McGill

THESIS NON-EXCLUSIVE LICENSE

AUTHOR'S NAME: PALANISAMY (*LAST NAME*) RAVISHANKAR (*FIRST NAME*) _____ (*INITIALS*)

AUTHOR'S DATE OF BIRTH: 03 (*MONTH*) / 28 (*DAY*) / 1978 (*YEAR*)

MCGILL EMAIL ADDRESS: RAVISHANKAR.PALANISAMY@MAIL.MCGILL.CA **STUDENT NUMBER:** 260190294

PERMANENT ADDRESS: 266, NAGARACHI NAGAR, SOLAR PUDUR, ERODE, INDIA (*COUNTRY*)

MCGILL UNIT: INTEGRATED PROGRAM IN NEUROSCIENCE

FACULTY: MEDICINE

DEGREE SOUGHT: MSc

TITLE OF THESIS: The role of Histone 4 in regulating wild type and mutant isocitrate dehydrogenase 1 tumor proliferation

I hereby promise that I am author of the thesis above cited.

I confirm that my thesis is my original work, does not infringe any rights of others, and that I have the right to make the grant conferred by this non-exclusive license. I also confirm that if third party copyrighted material was included in my thesis for which, under the terms of the Copyright Act, written permission from the copyright owners is required, I have obtained such permission from the copyright owners and may grant such permission for the full term of copyright protection.

I hereby grant to McGill University a non-exclusive, worldwide, irrevocable, royalty free license, in respect of my thesis, to reproduce, convert, publish, archive, preserve, conserve, communicate and distribute, and loan, in paper form, in microform, electronically by telecommunication or on the internet, and/or any other formats as may be adopted for such use from time to time. I also authorize McGill University to sub-license, sub-contract for any of the acts mentioned.

I hereby grant permission and authorize McGill University to permit access to my thesis and make it available to interested persons in paper or electronic form, through the library, interlibrary and public loan.

I understand that I retain copyright ownership and moral rights in my thesis, and that I may deal with the copyright in my thesis consistent with these rights. I promise to inform any person to whom I may hereafter assign or license my copyright in my thesis of the rights granted by me to **McGill University** and to **Library and Archives Canada.**

I confirm that I have executed a **Non-exclusive license with the Library and Archives Canada** and hereby grant permission to McGill University to submit the abstract and my thesis to the **Library and Archives Canada** in full compliance with these non-exclusive licenses.

The authorization contained in these non-exclusive licenses are to have effect on the date given below (**Effective Date**) unless a deferral of one year from the date has been expressly requested by me, the author, on submitting the thesis, and acknowledged by McGill University **GRADUATE AND POSTDOCTORAL STUDIES OFFICE.**

I agree to indemnify and hold McGill University harmless against any claim and any loss, damage, settlement cost or expense (including legal fees) incurred by McGill University and arising out of, or in connection with, my statements and representations or this license.

SIGNATURE OF AUTHOR:  **EFFECTIVE DATE:** 20th April 2018

Library and Archives Canada | Bibliothèque et Archives Canada

Canada

THESES NON-EXCLUSIVE LICENSE

Surname: PALANISAMY	Given Names: RAVISHANKAR
Full Name of University: McGill University	
Faculty, Department, School: Medicine, Integrated Program in Neuroscience	
Degree for which thesis was presented: MSc	Date Degree Awarded:
Thesis Title: The role of Histone 4 in regulating wild type and mutant isocitrate dehydrogenase 1 tumor proliferation	
Supervisor: Dr.Kevin Petrecca	
Date of Birth. It is **optional** to supply your date of birth. If you choose to do so please note that the information will be included in the bibliographic record for your thesis.	
E-mail: please provide your e-mail address if you are interested in receiving royalties on sales of your thesis. ravishankar.palanisamy@mail.mcgill.ca	
Permanent Address: please provide your permanent address if you are interested in receiving royalties on sales of your thesis. 266, Nagarachi Nagar, Solar pudur, Erode 638002	

In consideration of Library and Archives Canada making my thesis available to interested persons, I,

Ravishankar Palanisamy

hereby grant a non-exclusive, for the full term of copyright protection, license to Library and Archives Canada:

(a) to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell my thesis (the title of which is set forth above) worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats;

(b) to authorize, sub-license, sub-contract or procure any of the acts mentioned in paragraph (a).

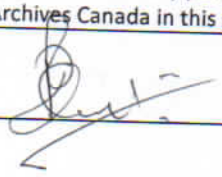
I undertake to submit my thesis, through my university, to Library and Archives Canada. Any abstract submitted with the thesis will be considered to form part of the thesis.

I represent and promise that my thesis is my original work, does not infringe any rights of others, and that I have the right to make the grant conferred by this non-exclusive license.

If third party copyrighted material was included in my thesis for which, under the terms of the Copyright Act, written permission from the copyright owners is required I have obtained such permission from the copyright owners to do the acts mentioned in paragraph (a) above for the full term of copyright protection

I retain copyright ownership and moral rights in my thesis, and may deal with the copyright in my thesis, in any way consistent with rights granted by me to Library and Archives Canada in this non-exclusive license.

I further promise to inform any person to whom I may hereafter assign or license my copyright in my thesis of the rights granted by me to Library and Archives Canada in this non-exclusive license.

Signature 	**Date** 20th April 2018

The role of Histone 4 in regulating wild type and mutant isocitrate dehydrogenase 1 tumor proliferation

Ravishankar Palanisamy

Integrated program in Neuroscience

McGill University, Montreal.

April 2018

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

ACKNOWLEDGEMENTS

I dedicate my thesis to my Parents for the support throughout this program. A special dedication goes to my soulmate MJ who is not with me anymore. I would like to thank to my MSc advisor, Dr. Kevin Petrecca for supporting me for giving me my dream project in epigenomics. Dr. Petrecca guided throughout my experiments by providing insightful discussions during lab meetings. I am also very grateful to Dr. Phuong ULe for her scientific advice and knowledge and many insightful discussions and suggestions. Dr.ULe was my primary resource for lab orders and also, I also would like to thank her for helping me to make beautiful graphs using the data I obtained. I also would like to thank Ms. Maryse Thomas for repeating some of my experiments in cell culture which basically helped to confirm my results obtained earlier. I also would like to thank members of the Petrecca lab including Dr.Mohammad Syed Sadr, Dr. Faisal Farrash and Dr. Chris Kent for their suggestions regarding my professional and personal life. I also would like to thank Mrs.Carmen Sabau who was very considerate and caring all times and especially when I was sick with flu. I enjoyed friendly discussions with Alex Dudley who gave valuable suggestions at the very beginning of the program. I acknowledge Ryan, Dr. Mai Sater and Gina from Petrecca lab for time away from work which helped me to refresh on a timely basis. Members of Jabado lab helped me out in preparing tissue samples, obtaining data from methylation platform. I would like to thank Dr.Dong Anh, Dr.Tenzing Gayden and Dr.Nada Jabado who allowed me to to interact freely with their lab members for research discussions. I will forever be thankful to our past program director and present Post graduate Dean, Dr.Josephine Nalbantoglu for helping me to identify Petrecca lab, mentoring me in the program and for regularly motivating me in the program during my down times. I can't forget timely guidance and advice by Dr. Jacques Lebrun, the Associate Dean, Faculty of

Medicine and the Program director, Dr.Joseph Rochford. Last but not the least, I thank my friends Maria, Gaby, Mathilde, Jason, Han, Maryam, Gopi, Vincent, GG, Fred, Red, Andrea and Horia for keeping me in good spirits. I would like to thank my friends Almira, Wafaa, Akash Bharat and Rajesh for taking time to spend with me outside lab work and keeping me cheerful. Some special thanks to Dr.Kiruba Christie for all her prayers and love. I thank all my dearest brothers and sisters of my Truro family for showering unlimited love and care.

Preface:

Experimental design, data collection and experimental procedures were carried out by Ravishankar Palanisamy. Mrs. Josee Bergeron did animal (mice) work which involved maintaining mice colonies, did glioma stem cell injections in the mice brains and obtained pictures of mouse brains using confocal microscope.Also, Mrs.Josee standardized immunolabelling conditions for patient tissue microarrays that will be carried out in the future by the Petrecca lab. Dr. Phuong Ule isolated glioma stem cells from brain tumor tissue supplied by Dr. Kevin Petrecca and helped coordinate mouse work and obtaining images from brain sections.

Abstract

Glioblastomas are aggressive brain tumors that invade normal brain parenchyma and spread wider quickly. Primary GBMs, which are usually detectable in elderly patients, grow rapidly *de nova* without precursor pathological lesions or and clinical symptoms[1].Secondary GBMs are diagnosable and develop from low-grade astrocytoma for about 5-10 years. Patient cohort astrocytoma samples were used for initial promoter DNA methylation analysis using Infinium ®HumanMethylation450 BeadChip. Methylation states on 12 wild type Isocitrate dehydrogenase 1 (WT-IDH1) and 9 mutant Isocitrate dehydrogenase 1(MUT-IDH1) astrocytomas were compared. Results displayed in a heat map showed that the MUT-IDH1 tumors are hypermethylated relative to WT-IDH1 tumors. Mann-Whitney-Wilcoxon Test to assess similarity between MUT-IDH1 and WT-IDH1 tumors showed that they are significantly different ($p \leq 0.001$) in their methylation states. Histone H4B gene expression was upregulated in WT IDH tumors relative to MUT-IDH tumors. To validate methylation array results, we compared H4B mRNA expression from a different set of WT and MUT IDH1 astrocytomas and found that indeed H4B mRNA levels were high in 8 of 10 WT IDH1 astrocytomas compared to 2 of 11 MUT IDH1 astrocytomas. Also, total protein levels of H4 were increased in WT IDH1 tumors compared to MUT-IDH1 tumors. IDH tumors are heterogeneous in nature, and exome sequencing of WT-IDH tumors showed alterations. Overexpression of H4B using a plasmid vector in tumor-derived glioma stem cells (GSC) proliferation rate in MUT-IDH GSCs was three times than in control. Results from *in vitro* studies outline a potential trigger for gliomagenesis and there is a need to validate *in vitro* results using mouse xenograft models. Preliminary work on implanting glioma stem cells in mice wassuccessful and steps to evaluate physiological and genetics of implanted tumors are in progress. To further tease out unique properties of H4B

induced proliferation in IDH Mutant GSCs overexpressing H4, DNA samples were prepared and submitted for whole exome sequencing. Analysis and validation of exome sequencing data are in progress.

Résumé

Les glioblastomes sont des tumeurs du cerveau agressives qui envahissent le parenchyme cérébral et se propagent rapidement. Les glioblastomes primaires ne présentent pas de lésion pathologique et les symptômes cliniques conduisant à un diagnostique apparaissent moins de trois mois avant leur développement complet. Les glioblastomes secondaires sont diagnostiquables et se développent durant 5 à 10 ans à partir d'un astrocytome de bas grade. Des analyses de méthylation primaire de l'ADN ont été menées sur des échantillons d'astrocytomes issus d'une cohorte de patients en utilisant la puce à ADN Infinium ®HumanMethylation450 BeadChip. Les niveaux de méthylation de l'ADN issu de 12 astrocytomes WT-IDH1et de 9 astrocytomes MUT-IDH1 ont été comparés. Les résultats, présentés sur une carte thermique, ont montré que les tumeurs MUT sont généralement hyperméthylées comparativement aux tumeurs WT-IDH. Les tumeurs MUT-IDH1 et WT-IDH1 forment des groupes nettement distincts, confirmant que leur niveau de méthylation est significativement différent. De plus, l'expression du gène Histone H4b est régulée à la hausse dans les tumeurs WT-IDH comparativement aux tumeurs MUT-IDH. Afin de valider les résultats de méthylation obtenus à l'aide des puces à ADN, nous avons ensuite comparé l'expression de l'ARNm H4b issus d'échantillons d'astrocytomes WT- et MUT-IDH1, distincts et nous avons alors mis en évidence des niveaux d'expression d'ARNm H4b élevés dans 8 des 10 échantillons d'astrocytomes WT-IDH1 et dans seulement 2 des 11 échantillons d'astrocytomes MUT-IDH1. Dans les tumeurs WT-IDH1, le niveau d'expression de la protéine H4 est également régulé à la hausse par rapport aux tumeurs MUT-IDH1. Les tumeurs IDH sont de nature hétérogène et le séquençage de l'ensemble des séquences codantes des tumeurs WT-IDH1 a mis en évidence des altérations de leurs génomes. La surexpression de H4b dans des cellules souches de gliome dérivées de tumeurs (GSC) MUT-

IDH1 ou WT-IDH1 a conduit à un taux de prolifération cellulaire trois fois plus importants des cellules MUT-IDH comparativement aux cellules issus de tumeurs contrôles. Ces résultats d'études *in vitro* sont encourageants pour l'étude du développement des gliomes *in vivo*, dans des modèles de xénogreffes chez la souris. Des résultats préliminaires ont montrés une implantation réussie des cellules souches de gliomes et l'analyse physiologique et génétique des tumeurs en résultant est en cours. Afin de déterminer les propriétés uniques des cellules souches de gliomes induits par H4b, le séquençage de l'ARN de cellules individuelles et le séquençage des séquences codantes des mutants IDH surexprimant H4 est en cours.

Table of contents Page Numbers

1. Introduction

The Brain Tumor Foundation of Canada has reported thatapproximately 55,000 Canadians have a brain tumor and every day, 27 are diagnosed. The World Health Organization classified grade I gliomas as benign, whereas grade II (low-grade diffuse astrocytomas) and grade III (anaplastic astrocytomas) areclassified as invasive tumors. Low-grade astrocytomas are primary tumors of the central nervous system derived from glial cells. Grade II and III tumors often progress to malignant glioblastoma (GBM) (grade IV)[2]. Glioblastomas are aggressive brain tumors which invade normal brain parenchyma and spread wider quickly.Primary GBMs which are usually detectable in elderly patients, grow rapidly *de nova* without precursor pathological lesions or and clinical symptoms[1].Secondary GBMs are diagnosable and develop from low-grade astrocytoma for about 5-10 years[3]. In-depth molecular analysis from several studies show that the primary GBMs have four classes namely the proneural (PN), the mesenchymal (MES), classical (CL), and neural (NL). The significantly younger PN type has been shown to overexpress the tumor suppressor gene, p53 and may haveIDH1 mutation. The normal like NL subgroup has overexpressing neuronal genes. Epidermal growth factor receptor (EGFR) expression and EGFR vIII mutations have been found predominantly in the 'classical' subgroups. These classes are also associated with canonical mutations such as *PDGFRA* amplification in PN GBMs, loss of *NF1* in MES GBMs, and amplification of Epidermal growth factor receptor (*EGFR)* in CL GBMs. However, some tumors have characteristics of more than one subtype[4-6]. For example, abnormalities in NL GBMs are not clearly known yet. The overlap that exists between different subtypes raises questions on the temporal sequence of the evolution of the subtypes and its corresponding alterations[5,7,8]. Also, it is unknown whether the subtypes arise from a common precursor. For example, abrupt transition of PN to MES types and subtype mosaicism suggests

that the subtypes might have a different lineage. Provided many subtype-defining genes are mostly late events, it is unknown whether the genes can be therapeutic targets[9]. Epithelial to mesenchymal transition (EMT) is a typical process that causes loss of epithelial features leading to resistance to targeted tumor therapy [10]. The 'mesenchymal' subtype is made up of EMT and mutations in NF1 tumor suppressor gene. All of the above subtypes failed to correlate with survival rates among patients nor was a better prognosis predictable from the signaling pathways associated. The subtypes are limited by their sample size, signature variations between the subtypes, and overall regarded as insufficient tumor classifications[56,11,12].

Specific biomarkers have shown to be present in primary and secondary GBM tissue, which includes amplification of epidermal growth factor receptor and loss of tumor suppressor protein TP53, respectively.The amounts of proteins that regulate cell cycle and tumor suppression have been shown to be modified in both GBM subtypes. Loss of tumor suppressors, phosphatase and tensinhomolog (PTEN) and cyclin-dependent kinase (CDK) inhibitor2A (CDKN2A) or overexpressed CDK4 regulate the G1 phase of the cell cycle.Furthermore, overexpression of mouse double minute 2 (MDM2), a protein that targets tumor suppressor proteins such as TP53[3], has been shown to be regulated in GBM subtypes. Recent studies suggest that genetic lesions occurring in GBM subtypes may be due to mutations in genes involved in GBM pathogenesis with the altered transcriptional rates. Because GBMs resist apoptosis and are resistant to chemo and or radiotherapies, it is important to study the biology and genetics to find novel targets for therapies.

Otto Warburg hypothesized that cancer cells utilize the less effective glycolysis pathway over a highly efficient oxidative phosphorylation pathway to generate lactate from glucose[13,14]. This metabolic switch occurring in the cancer cells is similar to healthy cells that adapt aerobic

glycolysis during times of increased proliferation rates[15,16]. Also, intermediate products of aerobic glycolyticcycle help for the biosynthesis of lipids, amino acids and nucleic acids necessary for the rapidly proliferating cells[17,18]According to the World Health Organization (WHO), brain tumors are classified into four grades depending on the malignancy. Grade II and III astrocytomas, oligodendrogliomas and secondary GBMs (grade IV) mostly contain IDH1 mutation and are rarely found in pilocytic (grade I) astrocytomas[19-21]. In 2008, a genome-wide study by Balss *et al.* (2008) identified a high percentage of IDH1 mutations in low-grade gliomas suggesting IDH1 mutations regulate early tumor development[19]. Also, an IDH1 mutation is prominent among all other mutations linked to gliomas. Also, IDHI Arginine 132 mutations (IDHR132)occur in more than 70% of adult grade II and III gliomas and more than 80% of adult secondary GBMs. However, IDH mutations occur in less than 10% of primary and pediatric GBMs.

2. Background

2.1. Importance of Isocitrate Dehydrogenase 1 (IDH1) mutation

In 2008, a genome-wide study by Balss *et al* (2008) identified a high percentage of IDH1 mutations in low-grade gliomas suggesting IDH1 mutations regulate early tumor development[19]. MUT-IDH1 tumors have shown to be less malignant compared to the WT-IDH1 tumors and progress rapidly[22]. WT-IDH1 tumors exhibit poor prognosis compared to MUT-IDH1 tumors and therefore are classified as aggressive GBMS independent of WHO classification[23]. Wild-type IDH1 tumors are less sensitive to radiation and chemotherapy than MUT-tumors[24]. Therefore, it's important to understand progression, development, and differentiation of aggressive gliomas. IDH1 enzyme is present in the cytoplasm and peroxisomes and regulates

glucose sensing and lipid metabolism in respective cellular compartments. WT-IDH1 enzyme catalyzes a reversible reaction in which isocitrate is converted to 2-oxogluarate (2-OG) and nicotinamide adenine dinucleotide phosphate (NADPH) by oxidative decarboxylation. Cytosolic NADPH serves as an important source for the synthesis of reduced glutathione, an important antioxidant for mammalian cells[25] and is also used by the hexose monophosphate shunt pathway to synthesize deoxyribose and ribose sugars for nucleic acid biosynthesis[19]. The antioxidant role of IDH1 has been well demonstrated by IDPc, the mouse homolog which has shown to be overexpressed in NIH3T3 cells when exposed hydrogen peroxide[26]. Also, under oxidative stress, IDPc deficiency has been shown to induce apoptosis in HL-60 cells by supporting caspase3 activation[27]. In cytosolic IDH1, the evolutionarily conserved region of the residue R132 that forms hydrophilic interactions with isocitrate has been shown to be conserved among several species[25]. Mutant-IDH enzymes are heterodimers consisting of mutant and wild-type proteins and catalyze isocitrate to D-2-hydroxy glutarate (D-2HG) instead of alpha-keto glutarate and NADPH[2]. Alpha-keto glutarate is therefore frequently reduced in IDH mutant cells whereas D-2HG levels increase for about ~ 100 folds in tumors compared to normal cells[2]. A study that substituted glutamate for an arginine residue at 132 of IDH1 resulted in the loss of enzyme catalytic activity[28] but surprisingly, the IDH neomorphic enzyme activity induced gliomagenesis or leukemogenesis probably by using alternative substrates [29-31]. In gliomas, IDH mutation has been associated with genetic alterations in other genes including tumor suppressors and oncogenes[31]. Interestingly, IDH mutations help increase the survival rates of grade II and III astrocytomas patients, and this may be due to the influence of the tumor suppressor gene, TP53 [31,32]. Themutant-IDH1 enzyme has been shown to induce gliomagenesis when expressed in early progenitor neural stem cells and therefore help retain self-renewal capacity of glioma stem cells

(GSC)[33]. Cells capable of self–renewal, initiation and differentiation within a heterogeneous tumor tissue are identified as cancer stem cells. Tumor recurrence after therapy hasbeen shown to occur with the support of cancer stem cells (CSC),which expresses genes responsible for EMT, tumor aggressiveness and pluripotency [34,35].

2.2. Isocitrate Dehydrogenase mutations in Glioma

Gliomas of glial subtypes frequently carry mutations in the isocitrate dehydrogenase genes, 1 and 2 (IDH1 and IDH2)[19,30,31,36]and these mutations are confined to a single codon in each gene (R132 in IDH1 and R172 in IDH2). Therefore, detecting IDH mutations has served as an important GBM diagnostic tool. Unlike other cancer-associated enzyme mutations, mutated IDH genes give rise to proteins that lose their enzyme activities, but gain neomorphic enzyme activities. This neo-morphism occurs when α-ketoglutarate gets converted to 2-hydroxyglutarate in a Nicotinamide adenine dinucleotide (NAD)-dependent manner. A missense mutation in a single copy of cytoplasmic IDH1 at codon 132 places histidine (H) instead of arginine (R)[36], which in turn decreases the binding affinity of IDH1 for the substrate. The mutation increases the binding affinity of IDH1 to NADPH which subsequently abolishes the oxidative decarboxylation activity of IDH1[2]. The heterodimeric form of the mutant protein is required in the cytoplasm for 2HG synthesis, and Jin *et al.* (2013) demonstrated the requirement of wild-type IDH1 for higher levels of 2HG in gliomas. The study compared 2HG levels of IDH1 mutant expressing grade III anaplastic astrocytomas, to heterozygous mutants. It was shown that 2HG levels were reduced to 14 folds when they had loss of WT-IDH1. Also, the group developed an IDH1 mutated astrocytoma (grade III) (IMA) stem cell line and showed that targeted depletion of WT-IDH1 led to an 87 fold increase in D-2HG compared to the IMA cell line[37].

