplan-Meier curves showing OS of patients separated by Classical, Mesenchymal, Neural, and Proneural glioblastoma subtype tumors from the TCGA based on normalized CDC20 expression and stratification of patients into two groups using the same CDC20 expression cutpoint derived in (A). Consistent with the results using a 2-fold CDC20 expression cut-off (Figure 5), the patient group with higher CDC20 expression was associated with decreased OS specifically in patients with Proneural subtype tumors (log-rank test, P = 0.004).

C) Kaplan-Meier curves showing overall survival (OS) of all 49 Proneural glioblastoma patients with available Agilent microarray and mutation data (both IDH1 MUT and WT) (left) and the 38 IDH1 WT patients only (right) stratified by an optimal CDC20 expression cutoff identified by the derived linear predictor from a Cox model applied to Proneural patients only. When IDH1 MUT patients were included, patients with high CDC20 exhibited a worse prognosis (HR=3.62, 95% CI:1.35~9.69, log rank test P = 0.007), but the survival curves of patients with CDC20 high and low tumors were not significantly different when excluding IDH1 MUT patients (log rank test, P = 0.781).

D) Box plot (median and middle 50% of data represented in each box) of CDC20 expression in the 49 Proneural glioblastomas in the TCGA with available Agilent microarray and mutation data. IDH1 MUT tumors demonstrated significantly lower expression of CDC20 compared to that of IDH1 WT tumors in the Proneural subgroup (mean IDH1 WT/IDH1 MUT CDC20 mean = -1.38/-2.20, two sample t-test, P = 0.005, Wilcoxon rank rum test, P = 0.001)
Chapter 4:

The TRIM26/WWP2 Ubiquitin Pathway Regulates SOX2 Proteostasis and Stem Cell Maintenance in Glioblastoma
4.1 Summary:

Ubiquitin-dependent signaling is critical for the regulation of glioblastoma stem-like cells (GSCs), a subpopulation of glioblastoma cells responsible for tumor growth and therapy resistance. TRIM26, an E3-ubiquitin ligase with immune-related functions, is highly expressed in glioblastoma tumors compared to normal brain. Here, we found that TRIM26 directly binds to SOX2 via TRIM26’s C-terminal PRY-SPRY domain. Unexpectedly, we found that TRIM26 knockdown decreased SOX2 protein stability and conversely increased SOX2 polyubiquitination in primary GSCs. Accordingly, TRIM26 knockdown reduced SOX2 transcriptional activity, self-renewal, and in vivo tumorigenicity in genetically divergent GSC lines. Mechanistically, we discovered TRIM26 stabilizes SOX2 protein by competitively reducing the interaction of SOX2 with WWP2, a bonafide SOX2 E3 ligase in GSCs. Consistent with this hypothesis, WWP2 depletion in the setting of TRIM26 knockdown rescued SOX2 protein levels, self-renewal, and in vivo tumorigenicity in GSCs.

4.2 Introduction:
Despite advances in surgery and chemoradiation therapies, the prognosis of the brain cancer glioblastoma (GBM) remains challenging (Stupp et al., 2009; Stupp et al., 2005; Thakkar et al., 2014). Therapy resistance in glioblastoma has been attributed to the intrinsic heterogeneity of these tumors. In addition to molecular and genetic heterogeneity, GBM tumors also exhibit appreciable cellular heterogeneity. One particular cell type, which has generated significant research interest is the tumor-initiating cancer stem cell, also known as the glioblastoma stem cell (GSC) (Aum et al., 2014; Matchett and Lappin, 2014). GSCs are thought to be therapy resistant and drive tumor recurrence post-treatment (Bao et al., 2006; Chen et al., 2012b).

The core pluripotency-related transcription factor SRY (sex determining region Y)-box 2 (SOX2) is highly expressed in glioblastoma relative to the normal brain and has been shown to be indispensable for GSC self-renewal and tumorigenicity (Gangemi et al., 2009; Suva et al., 2014). However, there has been limited success in the development of direct small molecule and nucleotide-based SOX2 inhibitors that demonstrate promise for clinical application (Garros-Regulez et al., 2016; Taniguchi et al., 2017). We and others have shown that proteasomal regulation of SOX2 is critical to the maintenance and function of stem cells in both cancer and normal development (Cui et al., 2018; Mao et al., 2015; Wang et al., 2019). Given the recent renaissance in and significant progress of proteolysis-targeted chimera (PROTAC) technology towards clinical applications in oncology (Mullard, 2019), there is impetus for both deepening our understanding of the relevance of ubiquitin-dependent proteasomal regulation of SOX2 to GSC biology and for identifying novel regulators of this process.

Herein, we report that ambient proteasomal turnover of SOX2 protein is a shared feature of GSCs from divergent genetic backgrounds, rendering this phenomenon an attractive therapeutic target. We identify the E3 ubiquitin ligase TRIM26 as a SOX2-interacting protein and show that
TRIM26 protects SOX2 from polyubiquitination and proteasomal degradation. We further find that TRIM26 regulates GSC self-renewal in a SOX2-dependent manner and is required for GSC tumorigenicity. Finally, we report that TRIM26 outcompetes WWP2, a bona fide SOX2-targeting E3 ubiquitin ligase not previously described in glioblastoma, for binding to SOX2 resulting in non-catalytic inhibition of WWP2-mediated SOX2 polyubiquitination and subsequent proteasomal degradation.

4.3 Results:

We profiled a number of primary patient derived GSC lines by exome sequencing, microarray, and immunocytochemistry. The mesenchymal subtype cell line, B36, had canonical mutations in PTEN and TP53 (PTEN_del, TP53_R273C) and five predicted major genetic subclones (Fig 1A, S1). The proneural line, B67, had a canonical mutation in PIK3CA (E542K) and four major genetic subclones. However, both cell lines had >90% SOX2+ cells by ICC staining (Fig 1A, S1). Together, these data suggest that a high frequency of SOX2+ patient-derived GSCs is maintained independent of intratumoral genetic subclonality, intertumoral genetic heterogeneity, and intertumoral transcriptional subtype heterogeneity.

We next assessed the regulation of SOX2 proteostasis in GSCs. Treatment of two GSC lines with 10µM MG132, a proteasome inhibitor, resulted in increased SOX2 protein expression relative to cells treated with vehicle (DMS0) (Fig 1C). Conversely, treatment with cycloheximide to inhibit de novo protein synthesis resulted in a time-dependent decrease in SOX2 protein levels (Fig 1D). The degradative half-life of SOX2 was 10.7hrs in B67 GSCs and 31hrs in BT87 GSCs (Fig 1D). Having established the phenomenon of dynamic regulation of SOX2 protein stability in
GSCs, we next sought to identify potential regulators of this process. We prepared GSC protein lysates and subjected them to immunoprecipitation with either SOX2 or control IgG antibody. The IP samples were resolved on SDS-PAGE gel, digested, and submitted for mass spectrometry analysis. We identified 734 SOX2-interacting proteins (Fig 1E). Gene ontology analysis of the SOX2 interactors revealed proteasome and ribosome terms as the top two enriched biological processes, further suggesting the importance of dynamic proteostatic regulation of SOX2 in GSCs.

Next, we sought to specifically identify E3 ubiquitin ligases that interact with SOX2 in GSCs. We intersected the SOX2 interactors identified in the IP-Mass spectrometry experiment with a comprehensive list of predicted human E3 ubiquitin ligases (Li et al., 2008). We identified 10 predicted E3 ligases that interact with SOX2, of which 5 (TTC3, TRIM26, KLHL22, ITCH, and TRIP12) had established functions as degradative E3 ubiquitin ligases for other known substrates (Chen et al., 2018; Keppler and Archer, 2010; Marchese et al., 2003; Suizu et al., 2009; Wang et al., 2015). We focused on the TRIM26 E3 ubiquitin ligase, which has well-characterized immune-related ubiquitin-dependent functions (Ran et al., 2016; Wang et al., 2015). We next sought to validate the mass spectrometry results for the interaction between SOX2 and TRIM26. Epitope-tagged SOX2 and TRIM26 were found to exist in a complex by immunoprecipitation from transfected HEK293 cells (Fig 2A). Further, we found that endogenous SOX2 and TRIM26 proteins interact in two distinct GSC lines (Fig 2B).

To investigate the biochemical roles of TRIM26 in regulating SOX2 in GSCs, we subjected GSCs to shRNA-mediated TRIM26 knockdown. Unexpectedly, TRIM26 knockdown decreased SOX2 protein levels in two distinct GSC lines. In line with this observation, we found that TRIM26 knockdown decreased SOX2 transcriptional activity, as measured using an established human SOX2 regulatory region 2 enhancer (hSRR2) luciferase-based reporter system (Mao et al., 2015;
Sikorska et al., 2008) (Fig 2D). We then interrogated The Cancer Genome Atlas (TCGA) glioblastoma dataset and found that, similar to SOX2, TRIM26 mRNA is highly expressed in glioblastoma tumor specimens relative to the normal brain (Fig 2E). Further, we observed a significant correlation between TRIM26 and SOX2 mRNA expression levels in TCGA glioblastoma samples (Fig 2F). In total, these data suggest that TRIM26 is required for the maintenance of SOX2 expression and transcriptional activity in GSCs.