2.3.The Oncogenic ability of R-2HG

In another study, patients with anaplastic astrocytomas had a median survival rate of 65 months for mutated IDH1 or IDH2 compared with 20 months for patients with wild-type IDH1 or IDH2[31]. Several studies emphasize tumor invasion in gliomas[38], but mechanisms of tumor proliferation and development influencing survival rates in patients with diffuse astrocytomas are unknown. Particularly the role of IDH in modulating grade II and III gliomas is yet to be investigated. The mutant IDH1 enzyme hasbeen shown to induce cell proliferation when expressed in early progenitor neural stem cells and therefore helps stem cells retain self-renewal capacity[33]. A stem cell can undergo either symmetric (two daughter cells are identical) or asymmetric (dissimilar daughter cells) cell division. Self-renewal capacity of a stem cell depends on symmetric division whereas asymmetrical dividing leads to cell differentiation[33]. Although the mutant IDH enzyme couldn't immortalize self-renewal capacity or induce differentiation in a committed cell type, it favors self-renewal and blocks primary cell differentiation to a limited extent leading to a leukemogenic phenotype. The self-renewal ability of the mutant IDH enzyme[39]was observed both *in vitro* as well *in vivo*(using IDH1 mutant knock-in mouse model) [40].When expressed in differentiated cells, the mutant IDH1 enzyme was not able to induce self-renewal capacities, but inhibited differentiation. However, the mutant IDH1 enzyme promoted cellular differentiation in brain cells during early embryogenesis and caused perinatal lethality when expressed *in vivo*. When expressed in the hematopoietic system, the mice didn't develop malignancy, but expanded healthy hematopoietic cells, but not the committed myeloid of lymphoid cells. Therefore, IDH1 specifically promotes proliferation of primary cells and further malignancy in committed cells may be achieved by the interaction of yet unidentified driver genes.The mutant IDH tumors mainly accumulate D-2HG or R-2HG enantiomer of 2HG in the

range of 2mM to 30mM concentration. Continuous and high levels of cell-permeable intracellular 2HG transformed TF-1 cells to an IDH1 mutant phenotype by blocking differentiation of TF-1 cells[41]. Also, R-2HG administration was able to inhibit differentiation of murine 3T3-L1 fibroblast and immortalized myeloid progenitor cells [41,42] further confirming the oncogenic role of R-2HG. Withdrawal of cell permeable R-2HG or inhibiting IDH1 mutant enzyme activity by a potential inhibitor caused differentiation of TF-1 cells and restored characteristics of suppressed IDH mutant phenotype [41,43].

2.4. Origin of Glioma cells

Neurons and different types of glial cells namely, astrocytes, oligodendrocytes, and ependymocytes are mature differentiated cells of the central nervous system (CNS) which develop from CNS stem cells that express Nestin (a biomarker for CNS stem cells). In early embryonic development, cerebral ventricles are lined by pseudostratified neuroepithelial tissue that gives rise to CNS stem cells. Radial glial cells are neural stem cells that arise from neuroepithelium [44,45]that express BLBP (brain lipid binding protein)[44,45.] During adulthood, early progenitors disappear but neural stem cells, which are self-renewing and multipotent cells, concentrate and reside in the subventricular zone (SVZ) that lines the lateral ventricle in the dentate gyrus. The dentate gyrus lies in the subgranular zone (SGZ). Neuraxis encompasses SGZ, the sub-cortical whitematter region and the corpus callosum[46].Radial glial cells have self-renewing capacity and transform to type B cells that label positive for glial fibrillary acidic protein(GFAP) retaining the properties of radial glia and are concentrated in SVZ which is the largest germinal region or neural stem cell origin in humans. Type B cells are slowly dividing astrocytes or primary progenitors that transform to actively dividing astrocytes or intermediate progenitors called type C cells. Furthermore, type C cells give rise to type ‘A’ cells that migrate

to the olfactory bulb where they differentiate to form glial subtypes[47,45]. Apart from SVZ, the SGZ also comprises neural stem cells of radial glial origin and the slower dividing primary progenitors give rise to intermediate actively dividing progenitors called type D cells[48].In humans and other adult mammalsneural stem cells, which can generate either neuronal or glial lineages when exposed to disturbances in cellular functions give rise to tumors. Neural stem cells are regulated by pathways similar to adult brain tumors and exhibit high cellular motility, blood vessel association, generate whitematter tracts and immature antigenic phenotypes [4,49,50].Gliomas have been shown to originate and be contiguous with subventricular zone or periventricular in occurrence[45, 51].Studies on animal models have shown that neural stem cells have a high potential for oncogenesis and the regions of brains with high stem cell proliferation were susceptible to chemical or viral oncogenesis; *e.g*., Tumors occurred in subventricular regions in the brains of dogs and rodents following systemic exposure to the carcinogen, N-ethyl-N-nitrosourea[48,52]. Furthermore, a population of cells in the CNS that are positive for NG2 proteoglycans are found in the developing brain as progenitors of oligodendrocytes as well as in the adult brain regions. Unlike SVZ and SGZ, these cells give rise to neurons and macroglia [50,53].Further studies are needed to explore proliferative mechanisms of glial progenitors especially in the presence of driver genes stimulating proliferation.

2.5. Tumor Heterogeneity

Tumor heterogeneity in gliomas varies spatially and temporarily within an individual tumor mass. Spatial tumor heterogeneity complicates histopathological classifications. A study conducted by Sottoriva *et al*[8]using a fluorescence-guided method, collected spatially distinct tumor fragments from 11 glioblastomas and showed that an individual tumor mass presents with

multiple regional transcriptional profiles. Point mutations in EGFR, PIK3CA, PTEN, and TP53 genes over time from diagnosis to the second recurrence have been found on seven primary glioblastomas suggesting the influence of temporal heterogeneity in tumor development[54]. Driver mutations of genes occurring in tumor cells undergo natural selection over multiple generations and cause tumor progression. Non-random mutations occurring at the same amino acid position (e.g. *IDH*) is called "Mut-driver genes" whereas randomly occurring mutations due to gene amplification/loss or epigenetic alterations are referred to as "Epi-driver genes". Targeting non-random universal genetic modifications as well as exploring epigenetic mechanisms producing non-random mutations in tumor cells (collectively called as "trunk events") during clonal evolution can lead to better outcomes[55,56]. Subclonal variation in tumor cells due to trunk events compared with other tumor cells cause tumor heterogeneity[57]. The driver mutation in IDH has been shown to drive the progression of low-grade lesions to invasive secondary GBM and therefore the mechanisms of development of secondary GBMs can be studied. However, primary GBMs that develop rapidly within few months of normal MRI scans are a challenge to study.

According to cancer stem cell theory, not all cells in a clone carry genetic variations in them, but those that possess self-renewal capacity proliferate with varied genetic profiles causing heterogeneity within the same tumor[58,59]. Neurochemistry, metabolic energy, and surrounding architecture are some of the factors that cause clonal variations in tumors arising from different anatomical location[60]. Sanai *et al*. (2005)[46] suggested that gliomas originate in the areas of the brain with the densest astrocyte population (*e.g*., subventricular zone) that might promote malignancy. Genetic variations in tumors promote invasion through blood vessels and white matter tracts and induce high tumor frequency in frontal lobes than temporal, parietal and occipital lobes in order [60]. The tumor microenvironment contains reactive astrocytes, microglial

cells and immune infiltrates, which can drive tumor cell progression and malignancy. Also, clonal subpopulations carrying genetic variations have also proved to be nourished by aberrantly proliferating microvascular tissues[61]. Glioblastoma cancer stem cells have been shown to take advantage of the hypoxic conditions from the perivascular zone to maintain self-renewal capacity suggesting that further studies on hypoxic microenvironment influencing tumor heterogeneity is needed[62]. Three different receptor tyrosine kinase (RTK) genes (*EGFR*, *MET*, and *PDGFRA*) in gliomas were found to constitute intra-tumoral heterogeneity[63], and these genes maintained different clonal subpopulations of GBM[64]. Interestingly Snuderl *et al*. (2011) demonstrated that *EGFR*, *MET*, and *PDGFRA* were differently amplified in clones of same tumor tissue that arose from a common tumor precursor cell.Brennan *et al*. (2013)[65]classified genetic alterations in GBMs into RTK/RAS/PI3K axis, p53/MDM2/MDM4 axis, and RB/CDK4/INK4A axis pathways. However, drugs that targeted these pathways, *e.g*., the EGFR kinase inhibitors such as Gefitinib® and Erlotinib® didn't show significant benefits to therapy [66,67]because inhibition of RTKs activated other compensatory pathways [68,69]. Since tumor heterogeneity is spatially and temporarily controlled, it is important to obtain information at the single cell level from clonal subpopulations over time. Molecular biomarkers that help prognosis or diagnosis of initial tumor occurrence do not predict prognosis of recurrent tumors. This was evident from TCGA transcriptional subtypes that differed between primary and recurrent tumors. Therefore before targeting drug resistance, it becomes important to monitor the evolution of clonal subpopulation (clonal dynamics) by obtaining the spatial orientation of patient samples[70] or debulking of tumors serially from patients.Strategies such as RNA-seq can be used to compare tumor heterogeneity and survival rate as described before by Patel et al. (2014). According to this study, tumor heterogeneity correlates with decreased survival rate in patients with GBM [71]

2.6. DNA Methylation

Among several mechanisms involved in glioma development and progression, DNA methylation has been shown to modulate cancer gene expression. DNA methylation plays a major role in the loss of pluripotency and specified cell growth during development[72,73]. Methylation of CpG islands in the promoter and non-promoter regions of genes has been shown to influence gliomagenesis[74]. Methylation, which occurs on cytosine residues present in cytosine-phosphate-guanine (CpG) dinucleotide environment, is considered as an important epigenetic mark regulating oncogene expression. CpG islands are regions of DNA enriched with CpG dinucleotides and approximately 1% of the human DNA contains CpG islands and nearly 70-80% of the CpGs are methylated in adult cells[75-77]. Transcriptional start sites (TSS) of 60-70% promoters of housekeeping and tissue-specific developmental genes are located in CpG islands[78,79]. To enable gene expression, CpG islands of promoter regions are usually unmethylated except in females where the X-chromosome is maintained inactive and CpG islands of imprinted genes where methylation helps to maintain a silenced state. Hypermethylation of CPG islands generally serves as an important early human cancer biomarker in blood and other fluids of the body[80]. Specific hyper and hypo DNA methylations of CPG islands frequently and concurrently occur in human cancers[81]. Promoter-specific hypermethylation, which can cause abrupt gene silencing, as well as global hypomethylation causing genomic stability, can lead to tumorigenesis[81].

By comparing DNA methylation in embryonic stem cells and differentiated tissues, two independent groups showed that *de nova* methylation occurs in a fraction of CpG islands of genes responsible for stem cell maintenance, germ line associated and tissue-specific genes during development. Methylated and unmethylated states of CpG islands have been shown to be

stably maintained by semiconservative DNA replication system through cell division[82-84]. Noushmehr *et al* (2010) and Ohgaki and Kleihues (2013) demonstrated that the PN tumors with CpG island methylator phenotype (CpGCIMP) were biologically different that no-GCIMP tumors by exhibiting hypermethylation, copy number variations (CNVs) and longer survival rates. The GCIMP tumors hence represent advanced stages of grade 2 and 3 diffuse gliomas and are therefore considered as secondary GBMs[1,85].

Among several mechanisms involved in glioma development and progression,DNA methylation has been shown to modulate cancer gene expressions. Methylation of CpG islands in the promoter and non-promoter regions of genes has been shown to influence gliomagenesis[74]. Heritable changes in gene expression with no alterations in DNA sequence are referred to as epigenetic changes. The CpG sites present in CpG islands are highly unmethylated but are methylated when present outside CpG islands, including intragenic regions, repetitive sequences, and mobile elements. This site-selective methylation pattern was studied by several groups and has been shown to occur throughout the genome[72,73,86-88]. It is believed that CpG sites present in the intergenic regions are methylated to prevent cryptic transcription which in turn favors gene expression. Genomic stability and host mechanisms are regulated by methylation patterns in repetitive and mobile elements[89]. Bogdanovic and Veenstra (2009) studied methylation of CpG sites in the promoter regions and showed that the binding of a MeCP2, a methyl binding protein or a transcriptional binding protein usually causes transcriptional silencing as a result of promoter methylation[90]. By comparing DNA methylation of embryonic stem cells and differentiated tissues, two independent groups showed that *de nova* methylation occurs in a fraction of CpG islands of genes responsible for stem cell maintenance, germ line associated and

tissue-specific genes during development. Hence DNA methylation plays a major role in the loss of pluripotency and specified cell growth during development[72,73].

Enzyme activities of DNMT3a and DNMT3b have shown to be affiliated with selective binding to CpG sites and are excluded from active promoter regions, enhancer sequences and nucleosomes thereby influencing transcriptional activities of certain genes[91]. In order to maintain the DNA methylation in a semiconservative pattern, mammalian Dnmt1, an enzyme which is readily available during S-phase of the cell cycle is present at the replication foci especially during embryo cleavage phase [9293]. At the replication foci, Dnmt1 preferentially methylates hemimethylated CpG islands in the presence of ubiquitin-like, containing PHD and RING finger domains 1 (UHRF1) that colocalize with Dnmt1 during S phase[94]. Also to the roles of DNA methylation in development, aberrant DNA methylation patterns cause tumorigenesis. Specific hyper and hypo DNA methylations of CPG islands frequently and concurrently occur in human cancers[81]. Promoter-specific hypermethylation which can cause abrupt gene silencing, as well as global hypomethylation causing genomic stability, can lead to tumorigenesis[81].Hypermethylation of CPG islands generally serves as an important early human cancer biomarker in blood and other fluids of the body. Also, Laird (2003) demonstrated that CPG island hypermethylation events can be used as a prognostic marker in recurrent tumors[80].

Noushmehr *et al* (2010) categorized gliomas into a separate group called glioma-CpG Island Methylator Phenotype (G-CIMP), which is characterized by hypermethylated DNA promoter regions. Toyota *et al.* (1999)[95] identified a cancer-specific CpG island hypermethylation phenotype (CIMP) in a subgroup of colorectal tumors, which was further characterized by Weisenberger et al. (2006).[96].In gliomas including human GBMs, the occurrence of CpG island hypermethylation in promoter regions has been reported by many

groups[97,98-100]. Secondary GBMs have been shown to be mostly regulated by promoter methylation than Primary GBMs[101]. Low-grade gliomas that belong to the proneural group had the GCIMP phenotype with unique copy number variations and Isocitrate dehydrogenase 1 (IDH1) somatic mutations. Response to treatments was better in G-CIMP tumors as they are diagnosed at a younger stage. Also, the G-CIMP tumors showed reduced copy number gains of Chr7 also to fewer Chr10 losses. The sample subset had characteristics of proneural subtype and had improved outcomes[85]. In contrast to existing evidence that C-GIMP subtype that has characteristic features of the proneural type with better outcomes, Verhaak *et al.* (2010) demonstrated that the classical subtype responded well to aggressive therapy and had better results than other subtypes[6]. Also, Kim *et al.* (2013)[102]showed that gliomas classified under G-CIMP phenotype with IDH mutations overlapped with classical subtype and in also, within the four glioma types there existed a category of subtype based on survival rates and showed better prognosis. Moreover, the other subtypes that showed poor prognosis showed more heterogeneity and expressed EMT-associated genes, and this pattern was found in several other types of tumors[103]. The transition towards a mesenchymal phenotype is usually aggressive and is caused by transcription factor networks and modifications at the epigenome level[104], which was evident by the fact that the overexpression of EMT occurs only in the primary, but not in the secondary GBMs, or low-grade astrocytoma or in normal brain tissues[105].This study suggests that there is a need for classifying tumors based on both methylation and survival subtypes.

2.7. Cell cycle

The human embryonic stem (hES) and the induced pluripotent stem (iPS) cellsmaintain an undifferentiated state and proliferate indefinitely.Also, the hES and the iPS cells can self-

renew and also develop into any cell type [106]. Reproduction of eukaryotic cell is led by a series of highly regulated events called cell cycle. During S-phase transition of the cell cycle, DNA synthesizing enzymes separate the double helix DNA at specific sites on the DNA called replication origins to complete DNA replication and chromosome duplication processes[107]. Mitotic (M) phase of a cell cycle follows S-phase and consists of a series of events such as chromosomal segregation, nuclear division, and cell division. Intra and extracellular signals help regulate cell cycle by acting between S and M phase and allows a 'gap' phase (G).During the gap phase, cells grow and prepare for the G1-phase[108]. Cells either continue further cell division or leavethe cycle during G1[109]. Cells exit the cycle when they encounter unfavorable conditions, they attain quiescent stage (G0) and normally terminally differentiated or senescent cells of the human body withdraw from the cell cycle[110]. Cancer cells have been shown to undergo repeated cell cycles without exit. Restriction (R) point in a cell cycle enables cells to pass from G1 phase to S-phase and hES and iPS cells normally lack the R point. R point is required for the hES and iPS cells to transform to a phenotypically committed or a differentiated state. However, using a transitory G1/S-phase, pluripotent cells can be regulated to enter S phase using growth factors to transform them to an active phase. Key cell cycle related gene activating events help pluripotent cells to enter G1/S-phase[111]. Mainly, chromatin remodeling genes regulate G_1/S-phase transition and help hES and iPS cells to undergo DNA replication in S-phase. S-phase entry of cells is regulated by nuclear microenvironments (e.g., histone locus bodies) essential for gene expression. During S-phase, following DNA replication, histone octamers or two heterodimers of H4-H3 and H2A-H2B direct newly replicated DNA for chromatin packaging. To accommodate dynamic chromatin remodeling events, pluripotent cells, in general, are interposed between M/G1 and G1/S phases temporarily and therefore, in general, remain in an abbreviated G1-

phase[112]. During self-renewal, hES cells exit mitosis accompanied by the assembly of nuclear microenvironments that favor chromatin decondensation is accelerated[113]. Thus, human hES and iPS cells achieve pluripotency by maintaining an open chromatin configuration.

2.8. Histone H4

Core histones, H2A, H2B, H3, and H4, make up the nucleosome which is the functional unit of chromatin. The nucleosome core is made of an octamer consisting of 2 copies of histones H2A, H2B, H3 and H4 wrapped around by 147 base pairs of DNA[114]. Histones not only package DNA into chromatin, but also regulate gene expression. A recent study showed the N-terminal region of H4 tail regulates chromatin state by controlling imitation switch (ISWI)-family remodeling enzymes, a family of chromatin remodeling proteins that promote gene silencing by regularly placing nucleosomes on the chromatin[114].Medina *et al.* (2012) showed that H4 locus presents a decondensed chromatin environment where TSS are free of nucleosomes rendering the loci hypersensitive for transcription[115].In eukaryotes, replication-dependent genes that encode canonical histone proteins have non-polyadenylated mRNAs[116]. In higher eukaryotes, there are 15 human H4 genes (H4A, B, C, D, E, F, G, H, I, J, K, L, M, N/O, P) with different proximal promoter organizations but translating the same protein. A recent study showed that Histone 3.3 plays a critical role in pediatric gliomas. During the S-phase, post-translationally modified histones are recruited or evicted to meet the demands of a continually remodeled chromatin of mammalian cells[117]. Studies have shown that H4 genes coordinate their expressions with cell cycle progression[118]. For example, following mitosis, when pluripotent cells undergo an abbreviated G1 phase to prepare to enter S phase (G1/S), H4 is recruited for chromatin decondensation[115]. Recent studies have shown that genes expressed in the G1/S phase drive

glioma development and also exhibit promoter hypermethylation [119]. Lower eukaryotes, however, have identical H4 promoter and coding regions that are usually organized as tandem repeats with other histone proteins[115]. Studies have shown that H4 genes coordinate their expression with cell cycle progression; *e.g*., temporal and functional regulation of histone H4/n gene proved to correlate with G1/S phase transition of cell cycle[118]. Among the histone nuclear factors (HiNF-M, -D, and -P), which interact with site II to regulate the S-phase gene transcription, HiNF-P has been shown to actually control H4 transcription with its association with the cyclin E/CDK2-responsive protein $p220^{NPAT}$[120]. RNA polymerase II (RNA pol II) that catalyzes the transcription of DNA to RNA binds H4 loci when H4 transcription factor (HINFP) activates H4. Since proximal promoter sequences vary in H4 genes, the levels of mRNA may also vary among the genes. Holmes et al. (2005), using quantitative PCR (qPCR) and chromatin immunoprecipitation (ChIP), identified variation in the relative contributions of 15 H4 mRNA expressions. Differences in the levels of mRNA expression correlated with varying promoter sequences to total histone pool in normal and tumor-derived cells. Also, Holmes et al. (2005), demonstrated that mRNAs of 11 out of 15 genes that contribute >95% of the H4 mRNA pool are cell cycle synchronized and adapt E/CDK2/$p220^{NPAT}$/HiNF-P signaling pathway[112].

Histone H4 genes didn't respond similarly to cell cycle-driven signaling events, particularly at the G1/S transition phase. H4 subtype mRNA expressions of T98G cells (human glioblastoma cells) from cells synchronized to G1, S, and G2 phases, showed that S phase cells have 5-20 folds higher mRNA expression than those from G1 and G2 phases. Genes, H4/b, H4/d, H4/e, H4/j, H4/n, and H4/o, were found maximally expressed contributing to 76% of the total H4 mRNA pool suggesting that these genes have a role to play in DNA replication and chromatin assembly. Histone H4 genes showed differential expression levels in fetal liver, fetal

colon, and IMR90 (normal diploid lung fibroblast) cells when compared with tumor-derived human cell lines (HCT116 colorectal carcinoma cells, T98G glioblastoma cells, SaOS osteosarcoma cells, and HeLa cervical carcinoma cells). In healthy cells, the percentages of highly expressed genes (H4/d, H4/e, H4/j, H4/n, and H4/o) and modestly expressed genes (H4/a, H4/b, H4/c, H4/k, H4/m, and H4/p) were 55 and 18, respectively, whereas 80 and 9 in tumor-derived cells. Holmes *et al* (2005) suggested that the differences in percentages of expression in tumor-derived cells might be because some of the modestly expressed genes (H4/b, H4/c, H4/k, and H4/m) are silenced in tumor-derived cells and hence produce undetectable levels of mRNA. H4 genes display heterogeneity beyond site II; there are mismatches in genes H4/a, H4/c, H4/k, and H4/l at the consensus H4 subtype-specific consensus element and a nucleotide change in TATA box of H4/m gene. The heterogeneity indeed led to a low expression of these genes confirming that site II significantly influences H4 transcription[120]. Holmes *et al*. (2005) besides using HiNF-P consensus oligonucleotide as a probe, performed competition electrophoretic mobility shift assay (EMSA)[112]. In the EMSA assay, increasing amounts of site II competitors of binding site II showed that mismatches or nucleotide changes in site II inhibited HINF-P binding[121].Among the H4 genes, H4/a and H4/k didn't bind HiNF-P whereas H4/c showed a weak binding. The genes that showed strong binding to HiNF-P were H4/g and H4/p.

To determine the level of enriched promoter sequences at each H4 gene locus compared to the amount of input DNA,Holmes *et al*. (2005) did ChIP assays using HiNF-P and RNA polymerase II (RNAP II) antibodies on T98G cells (human glioblastoma cells) and quantitative polymerase chain reaction (qPCR) on promoter regions of the H4 genes. Surprisingly, H4/b was higher in the input DNA suggesting that H4/b locus may enhance H4 gene expression in human glioblastoma cell line.Also, epigenetic histone marks that favor active chromatin such as

H3K4me3, H3K9ac or H4K12ac, H4K16ac were found specifically upstream of TSS at 5' regions in H4 locus. Furthermore, Ricardo *et al.* (2012)also found active chromatin histone marks at 3' region of H3 and H4 genes supporting mRNA maturation[115].There were no H3K27me3 that marks for transcriptionally inactive chromatin state suggesting that H4 gene is accessible for transcription substantiated by the increased nuclease sensitivity in the locus suggesting that the G1 phase progression in pluripotent cells may be accelerated by active chromatin topology. Histone genes of yeast to human are organized as clusters and have been shown to persist in groups throughout the evolution[122]. Histone gene expression increases to 2-5 folds during S phase of cell cycle and following DNA synthesis, expression increases within 1-3h[123]. There are 74 known histone genes located on chromosomes 6p21 and 1q21 as two major clusters[122,124]. During S-phase, *de novo* synthesis of histones, H2A, H2B, H3, and H4 and linker H1 occurs to pack nascent genomic DNA into chromatin[118,122]. During DNA replication synthesis of high levels of histones namely, H2A, H2B, H3, and H4 and the linker H1 is required for coordinated histone transcription at multiple loci[125]. Histone transcription initiates from compact promotersthat result in intron lacking primary transcripts that subsequently matures to mRNA containing non-polyadenylated mRNA at 3' end[125,126].

2.8.1. Regulation of H4 gene promoters

Following synthesis, histones get extensively modified, and during the S-phase, modified histones are recruited or evicted to meet the demands of a continuosly remodeled chromatin of mammalian cells[117]. In eukaryotes, replication-dependent genes that encode canonical histone proteins have non-polyadenylated mRNAs[116]. Mature mRNAs of histone coding genes are generated from pre-mRNAs by undergoing an endonucleolytic cleavage at their 3' end. The

result is the formation of an evolutionarily conserved stem-loop sequence that replaces conventional poly 'A' tail. Both replication-dependent and non-dependent histone genes have less than 100 nucleotides in the 3' untranslated region (UTR). A 26 nucleotide sequence called the stem-loop sequence is critical for histone mRNA stability[127]. In this region, there are 14 different nucleotides with 2 A's at the 5' end. Fourteen H4 genes in human and mouse have three consecutive CG base pairs in the stem-loop region instead of CUC bases commonly found in other histone genes. Stem-loop binding protein (SLBP) explicitly binds the stem-loop sequence to activate mRNA processing. The stem-loop structure replaces the role of poly (A) tails to regulate mRNAs of replication-dependent genes. High levels of mRNA are maintained throughout S-phase to coincide with the highly replicating DNA, and subsequently the levels go down at the end of S-phase [128]. Interestingly, all H4 genes regardless of their location encode the same protein[129].

2.9. Copy Number Alterations in Stem Cells

Cells capable of self –renewal, initiation and differentiation within a heterogeneous tumor tissue are identified as cancer stem cells. Tumor recurrence after therapy has been shown to occur with the support of cancer stem cells (CSC) that also possess overexpressing genes responsible for EMT, tumor aggressiveness and pluripotency [34,35]. Although in-depth molecular analysis from several studies shows that the primary GBMs are classified into different groups, some tumors have characteristics of more than one subtype[4-6]. The overlap that exists between different subtypes raises questions on the temporal sequence of evolution of the subtypes and its corresponding alterations[5,7,8]. Also, it is unknown whether the subtypes arise from a common precursor. For example, abrupt transition of PN to MES types and subtype mosaicism suggests

that the subtypes might have a different lineage. It is questionable whether the genes that define the subtypes can be used for therapy because most of the mutations that occur in these genes are late events[9].Although the transcriptomes of both GCIMP and non-GCIMP tumors are similar, Noushmehr et al (2010) and Ohgaki and Kleihues (2013) demonstrated that the PN tumors with GCIMP phenotypes were biologically different that non-GCIMP tumors by exhibiting hypermethylation, copy number variations (CNVs) and longer survival rates. The GCIMP tumors hence represent advanced stages of grade two or three diffuse gliomas and are therefore considered as secondary GBMs[1,85]. Copy number variations cause an imbalance to the diploid status of a genetic locus resulting in duplications or deletions of locus ranging between 50bp and 1Mbp.Copy number variations can alter the region or occurrence as well as nearby regions causing positional effects. Also, mainly CNVs can directly adjust the dosage of a gene in general and its fusion with other genes[130].Joining of non-homologous ends of a chromosome, recombination between homologous non-alleles, exchange of transposable elements of pseudogenes, abnormal tandem repeats and errors in DNA replication occurring due to fork stalling and alteration in DNA template can cause CNVs[131]. Development and progression of several human cancers have shown to possess copy number alterations (CNAs)[132]. Ozawa et al. (2014) havebeen shown that PDGFRA gene amplifications and loss of PTEN occur in PN tumors (no-GCIMP GBM subtype) suggesting that overexpression of PDGFA and a PTEN copy loss can cause gliomas. Also to PDGFRA and PTEN, other genes on chromosome seven and ten have also shown to gain and lose copies respectively. This suggests that copy number alterations is a regular event in chromosomes seven and ten[9]. Furthermore differences in copy number alterations between in GCIMP and no GCIMP tumors can help identify driver genes responsible for tumor cell proliferation.

Recent efforts to integrate genomic and epigenomic modifications to understand the mechanism of cancer progression and development delivered favorable results. Although arrays that measure CNAs have larger coverage of more than a million loci coupled with 4kb-2Mb genomic interrogating capacity [133], they lack the ability to screen small regions of the DNA[134]. Currently, arrays are produced to integrate methylation, and SNP analysis and efforts are underway through the International Cancer Genome Consortium (ICGC) and the Cancer Genome Atlas (TCGA) to develop an extensive integrated genomic and epigenomic database. Integrating methylation and SNP arrays is important to investigate the role of tumor heterogeneity on tumor proliferation and development.

Studies that analyzed GBMs have shown overlapping of glioma subtypes based on CNVs. Except for primary GBMs, DNA copy number changes are found to be widely distributed throughout the genome of low-grade astrocytomas and oligodendrogliomas and oligoastrocytomas carrying IDH1/2 mutations [19,36,135] suggesting the importance of detecting copy number variations in predicting patient outcomes. Copy number alterations of tumor suppressor genes namely PTEN and TP53 and CDKN2A/2B, have shown to correlate with EMT mediated tumor progression. The survival rates of patients with no deletions of tumor suppressor genes are longer than those with deletions suggesting that the biology and disease progression can correlate with glioma aggression. For example, high-grade gliomas had frequent deletions of CDKN2A in three independent genome wide association studies.CDKN2A deletions are biomarkers for sequential transfer of proneural type to an aggressive mesenchymal phenotype[9,136,137]. Loss of chromosome 10, PTEN locus, homozygous deletions of CDKN2A, amplification of chromosome 7/7p are some of the characteristics associated with poor prognosis of primary GBMs while better survival groups didn't have copy number alterations [138,139].

3.0. Hypothesis

Methylation is an early developmental event, and intracellular accumulation of 2HG due to mutation in IDH1 causes promoter hypermethylation and subsequently helps to drive gliomagenesis. Differences in promoter methylation between MUT-IDH1 and WT-IDH1 gliomas can provide good insight into the early development of WT-IDH1 gliomas. Since GSCs have both self-renewal and differentiating abilities, targeting GSCs will help identify a subpopulation of GSCs resistant to radiotherapy and chemotherapy.

4.0. Objective

The overall goal of my thesis proposal is to determine whether H4B is an important driver of WT- IDH1 gliomagenesis.

4.1. Specific Objectives

1. To identify drivers of early development especially the role of H4 genein WT-IDH1 gliomas due to differences in promoter methylation status between WT-IDH1 and MUT-IDH1
2. To determine if early drivers of WT-IDH1 gliomas increase GSC cell proliferation rates *in vitro* and *in vivo*
3. To determine if early drivers of WT-IDH1 gliomas modulate epigenomes of WT-IDH1 and MUT-IDH1 gliomas differently

5.0. Materials and Methods

Brain tumor samples (astrocytomas and glioblastomas) wereobtained from surgical resection material acquired from patients undergoing surgery at the Montreal Neurological Hospital on an Institutional Review Board approved protocol. The samples were genotyped for glioma grades at the Montreal Neurological Institute.Glioma stem cell lines (GSC) expressing WT-type IDH1 (84EF) or MUT-IDH1 (OPK 154, 183 and OPK211) were generated from the tumor samples and were genotyped at the McGill University and Génome Québec Innovation Centre. In total, there were 4 GSC lines (1 WT-IDH1 and 3 MUT-IDH1) used for *in vitro* study.

Briefly, tumour resection samples were mechanically dissociated and tumor spheres were established and propagated in Human NeuroCult NS-A Basal media (StemCell Technologies®) supplemented with EGF, FGFb and heparin sulphate. Furthermore, genomic DNA was extracted from the stem cells and submitted for genotyping[140].

5.1. DNA Methylation Analysis from Astrocytomas

Genome wide methylation states on a DNA can be assessed by powerful methylation tools using methylation bead chip *e.g.,* Infinium ®HumanMethylation450. Methylation states can influence candidate gene expression in low grade astrocytomas carrying normal (WT) and mutated (MUT) IDH genes. Analyzing the differences in the quantity and quality of DNA methylation between WT and MUT tumors will provide information on genes affected by methylation on a genome-wide scale. Whole epigenomic alterations in diffuse astrocytomas have yet to be identified. By using high-throughput DNA methylation sequencing, driver genes that are differentially methylated and responsible for glioma stem cell proliferation can be identified. To evaluate and characterize the driver genes that cause an aggressive phenotype between

GCIMP and non-GCIMP astrocytomas, we combined computational and experimental approach in our study. To interrogate CpG DNA methylation, DNA probes were designed to detect the bisulphite mediated conversion of unmethylated cytosine to uracil while methylated cytosines were protected from conversion. The locations of unmethylated cytosines and 5-methyl cytosines at single nucleotide resolution were determined using direct sequencing methods[141].

DNA methylation over thousands of CpG dinucleotides covering more than 480000 probes can be measured conveniently using Illumina Infinium 450 K Human Methylation Beadchip (II450kHMB). This Beadchip is widely used for Epigenome-Wide Association Studies (EWAS)[142-144]. This beadchip delivers a methylation value combining two different probe designs namely type 1 and 2. Type 1 probes have higher dynamic range compared to type 2 and the data are highly reproducible compared to type 1 probes. To avoid enrichment biases towards probe 1, peak based correction (PBC) was a method to normalize both probes for comparing differentially methylated regions[142]. We followed PBC method to measure differentially methylated regions as one single data set to avoid technical variations by combining unsupervised dimensional reduction as well as classification algorithms[145-148]. Also, to avoid pitfalls of PBC, we followed subset quantile normalization methods (SQN) that corrects for poorly defined peaks/modes and avoid type 2 bias which is referred to as Beta Mixture Quantile dilation (BMIQ). This normalization algorithm in addition to correcting type 2 biases, takes control of variations in technique, and type 1 enrichment bias [149]. The methylation data of each probe generated using II450kHMB follows a β-valued distribution between 0 as unmethylated locus to 1 as methylated. The value of β is defined at the ratio between methylated to methylated+ unmethylated values.

$$\beta = \frac{M}{U+M+e}$$

WhereU and M are, averaged bead replicated values of unmethylated and methylated probes and e = 100, is a correction applied at the low total signal intensity. For *e.g*., U+M ~0 after subtracting the background[144,150]. We didn't consider the background subtraction while obtaining the DNA methylation data. Considering these are patient tumor samples, background separation is usually done with normal brain tissue. Surgical resection of normal brain tissue for experimental procedures was not done considering the wellness of the patient.

WT-IDH1 and MUT-IDH1 grade 2 astrocytomas from patients were used in methylation analysis. Methylation states on 12 WT-IDH1 and 9 MUT-IDH1 astrocytomas were compared. Bisulphite converts unmethylated cytosine to uracil while methylated cytosines are protected from bisulphite conversion. The locations of unmethylated cytosines and 5mCs at single nucleotide resolution were determined using direct sequencing methods[141]. Methylation data from each probe was generated using II450kHMB follows a β-valued distribution between 0 and unmethylated locus to 1 as methylated. The value of β is defined at the ratio between methylated to unmethylated values using R statistical programming language (**http://www.r-project.org/**). Mann-Whitney-Wilcoxon Test was used to assess similarity between two independent data samples (IDH WT and MUT-IDH tumors).

5.2. Generation of Stable WT-IDH1 and MUT-IDH1 cells

To further confirm the results of the gene expression levels in the tumor samples, plasmids inserted with IDH1 WT and IDH1 R132 (MUT) were constructed and over expressed in glioma cells (U343) and gene expression levels were quantified using quantitative real time PCR (Biorad ®). GFP sequence from pRRL frame A vector was removed using endonuclease enzyme restriction sites, BamHI and EcoRV and was ligated to plasmid RC210582 carrying IDH1. Therefore, GFP was removed and replaced by cassette carrying IDH1 WT, IDH1R-132 (MUT), and H4 genes (Figure 2). Over expression of IDH1 and H4 genes in U343 (Glioblastoma cell line) and GSCs generated from either MUT-IDH1 or WT-IDH1 tumors were verified using western blots. H4B expression in infected glioblastoma cell lines and GSCs was also verified using real timePCR.

5.3. Protocol to Synchronize Cells

Stem cells (GSC) generated were synchronized to 'S' phase using a double dose of Aphidicolin (0.1 microgram/ml) as previously described [151]. Prior to synchronization to S phase, doubling times of cells were calculated. Aphidicolin treatment was timed accordingly to obtain S phase cells for isolating mRNA for qPCRs.

Phases required:

1. G1/S
2. S

Steps:

Before planning the experiment, doubling time of the cell line was determined.

e.g., BT142 doubles after 120h.

1. Stem cells were plated at a density ~400000-500000 cells/ml and Aphidicolin was added at a concentration of 0.1 µM/ml immediately after plating. Then cells were incubated for about 24h.

2. Following incubation, cells were washed with pre-warmed media (0.5ml) and was allowed to attach with no Aphidicolin for 32 h.

3.Steps 1 and 2 were repeated.

At the end of step 2, G1/S PHASE cells were lysed or pelleted for protein assays or PCRs respectively.To lyse cells,ice cold PBS 1X was used, and followed by centrifugation at 112gfor 5min. PBS was then aspirated and re-suspended using ice cold Np-20. To pellet cells for real time assay, ice cold PBS 1x was used to wash the cells and centrifugedat 112g for 5min

Note: Immediately after step 1, cells can be lysed or pelleted at 6, 12h, or more time points if necessary. To collect cells in S phase,steps 1, 2 and 3 were repeated once.

5.4. Tools to measure copy number variations

Copy number variations can be detected using exome data obtained from Whole Exome Sequencing (Agilent (http://www.genomics.agilent.com/)or using methylation arrays. Genetic mutations are prevalent in the coding regions of the genome, and therefore sequencing exons of the genome (exome) will help identify variations in coding sequences that lead to clinical phenotype. Current methods of exome sequencing are designed to generate exon sequence data from several hundred to millions of parallel DNA fragments. Protocols based on next generation sequencing emphasize enriching DNA libraries with specific coding or regulatory regions of interest. SNP arrays or comparative genome hybridization (CGH) arrays are widely used to detect copy number variations. By measuring fluorescence (from *in situ* hybridization) ratio between target and reference DNAs along the length of each chromosome, copy numbers were

calculated by Agilent (http://www.genomics.agilent.com/). Genome-wide or repetitive regions of interest can be probed using CGH arrays. Bacterial artificial chromosomes (BAC), cDNA or lengthy synthetic oligonucleotides are used to design probes in CGH arrays [134,152]. High-throughput SNP arrays (Illumina (San Diego, CA) (http://www.illumina.com/)) are based on existing SNP information Illumina has additional probes in the non-SNP region of the genome to obtain higher resolution [131,134]. For, e.g., Illumina 1M-Duo has nearly 1 million markers that include both SNP and non-SNP probes. Agilent (http://www.genomics.agilent.com/) probes incorporated custom designed CNV probes ~ 470K in number[153]. Infinium HumanMethylation450 BeadChips have been widely used by ICGC and TCGA to interrogate CNVs and CpG methylation. Infinium methylation arrays, in general, have proven to be sensitive to detect CNVs also to examining methylation states of individual loci. The ChAMP Bioconductor package also allows detection of CNVs along with methylation analysis[154]. Indeed Infinium methylation arrays are SNP arrays, and by using them, methylation and copy numbers can be detected from the same heterogeneous tumor sample or DNA specimen simultaneously [132]. Genomic DNA from WT-IDH1 samplesand GSCs were analyzed using both Infinium HumanMethylation450 BeadChips and Agilent probes. The algorithmic (bioinformatics) tools such as varscan2, DNACopy and GISTIC were combinely used to analyze copy number variations from the exome data

5.5. Steps used for CNV analysis

1: Run varscan2 on WT-IDH1-mpileup and make copy-number call.

A large output file with above threshold setting was obtained. This output file was further processed to find significant amplications and deletions and was used to detect single nucleotide

variations that are usually recurrent in a large cohort. CNV analysis were performed using standard log2 ratios of the population in comparison with reference set as we didn't submit normal set or mutant-IDH1 tumors for comparison.

2: CBS binary segmentation in R programming was used to find segmented region. To merge, mergeSegment.pl (provided by varscan) was used.

3. R programming language coding was used to analyze CNVs.

4.The output file was used to mergeSegments.pl (from Varscan2).

5.Varscan2 and GISTIC were used to further find recurrent CNVs among the samples.

GISTIC uses output from Varscan (copy numbers) and therefore the output from Varscan was fed to GISTIC. A segment file and marker position was created to run GISTIC.

6: Generation of output file (Appendix 1 (spread sheet attached).

Results were presented in a spreadsheet (260190294_Ravishankar_Palanisamy_IPN_supplementaryinfo.xls) shows large scale deletions and a fewer unchanged regions of the genome.

5.6. Quantitative Polymerase chain reaction (qPCR)

Copy number variations detected at single or multiple loci can be replicated or validated using methods such as quantitative PCR (qPCR), paralog-ratio testing (PRT) and molecular copy number counting (MCC). In qPCR, copy numbers are estimated by comparing threshold cycles (Ct) between target gene and reference genes and differences in the threshold cycles (ΔCt) are generated. Quantitative PCR method is often widely used to validate computationally detected CNVs on a genome wide scale. In PRT method, to probe for differences between CNV locus and dispersed repeats, a pair of primers is being used to scan sequence repeats[155]. In MCC, limited

dilution DNA aliquots are used to test (by PCR) target sequences. Multiple target sequences are compared by counting the number of aliquots that test positive within original DNA sample under interrogation. To detect CNVs in high resolutions, MCC can be used as it is widely adapted to scan particular regions of interest or probe scattered loci of interests[156]. Several other PCR based approaches are also used to validate computed CNVs detected from array platforms and some of them include, multiplex PCR-based real-time invader assay, quantitative multiplex PCR of short fluorescent fragments, multiplex amplicon quantification,multiplex amplifiable probe hybridization and multiplex ligation-dependent probe amplification[157]. Further validation of CNVs are being done using QPCR. The primer sequences are listed in Tables 1 and 2.

5.7. Westen Blot Analysis

Asynchronously grown cells or S-Phase arrested cells of total 2.5 X 10^7 were collected in 15ml Falcon® tubes. The cells were then pelleted by centrifugation at 112g and were frozen at − 80 °C for westernblotting. For immunoblotting,pelleted cells were dissolved in 200µl of sample buffer(50 mM Tris-Hydrochloric acid [pH 6.8], 2%Sodium Dodecyl Sulfate (SDS), 12.5 mM Ethylenediaminetetraacetic acid, and 10% glycerol).Using a 15% SDS-Poly acrylamide gel electrophoresis, proteins in the cell extracts were separated and transferred to a nitrocellulose membranes using a semidry system. Using a PBST buffer (10 mM Na2HPO4, 137 mM NaCl, 2.7 mM KCl, 1.76 mM KH2PO4, and 0.05%Tween-20) containing 5% skim milk, the membranes were incubated overnight at 4°C. Then, the membranes were incubated with anti-actin antibody (Abcam) or H4 monoclonal antibody (Pierce) in PBST containing 1% skim milk at 4°C overnight. The membranes were then washed three times with PBST and were incubated with horseradish peroxidase-conjugated secondary antibodies (Millipore) in PBST containing

1% skim milk at room temperature for 1 h. Using ImmuneStar LD (Wako), the signals were visualized and detected using Chemidoc MP imaging system (Bio-Rad).