To begin to elucidate the biological consequences of TRIM26 regulation of SOX2 in GSCs, we performed bulk mRNA sequencing of GSCs following knockdown of TRIM26. Gene ontology analysis of the differentially downregulated genes between TRIM26 knockdown and control conditions revealed an enrichment of terms associated with nervous system development (Fig 3A). Given this observation, we next sought to examine the role of TRIM26 in regulating GSC self-renewal, a cancer stem cell phenotype that is associated with tumorigenic potential (Suva et al., 2014). We measured the frequency of self-renewing cells in control versus TRIM26 knockdown GSCs using the extreme limiting dilution assay and observed that TRIM26 knockdown reduced GSC self-renewal capacity in two different GSC lines (Fig 3B). Next, we asked whether TRIM26 functions through SOX2 to regulate GSC self-renewal. Overexpression of SOX2 in the setting of TRIM26 knockdown restored 50% of the self-renewal capacity of GSCs (Fig 3C), suggesting that SOX2 is downstream of TRIM26 in the regulation of GSC self-renewal and further that TRIM26 may have additional SOX2 independent roles in GSC maintenance. To assess the impact of TRIM26 modulation on GSC tumorigenicity, we used a GSC line transduced with a GFP T2A luciferase lentiviral plasmid to allow for bioluminescent live imaging (BLI) to monitor tumor burden in vivo. The GSC line was infected with TRIM26 shRNA constructs or control shRNA and injected into the right putamen of NOD-SCIDy mice. BLI over a period of 8 weeks demonstrated
that TRIM26 knockdown resulted in significantly diminished GSC tumorigenicity (Fig 3D,E). Further, mice injected with TRIM26 knockdown GSCs had an increased neurologic deficit-free survival compared to control GSCs (Fig 3F). We repeated the survival experiment with a second distinct GSC line and observed a concordant decrease in GSC in vivo tumorigenicity upon TRIM26 knockdown (Fig 3G). Taken together, these biologic functional data indicate that TRIM26 is required for GSC self-renewal and in vivo tumorigenicity, both key defining phenotypes of this cellular state.

Having established the importance of TRIM26 regulation of SOX2 to GSC biology, we next sought to delineate the mechanism by which this regulation occurs. First, we observed that the decrease in SOX2 protein expression upon TRIM26 knockdown was reversed by treatment with the proteasome inhibitor MG132 (Fig 4A). Second, we found that TRIM26 knockdown led to decreased expression of exogenous, EF1-alpha promoter-driven, epitope-tagged SOX2 protein. Next, we sought to examine the impact of TRIM26 modulation on SOX2 polyubiquitination. We generated a stable GSC line expressing FLAG-UBIQUITIN and SOX2-6x-HIS. The GSC line was then transduced with TRIM26 shRNA constructs or a non-targeting control, followed by immunoprecipitation of SOX2-6x-HIS using Nickel-NTA beads under denaturing conditions. We observed that TRIM26 knockdown resulted in a robust increase in SOX2 polyubiquitination in GSCs (Fig 4C). We next performed cycloheximide chase experiments and found that treatment of GSCs with TRIM26-targeting shRNA significantly increased the rate of degradation of SOX2 protein relative to non-targeting control (Fig 4D). Taken together, these observations indicate that TRIM26 regulates SOX2 at the post-transcriptional level, specifically by protecting SOX2 from polyubiquitination and subsequent proteasomal degradation.
To further delineate the mechanism by which TRIM26 regulates SOX2 protein in GSCs, we performed structure-function studies. We generated two TRIM26 ring-domain loss of catalytic function mutants, I18A and C16AC36A (Wang et al., 2015) that disrupt the zinc finger domain of TRIM26. We then introduced synonymous mutations at the RNAi binding sites of each of the three TRIM26 variants (WT, I18A, and C16AC36A) to generate rescue constructs insensitive to RNAi. Overexpression of each of these three rescue constructs in the setting of TRIM26 RNAi was sufficient to rescue the decrease in SOX2 protein levels observed upon TRIM26 knockdown (Fig 5A). Next, we asked whether TRIM26 directly binds SOX2 and what the structural determinants of this interaction may be. First, we performed GST pulldown experiments using GST-SOX2 recombinant fusion protein produced in bacteria and in vitro translated TRIM26 and found that SOX2 directly interacts with SOX2 (Fig 5B). Second, we used in vitro synthesized domain deletion mutants of TRIM26 in the GST pulldown paradigm and found that the Ring, B-box, and Coiled-coil (CC) domains of TRIM26 were dispensable for the interaction with GST-SOX2, thus localizing the necessary domains to the C-terminus (Fig 5C). Third, we further employed the GST pulldown paradigm to demonstrate that the C-terminal PRYSPRY domain of TRIM26 was sufficient for robust direct binding to SOX2 (Fig 5D). Having mapped the domain of TRIM26 sufficient for interacting with SOX2, we asked if this domain alone had a biological function in GSCs. Remarkably, overexpression of the epitope-tagged TRIM26 PRYSPRY domain in the setting of TRIM26 RNAi was sufficient to rescue the decrease in SOX2 protein levels observed upon TRIM26 knockdown (Fig 5E). Taken together, these data suggest that TRIM26 regulates SOX2 protein stability in a non-catalytic, ring-domain independent manner by directly interacting with SOX2 via the C-terminal PRYSPRY domain.
Next, we sought to identify the downstream SOX2-modifying E3 ubiquitin ligase that is inhibited by the binding of TRIM26 to SOX2. Since no SOX2-targeting E3 ubiquitin ligase has been identified in glioblastoma, we performed an RNAi-based mini-screen in GSCs of four E3 ubiquitin ligases that have been demonstrated to have SOX2-specific activity in different cellular contexts (Cui et al., 2018; Fang et al., 2014; Fukushima et al., 2013; Wang et al., 2019). As assessed by a consistent increase in ambient SOX2 protein levels upon knockdown of the E3 ubiquitin ligase, WWP2 was found to be a potential active SOX2-targeting E3 ubiquitin ligase in GSCs (Fig 6A). The increase in ambient SOX2 protein expression upon WWP2 knockdown was further confirmed in two additional GSC lines (Fig 6B). To further evaluate the biochemical role of WWP2 in regulating SOX2 in GSCs, we first performed immunoprecipitation experiments and found that endogenous WWP2 and SOX2 co-exist in a complex in two distinct GSC lines (Fig 6C). Second, we used a GSC line expressing Flag-UBIQUITIN and SOX2-6x-His to assess the changes in SOX2 polyubiquitination upon WWP2 knockdown by pulling down SOX2-6x-His from GSC lysates under denaturing conditions using nickel-NTA beads. We found that knockdown of WWP2 with two distinct shRNA constructs in GSCs led to a robust decrease in SOX2 polyubiquitination relative to a non-targeting control shRNA (Fig 6D). To test the specificity of the observed biochemical phenotype upon WWP2 knockdown, we generated a rescue mutant of WWP2 with synonymous mutations that abrogate binding to RNAi. We found that overexpression of the WWP2 rescue construct in the setting of WWP2 RNAi reversed the increase in SOX2 protein levels observed upon WWP2 knockdown back to near ambient levels (Fig S3). In sum, these biochemical observations indicate that WWP2 is an active E3 ubiquitin ligase in GSCs that targets SOX2 for polyubiquitination and subsequent proteasomal degradation.
Having determined that WWP2 is a SOX2-targeting E3 ubiquitin ligase, we next asked whether WWP2 plays a biological role in the maintenance of GSCs. Given the critical role of SOX2 in GSC self-renewal and tumorigenicity, we hypothesized that WWP2 knockdown may increase stem-like behavior of GSCs by increasing ambient SOX2 protein levels. We used the WWP2 shRNA construct that we previously rescued (Fig S3) for this set of functional experiments. First, we assayed the frequency of self-renewing cells in control versus WWP2 knockdown GSCs using the extreme limiting dilution assay and observed that WWP2 knockdown increased GSC self-renewal capacity in two different GSC lines (Fig 6E). Second, GSC line MGG8 was infected with WWP2 shRNA or non-targeting control shRNA and injected into the right putamen of NOD-SCIDy mice. We found that mice injected with WWP2 knockdown GSCs exhibited decreased neurologic deficit-free survival compared to control GSCs (Fig 6F). Overall, these biologic data suggest that WWP2 negatively regulates GSC self-renewal and in vivo tumorigenicity, potentially due to WWP2’s role as an E3 ubiquitin ligase that targets SOX2 protein for proteasomal degradation.