6.0. Results

6.1. DNA Methylation Analysis from Astrocytomas

To identify drivers of early development of WT-IDH1 gliomas, using DNA methylation array, differences in promoter DNA methylation beween WT-IDH1 and MUT-IDH1 gliomas (6 WT-IDH1 and 4MUT-IDH1) were identified. Results were displayed in a heat map which showed that the MUT-IDH1 tumors were hypermethylated compared to the WT-IDH1 tumors (hypomethylated) (Figure 1). Mean β-values from WT and MUT tumors were first determined, and the differences between them were sorted in descending order.MUT-IDH1 and WT-IDH1 tumors clustered separately confirming that they were significantly different in their methylation states. From the annotated list, TET-1, OSBP2, RASA3, LRRC, BCAT1 and Histone H4 B (H4B) were picked and validated using quantitative real time PCR. Therefore, at least one among the significantly methylated genes (between WT and MUT –IDH1 tumors) was believed to be a candidate gene influencing tumor proliferation in low-grade astrocytomas. Since methylation causes aberrant gene expression, it is, therefore, important to quantify the gene expression levels of the chosen candidate genes in WT and MUT tumors. Hence gene expression was quantified using CFX-Connect™ Real time system. The results showed that the expression levels of H4 were significantly reduced in the MUT-IDH1 compared to the WT-IDH1 tumors (Figure 2). Also, by real time PCR, H4B detected by methylation analysis was significantly upregulated among other H4s (supplementary Figure 1). To further confirm the results of the gene expression levels in the tumor samples, plasmids carrying WT-IDH1 and MUT-IDH1 were constructed

(Figure 3).The results further confirmed that H4 mRNA expression was downregulated in the cells overexpressing MUT-IDH1 compared to the WT-IDH1 gene constructs (Fig 4).

6.2. Quantitative PCR results

Since methylation regulates gene expression, levels of expression were quantified using real time PCR (CFX-Connect™ Real-time system). In real time PCR, gene expression is estimated by comparing threshold cycles (Ct) between the target gene and reference genes and differences in the threshold cycles (ΔCt) are generated.InterestinglyH4Balone significantly differed between WT-IDH1 and MUT-IDH1 in many samples. To validate methylation array results, we compared H4B mRNA expression from a different set of WT and MUT IDH1 astrocytomas and found that indeed H4B mRNA levels are high in 8 of 10 WT IDH1 astrocytomas compared to 2 of 11 MUT IDH1astrocytomas (Figure 2).

6.3. *In vitro* glioma stem cell analysis

To determine if drivers of WT-IDH1 gliomas increase GSC cell proliferation rates *in vitro*, real time PCR analysis was performed in GSCs over expressingH4B. Glioma cell line was infected with viruses encoding WT-IDH1 or MUT-IDH1. The percentage of infected cells were estimated to be more than 90%.We tested if H4B expression was differentially expressed in cell cycle synchronized WT-IDH1 and MUT-IDH1 human derived GSCs. Both Histone4 (total pool) protein (Figure 5a, 5b, 5c) and H4B mRNA (Figure 6) expression were elevated in WT-IDH1 compared to the MUT-IDH1 GSCs. In comparison to the control and MUT-IDH1 infected cells, WT-IDH1 infected cells expressed ~5-fold more H4B mRNA as well as protein supporting a link between IDH status and H4B expression. Thus, H4B was highly elevated in both the GSCs and differentiated cell populations of WT-IDH1 gliomas.Western blotting against H4 antibody was

done to detect basal levels of H4 (Figure 5b) andalso to confirm maximal H4 protein accumulation in 'S' phase (Figure 5c).

6.4. Cell proliferation rate

Cell proliferation rate was calculated using doubling times of individual stem cells. By calculating cell doubling time of WT-IDH1 or MUT-IDH1 infected cells, we found that the MUT-IDH1 GSCs infected with H4B propagated more rapidly than the control infected GSCs (Figure 7). Alteration in cell proliferation rate suggests that H4B overexpression can affect self-renewal and stemness properties of MUT-IDH1 GSCs.

6.5. Modulators of WT and MUT-IDH1 epigenomes

R)-2HG has a significant role in MUT-IDH1 human cancers, and by inhibiting (R)-2HG in MUT-IDH1 derived GSCs, increase in H4B expression was expected. We expected a decline in the H4B mRNA expression probably due to induced hypermethylation. To determine if early drivers of WT-IDH1 gliomas modulate epigenomes of WT-IDH1 and MUT-IDH1 gliomas differently,(R)-2HG at 100mM concentration was used to culture WT-IDH1 GSCs. The GSCs were passaged for 15 times, and during each passage, less than a microgram of rmRNA was isolated and reverse transcribed to quantifyH4B mRNA expression. Results from R-2HG culturing of WT-IDH1 GSCs showed a slow and gradual decrease in H4B expression after 15 cell passages (Figure 9).

Intratumoral depletion of (R)-2HG was achieved by culturing MUT-IDH1 derived GSCs in 450mg/kg of AG1598 for 48h (Figure 10). AG5198 was previously demonstrated to block, (R)-2HG in TS603 glioma cells[43].By depleting (R)-2HG,H4B expression was expected to

increase in MUT-IDH1 GSCs. As expected, H4B expression significantly increased in MUT-IDH1 GSCs after 48h (Figure 10).

6.6.Copy number Variations from Whole exome sequencing

Copy number variations obtained from whole exome sequencing of WT-IDH1 tumor tissues showed variations in several genes indicative of the tumor phenotype (260190294_Ravishankar_Palanisamy_IPN_supplementaryinfo.xls). However, DNA isolated from GSCs failed in the sequencing process. Further steps are underway to setup ideal conditions to isolate DNA from the GSCs that suit the standards for whole exome sequencing. Further validation is necessary to identify potential biomarkers of WT-IDH1 tumors. Especially genes regulating cell development and cell cycle may be important for coordinated H4B regulation in WT-IDH1 tumor proliferation.

6.6. Xenograft model of brain cancer

To determine if elevated levels of H4B induces and proliferates gliomagenesis *in vivo,* we assessed H4B mediated proliferation rate in a xenograft model of brain cancer. We have shown that elevated H4B expression increases the rate of GSC proliferation *in vitro*. To determine if augmented H4B leads to increased cancer proliferation in a human derived GSC xenograft model of brain cancer, 100000 control infected or H4B infected MUT-IDH GSCs were successfully implanted into the basal ganglia of ten NSG mice per cohort (Figure 11). Analyzing these mice will be carried out as a separate study.

7.0. Discussion

Candidate genes that contribute to glioma progression and epigenetic silencing of oncogenes were previously studied [158].Secondary GBMs develop from low-grade astrocytomas[3]. Methylation states can influence candidate gene expressions in low-grade astrocytomas carrying normal (WT) and mutated (MUT) IDH genes. Analyzing the differences in the quantity and quality of DNA methylation between WT and MUT tumors provides information on genes affected by methylation on a genome- wide scale. The results from the methylation array show WT and MUT IDH1 tumors segregated into two different groups differing in methylation of promoter regions. Similar to our results, the occurrence of CpG island hypermethylation in promoter regions has been reported by many groups[97,98-100].Elevated H4B mRNA expression in differentiated WT-IDH1 tumors suggest that H4B can transform an undifferentiated tumor stem cell to a differentiated phenotype. The result correlates with an earlier study[112] where they demonstrated H4B gene promoter regions greatly enhanced gene expression in a glioblastoma cell line. The differences in H4B expression within WT tumors (Figure 3) could be due to tumor heterogeneity.

Earlier studies that profiled MUT-IDH1 human oligodendroglioma cells showed altered levels of amino acids, glutathione, choline derivatives, and tricarboxylic acid (TCA) cycle intermediates and metabolites upsetting cell cycle metabolism of human oligodendroglioma cells[159]. Also, culturing of human oligodendroglioma cells in (R)-2HG produced similar effects[41]. As discussed in earlier sections, chromatin modifying, 2-OG dependent enzymes were shown to be inhibited by (R)-2HG especially, histone demethylases and 5′-methylcytosine hydroxylase that help regulate nucleosomal complexes including histone proteins[160,161]. In our

study, H4B expression levels were downregulated and this phenomenon correlates with previous studies where histone modifying complexes were inhibited in the presence of 2HG.

Coding and non-coding sequences in a human genome have ~3 × 109 nucleotide bases out of which 1% bases (30 Mb) code for proteins. Sequencing whole genome provided only 10% genomic information which is insufficient genetic information about disease-causing mutations. Genetic mutations are prevalent in the coding regions of the genome, and therefore sequencing exons of the genome (exome) will help identify variations in coding sequences that lead to clinical phenotype. Current methods of exome sequencing are designed to generate exon sequence data from several hundred to millions of parallel DNA fragments. Johnson et al. (2014)[162], used whole genome exome sequencing to confer mutations in the primary tumor samples. The study found that the tumor initiation site possesses recurrent tumor initiating cells which had 50% mutations in slightly less than half of recurrent tumor specimens. The mutations in the tumor initiation site have been shown to evolve into distinct clonal subpopulations. However, the contribution of microenvironment on these recurrent tumors initiating cells is not well studied. Current protocols based on next generation sequencing emphasize enriching DNA libraries with specific coding or regulatory regions of interest. With limited scientific literature on the H4B function that explains a role of H4 in open chromatin state, we would like to know whether elevated H4B expression through chromatin modifications drive a malignant phenotype. Since eliminating H4B expression can be lethal to our study models, dysregulated chromatin modifications can be assessed by determining copy number variability (CNV) and mutagenesis on MUT-IDH1 GSCs overexpressing H4B. Using next generation sequencing techniques *e.g.,* whole exome sequencing (WES), copy number variations and mutations from stable versions of GSCs lines can be assessed. Our results showed different phenotypic traits in tumors which are

heterogeneous in nature. Also, this suggests that they may possess subtle areas of genetic and epigenetic variations as described previously[163].

Tumor heterogeneity in gliomas varies spatially and temporarily within a tumor mass. Spatial tumor heterogeneity complicates histopathological classifications. Sottoriva et al. (2013)[8] used the fluorescence guided method to collect spatially distinct tumor fragments from 11 glioblastomas. This study demonstrated that a single tumor mass carried multiple transcriptional profiles. Changes in base level substitutions of specific *EGFR*, *PIK3CA*, *PTEN*, and *TP53* genes over time from diagnosis to the second recurrence have been found on seven primary GBMs suggesting the influence of temporal heterogeneity in tumor development[54]. In our study, overall H4 expression was temporally regulated (Figure 5C). It is possible mutations may occur in H4 genes from diagnosis to the second recurrence in WT-IDH1 tumors.

In contrast to existing evidence that C-GIMP subtype has characteristic features of PN-type with better outcomes, Verhaak *et al.* (2010) demonstrated that the classical subtype responded well to aggressive therapy and had better results than other subtypes[6]. Also, Kim et al (2013)[102] showed that gliomas classified under G-CIMP phenotype with IDH mutations overlapped with classical subtype. Furthermore, within the four glioma types there existed a category of subtype based on survival rates that correlated with prognosis. Other subtypes that showed poor prognosis showed more heterogeneity and expressed EMT-associated genes, and this pattern was also found in other cancers[103]. The transition towards a mesenchymal phenotype is usually aggressive, and modifications occur at epigenome level[104]. Aggressive phenotype was evident by fact that the overexpression of EMT genes are expressed only in the primary, but not in the secondary GBMs or low-grade astrocytomas or normal brain tissues[105]. Our methylation array results suggest thatH4B mediates transition of G-CIMP to a mesenchymal phenotype. In

MUT-IDH1 GSCs infected with a lentivirus encoding H4B, we observed a differentiated or transitional phenotype where the GSCs attach to the culture dish compared to uninfected spheroid phenotype (Figure 8).

Prolonged treatment of WT-IDH1 GSCs with R-2HG lowered the expression of H4B suggesting that 2HG can regulateH4B expression through IDH1 mutational status. Mutant IDH tumors mainly accumulate R-2HG in the range of 2mM to 30mM concentration. Continuous and high levels of cell permeable intracellular 2HG transformed TF-1 cells to an IDH1 mutant phenotype by blocking differentiation of TF-1 cells[41]. Also, R-2HG administration was able to inhibit differentiation of murine 3T3-L1 fibroblast and immortalized myeloid progenitor cells[41,42] further confirming the oncogenic role of R-2HG. Withdrawal of cell permeable R-2HG or inhibiting IDH1 mutant enzyme activity by a potential inhibitor caused differentiation of TF-1 cells and restored characteristics of suppressed IDH mutant phenotype [41,43].We observed elevated levels of H4B in MUT-IDH1 GSCs following AG5198 treatment. Hypermethylation of promoter regions of genes in MUT-IDH1 tumors (the CpG island methylator phenotype) arises as a result of the action of IDH1 mutation generating a neoenzyme that accumulatesR-2HG in cells[164].The MUT-IDH1 inhibitor, AG5198 has been previously shown to inhibit mutant IDH1 enzyme activity[43]. We examined the effect of AG5198 in modulating H4B expression, and we found that treatment of human-derived MUT-IDH1 GCSs with AG5198 lead to a significant elevation of H4BmRNA expression. Elevated H4B expression under depleted 2HG conditions emphasize that mutant IDH1 status regulates H4B expression.

In summary, we found elevated H4Bexpression in WT-IDH1 gliomas compared to MUT-IDH1 gliomas. Also, there was an increase in H4B expression found in WT-IDH1 human derived GSCs and in differentiated cancer cells. An increase in proliferation rate and epithelial to

mesenchymal transition (phenotype) was observed in MUT-IDH1 GSCs following overexpression of H4Bsuggesting that elevated H4B expression may trigger early gliomagenesis in WT-IDH1 tumors.

Our results correlate with the clinical nature of WT-IDH1 and MUT-IDH1 gliomas especially WT-IDH1 being more aggressive than MUT-IDH1 gliomas.Identifying early driver genes of gliomagenesis will become useful to target GSCs that encompasses both self-renewing and differentiating cancer cells. To substantiate, for example, an IDH1 mutation occurs early, and prevails in GSCs of an MUT-IDH1 tumor mass. Ongoing clinical trials have shown a significant effect on targeting IDH mutant neoenzymes. However, existing gene therapies that target WT-IDH1 gliomas or mutations of WT-IDH1 glioma,*e.g*., EGFR amplifications are not effective because downstream oncogenic mutations are not expressed in all cells within a tumor and therefore*e.g.,* cells that don't express EGFRv3 continue to proliferate leading to treatment failures.

Copy number variations cause an imbalance to the diploid status of a genetic locus resulting in duplications or deletions of loci ranging between 50bp and 1Mbp.Copy number variations can alter the region or occurrence as well as nearby regions causing positional effects. Joining of non-homologous ends of a chromosome, recombination between homologous non-alleles, exchange of transposable elements of pseudogenes, abnormal tandem repeats and errors in DNA replication occurring due to fork stalling and alteration in DNA template can cause CNVs[131]. In our results, heterogeneous tumors presented themselves with different phenotypic traits probably due to subtle areas that possess both genetic and epigenetic variations.

Neurons and various types of glial cells namely, astrocytes, oligodendrocytes, and ependymocytes are mature differentiated cells of the central nervous system (CNS) that develop

from stem cells expressing Nestin (a biomarker for CNS stem cells). In early embryonic development, cerebral ventricles are lined by pseudostratified neuroepithelial tissue which gives rise to CNS stem cells[51]. Radial glial cells are neural stem cells that arise from neuroepithelium[44,45]and express BLBP (brain lipid binding protein) which marks a pathological profile[44,45,165]. During adulthood, early progenitors disappear but neural stem cells which are self-renewing and multipotent, concentrate and reside in the subventricular zone (SVZ) (lines the lateral ventricle), in the dentate gyrus (lies within hippocampus called the subgranular zone (SGZ)), in the sub-cortical whitematter region, throughout the cortex, the corpus callosum, and in the periventricular whitematter regions (collectively referred to as Neuraxis)[46].

8.0. Conclusions and Future Directions

Neural stem lines infected with lentivirus encoding H4B GSCs was stereotactically injected into mouse dentate gyrus to monitor glioma development. To verify targeting to mouse dentate gyrus, lentivirus carrying a GFP cassette was used (Fig 11). Using intravital imaging system tumor development was assessed. Experiments are in progress to do bilateral injections of nearly 40 mice at dentate gyri to image for tumor development at regular intervals over a span of 1.5 years. Our results show that we were able to target dentate gyrus (SGZ) and further studies are underway to investigate the role of H4B overexpressing GSCs in tumor development in the SGZ (Fig 11).

Tumor cell proliferation rate, sphere-forming ability, mesenchymal morphology, GSC markers and tumor- generating ability of GSCs were propose in a mouse xenograft model overexpressing H4B. Steps are underway to estimate overall survival, or the time at which an animal develop signs of disease as defined by animal care standards. Once these endpoints have

been reached animals were proposed to be sacrificed and harvested tumors will be histopathologically analyzed to determine tumor size, assess invasion, and the presence of high-grade features such as necrosis and endothelial proliferation.

Since gliomas are heterogeneous in nature, GSCs derived from any tumor would have a mixed population of primordial and differentiated cells. Using RNA sequencing, plans are in progress to derive a hierarchical gene expression signature from neural stem cell clones to differentiated GSC clones. The transcriptome results that will primarily encompass, gene and fusion gene expression signatures will be verified using real time PCR and Fluorescent in situ hybridization methods. Furthermore, using flow cytometer, steps are in progress to quantify the proportion of cells expressing stem cell biomarkers at different cell cycle phases. Also, experiments are in progress to determine stem cell markers at the S-phase under H4B overexpression.

To determine the frequency of elevated H4B expression in WT-IDH GSCs and gliomas, 100 WTGSCs from GSC bank at the Montreal Neurological Institute are being used for analysis. The GSCs are cell cycle synchronized and are being quantif for H4B expression using QPCR. Furthermore, steps are in progress to compare WT GSC H4B expression to at least 20 MUT-GSCs. Also, to determine the frequency of elevated H4 in WT-IDH gliomas, a large cohort of patient WT-IDH tissue microarrays (TMAs) are being tested. Since all H4 isoforms have the same amino acid sequence, we cannot distinguish isoforms at the protein level. We have clarified the optimal conditions for H4 immunolabeling of our TMAs. These GSC lines and TMAs were obtained from Montreal Neurological Hospital patients and clinical annotation was performed.

Using GSC lines and TMAs, steps are in progress to correlate of progression-free survival and overall survival with H4B expression.

By injecting isolated fusion proteins to the mouse dentate gyrus, gliomagenesis and tumor proliferation rate can be tested *in vivo*.Oligonucleotide probes can be used to detect mRNA transcripts before and after mRNA splicing. By using different fluorescent labels, simultaneous detection of many genes between tumor cells and tumors can be achieved. Also, using a fluorescent microscope, the method can be use to probe FFPE samples routinely in a clinical setting to distinguish neoplastic from non-neoplastic cell types. In small biopsies from diffuse gliomas, fluorescent *in situ* hybridization (FISH) has been used to identify diagnosis and prognosis biomarkers such as 1p19q co-deletion, *EGFR*, *PTEN10q*, *CDKN2A*[166]. Unique cells in a population have shown to bias average expression level of a population, and also, the average expression level doesn't represent the particular cell population under study[167]. Therefore it is possible to detect mature mRNA by incorporating bromouridine (BrUTP) into single living cells[167] or RNA can also be visualized in frozen sections to compare the spatial orientation of expression patterns across different tissues[168,169]. Using FISH RNA transcripts, it has been shown that nearly1000 fold levels of mRNA from single cells can be estimated[170].

Gene transcriptional networks have shown to control the transition of neural stem cell to a mesenchymal phenotype which determines the aggressiveness of a malignant glioma [104]. Carro *et al.* (2010) have shown that the two transcription factors C/EBPβ and STAT3 act as synergistic initiators and primary regulators of reprogramming neural stem cells to mesenchymal lineage. Expression of C/EBPβ and STAT3 correlate with a poor clinical outcome and suggests a need for a regulatory mechanism to control this abnormal phenotype in glioma. Overexpression of genes, MMP9, TGFB2, FGF1, SERPINB1 have been shown to induce the EMT phenotype

particularly in older patients suggesting that genetic and epigenetic possibilities accumulate over time leading to poor prognosis compared with younger patients with similar overexpression of genes[171].

By using chromosome immunoprecipitation assay, the DNA binding sites of H4 and other candidate genes can be interrogated. These DNA binding sites could be potential transcriptional regulation sites on IDH1 which could influence apoptosis and or survival rates in patients with IDH1MUTs. Fluorescence recovery after photobleaching (FRAP)can be used to detect IDH protein interactions with candidate proteins. Primary brain cancers encompass 80% of gliomas with an incidence rate of 6-8/10000. Identifying early drivers of gliomagenesis will help to develop targeted glioma therapies.To deliver a personalized gene-targeted therapy based on characteristics of a tumor tissue, we believe understanding the role of H4B in tumor proliferation, and the mesenchymal transition would pinpoint nature of aggressive gliomas ingeneral.

CRISPR technology will be used to delete MUT-IDH1 expression to evaluate the H4B expression in GSCs. Previously FGFR3-TACC fusion, which occurs in less than 3.5% WT gliomas, has been shown to initiate gliomagenesis[172]. Using bioinformatic tools, H4 fusion proteins can be isolated and validated.

9.0. Figures

Figure1. Heat map of agglomerative hierarchical clustering of WT and MUT-IDH1 tumors using average methylation (β) values. MUT-IDH1tumors are hypermethylated (dark red) compared to hypomethylated (light blue) WT-IDH1 tumors. Methylation profile classes are stacked in rows separated by tumor numbers (T1 -T21), and class height corresponds to the number of samples in each class. Class methylation at each CpG locus is a mean of methylation for all samples within a class. Each row represents a sample and each column represents a CpG locus.Mann-Whitney-Wilcoxon Test was used for statistical analyses. Statistical significance was determined by $p<0.05$. *$p \leq 0.001$ ***$p \geq 0.05$.

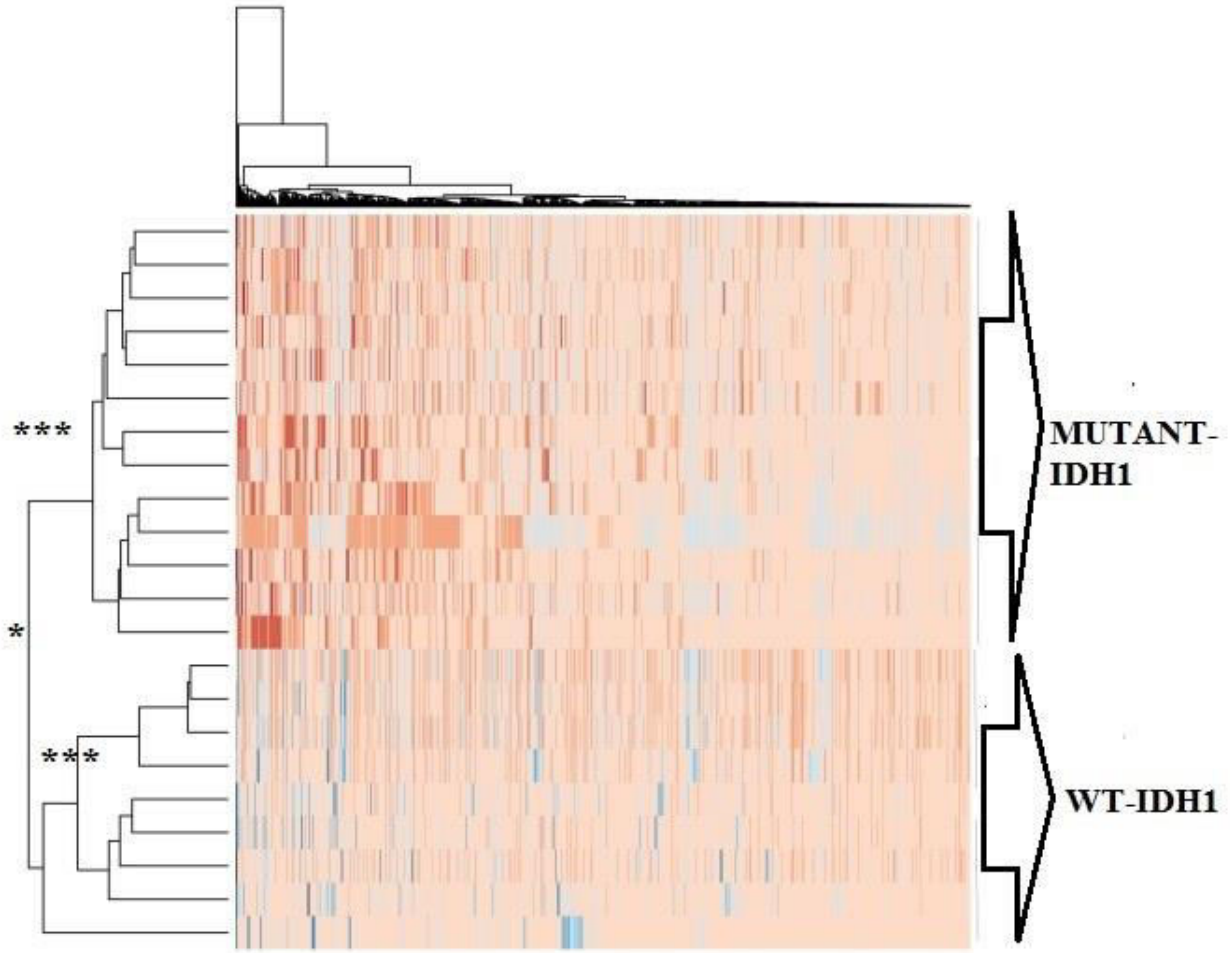

***
*
***
MUTANT-
IDH1
WT-IDH1

Figure 2. H4B mRNA expression levels in MUT and WT-IDH1 tumors. H4B levels were normalized using the mean delta CT values of the reference genes, TBP, B2M and ARF-1 genes. Wilcoxon Two-Sample Tests was used for statistical analyses. Statistical significance was determined by $p<0.05$. *$p \leq 0.001$ ***$p \geq 0.05$.

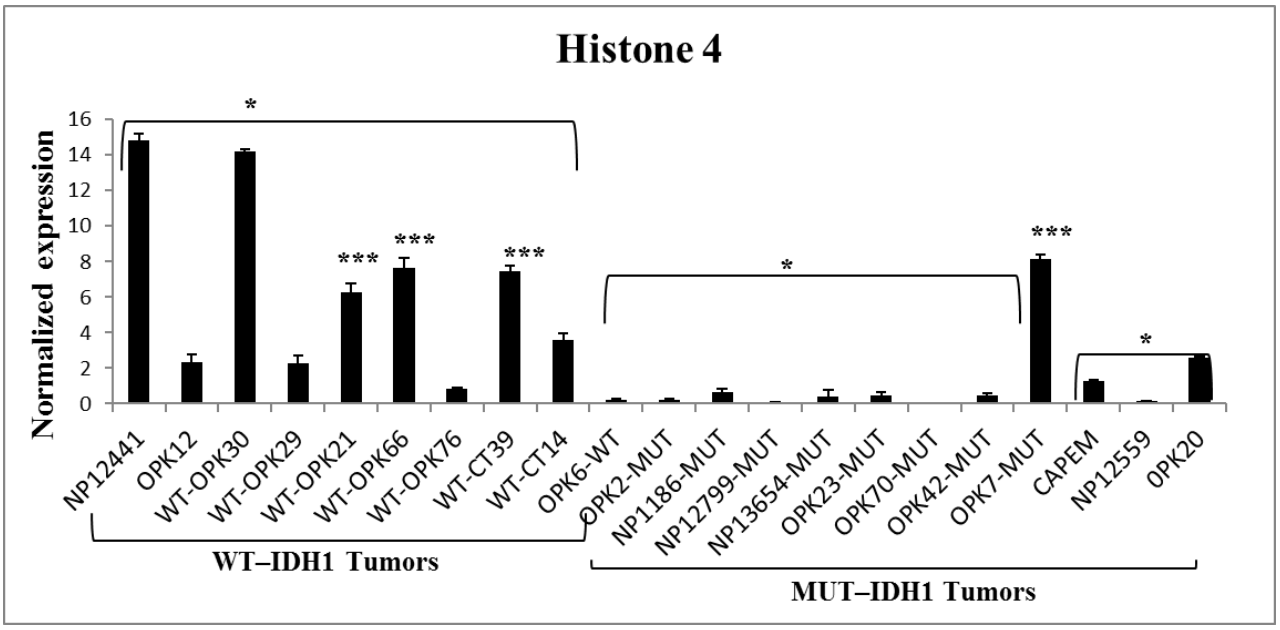


Figure 3. Plasmid showing IDH1 insertion site.

(Figure generated by Origene® technologies, USA and reproduced with permission)

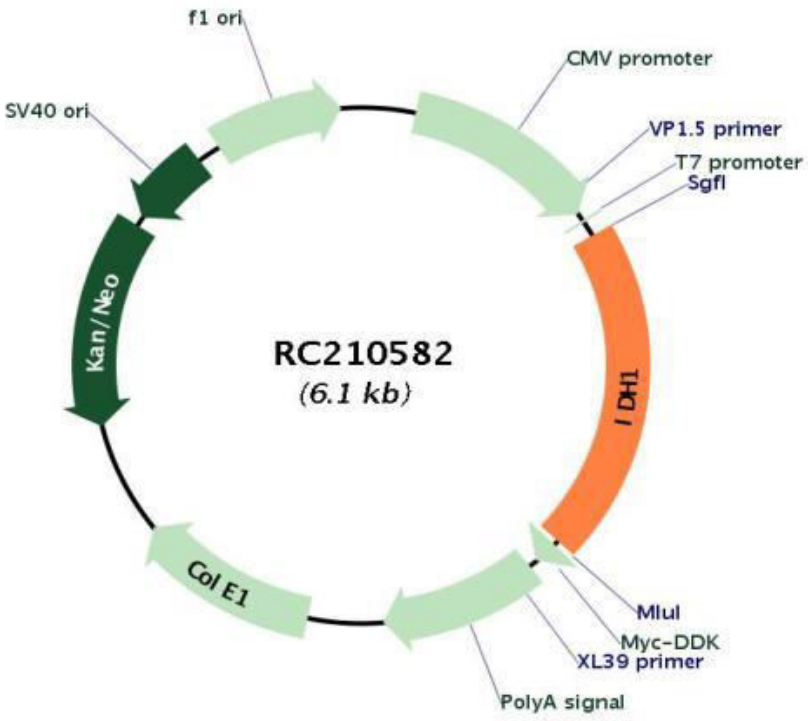


Figure 4. H4B mRNA expression levels over expressing MUT and WT-IDH1 in a glioblastoma cell line (U343).Wilcoxon Two-Sample Tests was used for statistical analyses. Statistical significance was determined by $p<0.05$. *$p \leq 0.001$ ***$p \geq 0.05$.


Histone 4
Normalized expression
6
5
4
3
2
1
0
*
***
Control
Over Expressed WT-IDH1
Over Expressed R132H-IDH1

Figure 5a. Total H4 expression in S phase WT-IDH1 and MUT-IDH1 human derived GSCs. There were three MUT-IDH1 GSCs and one WT-IDH1 GSCs isolated from the tumors and were genotyped and characterized. Anti-H4 antibody (Pierce) that we used generated a 11 kDa band in the western blots. The 17kDa band was a non-specific band. Wilcoxon Two-Sample Tests was used for statistical analyses. Statistical significance was determined by $p<0.05$. *$p \leq 0.001$ ***$p \geq 0.05$.

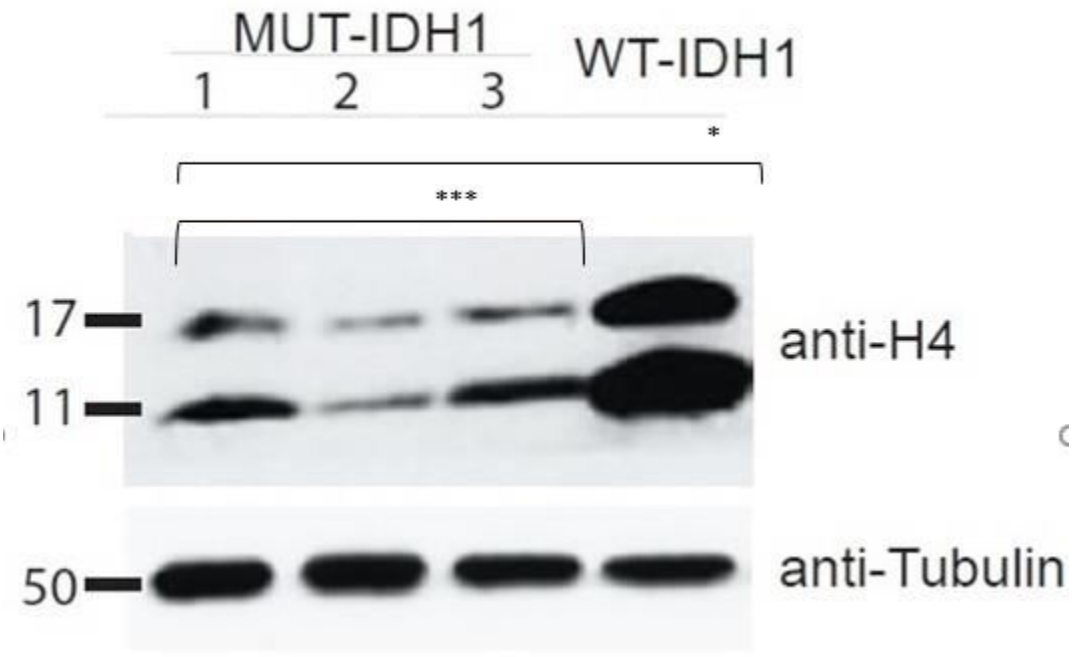


Figure 5b. Western blots showing basal H4 levels before cell synchronization in brain tumor derived stem cells (46EF, MNI, 48EF, 50EF, 84EF, 25M, 12M, OPK61, OPK49 and BT142). Tubulin antibody was used as a loading control for samples. Anti-H4 antibody (Pierce) that we used generated a 11 kDa band in the western blots. The 17kDa band was a non-specific band.

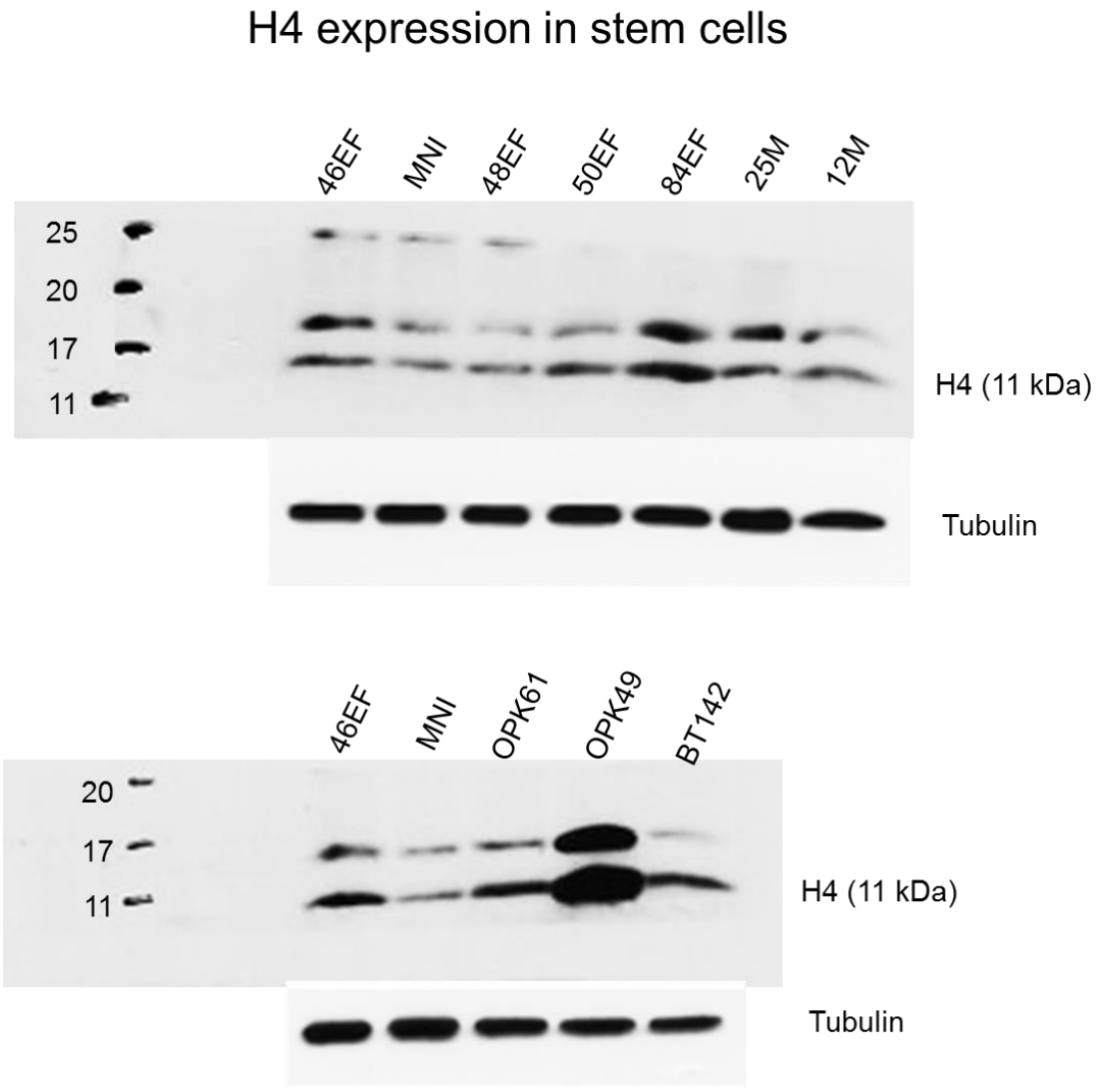


Figure 5c. **Western blots showing H4 levels before and after cell synchronization in brain tumor derived stem cells (84EF, 25M, and BT142). Tubulin antibody was used as a loading control for samples. Maximal H4 accumulation was seen after 12h after synchronization using Aphidicolin a 0.1 µM/ml**

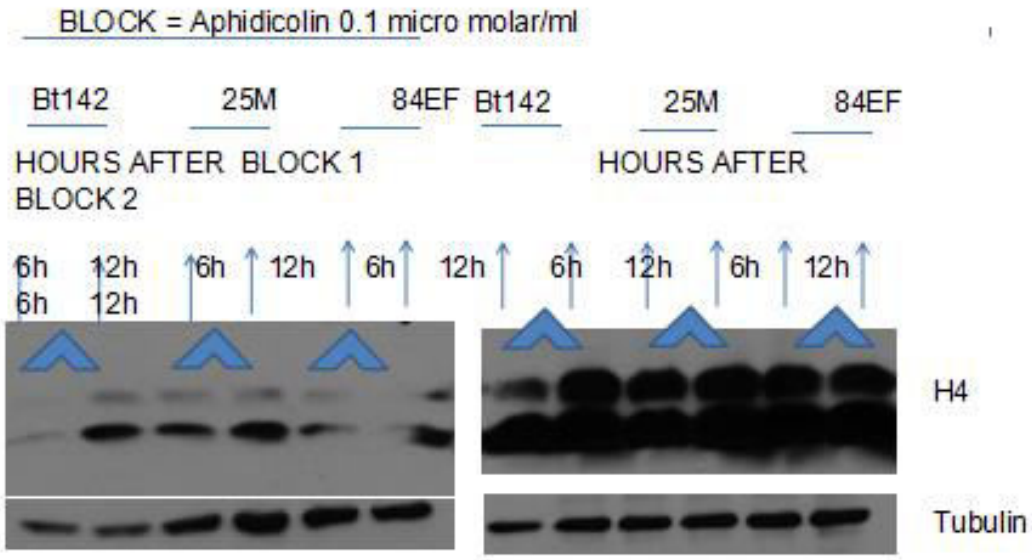


Figure 6: H4B mRNA expression in S phase WT-IDH1 and MUT-IDH1 human derived GSCs.H4B levels were normalized using the mean delta CT values of the reference genes, TBP, B2M and ARF-1 genes. Wilcoxon Two-Sample Tests was used for statistical analyses. Statistical significance was determined by $p<0.05$. *$p \leq 0.001$ ***$p \geq 0.05$.

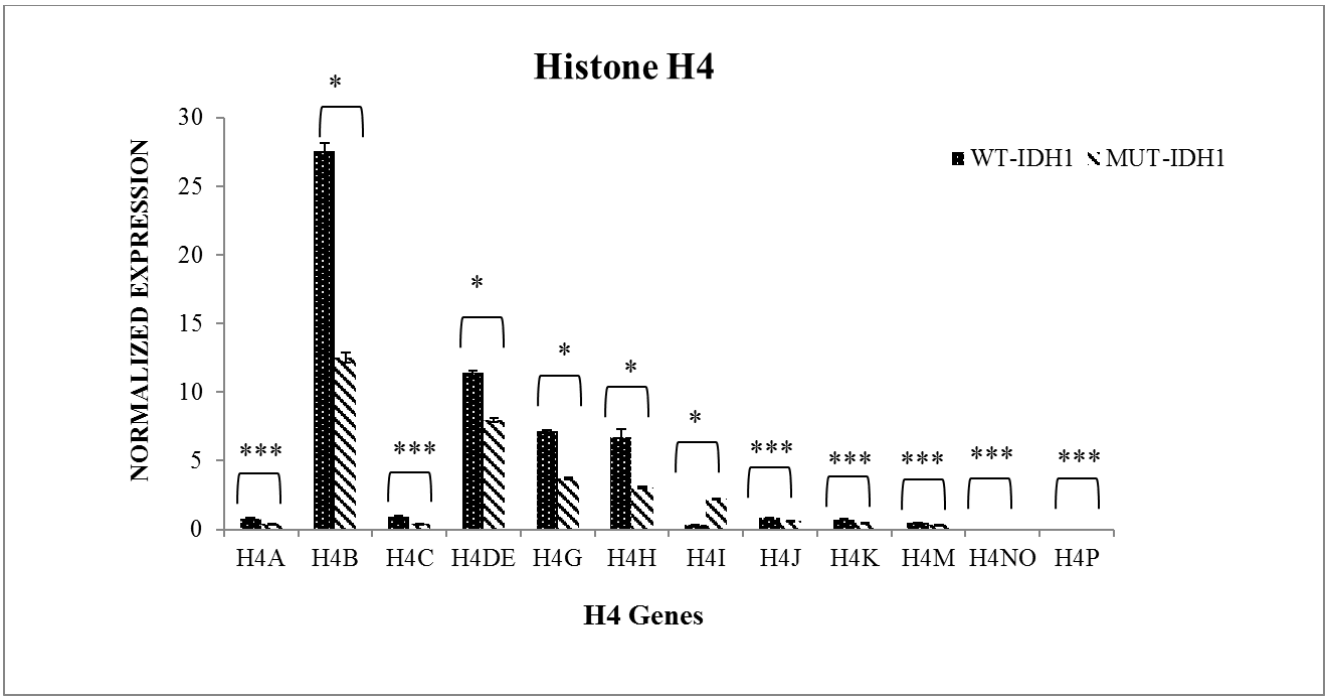


Figure 7: Proliferation rates of MUT-IDH1 and WT-IDH1 GSCs overexpressing H4 measured 8 days after plating. A total number of $1x10^5$ MUT-IDH1 and $2x10^5$WT-IDH1 GSCs were plated. Wilcoxon Two-Sample Tests was used for statistical analyses. Statistical significance was determined by $p<0.05$. *$p \leq 0.001$ ***$p \geq 0.05$.