Next, we asked whether WWP2 acts downstream of TRIM26 to regulate SOX2 protein stability. We hypothesized, based on our structure-function studies, that TRIM26 inhibits WWP2-mediated polyubiquitination of SOX2 by outcompeting WWP2 for binding to SOX2. First, we transfected HEK293 cells separately with expression plasmids for HA-TRIM26, HA-WWP2, and GFP-SOX2. We mixed HA-WWP2 and GFP-SOX2 lysates in vitro, together with increasing amounts of HA-TRIM26 lysates. Following immunoprecipitation of GFP-SOX2, immunoblot analysis revealed that increasing amounts of HA-TRIM26 in the IP conditions resulted in decreased WWP2 co-immunoprecipitation with GFP-SOX2 (Fig 7A). Second, we performed in vitro ubiquitination experiments and found that addition of TRIM26 ring domain mutant (TRIM26
I18A) to the reaction blocked WWP2-mediated polyubiquitination of SOX2 (Fig 7B). Third, we examined the biochemical consequence of the TRIM26/WWP2 genetic interaction on SOX2 protein expression in GSCs. We found that knockdown of WWP2 in the setting of TRIM26 RNAi reversed the decrease in SOX2 protein that was observed upon TRIM26 knockdown in two distinct GSC lines (Fig 7C, D). In total, these observations indicate that WWP2 is a major downstream SOX2-specific E3 ubiquitin ligase that is inhibited by TRIM26 binding to SOX2, leading to maintenance of ambient SOX2 protein levels GSCs.

Having established the biochemical consequences of the TRIM26/WWP2 genetic interaction on SOX2 protein expression in GSCs, we asked whether this interaction was relevant to the regulation of GSC biology and function. First, we found that knockdown of WWP2 in the setting of TRIM26 RNAi significantly reversed the decrease in the frequency of self-renewing GSCs observed upon TRIM26 knockdown, as measured by the extreme limiting dilution assay (Fig 7E). Second, we used GSC line MGG8 to assess the role of the TRIM26/WWP2 ubiquitin signaling pathway on the tumorigenicity of GSCs. We injected RNAi-treated GSCs into the right putamen of NOD-SCIDy mice and monitored the mice for neurologic deficit-free survival. We found that mice injected with GSCs treated with WWP2 RNAi in addition to TRIM26 RNAi had significantly shorter survival than mice treated with TRIM26 RNAi alone. Remarkably, mice injected with the dual RNAi-treated GSCs (TRIM26i + WWP2i) had similar survival to those injected with the control RNAi-treated GSCs, thus demonstrating complete rescue of the reduction in GSC tumorigenic potential observed upon treatment with TRIM26 RNAi alone (Fig 7F). In summary, these results suggest that the E3 ubiquitin ligase WWP2 functions directly downstream of TRIM26 in regulating GSC self-renewal and in vivo tumorigenicity, due to the counter-regulatory activities of the two E3 ubiquitin ligases on SOX2 protein stability in GSCs.
4.4 Discussion:

In this study, we report that SOX2 is expressed by a majority of cells in GSC cultures from divergent genetic backgrounds and is subject to proteasomal regulation, thus presenting a potential shared targetable vulnerability. Additionally, we identify the E3 ubiquitin ligase TRIM26 as a novel regulator of SOX2 proteasomal degradation in GSCs and demonstrate that TRIM26 non-catalytically inhibits WWP2, a SOX2-targeting E3 ubiquitin ligase in GSCs, thus protecting SOX2 from polyubiquitination and subsequent proteasomal degradation. We find that modulation of TRIM26 and WWP2 expression levels impacts GSC self-renewal and in vivo tumorigenicity in xenotransplantation mouse models.

The core pluripotency-related transcription factor SRY (sex determining region Y)-box 2 (SOX2) has been extensively studied in glioblastoma. SOX2 has been reported to be highly expressed in a majority of glioblastoma biopsy specimens, with SOX2 expression levels correlating with increased tumor grade and aggressiveness (Garros-Regulez et al., 2016). Similar to SOX2’s role in normal tissue stem cell homeostasis, various studies have demonstrated that SOX2 expression and transcriptional activity is indispensable for the self-renewal and tumorigenicity of GSCs (Gangemi et al., 2009; Suva et al., 2014). Remarkably, it has been reported that while oncogenic receptor tyrosine kinase (RTK) signaling induces SOX2 activity during transformation in engineered glioblastoma mouse models, the maintenance of SOX2 expression and tumor-driving transcriptional networks is independent of aberrant RTK activity (Singh et al., 2017). Further, single-cell RNA sequencing analysis of IDH-wild-type glioblastomas revealed that
SOX2 expression is maintained, albeit at differing levels, across all four defined cellular states that exhibit substantial plasticity and associate with divergent oncogenic drivers (Neftel et al., 2019; Suva and Tirosh, 2020). Consistent with these studies, our profiling of low-passage primary patient-derived GSCs by immunocytochemistry, whole exome sequencing, and gene expression microarray revealed a high frequency of SOX2+ cells in all GSC cultures, independent of the divergence in both their genetic and dominant transcriptional subtype signatures. Taken together with SOX2’s critical role in GSC maintenance, the high-level expression of SOX2 across heterogenous glioblastomas makes SOX2 a potentially important therapeutic target.

We found that SOX2 protein stability is dynamically regulated across a number of primary patient-derived GSC samples harboring heterogenous genetic and transcriptional subtype backgrounds. In the context of cancer, and glioblastoma in particular, the functional biological consequences of altering SOX2 proteostasis have not been well established. We demonstrated, by RNAi-based combinatorial alteration of expression levels of the counter-regulatory TRIM26 and WWP2 E3 ubiquitin ligases, that modulation of SOX2 protein stability had significant impact on GSC self-renewal in vitro and tumorigenicity in vivo in xenotransplantation mouse models. These findings add to accumulating evidence in normal development that SOX2 protein stability is actively regulated by the ubiquitin proteasome system (UPS) in both embryonic and neural stem cells, with functional implications for stem cell maintenance and fate commitment (Cui et al., 2018; Fang et al., 2014). Further, our observations in glioblastoma lend urgent impetus for follow-up studies to elucidate additional regulators of SOX2 protein stability in GSCs and suggest that SOX2-specific UPS-based strategies represent a potentially viable avenue to therapeutically target GSCs.
Mechanistically, we found that TRIM26 was able to non-catalytically inhibit WWP2 by outcompeting WWP2 for binding to SOX2. To the best of our knowledge, this represents a novel description of direct and completely non-enzymatic inhibitory counter-regulation between two non-paralogous E3 ubiquitin ligases, adding a new level of complexity to mammalian ubiquitin signaling. It has been shown that the E3 ubiquitin ligase TRIM67 outcompetes TRIM9 for binding to the actin polymerase VASP and prevents TRIM9-mediated polyubiquitination and degradation of VASP, with consequences for filopodial dynamics and axon guidance (Boyer et al., 2020). TRIM67 and TRIM9 are paralogs, with identical domains and 63% sequence identity. In addition to direct binding competition, the regulation of TRIM9 by TRIM67 also required TRIM67 ligase activity by an as yet undefined mechanism. A second study described competition between paralogous SCF E3 ligase adaptor proteins FBXL3 and FBXL21 for ubiquitination and degradation of the circadian clock regulator CRY. While both SCF-FBXL3 and SCF-FBXL21 were able to ubiquitinate CRY, SCF-FBXL3 was found to be more potent at this activity, such that SCF-FBXL21-bound CRY in the nucleus was degraded at a slower rate and thus protected compared to SCF-FBXL21-bound CRY (Yoo et al., 2013). However, SCF-FBXL21 was the major degrader of CRY in the cytoplasm. Our study, by contrast, involves direct inhibitory competition between two non-paralogous E3 ubiquitin ligases that share essentially no sequence homology and have opposite roles in regulating SOX2 ubiquitination. TRIM26 is a RING-domain E3 ligase, while WWP2 is a HECT-type E3 ligase. Further, our data revealed that the inhibitory interaction between TRIM26 and WWP2 could be fully attributed to competitive binding to SOX2 and did not require TRIM26’s ubiquitin ligase activity. Further, TRIM26 demonstrated no E3 ubiquitin ligase activity towards SOX2 and has a primary function of inhibiting SOX2 ubiquitination.
Our structure-function analysis of TRIM26’s direct binding to SOX2 identified the C-terminal PRYSPRY domain of TRIM26 as an avid SOX2 interactor with cellular consequences for SOX2 protein stability. We anticipate building on our structure-function findings with future crystallographic and finer interaction mapping studies to define the minimal amino acid residues of the TRIM26 PRYSPRY domain that are essential for SOX2 binding. This information, coupled with single cell RNA sequencing data to identify ubiquitous highly expressed E3 ubiquitin ligases in GSCs, can be leveraged to create novel peptide-based SOX2 proteolysis-targeted chimeras (PROTACs) for potential use in combination therapies for glioblastoma. There has been significant progress recently in the development of small molecule PROTAC technology, with small-molecule-based PROTACS specific for the androgen and estrogen receptors already in human clinical trials (Mullard, 2019; Wang et al., 2020). Additionally, inroads have been made in developing peptide-based PROTACS by leveraging variable chemical approaches to develop stabilized and cell-permeable peptides. A heterobifunctional peptide PROTAC was recently shown to have significant \textit{in vivo} efficacy for targeted estrogen receptor alpha degradation in a breast cancer preclinical model (Jiang et al., 2018). Together with our findings, these developments provide rationale and are critical to the innovation of efficacious peptide-based PROTACs to target SOX2 for degradation, given that SOX2 does not have well-defined small-molecule binding domains.

While our results delineate a clear and biologically important role for TRIM26 in regulating SOX2 protein stability in GSCs, whether TRIM26 has other SOX2-dependent or independent functions in GSC biology remains an open question. Indeed, re-expression of exogenous SOX2 to rescue the decrease in SOX2 protein levels upon TRIM26 knockdown did not fully restore GSC self-renewal capacity. This observation suggests that either 1) TRIM26 has additional SOX2-
independent functions that support GSC self-renewal or 2) SOX2 transcriptional activity is partially impaired in the absence of TRIM26. Further experiments with SOX2 mutants that are transcriptionally active but insensitive to proteasomal degradation will be required to address this question. It will also be interesting to elucidate the upstream signaling pathways that lead to increased TRIM26 expression in glioblastoma compared to the normal brain, including the potential role of interferon signaling as has been described in other cellular contexts.

4.5 Methods:

Cell Culture:

The generation of adherent human GSC cultures has been described (Pollard et al., 2009). In brief, surgical tumor samples were dissociated mechanically and by incubation in Accutase (SIGMA) for 20-60 minutes at 37°C. Cell suspensions were passed through a 70 micron cell strainer (Falcon) and plated using NeuroCult NS-A Basal Medium (Human) supplemented with N2 supplement, B27 supplement, Glutamax, 75µg/ml BSA, and EGF and FGF-2 (Peprotech) (hereafter “GSC media”), each at 20 ng/ml, on polyornithine and laminin (SIGMA)-coated Primaria dishes/flasks (BD Bioscience). Growth factors (EGF/FGF-2) were replenished every 2-3 days. Cells were routinely used between passages 5 and 20. Informed consent was obtained from patients for use of human tissue and cells, and all human tissue-related protocols used in this study were approved by the Institutional Review Board (Washington University). BT87 GSCs were a kind gift from Dr. Sunit Das at University of Toronto and were cultured as described above. MGG8 GSCs were the kind gift of Dr. Daniel Cahill at the Massachusetts General Hospital. MGG8 GSCs
were cultured as non-adherent spheres in ultralow attachment culture dishes (Corning) using Neurobasal Media supplemented with N2 supplement, B27 supplement, sodium pyruvate, Non-essential amino acids (NEAA), and EGF and FGF-2 at 20ng/ml each. Human embryonic kidney 293 (HEK293) cells were cultured in Dulbecco's modified Eagle's medium with 10% fetal bovine serum (FBS) and penicillin/streptomycin (Life Technologies). All cell lines were incubated at 37°C with 5% CO2. Lentiviral transduction was performed by adding virus with 4 µg/mL of polybrene for 4 hours to cells. For rescue and epistasis experiments, GSCs were transduced with TRIM26 or WWP2 RNAi or control lentivirus one day after plating and then transduced with TRIM26, WWP2, or SOX2 expression virus or control virus on the second day after the initial transduction. Cells were selected in 2 µg/mL of puromycin 1-2 days after infection. For self-renewal and in vivo tumorigenicity experiments, GSCs were utilized 5 days following indicated viral infections.

**Extreme limiting dilution analysis**

Cells were plated at five-fold dilutions (3000, 600, 120, 24, 5 or 1 cell/well) in Corning ultra-low attachment 96-well plates. 14 days later, the number of wells containing spheres was counted and used to calculate the frequency of self-renewing GSCs by online software ([http://bioinf.wehi.edu.au/software/elda/](http://bioinf.wehi.edu.au/software/elda/)) (Hu and Smyth, 2009; Singh et al., 2004).

**Xenotransplantation**

Animals were used in accordance with a protocol approved by the Animal Studies Committee of the Washington University School of Medicine per the recommendations of the Guide for the Care and Use of Laboratory Animals (NIH). 250,000 cells per animal (BT87 GSCs TRIM26 RNAi experiment), 10,000 cells per animal (MGG8 GSCs TRIM26 RNAi experiment),
or 50,000 cells per animal (MGG8 GSCs WWP2 RNAi experiment and MGG8 TRIM26/WWP2 RNAi epistasis experiment) were injected stereotactically into the right putamen of approximately 6-week-old female NOD-SCIDγ mice. The coordinates used were: 1 mm rostral to bregma, 2 mm lateral, and 2.5 mm deep.

**Live bioluminescence imaging**

BT87 GSCs stably expressing GFP T2A luciferase by lentiviral transduction were infected with indicated lentiviruses and injected into the brains of NOD-SCIDγ mice as above. For bioluminescence imaging, animals were anesthetized with 2.5% (vol/vol) isoflurane, injected i.p. with 150 µg/mL D-luciferin (Biosynth) in PBS, and imaged with a charge-coupled device (CCD) camera-based bioluminescence imaging system (IVIS 100; Perkin-Elmer; exposure time 3–5 min, binning 16, field of view 12, f/stop 1, open filter; Molecular Imaging Center, Washington University). Signal was displayed as photons per s/cm² per steradian.

**Transient transfection**

Polyethyleneimine (PEI, Polysciences #24765-2) was dissolved in ddH2O to a concentration of 1mg/mL. The solution was then adjusted to pH 7.0 and filter sterilized. A mixture of plasmid DNA and PEI solution was made in OPTIMEM (Life Technologies) and incubated at room temperature for 15 minutes. DNA/PEI complexes were applied to cells, and media was changed after 12-16 hours. Experiments were performed 36 hours following transfection.

**Lentiviral production**
HEK293 cells were plated with a goal density of 70-80% after 1 day. The next day, transfection was performed using the PEI transfection method to introduce the plasmid of interest along with packaging plasmid psPAX2 and envelope plasmid pCMV-VSVG to Opti-MEM (Life Technologies). On day 6, medium from the plates was collected and spun down at 1200 x g for 5 min at 4°C. Supernatant was filtered through 0.45micron filters. Lenti-X Concentrator (Clontech) was then added to the filtrate and mixed, and the tubes were incubated at 4°C for 6-7 hours. Lentiviruses were then centrifuged at 1500 x g for 45 minutes at 4°C. Supernatant was aspirated; pellets were re-suspended in one tenth of the original medium volume of cold PBS, and stored at -80°C in aliquots. Viral copy number was adjusted for transduction of GSCs on the basis of titer measured using the Lenti-X qRT-PCR titration kit (Clontech).

Co-immunoprecipitation and immunoblot analysis

Cells were lysed in 1% NP-40 lysis buffer containing 20 mM Tris [pH 8], 200 mM NaCl, 10% glycerol, 1 mM EDTA, 10 mM NaF, 1mM sodium orthovanadate, and a protease inhibitor cocktail (Calbiochem). Clarified lysates were precleared with protein A/G-sepharose (Life Technologies), and immunoprecipitations were performed overnight at 4° with indicated antibodies followed by protein A/G-sepharose beads. Beads were washed 6 times with lysis buffer and boiled in Laemmli sample buffer. Samples with were separated by SDS-PAGE and transferred to 0.45 µm Immobilon-P PVDF membrane (EMD Millipore). Membranes were blocked in 5% Milk in Tris-buffered saline with Tween-20 (TBST) at room temperature and incubated with primary antibodies at 4°C overnight or at room temperature for 4 hours. Membranes were washed with TBST and incubated with appropriate horseradish peroxidase-conjugated secondary
antibodies for 1 hour at room temperature. Membranes were then washed with TBST and developed using Pierce ECL western blotting substrate (Thermo Scientific).