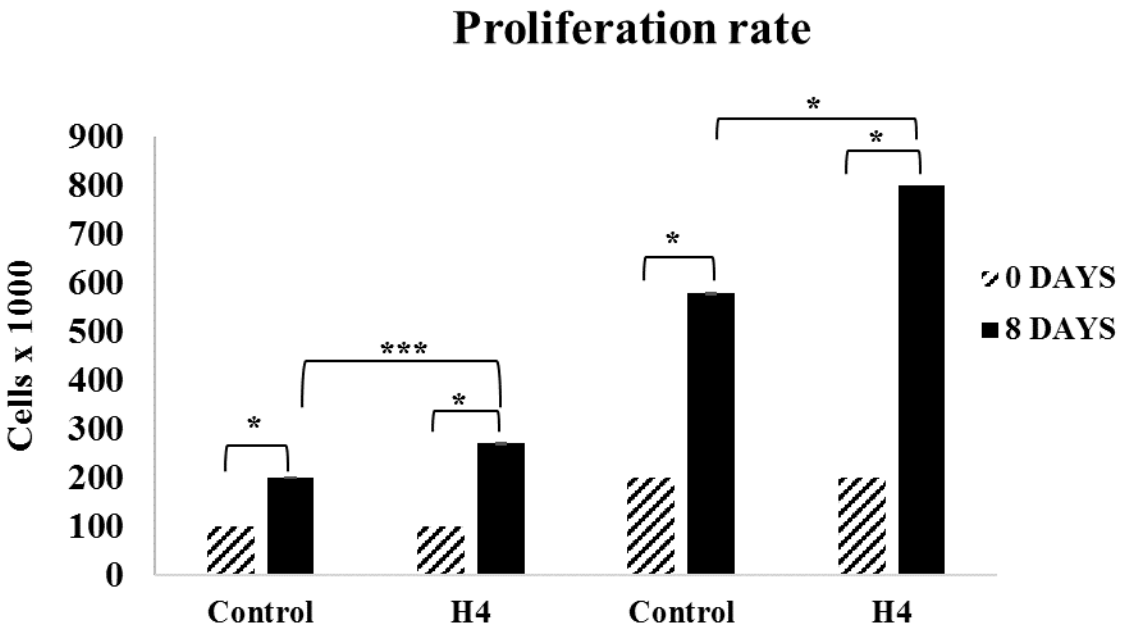


Figure 8: Bright field images showing mesenchymal phenotype (differentiating structures) of MUT-IDH1GSCs overexpressing H4 (B) compared to spheroid structures of uninfected GSCs (A)

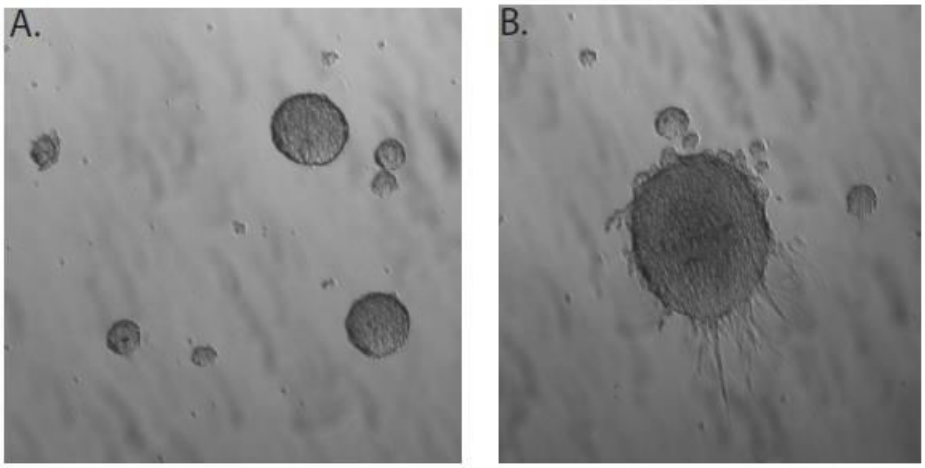


Figure 9: H4 mRNA expression levels in WT-IDH1GSCs cultured in (R) 2HG (100mM) for 15 cell passages. WT-IDH1 GSCs cultured at concentrations above 100mM 2HG had cell death and 100mM was found to be the optimal 2HG concentration. A total number of $2x10^5$ WT-IDH1 GSCs were plated. Wilcoxon Two-Sample Tests was used for statistical analyses. Statistical significance was determined by $p<0.05$. *$p \leq 0.001$ ***$p \geq 0.05$.

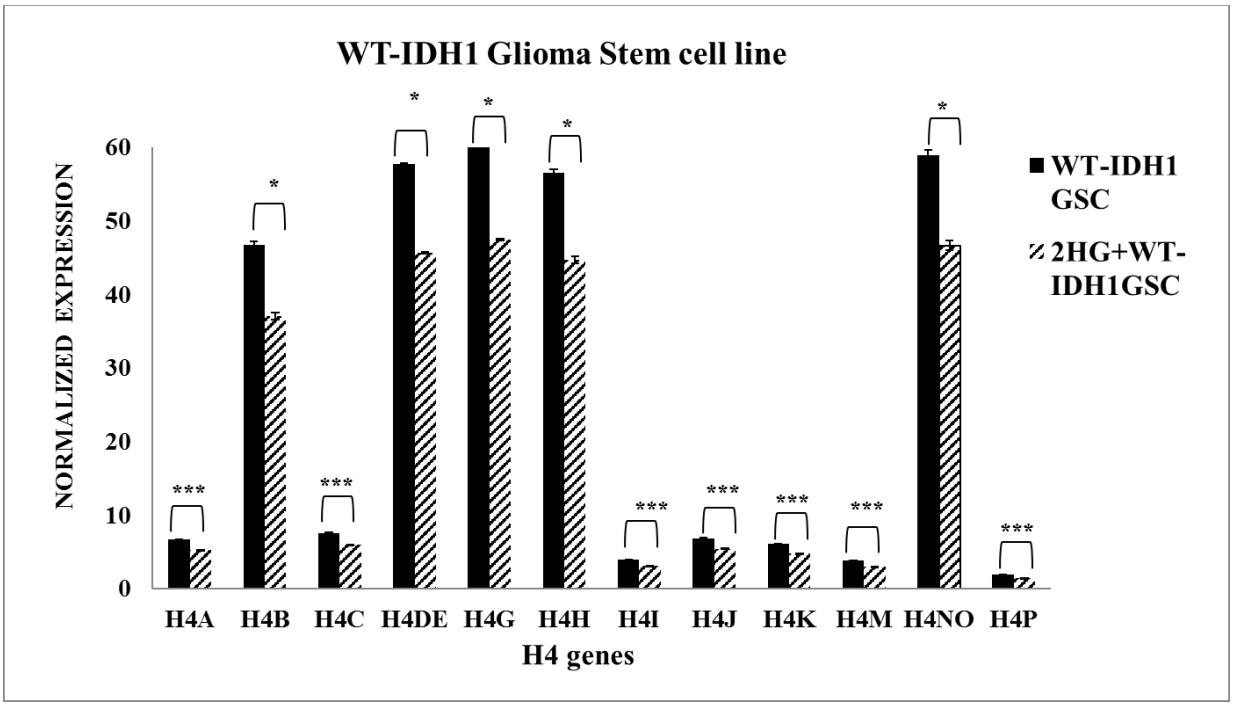


Figure 10: H4 mRNA expression levels in MUT-IDH1GSCs cultured in AG1598 (450mg/kg) for 48h.MUT-IDH1 GSCs cultured at concentrations above 450mg/kgAG5198 had cell death and 450mg/kgwas found to be the optimal AG1598 concentration. A total number of 1x10^5 WT-IDH1 GSCs were plated. Wilcoxon Two-Sample Tests was used for statistical analyses. Statistical significance was determined by $p<0.05$. *$p \leq 0.001$ ***$p \geq 0.05$.

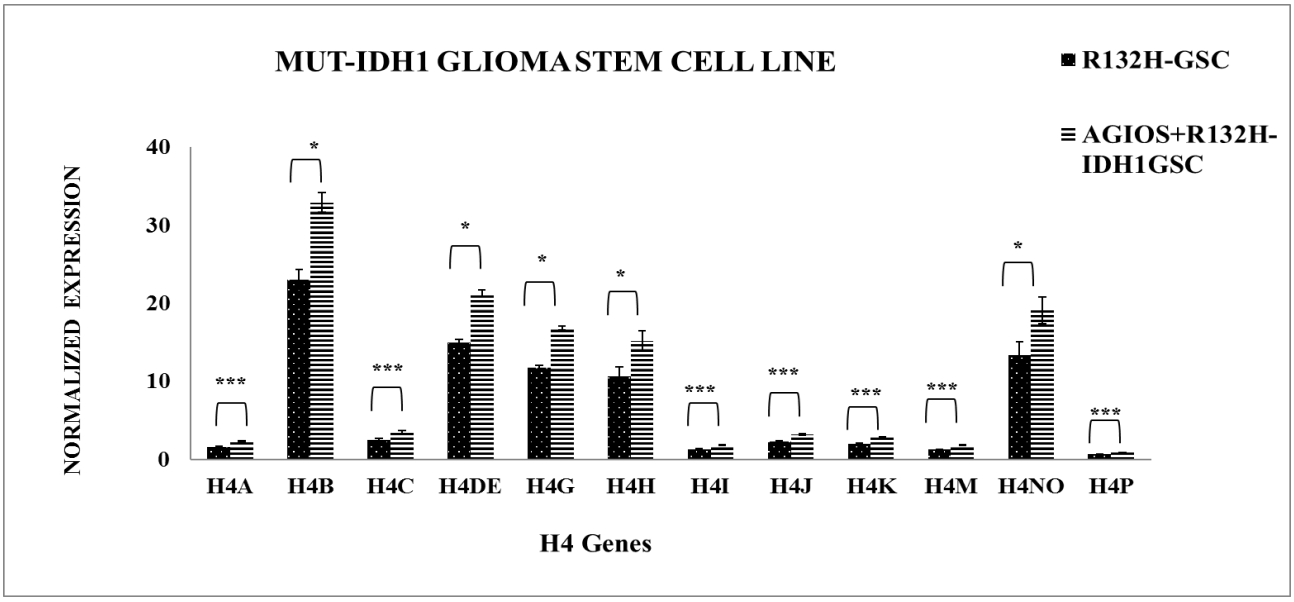
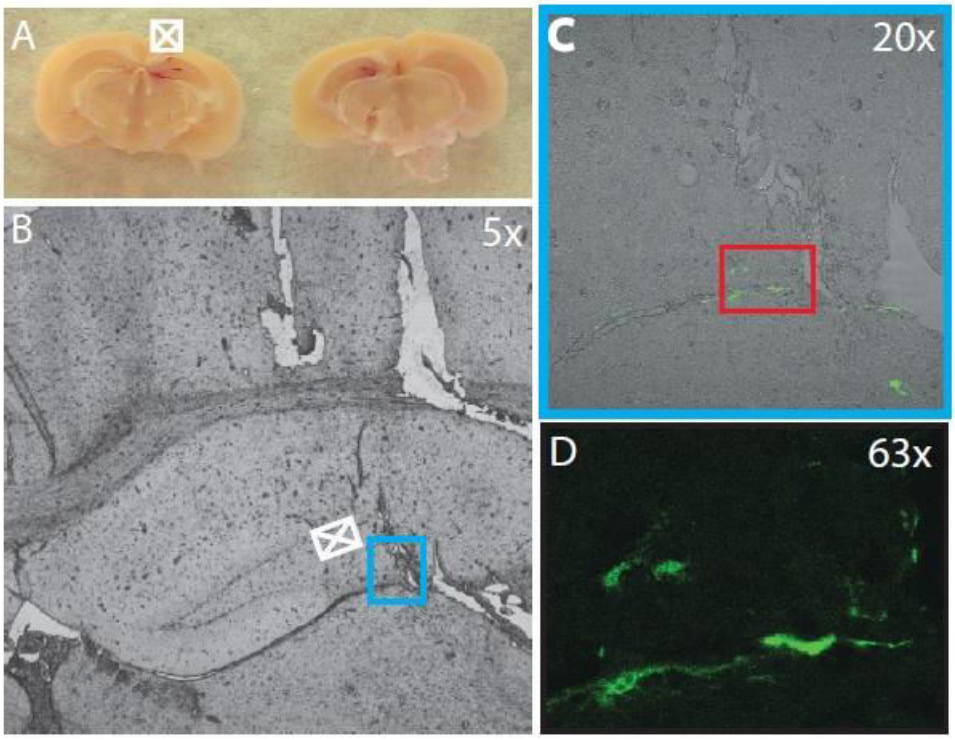
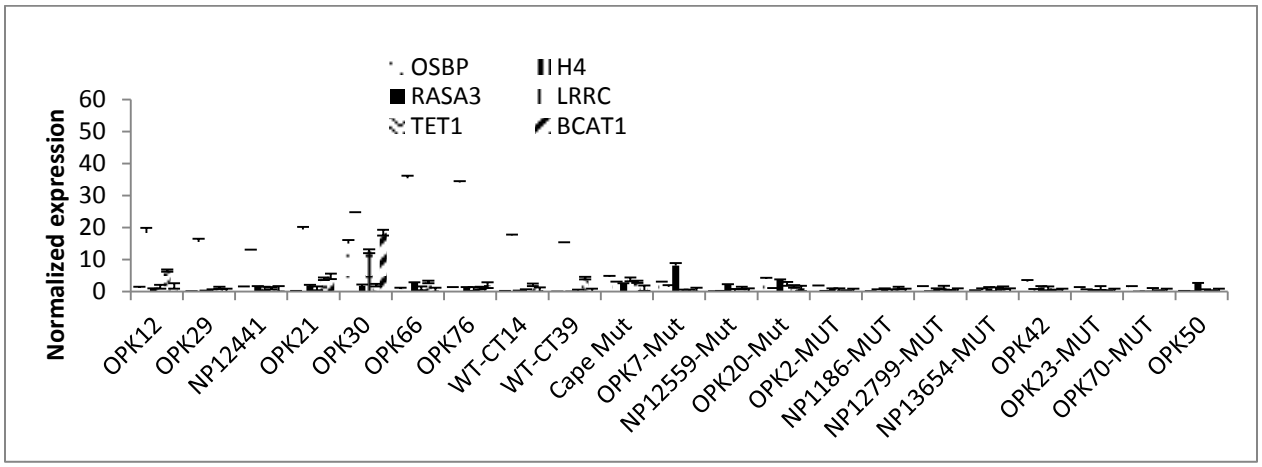


Figure 11: Mouse dentate gyrus expressing GFP (using lentivirus vector). Injection tract on mouse brain (coronal section) (arrows in A and B) showing glioma stem cells expressing GFP (C and D).



Supplementary Figure1. Comparitive expression differentially methylated genes (H4B, OSBP, LRRC, TET1, RASA3 and BCAT1) in WT and Mut-IDH1 tumors.



Tables

Table 1. List of primers used for validating top reference genes

PRIMERS	5’ ----------3’
lpOSBP2030758	gtactagatggtgcctcgct
rpOSBP2 030758	atggacaggggctcattgaa
lpaOSBP2030758	acatctccatcatgccgcta
rpaOSBP2030758	tcactcaccactcctgtcac
lpMYADM001020818	cccaccacctatgtccagtt
rpMYADM 001020818	tgaacgcgaagatgatgcag
lpaMYADM001020818	cgtgaccctgatcatcctca
rpaMYADM001020818	aactggacataggtggtggg
lpCHST8001127895	agcacctacaccaagatgct
rpCHST8001127895	tcccagtgaatgtccatccc
lpaCHST8001127895	caagtttgagcaccccaaca
rpaCHST8001127895	ctccatgctctcgaacttgc
lpHIST14D003539	ggcggaaagggtctaggtaa
rpHIST14D003539	tgtgactgtcttgcgtttgg
lpaHIST14D003539	gtaagggcggaaagggtcta
rpaHIST14D003539	tccatggctgtgactgtctt
lpPLOD1000302	actggaatgtggagaagggg
rpPLOD1000302	tcagcagagaagaccacctg
lpaPLOD1000302	aacaagaacgtcattgcccc
rpaPLOD1000302	agatctgaggactgcagctc
lpIGSF21032880	catggagaactaccgcaagc
rpIGSF21032880	agccatgacgttgaggaaga
lpaIGSF21032880	gaggaaaaccagcacccatg
rpaIGSF21032880	gatgttgggatctggggtca
lpRASA3007368	aaaccttccctcttacccgg
rpRASA3007368	tctgcaagtcctccttctgg
lpaRASA3007368	attcctcggagctttcgtca
rpaRASA3007368	gaccccagtgtctgtgatga
lpZFR2001145640	ggcgacgagtcagtatttcg
rpZFR2 001145640	aggcagggtcccataagatg

lpaZFR2001145640	ccgcagtacagtccagaagt
rpaZFR2 001145640	ccaggcagggtcccataag
lpZFR2015174lp	agaggaagagggtgacaagc
rpZFR2 015174	tcatcctcggtcaccatctg
lpaZFR2015174	aaggaggctcaaaggaagct
rpaZFR2015174	ctttcttccggtactgcagc
lpNBL1182744	tcaacaagctggcactgttc
rpNBL1182744rp	gactctgtggactgtgggaa
lpaNBL1182744	cccagataagagtgcctggt
rpaNBL1182744	gagtcacagtgaaccaggga
lpNBL1005380	tcaacaagctggcactgttc
rpNBL1005380	gactctgtggactgtgggaa
lpaNBL1005380	cccagataagagtgcctggt
rpaNBL1005380	gagtcacagtgaaccaggga
lpLRRC8C032270	tgtcaaaatccacagtgcgg
rpLRRC8C032270	gcttttgagatcccgcagag
lpaLRRC8C032270	tggtgggatgtgtttaccga
rpaLRRC8C032270	tttaggtggaggcagtggag
lpTET1030625	gcaccccaaccgtaatcatc
rpTET1 030625	accttcattagctgcctggt
lpaTET1030625	actcaactctcagaagcccc
rpaTET1 030625	ggtggctctccctgtaagtt
lpFAM208B 017782	cacccgagtttcagtgcaaa
rpFAM208B 017782	tggagatctgggttggtgac
lpaFAM208B017782	gtttcctccctttgtgctgg
rpaFAM208B 017782	ccctccccactgttctcaat

Table 2. Primer sequences used for H4 genes.

Primer sequences were obtained from Holmes et al (2005) [112]

Gene	Forward primer	Reverse primer
H4s/a	TCTCACTTGCTCTTGGTTCACTTC	CCCTTCCCGCCCTTACC
H4/b	CAGATTTAACAGCTGTGGTTTCA	TAGACCCTTTCCGCCCTTAC
H4/c	AAGTTAAGAGTTGTTGTTTGTCTTCG	CCACCTTTGCCTCTACCAGA
H4/d and H4/e	TGGGTGAGACTCCTCTTGCT	AAGACCCTTCCCGCCTTT
H4/g	TACCTTCCACTGCGATAGGA	AAGCCTTTTCCGCCTTTG
H4/h	TCTTAAGTTGGTTTTAGAAGTTGCTT	CCCAAACCTTTTCCACCTTT
H4/i	TGGTTTTTGTTGTCAGCTGGT	CACCTCCCTTACCCAAACCT
H4/j	TTTTGCGGCTATTTTCGTTG	CCTTTACCCAGTCCCTTTCC
H4/k	GCCGTGTTGCATTTTGTTTA	CTTTGCCCAGACCCTTCC
H4/m	GCAGACCTTTGTTCTCTGACC	CAGGCCCTTACCTCCTTTG
H4/n and H4/o	AGCTGTCTATCGGGCTCCAG	CCTTTGCCTAAGCCTTTTCC
H4/n, H4/o, and H4/a	GAGACAACATTCAGGGCATCAC	GAGGCCAGAGATCCGCTTAA
H4/p	GCTGAGTAGGCGCTGTGATT	CCCTTGCCACCTTTACCTC

Supplementary information

Supplementary1. Exome data attached (Spreadsheet)

Supplementary 2. H4 genes and their coding regions

1. **H4/a** ----Accession No: **X60481** (Chr location: 6p21)

1 gctcgtcccg ccatttcctg gggcttggag gaggggttaa aggagcggac tgtaggcgtc
61 acatttcctg cgcgcttttc agtctctgtg tccgctggag gtgggggcag gggtaacgta
121 ga**tatataaa** gatcggtttc ctattctctc acttgctctt ggttcacttc ttgggaagtc
181 atgtctggac gtggtaaggg cgggaagggt ttgggtaagg ggggtgccaa gcgccaccgc
241 aaggtgttgc gtgacaacat ccagggcatc accaagccag ccatccggcg tctggcccgg
301 cgtggcggtg tgaagcggat ctctggtctg atctacgagg agactcgcgg ggtgctcaag
361 gtgtttttgg agaacgtgat ccgtgacgct gtcacctata cggagcacgc caagcgcaag
421 acagtcactg ccatggacgt ggtctacgcg cttaagcgcc agggacgcac cctttatggc
481 tttggcggtt aaggttgctg atttctccac agcttgcatt tctgaaccaa aggccctttt
541 cagggccgcc ctaactaaac aaaagaagag cttgtatcca ttaagtcaag aagctcaatg
601 tgtaattaag atgaatgata ctgagctgac atcctaaaaa ggaaaattag gggaactcca
661agtttgtcct cc

1. **H4/b** ----Accession No:**X60482** (Chr location: 6p21)

1 gatcgcgcca ctgcattcca gcctgggcaa cagagcaaga acccgtctca aaaaaaaaaa
61 aaaaaaaaaa aaaaagcacc ctgtgaggaa acagtactaa ttatttgata ttctgggaaa
121 agtgggggac aactgtcagg cttctttgtc gaaagtttat gaactgatgg ctcagttaat
181 ggctgcaagt atagtgtgtg tgtatatata **tatatata**cc tagcagtatt tattaaatcc
241 cagctgtggt ttcaagatgt ctggccgcgg taagggcgga aagggtctag gtaagggtgg
301 cgccaagcgt caccgtaagg tattgcgtga caatatccaa ggaatcacca agcccgctat
361 ccgccgcctg gctcgccgcg gcggcgtcaa gcgtatttct ggcctcattt atgaggaaac
421 tcgcggagtg ctgaaagttt tcctggaaaa tgtaatccgc gatgctgtca cctacacgga
481 acacgccaaa cgcaagacag tcacagccat ggacgtggtg tacgcgctca agcgccaggg
541 acgcactctt tatggcttcg gcggctgagc ttacctctac agtacactac cgcaaaacca
601 acggcccttt tcagggccac ctatccactc aggagaaaga gtagtagtca ctgctaaaag
661 tgtagtttca cgtgtttagt agctccggtt ttcaagttaa atggtcttat tacgccttgg
721 cttcatatct tactggccgg tgaggcatta gtgtattaaa gtttattttc actcttgctg
781 tgtcgcccat gctggagtaa tcaatggcgc gatc

3.**H4/c** ----Accession No **M60749** (Chr location: 6p21)

ctgcaggggc tgaggcagaa gaatcgctcg aacccgagag gcggaggttg cagtgagccg
61 agatcacgcc actgcactcc agtctggggc agagcaaaac tgtttcaaaa gaaataataa