**Cellular ubiquitination assays**

GSCs were transduced with PCDH-CMV-Flag-Ubiquitin and pLV-EF1a-SOX2-6X-Histidine plasmids to stably express Flag-Ubiquitin and C-terminal 6X-His tagged human SOX2 (SOX2-6X-His). GSCs were then transduced with either of 2 TRIM26- or WWP2-targeting RNAi constructs or the respective controls. On day 5 after RNAi transduction, cells were treated with 20µM MG132 for 8 hours. Cells were washed with cold PBS and detached from culture dishes using Accutase (sigma). 10% of the cell pellet (for input) was lysed in 1% NP-40 lysis buffer containing 20 mM Tris [pH 8], 200 mM NaCl, 10% glycerol, 1 mM EDTA, 10 mM NaF, 1mM sodium orthovanadate, and a protease inhibitor cocktail (Calbiochem). Input lysates were clarified by centrifugation, subjected to Bradford assay to quantitate protein concentrations, and boiled in Laemmlli sample buffer for immunoblot analysis. The rest of the cell pellet (90%) was resuspended in guanidium lysis buffer (6M Guadinium-HCl, 0.1M Na$_2$HPO$_4$/NaH$_2$PO$_4$, 10mM Tris-HCl (pH 8), 0.005M imidazole, 0.01M -mercaptoethanol, and EDTA-free protease inhibitor cocktail). The cells were sonicated, centrifuged, and the supernatant was transferred to a 15ml conical containing 4ml of guanidium lysis buffer. Immunoprecipitation was performed using Ni$^{2+}$-NTA-agarose beads for 4 hours at room temperature. The agarose beads were then washed once with guanidium lysis buffer, once with Urea wash buffer (8M Urea, 0.1M Na$_2$HPO$_4$/NaH$_2$PO$_4$, 10mM Tris-HCl (pH 6.8), 0.005M Imidazole, and 0.01M -mercaptoethanol), twice with urea wash buffer containing 0.1% Triton-X 100, and twice with His-Protein Lysis buffer :50mM Tris (pH 7.3), 250mM NaCl, 0.05% Triton X-100, 20mM Imidazole, and 3-5 mM 2-mercaptoethanol. Bound
protein was eluted from beads using His-Protein Lysis Elution Buffer: 50mM Tris (pH 7.3), 250mM NaCl, 0.05% Triton X-100, 400mM Imidazole (add fresh: 3-5mM- 2-mercaptoethanol). Laemmli sample buffer was added and the eluate was boiled and subsequently analyzed by immunoblot.

**In vitro ubiquitination assays**

Radiolabeled substrate, $^{35}$S-myc-SOX2, was in vitro translated from pcdna3-myc-hSOX2 plasmid template using the TNT-coupled reticulocyte lysate system (Promega). SOX2 from this reaction was purified by immunoprecipitation with SOX2 antibody using the same protocol as above for co-immunoprecipitation experiments. HEK293 cells were transfected with empty vector (pcdna3) or with pcdna-HA-TRIM26-I18A-mut or HA-WWP2. 36 hours post-transfection, lysates from HEK293 cells were clarified and subjected to immunoprecipitation with HA-tag antibody. Immunoprecipitated HA-TRIM26-I18A and empty-vector (pcdna3) samples were peptide-eluted from IP beads by incubation with 5-fold excess (by weight) HA-peptide for 2 hours at room temperature in PBS. $^{35}$S-myc-SOX2 was preincubated with peptide eluted HA-TRIM26-I18A or eluted control (empty pcdna3) for 16 hours at 4 degrees. Control beads were added to the $^{35}$S-myc-SOX2/Contol-pcdna3 preincubated sample for the negative ubiquitination reaction control. HA-WWP2 IP beads were added to the $^{35}$S-myc-SOX2/Contol-pcdna3 and $^{35}$S-myc-SOX2/HA-TRIM26-I18A preincubated samples. 100µM ubiquitin, 100nM E1 enzyme, 1µM UBE2D3 E2 enzyme, and 1x E3 ligase reaction buffer (Boston biochem/R&D Systems) were added to each of the reactions. The ubiquitination reaction was initiated by addition of 100µM MgATP. The reaction was incubated at 37°C for 60 minutes with gentle agitation. The reaction was terminated by addition of Laemmelli sample buffer and boiling. Samples with were separated by SDS-PAGE.
using 4-12% gradient gels (Invitrogen). The gel was fixed and dried and subsequently analyzed by autoradiography to detect unmodified and polyubiquitinated $^{35}$S-myc-SOX2.

**In vitro binding**

GST-human SOX2 and the corresponding GST control proteins were bacterially expressed and isolated using Glutathione Sepharose 4 Fast Flow (GE Healthcare) according to manufacturer's instructions. Using pcDNA3-HA TRIM26 full length and deletion mutant plasmids, $^{35}$S-labeled products were *in vitro* translated using the TNT-coupled reticulocyte lysate system (Promega) and incubated with indicated GST-fusion proteins bound to glutathione-sepharose 4 Fast Flow beads in 1% NP-40 lysis buffer at 4°C for 16 hours (GE Healthcare). The beads were washed 6 times with lysis buffer and boiled in Laemmli sample buffer. Proteins were resolved by SDS-PAGE, and $^{35}$S-labeled proteins were visualized by autoradiography. GST-fusion proteins were assessed by Coomassie Brilliant Blue R-250 (Biorad).

**Immunocytochemistry**

Polyornithine and laminin-coated German glass coverslips (BelloGlass, Inc) were seeded with 2 x $10^4$ GSCs in a 24-well plate. Cells were fixed with 4% paraformaldehyde for 15 minutes at room temperature and subjected to immunofluorescence analysis with indicated antibodies according to standard protocols.

**In vitro luciferase assay**
Cells seeded on PLO and laminin-coated -96 well plates stably infected with SOX2 reporter pGreenfire1 2X-hSRR2/mCMV-GFP-T2A-Firefly-luciferase and control pLV-EF1α-Renilla-Luciferase lentivirus was transduced with indicated RNAi viruses and selected with puromycin. Six days later, cells were assayed for luciferase activity using the One-Glo Dual Luciferase reporter system (Promega). Luciferase signal was quantified by luminometry using a Biotek Cytaion 5 Multiwell plate reader. Firefly luciferase signal from SOX2 reporter activity was divided by the corresponding Renilla Luciferase reading to control for cell number. Data were normalized to the corresponding control RNAi condition.

**RNA Sequencing, Data Acquisition, Quality Control, and Processing**

GSCs were infected with either of 2 TRIM26-targeting RNAi constructs or a non-targeting control (3 biological replicates per RNAi construct). Cells were harvested on Day 4 after transduction and total RNA was isolated using the RNeasy Plus Mini Kit (Qiagen). Ribosomal RNA was removed by a hybridization method using Ribo-ZERO kits (Illumina). mRNA was then fragmented and reverse transcribed to yield double-stranded cDNA. cDNA was blunt ended, had an A base added to the 3’ ends, and then had Illumina sequencing adapters ligated to the ends. Ligated fragments were then amplified for 12 cycles using primers incorporating unique index tags. Fragments were sequenced on an Illumina HiSeq-2500 using single reads extending 50 bases. RNA-seq reads were aligned to the Ensembl release 76 assembly with STAR version 2.0.4b. Gene counts were derived from the number of uniquely aligned unambiguous reads by Subread:featureCount version 1.4.5. Transcript counts were produced by Sailfish version 0.6.3. Sequencing performance was assessed for total number of aligned reads, total number of uniquely aligned reads, genes and transcripts detected, ribosomal fraction known junction saturation, and
read distribution over known gene models with RSeQC version 2.3. All gene-level and transcript counts were then imported into the R/Bioconductor package EdgeR and TMMinormalization size factors were calculated to adjust for samples for differences in library size. Genes or transcripts not expressed in any sample were excluded from further analysis. The TMM size factors and the matrix of counts were then imported into R/Bioconductor package limma, and weighted likelihoods based on the observed mean–variance relationship of every gene/transcript were then calculated for all samples with the Voom function. Performance of the samples was assessed with a Spearman correlation matrix and multidimensional scaling plots. Gene/transcript performance was assessed with plots of residual SD of every gene to their average log-count with a robustly fitted trend line of the residuals. Generalized linear models with robust dispersion estimates were then created to test for gene/transcript level differential expression. Differentially expressed genes and transcripts were then filtered for FDR adjusted P ≤ 0.05. Gene Ontology analysis was performed by using the Enrichr gene analysis tool (https://amp.pharm.mssm.edu/Enrichr/)(Chen et al., 2013; Kuleshov et al., 2016). Transcription factor enrichment analysis was performed using the X2Kweb platform (https://amp.pharm.mssm.edu/X2K/)(Chen et al., 2012a; Clarke et al., 2018).

Statistics

All images are representative of results from 3 independent experiments unless otherwise stated. Statistical analyses were performed with Prism Version 8. The unpaired Student’s t-test was used for comparisons in experiments with only two groups. In experiments with more than two comparison groups, analysis of variance (ANOVA) was performed followed by the Bonferroni test for pairwise comparisons.
4.6 References:


Figure 4.1: SOX2 is widely expressed among intra- and intertumoral GSC subclones and is subject to proteasomal regulation.

A. GSCs were subjected to immunofluorescence with SOX2 antibody and Hoechst nuclear stain (left). Exome sequencing data for GSC lines was used to infer global copy number alterations (right).

B. GSCs were treated with vehicle (DMSO) or 10µM proteasome inhibitor MG132 for 8 hours. Protein lysates were harvested and subjected to immunoblotting with the indicated antibodies.

C. Lysates from GSCs treated with 100µM cycloheximide for the indicated durations were subjected to immunoblotting with SOX2 and Alpha-tubulin antibodies.

D. Quantification of protein expression levels for the cycloheximide time course experiments in D for B67 and BT87 GSC lines.

E. Experimental workflow schematic for SOX2 immunoprecipitation followed by mass spectrometry (SOX2 IP-MS) (top). Gene ontology analysis for proteins unique to the SOX2 IP condition was performed using the Enrichr online platform (bottom).
Figure 4.2: TRIM26 maintains SOX2 expression and transcriptional activity in GSCs.

A. HEK293 cells were transfected with the indicated DNA plasmids. Protein lysates were harvested 36 hours post-transfection and subjected to anti-GFP immunoprecipitation (IP) followed by immunoblotting (WB) with the indicated antibodies.
B. GSC protein lysates were subjected to immunoprecipitation with SOX2 IgG antibody or control IgG. Input lysates and IP samples were subjected to immunoblotting with the indicated antibodies.

C. GSCs were transduced with either of 2 lentiviral shRNA constructs targeting TRIM26 or a non-targeting control shRNA. Protein lysates were harvested on day 5 post-transduction and analyzed by immunoblotting with the indicated antibodies.

D. GSCs were transduced with a firefly luciferase-based SOX2 reporter lentiviral plasmid as well as a lentiviral plasmid for constitutive expression of renilla luciferase. The stable lines produced were subsequently treated with either of the 2 TRIM26-targeting shRNA constructs or a non-targeting control shRNA. On day 5 after shRNA transduction, both firefly and renilla luciferase activities were measured using a dual-luciferase assay. The firefly luciferase activity was divided by the renilla luciferase baseline to obtain normalized SOX2 reporter activity. Statistical analysis was performed using ordinary one-way ANOVA with Bonferroni correction for multiple comparisons.

E. Box plots (median and middle 50% of data represented in each box) for TRIM26 mRNA expression in TCGA glioblastoma (n = 156) and normal brain tissue (n = 4). TRIM26 expression is higher in glioblastoma specimens compared to normal brain tissue (pairwise t test, p = 0.0096).

F. Correlation analysis for TRIM26 and SOX2 mRNA in TCGA glioblastoma (n = 152). TRIM26 mRNA expression is positively correlated with SOX2 mRNA in glioblastoma specimens (Pearson r = 0.58, p = 2.63 x 10^{-19})
**Figure 4.3: TRIM26 knockdown decreases GSC self-renewal and in vivo tumorigenicity.**

A. GSCs were transduced with either of 2 TRIM26-targeting shRNA constructs or a non-targeting control. Total mRNA was harvested on day 4 post-transduction and submitted for bulk mRNA sequencing. Common differentially expressed genes between each of the two TRIM26 shRNA constructs compared to the non-targeting control were subjected to gene ontology analysis using the Enrichr platform.

B. Following transduction with TRIM26 shRNA or a non-targeting control, GSCs were subjected to the neurosphere extreme limiting dilution assay (ELDA) assay to determine the stem-cell frequency for each condition. Statistical analysis was performed using the ordinary one-way ANOVA test with Bonferroni correction for multiple comparisons.

C. GSCs were transduced with a TRIM26-targeting shRNA construct or a non-targeting control. On day 2 post-shRNA transduction, the GSCs were further transduced with a SOX2 overexpression plasmid or empty vector control. GSCs were subsequently subjected to the neurosphere formation extreme limiting dilution assay to determine stem-cell frequency. Statistical analysis was performed using ordinary one-way ANOVA with Bonferroni correction for multiple comparisons.

D. BT87 GSCs stably expressing firefly luciferase were transduced with either of 2 TRIM26-targeting shRNA constructs or a non-targeting control. 6 days post-transduction, the GSCs were stereotactically injected into the right putamen of immunocompromised mice. Bioluminescent live imaging was performed at the indicated time points after injection. Four animals per group are shown for the 8 week time point.

E. Quantification of bioluminescence live imaging for the animals treated as in (E). Data represent mean ± SEM. Statistical analysis was performed using ordinary one-way ANOVA with Bonferroni correction for multiple comparisons (n = 6, p = 0.006 for Control vs TRIM26i.1, and 0.0042 for Control vs TRIM26i.2 at the 8 week timepoint).

F. Kaplan-Meier curves showing neurological deficit-free survival of mice treated as in (E). Curve comparisons were performed using the Log-Rank (Mantel-Cox) test (n = 6, p = 0.0008).

G. MGG8 GSCs stably expressing firefly luciferase were transduced with either of 2 TRIM26-targeting shRNA constructs or a non-targeting control. 6 days post-transduction, the GSCs were stereotactically injected into the right putamen of immunocompromised mice. Kaplan-Meier curves showing neurological deficit-free survival of mice treated as in (E). Curve comparisons were performed using the Log-Rank (Mantel-Cox) test (n = 6, p = 0.0027).
Figure 4.4: TRIM26 protects SOX2 from polyubiquitination and subsequent proteasomal degradation.

A. GSCs were transduced with a TRIM26-targeting shRNA construct or a non-targeting control. On day 6 post-transduction, GSCs were treated with vehicle (DMSO) or 10µM proteasome inhibitor MG132 for 8 hours. Lysates were harvested and subjected to immunoblotting with the indicated antibodies.
B. GSCs stably expressing exogenous, EF1-alpha promoter-driven, HA-SOX2 were transduced with either of 2 TRIM26-targeting shRNA constructs or a non-targeting control. On day 6 post-transduction, protein lysates were harvested and immunoblotted with the indicated antibodies.

C. GSCs were first transduced with lentiviral plasmids for stable expression of EF1-alpha-driven C-terminal 6X-His tagged SOX2 and CMV-driven Flag-ubiquitin. The GSCs were subsequently transduced with either of 2 TRIM26-targeting shRNA constructs or a non-targeting control. Lysates were harvested and subjected to immunoprecipitation (IP) under denaturing conditions using nickel-NTA beads. IP and input samples were subjected to immunoblotting using the indicated antibodies.

D. GSCs were transduced with a TRIM26-targeting shRNA construct or a non-targeting control. On day 6 post-transduction, GSCs were treated with 100µM cycloheximide for the indicated time points. Lysates were harvested and subjected to immunoblotting with the indicated antibodies (top). Protein amounts on immunoblots were quantified by pixel intensity measurement using ImageJ software and SOX2 levels normalized to the respective tubulin loading control were calculated. Normalized SOX2 protein levels relative to the untreated control (time 0) were calculated and plotted for each of the 2 conditions (bottom) (n = 3). Statistical analysis was performed using unpaired t test at each timepoint (*p = 0.047).
Figure 4.5: TRIM26 regulates SOX2 protein expression non-catalytically via TRIM26’s C-terminal PRYSPRY domain.

A. GSCs were transduced with a TRIM26-targeting shRNA construct or a non-targeting control. On day 2 post-transduction, the cells were further transduced with either empty vector, wildtype (WT) TRIM26, or ring-domain mutant TRIM26 (I18A and C16AC36A) lentiviral plasmids. Lysates were harvested on day 4 after the second round of transductions and subjected to immunoblotting using the indicated antibodies.

B. In-vitro synthesized, radioactively-labelled TRIM26 (\(^{35}\)S) was incubated with GST-control or GST-tagged SOX2 produced in E-coli. The samples were subjected to
immunoprecipitation using glutathione beads, followed by gel electrophoresis. Radioactively-labeled proteins in the IP and input samples were visualized using autoradiography. Input GST and GST-SOX2 were visualized by Coomassie (CBB) staining.

C. Radioactively-labelled full-length (FL) TRIM26 and TRIM26 domain deletion mutants (1-138, 1-223, and 1-363) were synthesized in vitro and incubated with GST-control or GST-SOX2 produced in E-coli. The samples were subjected to immunoprecipitation using glutathione beads, followed by gel electrophoresis. Radioactively-labeled proteins in the IP and input samples were visualized using autoradiography. Input GST and GST-SOX2 were visualized by Coomassie (CBB) staining. An illustration of TRIM26 domains is depicted in the rightmost panel.

D. Radioactively-labelled full-length (FL) TRIM26, TRIM26 1-363, and TRIM26 PRY-SPRY domain were synthesized in vitro and incubated with GST-control or GST-SOX2 produced in E-coli. The samples were subjected to immunoprecipitation using glutathione beads, followed by gel electrophoresis. Radioactively-labeled proteins in the IP and input samples were visualized using autoradiography. Input GST and GST-SOX2 were visualized by Coomassie (CBB) staining.

E. GSCs were transduced with either of 2 lentiviral shRNA constructs targeting TRIM26 or a non-targeting control shRNA. On day 2 post-transduction, the GSCs were further transduced with either empty vector control or TRIM26-GFP-PRYSPRY expression plasmid. Protein lysates were harvested on day 4 following the second round of transductions and subjected to immunoblotting with the indicated antibodies.
Figure 4.6: WWP2 is an active SOX2 E3 ubiquitin ligase in glioblastoma stem cells.