121 ttttattctc ttgtctatcc attcggtaga caagtgtagc gttactgcct ccttttgca
181 aataataagg atttaggaga cagacattat cagttcgaga gcccacaggc actggactag
241 agacacagaa cctacactca gacccaaaca ggaaacggag gagtcaaatc aacttgcatt
301 tcccctaccc tcaccgcccc ctttccttcc caaggcaata gtgtagggga cgcccagtaa
361 gttacggaaa aggcggagac agaggtttcg ttcccgcccc tccaattcag tctccaaaag
421 gtccgcataa ttga**tatata** aggggcttca gtgtgtagca aagttgcaaa agttaagagt
481 tgttgtttgt cttcgatcat gtctggtaga ggcaaaggtg gtaaaggttt aggaaaggga
541 ggcgccaagc gccatcgcaa agtgctgcgt gacaacatac agggcatcac gaagcccgcc
601 atccgtcgct tggcccgacg cggcggcgtg aaacgcattt cgggcctcat ttatgaggag
661 acccgcggtg ttcttaaggt gttcctggag aatgtgatac gggacgccgt aacctacacg
721 gagcacgcca agcgtaagac agtcactgca atggatgttg tctacgcgct caagcgccag
781 ggacgcactc tgtacggctt tggtggctga gcctcacccc ggctttttat ttaacagctc
841 acccataaaa ggccctttc agggccacct ccttcgtcac acgaagggct gtaactgatg
901 acgacttggg tttcgttttg taaatttggg attctaactg agttaaaccg agccgttttt
961 agcgatcttc ctaagatggc ggatgtgcta aggagaaagg gaaggcgaaa cattagaaac
1021 ttgttcaggt atttcgatcg caaacattca caggaaaaga tggtggaaat ggtttgggaa
1081 taccagcagc cttttatgac gtatgtgatg tgttttattc ccacgtttct gaaatttact
1141 cctatgttag gtggctggga cttaactaca ctagcattac atgtttagtt tttgagatgg
1201 agttgctctt gtccaggctg tagtgcaatg acgcaatctc ggctcaccgc aaccaccccc
1261tcccgggttt aagcggttct cctgactcaa cctcccgagt aactggggtt acagcatgcg

4. H4/d ----Accession No **X60483** (Chr location: 6p21)

1 gtcgtcccaa gtatagagaa aaagtacatt tatgttgaat gtggaaaaca ttatttttaa
61 aatacagccg attaaaaaat tgctggggaa ctgattataa ctattgataa ttcctaagaa
121 acatagacac caaaacatca cttttatctc tttaatagaa attggttaaa ttatatcaat
181 ataaggaggt tacaaaactc catataacac tgtatacatt ttcttggaaa aatatgtgca
241 actgctctgc aaaatacatg attaactagc ctgtgtgatg ggataacatt gtagtttctt
301 aatttcttgt cctttcttgt atttcccaca gtattgatgt atatcttctg cgttcaaaag
361 caatttttta aagcctcata acgtggtaac agaatacttt gcacattaca aaattcagaa
421 cacggaaaca agaagctcgc ttttttttcc cccctatttc ggtttggccc tttagatttc
481 ccctccccac cggcggggac ttcccgccga cttctttcag gttctcagtt cggtccgcca
541 actgtcg**tata**aaggcgctg cctcaggcca gagtcctcac aaagcgttgg gtgagactcc
601 tcttgctcgt catgtctggc cgcggcaaag gcgggaaggg tcttggcaaa ggcggcgcta
661 agcgccaccg taaagtactg cgcgacaata tccagggcat caccaagccg gccatccggc
721 gccttgctcg ccgcggcggc gtgaagcgca tctccggcct catctacgag gagactcgcg
781 gggtgctgaa ggtgttcctg gagaacgtga tccgggacgc cgtgacctat acagagcacg
841 ccaagcgcaa gacggtcacc gccatggatg tggtctacgc gctcaagcgc cagggccgca
901 ccctctacgg tttcggtggt tgagcgtccc tttctatcaa caaaaggccc tttttagggc
961 ccccactttc tcagctgagg agccgtgatg cttgtttgcc cagtttctct ttaccatttg
1021 ttcgcgcctc ctgagtttgt gctgctatcg gagcatgcgt ctctagtcgc tgtaagtagg
1081 catcagtcac taatcgctct ccagtaaata aaatacccaa cacaacttgt ttcaggtgct
1141 ccagagctta ctgacatggg ccgtcgcgca gactgtagtg cagtgtcatc atggctcact

1201 gtagcctccc gattagccgg aacctcgcgc ccaccaccac gcaaggctaa ttatttttc
1261 caaagaccgg gtctccgtgt gtttcccagg ctagtttcga tatcctggcc tcaatcccct
1321 cgcctatgcc taatgttggt attacagtag tgagccacca cgcctggcca cgatactgtt
1381 gaggttttag ggttagttac atttaagggg caattttcgc agtgtagtgg ggaggaaagt
1441 caagcagtta taggtttctg gcgctgtgaa tccaactgct gaacatagca agaacttatt
1501atgcaacaac ctttaatcag tggacgac

5.H4/e ----Accession No **X60484** (Chr location: 6p21)

1 cccggtcact tttttgtatt ccccacagta ttgatgtata tcttctgcgt tcaaaagcaa
61 ttttttaaag cctcataacg tggtaacaga atactttgca cattacaaaa ttcagaacac
121 ggaaacaaga agctcgcttt ttttcccccc ctatttcggt ttggcccttt agatttcccc
181 tcccccaccg gggcgggact tcccgccgac ttctttcagg ttctcagttc ggtccgccaa
241 ctgtcg**tataaa**ggcgctgc ctcaggccag aggcctcaca aagcgttggg tgagactcct
301 cttgctcgtc atgtctggcc gcggcaaagg cgggaagggt cttggcaaag gcggcgctaa
361 gcgccaccgt aaagtactgc gcgacaatat ccagggcatc accaagccgg ccatccggcg
421 ccttgctcgc cgcggcggcg tgaagcgcat ctccggcctc atctacgagg agactcgcgg
481 ggtgctgaag gtgttcctgg agaacgtgat ccgggacgcc gtgacctata cagagcacgc
541 caagcgcaag acggtcaccg ccatggatgt ggtctacgcg ctcaagcgcc agggccgcac
601 cctctacggt ttcggtggtt gagcgtcctt ttctaccaat aaaaggccct tttcagggcc
661 accctacttt ctcagctgaa gagtggtaac actgaggagt ggttttggta ggtacggaat
721 tttgcttggt tctgagtcag ttctgggggg aacagttttt tgaacacagc ggcacacgtg
781 tggccattca cccggggtca ctgtaggcag gactaattac gagatgtaat gtctaaactt
841 gctcaaaatt cgtaagctt

6.**H4/g** ----Accession No **X60486** (Chr location: 6p21)

ORIGIN
1 aatacagcgc attcaacttg caaacaccct tccactccca caaagagcaa gctgtcactg
61 gccaatcaaa acaatgaacc ataatgaaac agtttttctt gctccaccca ctcggtgacc
121 aaatttgaaa aaaaaaaaaa accgcgccaa ctcatgttgt tttcaatcag gtccgccaag
181 tttg**tattta** aggaactgtt tcagttcata ccttccactg cgataggaat catgtctggt
241 cgcggcaaag gcggaaaagg cttggggaag ggtggtgcta agcgccatcg taaggtgctc
301 cgggataaca tccagggcat tacaaaaccg gctatccgcc gtttggctcg gcgcggtggg
361 gtcaagcgca tttccggtct tatctatgag gagactcgag gtgtgcttaa ggttttctta
421 gagaacgtta ttcgagacgc cgtcacctat acggagcacg ccaagcgcaa aactgtcaca
481 gccatggatg tagtatatgc cctaaaacgt caggggcgca ctctgtatgg cttcggcggc
541 tgaatctaag aatacgcggt ctcctgagaa cttcaaaaaa caaaaaaacc caaaggccct
601 tttcagggcc gctcacaaag tcgtttaaag agctgaaatg cgttgcgaga atgagtttgg
661 atgacagaaa taaccgtgac agcctgcata agaatgaatt gtgtttgcca tgaccggcca
721 cactgtgaca aaatttca

7. **H4/h** ----Accession No **X60487** (Chr location: 6p21)

1 aagctgcact cctcaggctt ccggtcttgt ggaccgaggc tcctctccaa tcgtcctaaa
61 gcaactgaag gaacaattga taacaattta aacatgttaa tttatttaat ctgatttggt
121 tgttaaaaca gccttaaaac tctcattcct atttcccagc aattgcaaga gacgcctccg
181 cggagccggc gctcagatgc caaaggctgg agtcgttggt gacgtcatcc agggctgggc
241 tatgcggggt ggggccaaga agcaggacca aagttccggt atgcgcgctt tcagtcttca
301 attaggtccg aattcccggc a**tataa**gggc gttgctttgc cttggcgttt aggtttctta
361 agttggtttt agaagttgct tagtcatgtc tggccgtggt aaaggtggaa aaggtttggg
421 taagggagga gctaagcgtc atcgcaaggt tttgcgcgat aacatccagg gcatcactaa
481 gccagctatc cggcgccttg ctcgtcgcgg cggtgtcaag cgaatttctg gccttatcta
541 tgaggagact cgtggtgttc tgaaggtgtt cctggagaac gtgattcgtg acgctgtcac
601 ttacacagag cacgccaaac gcaagaccgt gacagcaatg gatgtggtct acgcgctgaa
661 gcgacaggga cgcactcttt acggcttcgg tggctaaggc tcctgcttgc tgcactctta
721 ttttcatttt caaccaaagg ccctttcag ggccgcccac ttttttcata aaagagcaga
781 catcttgtta tcctgctttg ttccagattt ttgctgagaa gtagttactg tgcacatgtg
841 tagatttgca ggtaacgttt aattactcgt gcttaatctc gcatgtttag gatttgggca
901 taaagtagtt aaagccactg ttttctgctt gagggaagtt ggtatgttga atatgtttag
961 ctttcttgtg gcattttaac actagatatc agtgagtaag ctggattgtg acagtttata
1021 attcccgtgg gccttcctcc ttcagtttaa gctgac

8. **H4/ i**----Accession No **X67081** (Chr location: 6p21)

ORIGIN
1 gtcgagggta taaattcaga ttcaaatagt agtgtgctaa acgggaggga aaaactaaag
61 ggattaaaaa gtgaaactat tgtgttctcc ctcgcagtcc ttaggtcact gcccctcgag
121 gggcggagca aaaagtgagg cagcaacgcc tccttatcct cgctcccgct ttcagttctc
181 aataaggtcc gatgttcgtg **tataaa**tgct cgtggcttgc tttcttttcg cgtacctggt
241 ttttgttgtc agctggttag acatgtctgg tcgcggcaaa ggcggtaaag gtttgggtaa
301 gggaggtgcc aagcgtcacc gaaaagtgct gcgggataac atccaaggca tcaccaaacc
361 ggccattcgg cgccttgcta ggcgtggtgg ggttaagcga atttccggtt tgatttatga
421 ggagactcgt ggcgttctca aggtgtttct ggagaacgtg atccgggacg ccgtgaccta
481 cacggagcac gccaagcgca agactgtcac tgccatggat gtggtttacg cgctcaagcg
541 tcaaggacgc actctgtacg gcttcggcgg ttaatctttt cgtcagtttt cttccaatgg
601 ccctttcag ggccgcccac tccctctcag aaagagctgt gattgtattc tttcggatgg
661 taacatctca atggctttac tcggctattc tgcctagtat gtagaactat tataaaccag
721 ttgggagaga ccaggttgtt tggtctgagt ggctgctaaa gcagaaatca gctaagtaaa
781 cgaggtctcc gagataagtg agctataaac ttcaatgcta tagttttgac atgtcaagca
841 acttaacgtg cagcgcgagt ccgataaatg agtagctcag ctttttagtt ttaaaacgag
901 ttgtgcgtta tttgtac

9. **H4/ j**----Accession No**: Z80787** (Chr location: 6p21)

1 gttaaaatat gcctaaattt cctctttggg aacgcaagac ttgcagagat gactccatgg
61 agagcggact ctgccggcgg gaactggagt cgttggtgac gtcatcccag tctgatctgt

121 gaagggtagg gccagcaggc agcaccaaag ttcccgtatg cgcgttttca gtcttcattt
181 aggtccgaat tcccggca**tataa**gaatact accgtcgctt gtttttcaga tttttgcggc
241 tattttcgtt ggtgtgttgg tcatgtctgg tcgcggcaaa ggcggaaagg gactgggtaa
301 aggaggcgct aagcgtcacc gtaaggtcct gcgagataac atccagggca ttaccaagcc
361 tgccatccgg cgccttgctc gtcgcggggg tgtcaagcgc atttctggtc tcatctacga
421 ggagactcgc ggggttctga aggtgtttct ggaaaacgtg attcgtgatg ctgtgactta
481 cacggagcac gccaaacgca agacagtgac agcgatggat gtggtctacg cgctgaagag
541 acagggacgc actctttacg gcttcggcgg ctaatgctac cgcttaaacg actcagcatc
601 tcgacttccc aaatcaaagg ccctttttcag ggccgcccac agttttccgc aaaagagctc
661 atgacttgtt agacgattgg ttagtctctt tataagttaa t

10. **H4/ k**----Accession No **X83548** (Chr location: 6p21)

1 tctagagatg gcgccatttg attccagcag ccacaaagca ctagaacaat cgatgctaag
61 aggtgacagg aaaaacaggc tgcaaagacc cagacaatgg aatgcagcgg tggtcagcct
121 aaaacactgt agaagggcaa gatgagctga gtaattttta actgggcatc atttttagaa
181 actggagttt aagtaccccc ttttccattt tttcctgaag tcgtgggcag ggcgcaaggt
241 ctgtgaatcg gccgaccgga tgcagctggt gtggagagtt cccaatcagg tccgatttat
301 tac**tatataa** agtactgctg cgaggcttgc cgtgttgcat tttgtttagt acaagacatg
361 tctgggcgcg gcaaaggcgg gaagggtctg ggcaaaggag gcgctaagcg ccaccgcaaa
421 gttctgcgcg acaacattca gggcatcacc aagcccgcca tccgacgcct ggcacggcgt
481 ggaggcgtta agcgcatctc aggccttata tacgaggaga cacgcggagt tcttaaagtg
541 tttttggaga atgtaatccg cgatgcagtt acctacacgg agcacgccaa acgcaagaca
601 gtcacagcca tggacgtggt ttacgcgctc aagcgccagg gccgcaccct gtatggcttt
661 ggcggctgag tgttttactt acttacacgg ttcctcaaag gcccttctca gggccaccca
721 tgaagtctgt gaaagagctg tagactaaag atagttaatt tcttaagaac acttaaacgt
781 atggcagttt tggcaaatta gcgattccac ataagcagtc gctgaagttt gaggttcggt
841 gcccctttca gcattactta gtggttaaaa

11. H4/ l----Accession No **Z80788** (Chr location: 6p21)

1 ctgcaggctc caccccacgg ggttcatgcc attcttctgc ctcagcctcc tgcatagctg
61 ggactacagg cgcccgccac ctcgccctgc taatttttg tatttttatt agagacaggg
121 tttcaccgtg ttagccagga tggtctcgat ctcctgagct cgtgatccgc ccgcctcggc
181 ctccaaaagt gctgggatta caggcgtgag ccaccgcgcc cggccaggag taaatcttta
241 aacacataca gacatcaaaa atcgagatac tttcctctag gtacatgttt gacactttcg
301 tttagagatc cttattctct aaatgccagc aaatctccct actacaagag aatttcttat
361 gtcttgttat tactagaagt agatgagcgg ggactgaata cgttttcccg cctttcctgg
421 tctttcagtt ttcaaaaagg tctactctac cgt**tatataa** aaggctgcgg aaagcgagta
481 atcgttttgt ttagagatag ttctgacttg tttatcatgt ctgttcgggg caaggccgga
541 aaaggccttg ggaaaggcgg tgccaagtgc catcgcaagg tactgagcga taatattcag
601 ggcattacca agtgcactat ccggcgcttg gcccggcatg gcggtgtcaa gcgcatcttg

661 ggcctcattt atgaggagac ccgccgggtg ttcaaggtgt tcctggaaaa tgtgatctgg
721 tacgccgtga ccaacacgga gcacgccaag cgcaagacgg tcaccgccat ggccgtggtc
781 tacgtgctca aacgccaggg aagaaccctg taaggctttg gcggttaagt gaggccaggc
841 ctttccttgc aatatcgtta atggcccttt tcagggccgc gtacgttttc ttccgaaggg
901 ctcacgtttg acattgtgga aacttaacta acgcactcag caatattgtg tgttttattg
961 ctgtaggtgc ccaaaagctg ttaggtgtat aatgctgtaa agtataaatt tacgccaagg
1021 ctgacataca acatggaata caggtggtaa gttaatcatg tgttcagaga ttagaaatct
1081 ggtgggctgg ctgaggacat caacggctac attgtcactt aaaagctctc aaaatgtggt
1141 tatggtagtg tgtttttctc cttactgtta ggattaccct ccgtaattaa aattgtagtt
1201 ttagtgccaa gataacccga cttttatatt ctgcttaata agcaatttca ataaccagag
1261 aatacacaat aattagcaca aacttgccag gtgccgttct gacaaatgtg tctgcag

12. H4/ m----Accession No **AB000905** (Chr location: 6p21)

1 ctgcagtaga taatgaggta accgaaggca attgtgcttc ttttgataag aagctttctt
61 ggtcatatca ggaaattcca gagaaagtcc ctccctgtat ttggggaaga gaaacaggac
121 aaagttagag ggaccttgat tcttagactt gtttctgaga accctcaatt ttcaaaaaca
181 cccaccatta ccaagctcga tatttggggg gataattctc caccccaaac actagaaatg
241 aaaataagta gaaaagaact tagcaa**tata** cctgaacgat ctttaaattc tatgagtcta
301 tcttgttcta ttgctaggaa tttagcatat ggagtatatt tccattgtat attaaggaga
361 aaaatgccac aaaatagcat agttccatac aaataatgtt aaaggaaaac atgcatatat
421 gcagagataa cattcttttt gctttatgta tcttctataa tgttaacatt tttgttatga
481 ggcaggaatc atgttaaatg aaaaaaggaa aaacagccta ggtcttgagg attaaaaagg
541 actgaaggag aacaagaggg agtagagcac agcaggcctg tttccctttt aggtcccctc
601 ccccaatgca gagggacttc cgccaaagct cttccggttt tcagtctggt ccgcagaggt
661 tacccataaa agaaagctgc catcacaggc agcagacctt tgttctctga ccacttgata
721 atgtcaggac gcggcaaagg aggtaagggc ctggggaaag ggggtgccaa gcgccaccgc
781 aaggtgctgc gcgacaacat ccagggtatc accaagccag ccattcggcg ccttgctcgc
841 cgcggcggcg tgaagcgcat ttctggcctc atctatgagg agacccgcgg agtgttgaag
901 gtgttcctgg agaacgtgat ccgggacgcc gtgacctaca cggagcacgc caagcgcaag
961 acggtcaccg ccatggacgt ggtctacgcg ctcaagcgcc agggccgcac cctctatggc
1021 ttcggcggct aaatggcatt ttgaagccca gtcattctct aaaaaggccc tttttagggc
1081 ccctaagctt tcaacaaaag agttgaaatg actccaaact gagtctctta atagggccat
1141 tgtcagtgag ttctgtcatc ctattttaca agattaactc gacgccgaaa atgggctgat
1201 gactacaggt gaccttgggc cgagattttt ccaaggccag aagagcctct gctggccagt
1261 aacttctggc ggctgcctgg aaattgcctg cag

13. H4/ n----Accession No **AB000905** (Chr location: 6p21)

1 aattctcctg tgtgagctaa aatacagtgg ctcggtccaa caaaacagag cctggagcca
61 ggaattatgg cgaacctgct ccctccgtcc tccttcggcg aagatccctg gcgcgcgtcc
121 ttgaggtcgc cttcggtgtt gacctcatcg tcggaacggc gcttcctgaa gctt**tatata**
181 agcacggctc tgaatccgct cgtcggatta aatcctgcgc tggcgtcctg ccagtctctc

241 gctccatttg ctcttcctga ggctccctcc agagaccttt cccttagcct cagtgcgaat
301 gcttccgggc gtcctcagaa ccagagcaca gccaaagcca ctacagaatc cggaagcccg
361 gttgggatct gaattctccc ggggaccgtt gcgtaggcgt taaaaaaaaa aaagagtgag
421 agggacctga gcagagtgga ggaggaggga gaggaaaaca gaaaagaaat gacgaaatgt
481 cgagagggcg gggacaattg agaacgcttc ccgccggcgc gctttcggtt ttcaatctgg
541 tccgatactc ttgtatatca ggggaagacg gtgctcgcct tgacagaagc tgtctatcgg
601 gctccagcgg tcatgtccgg cagaggaaag ggcggaaaag gcttaggcaa agggggcgct
661 aagcgccacc gcaaggtctt gagagacaac attcagggca tcaccaagcc tgccattcgg
721 cgtctagctc ggcgtggcgg cgttaagcgg atctctggcc tcatttacga ggagacccgc
781 ggtgtgctga aagtgttctt ggagaatgtg attcgggacg cagtcaccta caccgagcac
841 gccaagcgca agaccgtcac agccatggat gtggtgtacg cgctcaagcg ccaggggaga
901 accctctacg gcttcggagg ctaggcgccg ctccagcttt gcacgtttcg atcccaaagg
961 ccctttttgg gccgaccact tgctcatcct gaggagttgg acacttgact gcgtaaagtg
1021 caacagtaac gatgttggaa ggtaactttg gcagtggggc gacaatcgga tctgaagtta
1081 acggaaagac ataaccgc

13b. H4/ n----Accession No **M11930** (Chr location: 6p21)

1 aattctcccg gggaccgttg cgtaggcgtt aaaaaaaaaa aagagtgaga gagggactga
61 gcagagtgga ggaggaggga gaggaaaaca gaaaagaaat gacgaaatgt cgagagggcg
121 gggacaattg agaacgcttc ccgccggcgc gctttcggtt ttcaatctgg tccgatatct
181 ctg**tata**tta cggggaagac ggtgacgctc cgatcgancn nctatcgggc tcctgcggtc
241 atg

13 c. H4/ n----Accession No **X00038** (Chr location: 6p21)