A. B67 GSCs were transduced with either a control non-targeting shRNA construct or 2 unique shRNA constructs targeting each of CDH1, CUL4A, UBR5, and WWP2. Protein
lysates were harvested on day 6 post-transduction and subjected to immunoblotting with the indicated antibodies.

B. B36 and BT87 GSCs were transduced with either of 2 WWP2-targeting shRNA constructs or a non-targeting shRNA construct. Protein lysates were harvested on day 6 post-transduction and subjected to immunoblotting with the indicated antibodies.

C. B51 and B67 GSC protein lysates were subjected to immunoprecipitation with SOX2 IgG antibody or control IgG. Input lysates and IP samples were subjected to immunoblotting with the indicated antibodies.

D. B67 GSCs were first transduced with lentiviral plasmids for stable expression of EF1-alpha-driven C-terminal 6X-His tagged SOX2 and CMV-driven Flag-ubiquitin. The GSCs were subsequently transduced with either of 2 WWP2-targeting shRNA constructs or a non-targeting control. Lysates were harvested and subjected to immunoprecipitation (IP) under denaturing conditions using nickel-NTA beads. IP and input samples were subjected to immunoblotting using the indicated antibodies.

E. Following transduction with a WWP2 shRNA construct or a non-targeting control, GSCs were subjected to the neurosphere extreme limiting dilution assay (ELDA) assay to determine the stem-cell frequency for each condition. Statistical analysis was performed using the unpaired t-test (p = 0.0067 for B67 GSCs and p = 0.0093 for BT87 GSCs).

F. MGG8 GSCs were transduced with either a WWP2-targeting or a non-targeting shRNA construct. GSCs were subsequently stereotactically injected into the right putamen of immunocompromised NOD-SCID-y mice. Kaplan-Meier curves showing neurological deficit-free survival of the mice for the 2 experimental conditions are shown. Curve comparisons were performed using the Log-Rank (Mantel-Cox) test (n = 12, p = 0.026)
Figure 4.7: TRIM26 competes with WWP2 for binding to SOX2 and blocks WWP2-dependent SOX2 polyubiquitination and proteasomal degradation.

A. HEK293 cells were independently transfected with expression plasmids for HA-TRIM26, HA-WWP2, and GFP-SOX2. Lysates were harvested 36 hours post-transfection. GFP-SOX2 and HA-WWP2 lysates were added to all 3 conditions in equal amounts. HA-
TRIM26 lysate was excluded from the first condition and added in incremental amounts to the last 2. Lysates were subjected to anti-GFP immunoprecipitation. IP samples were subjected to immunoblotting with anti-HA and anti-GFP antibodies.

B. HEK293 cells were independently transfected with HA-WWP2 and HA-TRIM26-I18A expression plasmids. Lysates were harvested and subjected to anti-HA immunoprecipitation. E1 and E2 (UBE2D3) were added to all 3 conditions. HA-WWP2 beads were added to the last 2 conditions. HA-TRIM26-I18A IP beads were added to the last condition. In vitro synthesized radioactively-labeled SOX2 was added to all conditions. In vitro ubiquitination was performed and the samples were then subjected to electrophoresis and subsequent autoradiography to detect modified and unmodified SOX2.

C. MGG8 GSCs were transduced with either a WWP2-targeting or a non-targeting shRNA construct. On day 2 post-transduction, cells were further transduced with either a TRIM26-targeting or a non-targeting shRNA construct as indicated. Protein lysates were harvested and subjected to immunoprecipitation with the indicated antibodies.

D. B67 GSCs were transduced with either a WWP2-targeting or a non-targeting shRNA construct. On day 2 post-transduction, cells were further transduced with either a TRIM26-targeting or a non-targeting shRNA construct as indicated. Protein lysates were harvested and subjected to immunoprecipitation with the indicated antibodies.

E. B67 GSCs were transduced with either a WWP2-targeting or a non-targeting shRNA construct. On day 2 post-transduction, cells were further transduced with either a TRIM26-targeting or a non-targeting shRNA construct as indicated. GSCs were subsequently subjected to the neurosphere extreme limiting dilution assay (ELDA) assay to determine the stem-cell frequency for each condition. Statistical analysis was performed using ordinary one-way ANOVA with Bonferroni correction for multiple comparisons (p = 0.0004 for Control vs TRIM26i.2 and p = 0.0387 for TRIM26i.2 vs TRIM26i.2 + WWP2i.1)

F. MGG8 GSCs were transduced with either a WWP2-targeting or a non-targeting shRNA construct. On day 2 post-transduction, cells were further transduced with either a TRIM26-targeting or a non-targeting shRNA construct. GSCs were subsequently stereotactically injected into the right putamen of immunocompromised NOD-SCID-y mice. Kaplan-Meier curves showing neurological deficit-free survival of the mice for the 3 experimental conditions are shown. Curve comparisons were performed using the Log-Rank (Mantel-Cox) test (n = 12, p = 0.0006 for Control vs TRIM26i, p = 0.2541 for Control vs TRIM26i + WWP2i, and p < 0.0001 for TRIM26i vs TRIM26i + WWP2i)
Chapter 5:

Conclusions and Future Directions
6.1 Summary and Future Directions: Biological and Therapeutic Implications of Multisector Sequencing in Newly Diagnosed Glioblastomas

In this study, we performed whole exome sequencing of 2 to 4 biopsy specimens for each of ten newly diagnosed IDH-WT glioblastomas. We found that 46% of mutations were private across all samples sequenced, suggesting that most diagnostic platforms that employ single site biopsies may routinely miss a significant proportion of genetic alterations in patient tumors. In two striking example of extensive intratumoral genetic heterogeneity, we described two patients who harbored tumors with sector-specific hypermutator phenotype. The implications of this finding are significant, as there is evidence that hypermutant tumors respond well to immunotherapies. If otherwise eligible, both of these patients may have been excluded from potentially beneficial treatment regimens based on data from the single non-hypermutant biopsy.

Our analysis of the clonal architecture of multisector samples suggested that 80% of patients had subclones that were restricted to one of the sectors sequenced. Focusing on potentially actionable variants, we found similarly that 80% of patients had druggable mutations that were private to one sector and absent in the other. We observed sector specific mutations in EGFR and PTEN, two major therapeutically relevant genes in GBM. Additionally, we found that the existence and variant allele frequencies (VAFs) of predicted neoantigens differed between sectors for individual patients.

A number of future considerations arise from this study. The major question is how many sectors of a tumor need to be sequenced to satisfactorily cover the majority of actionable variants in each patient’s tumor. Some authors, based on modeling in medulloblastoma, suggest that 5 biopsies would be needed to cover 80% of actionable mutations(Morrissey et al., 2017). In the
absence of data validating such approaches, the costs would most certainly be prohibitive and
difficult to justify. However, we hope that more investigators will incorporate multiregion
sequencing, at least from 2 sites, especially for biomarker-based randomization of patients to
treatment arms in clinical trials of molecular targeted therapies similar to the NOA-20 trial(Wick
et al., 2019). A logical extension of our study is to perform multiregion analysis of primary and
recurrent tumors by both whole exome sequencing and single nucleus RNA sequencing. In
particular, it would be helpful to accrue larger numbers of leading edge and infiltrative multisector
samples as these are likely to be more informative with regards to minimal residual disease that
usually seeds recurrent tumors (SPITERI).

6.2 Summary and Future Directions: A CDC20-APC/SOX2 Signaling Axis Regulates
Human Glioblastoma Stem-Like Cells

In this study we discovered that the mitotic E3 ubiquitin ligase, CDC20-APC, is required
for both GSC invasiveness and self-renewal. We modulated CDC20-APC activity using RNai as
well as a small molecule drug PROTAME, which inhibits the interaction between CDC20 and the
APC subunit CDC27. Mechanistically, we found that CDC20 directly interacts with SOX2 and is
required for maintenance of SOX2 protein levels in GSCs. Critically, PROTAME also led to
decreased SOX2 protein levels, suggesting that CDC20 cooperates with the APC to regulate SOX2
expression. Using a transcriptional reporter assay as well as measurement of mRNA levels of the
SOX2-regulated gene NESTIN, we found CDC20-APC positively regulates SOX2 transcriptional
activity.
Further, using epistasis experiments, we found that SOX2 acts downstream of CDC20 to control GSC invasiveness and self-renewal. Further, using a SOX2 mutant that does not bind CDC20, we found that SOX2 binding to SOX2 is required for the regulation of GSC invasion. Using a xenotransplantation mouse model we found that CDC20 knockdown reduced GSC in vivo tumorigenicity, while CDC20 overexpression led to augmentation of tumor growth in vivo. Analysis of TCGA data revealed that CDC20 is overexpressed in glioblastomas compared to the normal brain and was significantly highly expressed in the Proneural subtype compared to the other TCGA subtypes. On subtype analysis, high CDC20 expression was associated with shorter overall survival in Proneural GBM. Exclusion of IDH1 mutant tumors led to similar survival between CDC20 high and CDC20 low Proneural patients, suggesting an interaction between IDH1 mutation and CDC20.

In terms of future directions, the upstream regulation of CDC20 in glioblastomas remains an open question. A recent scRNA-seq has suggested that NPC-like and OPC-like subtypes are more proliferative than the AC-like and MES-like subtypes (Neftel et al., 2019). Given the strong association of CDC20 with the Proneural subtype (which corresponds to the newly defined NPC-like and OPC-like states) in survival analysis, it would be interesting to know if cell cycle regulation does contribute to CDC20 function in GSCs in a subtype-specific manner. Additionally, how CDC20-APC regulates SOX2 protein remains an open question. The requirement of both CDC20 and the APC for SOX2 stabilization suggests that the catalytic activity of the CDC20-APC E3 ubiquitin ligase may be required for this function. It is likely that the CDC20-APC mediates ubiquitination of SOX2 that results in SOX2 stabilization. Critically, the E3 ligase that functions downstream of CDC20-APC to mediate SOX2 degradation has not been identified.
6.3 Summary and Future Directions: The TRIM26/WWP2 Ubiquitin Pathway Regulates SOX2 Proteostasis and Stem Cell Maintenance in Glioblastoma

In this study we found that SOX2 proteostatic regulation is a conserved feature of GSCs harboring divergent genetic backgrounds. We identified the immune-related E3 ubiquitin ligase, as a SOX2-interacting protein in GSCs, which unexpectedly was required for the maintenance of SOX2 protein levels. Further, TRIM26 knockdown resulted in decreased SOX2 transcriptional activity. Interrogation of TCGA data revealed a positive correlation between TIM26 and SOX2 mRNA expression levels.

Gene ontology analysis revealed that downregulated genes upon TRIM26 knockdown were associated with neurodevelopment, suggesting a role for TRIM26 as a potential regulator of GSC stem-cell function. Indeed, SOX2 knockdown resulted in decreased GSC in vitro self-renewal and in vivo tumor initiating potential. Further, we found that SOX2 acts downstream of TRIM26 in the regulation of GSC self-renewal. Mechanistically, we determined that TRIM26 regulates SOX2 at the protein level, and further that TRIM26 knockdown led to increased SOX2 polyubiquitination and a concurrent decrease in SOX2 protein stability as measured by cycloheximide-chase experiments. Further, we found catalytically inactive ring-domain mutants of TRIM26 were able to rescue SOX2 protein expression following TRIM26 knockdown, suggesting that TRIM26 catalytic activity is dispensable for TRIM26’s protection of SOX2 from degradation. We mapped the domain of TRIM26 required to interact with SOX2 to the C-terminal PRYSPRY domain of TRIM26. Critically, we found that expression of the TRIM26 PRYSPRY domain was sufficient to rescue SOX2 protein levels following TRIM26 knockdown.
A miniscreen of candidate genes identified WWP2 as a potential E3 ubiquitin ligase for SOX2 in GSCs. WWP2 knockdown resulted in decreased SOX2 polyubiquitination and an increase in ambient SOX2 protein levels in GSCs, validating it as a SOX2-targeting E3 ligase in GSCs. Further, WWP2 knockdown increased GSC in vitro self-renewal and in vivo tumorigenicity. We found that TRIM26 outcompetes WWP2 for binding to SOX2 and further that TRIM26 inhibited WWP2-mediated SOX2 ubiquitination in vitro. In epistasis experiments, we found that knockdown of WWP2 in the setting of TRIM26 RNAi was able to restore SOX2 protein levels and rescued the decrease in both GSC self-renewal and in vivo tumorigenicity that was observed with TRIM26 knockdown alone.

A major outstanding question is what other roles TRIM26 may play in GSCs in addition to regulating SOX2 protein levels. Indeed, restoration of SOX2 expression following TRIM26 knockdown did not fully rescue GSC self-renewal. Additionally, it would be interesting to know what the upstream regulators are that lead to markedly increased expression of TRIM26 in glioblastoma compared to the normal brain. It has been shown in mouse and human macrophages that treatment with type I Interferons or Interferon gamma (IFNy) led to a 2 to 5 fold increase in TRIM26 mRNA levels(Carthagena et al., 2009). It will be interesting to see whether interferons have the same effect on TRIM26 expression in GSCs, what the functional implications of this upregulation would be, and further what role the tumor microenvironment might play in modulating the TRIM26/WWP2/SOX2 axis. Finally, our studies only studied the impact of TRIM26 on GSCs as a static state. Given the description of state plasticity in glioblastoma and the existence of multiple cell states with GSC behavior, it would be interesting to perform analysis of cell state distributions in xenotransplantation models in TRIM26 knockdown versus control GSCs.
6.4 Conclusions:

The genomics era in cancer research and clinical care has availed an avalanche of information on the genetic and transcriptomic features of many cancers, including glioblastoma, both at bulk tumor and at single-cell resolution. A theme that has emerged for glioblastoma is the staggering extent of genetic, transcriptional, and cellular heterogeneity in these aggressive tumors. The work presented in this dissertation provides some insights into how multi-regional biopsies can avail critical biologic and therapeutic information, often missed by single-site biopsies, for both clinicians and researchers. Further, this work leverages both in vitro and in vivo xenotransplantation models to explore the therapeutic utility of targeting master transcription factors, in this case SOX2, that are critical for the neurodevelopment signature that is commonly seen in glioblastomas. This work will hopefully avail findings that can inform future work aimed at implementing differentiation therapies to simultaneously target multiple heterogenous GSC states.
6.5 References:


TATENDA MAHLOKOZERA
Washington University in St Louis School of Medicine, Program in Neuroscience and Medical Scientist Training Program (MSTP) • (314)-761-5430 • mahlokozerat@wusm.wustl.edu
6 Baring Cres, Lundi Park, Gweru, Zimbabwe

EDUCATION

GRADUATE:

Washington University in St. Louis School of Medicine, Saint Louis MO

MD/PHD (Neurosciences): 2012- 2022 (expected).

Honors:

• Distinguished Young Scholar Award- Washington University MSTP- 2012
• Congress of Neurological Surgeons Best Basic Science Tumor Award, 2019 CNS Annual Meeting, San Francisco, CA.

UNDERGRADUATE:

Colby College, Waterville ME.

Bachelor of Arts: May 2009

Double Major: Chemistry:Biochemistry-A.C.S certified
Physics

Minor: Mathematics

Honors:

• Magna cum laude
• Honors and Distinction in Chemistry-Biochemistry
• Honors and Distinction in Physics
• Oak Foundation Scholar
• Elected to The Phi Beta Kappa National Honor Society
• Elected to The Sigma Pi Sigma National Physics Honor Society
• First Year Departmental Award in Chemistry
• Senior Departmental Award in Chemistry

RESEARCH EXPERIENCE

Research Assistant: Professor Dasan Thamattoor- Physical Organic Chemistry
• Synthesis of non-nitrogenous precursors to unusual carbenes and subsequent laser flash photolysis of the precursors.
• Synthesis of tosylhydrazone precursors to functionalized carbenes and their subsequent flash vacuum pyrolysis. Mass spectrometric, NMR, and theoretical computational analysis of photolysates- **Honors Thesis Project**

**Research Assistant: Professor Duncan Tate- Atomic Physics**  
• Measurement of the radiative lifetimes of Rydberg states of rubidium using thermally cold Rb atoms derived from a magneto optical trap (MOT)- **Honors Thesis Project**

**Research Assistant: Dr Sallie Permar**  
Beth Israel Deaconess Medical Center, Harvard Medical School Division of Viral Pathogenesis (July 2009-July 2011).
• Investigating immune correlates to infant protection from HIV transmission via breastfeeding.

**Research Technician II: Dr Sallie Permar**  
Duke University Medical Center/Duke Human Vaccine Institute (July 2011-July 2012).
• Investigating immune correlates to infant protection from HIV transmission via breastfeeding

**Graduate Thesis Research: Dr Albert H Kim (Department of Neurosurgery)**  
Washington University School of Medicine, Program in Neuroscience and Medical Scientist Training Program (July 2014-Present)
• Investigating common developmental signaling pathways in the normal brain and in Glioblastoma Multiforme (GBM).

**PRESENTATIONS AND ABSTRACTS:**

1) **Mahlokozera T.,** Thamattoor D.M. “Probing the electronic influence of bystander substituents on the 1,2-Hydride in carbenes (carbenoids).” Poster presentation at the 237th ACS National Meeting, March 2009, Salt Lake City, Utah.


6) **Tatenda Mahlokozera**, Rukayat Taiwo, Diane Mao, Afshin Salehi, Amit Gujar, and Albert Kim. *Non-canonical regulation of SOX2 by the TRIM26 E3 ubiquitin ligase in glioblastoma stem-like cells*. Oral presentation at the 23rd Society for Neuro-Oncology Annual Scientific Meeting, **November 2018**, New Orleans, LA.


---

**PEER-REVIEWED JOURNAL PUBLICATIONS**