1 aattctcccg gggaccgttg cgtaggcgtt aaaaaaaaaa aagagtgaga gagggactga
61 gcagagtgga ggaggaggga gaggaaaaca gaaaagaaat gacgaaatgt cgagagggcg
121 gggacaattg agaacgcttc ccgccggcgc gctttcggtt ttcaatctgg tccgatatct
181 ctg**tata**tta cggggaagac ggtgacgctc cgatcgancn nctatcgggc tcctgcggtc
241 atgtccggct gtggaaaggg cggaaagggc ttaggcaaag gtggcgctaa gcgccaccgc
301 aaggtcttga gagacaacat tcagggcatc accaagcctg ccattcggcg tntagctcgg
361 cgtggcggcg ttaagcggat ctctggcctc atttacgagg agacccgcgg tgtgctgaaa
421 gtgttcttgg agaatgtgat tcgggacgca gtcacctaca ccgagcacgc caagcgcaag
481 accgtcacag ccatggatgt ggtgtacgcg ctcaagcgnc aggggagnac cctctacggc
541 ttcggaggct aggccgccgc tccagctttg cacgtttcga tcccaaaggc ccttttgggg
601 ccgaccactt gctcatcctg aggagttgga cacttgactg cgtaaagtgc aacagtaacg
661 atgttggaag gtaactttgg cagtggggcg acaatcggat ctgaagttaa cggaaagaca
721 taaccgc

14. H4/ o----Accession No **AY648850** (Chr location: 6p21)

1 tgagaaaagg caa**tataaa**t tcttttagct catactgttt taactttaaa acgtatcacc
61 atgaaccttt ctccaaacat cagaaaaaat tttcccaaaa gagcaaacaa caaaaaaatc
121 aggctggttt cgctgactgc ttcctggact taacaactat agcatgtctc cagaggtggg
181 gcctagagct ccgcccactt ttttgatgtt ttcaaacagg tccgcatcga gagactataa
241 gccctggtct gcgactggca atcacagtgg gcagcccgat tttctgctga gtaggcgctg
301 tgatttcaga atgtctgggc gaggtaaagg tggcaagggg ctgggtaagg gaggcgccaa
361 gcgccaccgg aaggtgctgc gggacaatat ccaaggcatt acaaagccgg cgattcgccg
421 tctcgcccga cgtgggggcg tcaagcgcat ttctggtctc atctacgagg agacccgggg
481 agtcctcaaa gtcttcctgg agaacgtgat ccgtgacgcg gtgacttaca cggagcacgc
541 caagcgcaag accgtcacgg ccatggatgt ggtgtacgcg ctgaaacgcc agggtcgcac
601 cctttatggt ttcggcggtt gagctgtccc cacagcttct ctacagactc caaaaggccc
661 ttttcagggc ccccaaactg tcacagaaag agctgttaac acttcctaga taacgggtaa
721 atatcaaccc tttgatgtcg ccttccggat accggaaatg ggatattcgc ggcccgagca
781 tctgctcggg attagacttg gcgggcacgg ggagaactag acctcaagga gggggttccg
841 cctgtgctta ggtacgaccc tggaggtgcc gcgcaacgca gtagtgacct cccaaaatcc
901 agacaccgga ttttgtttag aagagtgctt

15. H4/ p----Accession No **AY128653** (Chr location: 6p21)

1 tgagaaaagg caa**tataa**at tcttttagct catactgttt taactttaaa acgtatcacc
61 atgaaccttt ctccaaacat cagaaaaaat tttcccaaaa gagcaaacaa caaaaaaatc
121 aggctggttt cgctgactgc ttcctggact taacaactat agcatgtctc cagaggtggg
181 gcctagagct ccgcccactt ttttgatgtt ttcaaacagg tccgcatcga gagactataa
241 gccctggtct gcgactggca atcacagtgg gcagcccgat tttctgctga gtaggcgctg
301 tgatttcaga atgtctgggc gaggtaaagg tggcaagggg ctgggtaagg gaggcgccaa
361 gcgccaccgg aaggtgctgc gggacaatat ccaaggcatt acaaagccgg cgattcgccg
421 tctcgcccga cgtgggggcg tcaagcgcat ttctggtctc atctacgagg agacccgggg
481 agtcctcaaa gtcttcctgg agaacgtgat ccgtgacgcg gtgacttaca cggagcacgc
541 caagcgcaag accgtcacgg ccatggatgt ggtgtacgcg ctgaaacgcc agggtcgcac
601 cctttatggt ttcggcggtt gagctgtccc cacagcttct ctacagactc caaaaggccc
661 ttttcagggc ccccaaactg tcacagaaag agctgttaac acttcctaga taacgggtaa
721 atatcaaccc tttgatgtcg ccttccggat accggaaatg ggatattcgc ggcccgagca
781 tctgctcggg attagacttg gcgggcacgg ggagaactag acctcaagga gggggttccg
841 cctgtgctta ggtacgaccc tggaggtgcc gcgcaacgca gtagtgacct cccaaaatcc
901 agacaccgga ttttgtttag aagagtgctt

9.0. References

1 Ohgaki, H. & Kleihues, P. The definition of primary and secondary glioblastoma. *Clinical cancer research : an official journal of the American Association for Cancer Research***19**, 764-772, doi:10.1158/1078-0432.ccr-12-3002 (2013).

2 Dang, L. *et al.* Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature***462**, 739-744, doi:10.1038/nature08617 (2009).

3 Kleihues, P. & Ohgaki, H. Primary and secondary glioblastomas: from concept to clinical diagnosis. *Neuro-oncology***1**, 44-51 (1999).

4 Brennan, C. *et al.* Glioblastoma subclasses can be defined by activity among signal transduction pathways and associated genomic alterations. *PloS one***4**, e7752, doi:10.1371/journal.pone.0007752 (2009).

5 Phillips, H. S. *et al.* Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. *Cancer cell***9**, 157-173, doi:10.1016/j.ccr.2006.02.019 (2006).

6 Verhaak, R. G. *et al.* Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer cell***17**, 98-110, doi:10.1016/j.ccr.2009.12.020 (2010).

7 Bhat, K. P. *et al.* Mesenchymal differentiation mediated by NF-kappaB promotes radiation resistance in glioblastoma. *Cancer cell***24**, 331-346, doi:10.1016/j.ccr.2013.08.001 (2013).

8 Sottoriva, A. *et al.* Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proceedings of the National Academy of Sciences of the United States of America***110**, 4009-4014, doi:10.1073/pnas.1219747110 (2013).

9 Ozawa, T. *et al.* Most human non-GCIMP glioblastoma subtypes evolve from a common proneural-like precursor glioma. *Cancer cell***26**, 288-300, doi:10.1016/j.ccr.2014.06.005 (2014).

10 Lo, H. W. EGFR-Targeted Therapy in Malignant Glioma: Novel Aspects and Mechanisms of Drug Resistance. *Current molecular pharmacology***3**, 37-52 (2010).

11 Colman, H. *et al.* A multigene predictor of outcome in glioblastoma. *Neuro-oncology***12**, 49-57, doi:10.1093/neuonc/nop007 (2010).

12 Yan, X. *et al.* A CD133-related gene expression signature identifies an aggressive glioblastoma subtype with excessive mutations. *Proceedings of the National Academy of Sciences of the United States of America***108**, 1591-1596, doi:10.1073/pnas.1018696108 (2011).
13 Dang, C. V. Links between metabolism and cancer. *Genes & development***26**, 877-890, doi:10.1101/gad.189365.112 (2012).
14 Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science (New York, N.Y.)***324**, 1029-1033, doi:10.1126/science.1160809 (2009).
15 Locasale, J. W. & Cantley, L. C. Metabolic flux and the regulation of mammalian cell growth. *Cell metabolism***14**, 443-451, doi:10.1016/j.cmet.2011.07.014 (2011).
16 Lunt, S. Y. & Vander Heiden, M. G. Aerobic glycolysis: meeting the metabolic requirements of cell proliferation. *Annual review of cell and developmental biology***27**, 441-464, doi:10.1146/annurev-cellbio-092910-154237 (2011).
17 Sebastian, C. *et al.* The histone deacetylase SIRT6 is a tumor suppressor that controls cancer metabolism. *Cell***151**, 1185-1199, doi:10.1016/j.cell.2012.10.047 (2012).
18 Lyssiotis, C. A. & Cantley, L. C. SIRT6 puts cancer metabolism in the driver's seat. *Cell***151**, 1155-1156, doi:10.1016/j.cell.2012.11.020 (2012).
19 Balss, J. *et al.* Analysis of the IDH1 codon 132 mutation in brain tumors. *Acta neuropathologica***116**, 597-602, doi:10.1007/s00401-008-0455-2 (2008).
20 Korshunov, A. *et al.* Combined molecular analysis of BRAF and IDH1 distinguishes pilocytic astrocytoma from diffuse astrocytoma. *Acta neuropathologica***118**, 401-405, doi:10.1007/s00401-009-0550-z (2009).
21 Yan, H. *et al.* IDH1 and IDH2 mutations in gliomas. *The New England journal of medicine***360**, 765-773, doi:10.1056/NEJMoa0808710 (2009).
22 Shibahara, I. *et al.* Malignant clinical features of anaplastic gliomas without IDH mutation. *Neuro-oncology***17**, 136-144, doi:10.1093/neuonc/nou112 (2015).
23 Hartmann, C. *et al.* Patients with IDH1 wild type anaplastic astrocytomas exhibit worse prognosis than IDH1-mutated glioblastomas, and IDH1 mutation status accounts for the unfavorable prognostic effect of higher age: implications for classification of gliomas. *Acta neuropathologica***120**, 707-718, doi:10.1007/s00401-010-0781-z (2010).
24 Tran, A. N. *et al.* Increased sensitivity to radiochemotherapy in IDH1 mutant glioblastoma as demonstrated by serial quantitative MR volumetry. *Neuro-oncology***16**, 414-420, doi:10.1093/neuonc/not198 (2014).
25 Sanson, M. *et al.* Isocitrate dehydrogenase 1 codon 132 mutation is an important prognostic biomarker in gliomas. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology***27**, 4150-4154, doi:10.1200/jco.2009.21.9832 (2009).
26 Lee, S. M. *et al.* Cytosolic NADP(+)-dependent isocitrate dehydrogenase status modulates oxidative damage to cells. *Free radical biology & medicine***32**, 1185-1196 (2002).
27 Kim, S. Y. *et al.* Regulation of singlet oxygen-induced apoptosis by cytosolic NADP+-dependent isocitrate dehydrogenase. *Molecular and cellular biochemistry***302**, 27-34, doi:10.1007/s11010-007-9421-x (2007).
28 Jennings, G. T., Minard, K. I. & McAlister-Henn, L. Expression and Mutagenesis of Mammalian Cytosolic NADP+-Specific Isocitrate Dehydrogenase†. *Biochemistry***36**, 13743-13747, doi:10.1021/bi970916r (1997).

29 Dang, L. *et al.* Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature***462**, 739-744, doi:http://www.nature.com/nature/journal/v462/n7274/suppinfo/nature08617_S1.html (2009).
30 Ward, P. S. *et al.* in *Cancer Cell* Vol. 17 225-234 (2010 Elsevier Inc, 2010).
31 Yan, H. *et al.* in *N Engl J Med* Vol. 360 765-773 (2009 Massachusetts Medical Society, 2009).
32 Sanson, M. *et al.* Isocitrate Dehydrogenase 1 Codon 132 Mutation Is an Important Prognostic Biomarker in Gliomas. *Journal of Clinical Oncology***27**, 4150-4154, doi:10.1200/jco.2009.21.9832 (2009).
33 Matsumoto, A. & Nakayama, K. I. Role of key regulators of the cell cycle in maintenance of hematopoietic stem cells. *Biochimica et biophysica acta***1830**, 2335-2344, doi:10.1016/j.bbagen.2012.07.004 (2013).
34 Kong, D., Li, Y., Wang, Z. & Sarkar, F. H. Cancer Stem Cells and Epithelial-to-Mesenchymal Transition (EMT)-Phenotypic Cells: Are They Cousins or Twins? *Cancers***3**, 716-729, doi:10.3390/cancers30100716 (2011).
35 Kashyap, V. *et al.* Regulation of stem cell pluripotency and differentiation involves a mutual regulatory circuit of the NANOG, OCT4, and SOX2 pluripotency transcription factors with polycomb repressive complexes and stem cell microRNAs. *Stem cells and development***18**, 1093-1108, doi:10.1089/scd.2009.0113 (2009).
36 Parsons, D. W. *et al.* An Integrated Genomic Analysis of Human Glioblastoma Multiforme. *Science (New York, N.Y.)***321**, 1807-1812, doi:10.1126/science.1164382 (2008).
37 Jin, G. *et al.* Disruption of wild-type IDH1 suppresses D-2-hydroxyglutarate production in IDH1-mutated gliomas. *Cancer research***73**, 496-501, doi:10.1158/0008-5472.can-12-2852 (2013).
38 Le, P. U. *et al.* in *Oncogene* Vol. 29 4636-4647 (2010).
39 Losman, J. A. & Kaelin, W. G., Jr. What a difference a hydroxyl makes: mutant IDH, (R)-2-hydroxyglutarate, and cancer. *Genes & development***27**, 836-852, doi:10.1101/gad.217406.113 (2013).
40 Sasaki, M. *et al.* D-2-hydroxyglutarate produced by mutant IDH1 perturbs collagen maturation and basement membrane function. *Genes & development***26**, 2038-2049, doi:10.1101/gad.198200.112 (2012).
41 Losman, J. A. *et al.* (R)-2-hydroxyglutarate is sufficient to promote leukemogenesis and its effects are reversible. *Science (New York, N.Y.)***339**, 1621-1625, doi:10.1126/science.1231677 (2013).
42 Lu, C. *et al.* IDH mutation impairs histone demethylation and results in a block to cell differentiation. *Nature***483**, 474-478, doi:10.1038/nature10860 (2012).
43 Rohle, D. *et al.* An inhibitor of mutant IDH1 delays growth and promotes differentiation of glioma cells. *Science (New York, N.Y.)***340**, 626-630, doi:10.1126/science.1236062 (2013).
44 Alvarez-Buylla, A., Garcia-Verdugo, J. M. & Tramontin, A. D. A unified hypothesis on the lineage of neural stem cells. *Nature reviews. Neuroscience***2**, 287-293, doi:10.1038/35067582 (2001).
45 Merkle, F. T., Tramontin, A. D., Garcia-Verdugo, J. M. & Alvarez-Buylla, A. Radial glia give rise to adult neural stem cells in the subventricular zone. *Proceedings of the National Academy of Sciences of the United States of America***101**, 17528-17532, doi:10.1073/pnas.0407893101 (2004).

46 Sanai, N., Alvarez-Buylla, A. & Berger, M. S. Neural stem cells and the origin of gliomas. *The New England journal of medicine***353**, 811-822, doi:10.1056/NEJMra043666 (2005).
47 Doetsch, F., Caille, I., Lim, D. A., Garcia-Verdugo, J. M. & Alvarez-Buylla, A. Subventricular zone astrocytes are neural stem cells in the adult mammalian brain. *Cell***97**, 703-716 (1999).
48 Seri, B., Garcia-Verdugo, J. M., Collado-Morente, L., McEwen, B. S. & Alvarez-Buylla, A. Cell types, lineage, and architecture of the germinal zone in the adult dentate gyrus. *The Journal of comparative neurology***478**, 359-378, doi:10.1002/cne.20288 (2004).
49 Altman, J. & Das, G. D. Autoradiographic and histological evidence of postnatal hippocampal neurogenesis in rats. *The Journal of comparative neurology***124**, 319-335 (1965).
50 Horner, P. J., Thallmair, M. & Gage, F. H. Defining the NG2-expressing cell of the adult CNS. *Journal of neurocytology***31**, 469-480 (2002).
51 Kriegstein, A. & Alvarez-Buylla, A. The glial nature of embryonic and adult neural stem cells. *Annual review of neuroscience***32**, 149-184, doi:10.1146/annurev.neuro.051508.135600 (2009).
52 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell***144**, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).
53 Nishiyama, A., Komitova, M., Suzuki, R. & Zhu, X. Polydendrocytes (NG2 cells): multifunctional cells with lineage plasticity. *Nature reviews. Neuroscience***10**, 9-22, doi:10.1038/nrn2495 (2009).
54 Nickel, G. C. *et al.* Characterizing mutational heterogeneity in a glioblastoma patient with double recurrence. *PloS one***7**, e35262, doi:10.1371/journal.pone.0035262 (2012).
55 Nowell, P. C. The clonal evolution of tumor cell populations. *Science (New York, N.Y.)***194**, 23-28 (1976).
56 Vogelstein, B. *et al.* Cancer genome landscapes. *Science (New York, N.Y.)***339**, 1546-1558, doi:10.1126/science.1235122 (2013).
57 Parker, N. R., Khong, P., Parkinson, J. F., Howell, V. M. & Wheeler, H. R. Molecular heterogeneity in glioblastoma: potential clinical implications. *Frontiers in oncology***5**, 55, doi:10.3389/fonc.2015.00055 (2015).
58 Cheshier, S. H. *et al.* A neurosurgeon's guide to stem cells, cancer stem cells, and brain tumor stem cells. *Neurosurgery***65**, 237-249; discussion 249-250; quiz N236, doi:10.1227/01.neu.0000349921.14519.2a (2009).
59 Reya, T., Morrison, S. J., Clarke, M. F. & Weissman, I. L. Stem cells, cancer, and cancer stem cells. *Nature***414**, 105-111, doi:10.1038/35102167 (2001).
60 Larjavaara, S. *et al.* Incidence of gliomas by anatomic location. *Neuro-oncology***9**, 319-325, doi:10.1215/15228517-2007-016 (2007).
61 Marusyk, A., Almendro, V. & Polyak, K. Intra-tumour heterogeneity: a looking glass for cancer? *Nature reviews. Cancer***12**, 323-334, doi:10.1038/nrc3261 (2012).
62 Hu, Y. Y. *et al.* Hif-1alpha and Hif-2alpha differentially regulate Notch signaling through competitive interaction with the intracellular domain of Notch receptors in glioma stem cells. *Cancer letters***349**, 67-76, doi:10.1016/j.canlet.2014.03.035 (2014).
63 Snuderl, M. *et al.* Mosaic amplification of multiple receptor tyrosine kinase genes in glioblastoma. *Cancer cell***20**, 810-817, doi:10.1016/j.ccr.2011.11.005 (2011).
64 Szerlip, N. J. *et al.* Intratumoral heterogeneity of receptor tyrosine kinases EGFR and PDGFRA amplification in glioblastoma defines subpopulations with distinct growth factor

response. *Proceedings of the National Academy of Sciences of the United States of America***109**, 3041-3046, doi:10.1073/pnas.1114033109 (2012).
65 Brennan, C. W. *et al.* The somatic genomic landscape of glioblastoma. *Cell***155**, 462-477, doi:10.1016/j.cell.2013.09.034 (2013).
66 Raymond, E. *et al.* Phase II study of imatinib in patients with recurrent gliomas of various histologies: a European Organisation for Research and Treatment of Cancer Brain Tumor Group Study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology***26**, 4659-4665, doi:10.1200/jco.2008.16.9235 (2008).
67 Reardon, D. A. *et al.* Phase 2 trial of erlotinib plus sirolimus in adults with recurrent glioblastoma. *Journal of neuro-oncology***96**, 219-230, doi:10.1007/s11060-009-9950-0 (2010).
68 Mellinghoff, I. K., Schultz, N., Mischel, P. S. & Cloughesy, T. F. Will kinase inhibitors make it as glioblastoma drugs? *Current topics in microbiology and immunology***355**, 135-169, doi:10.1007/82_2011_178 (2012).
69 Wen, P. Y. *et al.* Phase I/II study of erlotinib and temsirolimus for patients with recurrent malignant gliomas: North American Brain Tumor Consortium trial 04-02. *Neuro-oncology***16**, 567-578, doi:10.1093/neuonc/not247 (2014).
70 Bedard, P. L., Hansen, A. R., Ratain, M. J. & Siu, L. L. Tumour heterogeneity in the clinic. *Nature***501**, 355-364, doi:10.1038/nature12627 (2013).
71 Patel, A. P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science (New York, N.Y.)***344**, 1396-1401, doi:10.1126/science.1254257 (2014).
72 Meissner, A. *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature***454**, 766-770, doi:10.1038/nature07107 (2008).
73 Mohn, F. *et al.* Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Molecular cell***30**, 755-766, doi:10.1016/j.molcel.2008.05.007 (2008).
74 Christensen, B. C. *et al.* DNA methylation, isocitrate dehydrogenase mutation, and survival in glioma. *Journal of the National Cancer Institute***103**, 143-153, doi:10.1093/jnci/djq497 (2011).
75 Bird, A. DNA methylation patterns and epigenetic memory. *Genes & development***16**, 6-21, doi:10.1101/gad.947102 (2002).
76 Jabbari, K. & Bernardi, G. Cytosine methylation and CpG, TpG (CpA) and TpA frequencies. *Gene***333**, 143-149, doi:10.1016/j.gene.2004.02.043 (2004).
77 Jaenisch, R. & Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature genetics***33 Suppl**, 245-254, doi:10.1038/ng1089 (2003).
78 Blackledge, N. P. & Klose, R. CpG island chromatin: a platform for gene regulation. *Epigenetics : official journal of the DNA Methylation Society***6**, 147-152 (2011).
79 Deaton, A. M. & Bird, A. CpG islands and the regulation of transcription. *Genes & development***25**, 1010-1022, doi:10.1101/gad.2037511 (2011).
80 Laird, P. W. The power and the promise of DNA methylation markers. *Nature reviews. Cancer***3**, 253-266, doi:10.1038/nrc1045 (2003).
81 Jones, P. A. & Baylin, S. B. The epigenomics of cancer. *Cell***128**, 683-692, doi:10.1016/j.cell.2007.01.029 (2007).
82 Pollack, Y., Stein, R., Razin, A. & Cedar, H. Methylation of foreign DNA sequences in eukaryotic cells. *Proceedings of the National Academy of Sciences of the United States of America***77**, 6463-6467 (1980).

83 Wigler, M., Levy, D. & Perucho, M. The somatic replication of DNA methylation. *Cell***24**, 33-40 (1981).
84 Stein, R., Gruenbaum, Y., Pollack, Y., Razin, A. & Cedar, H. Clonal inheritance of the pattern of DNA methylation in mouse cells. *Proceedings of the National Academy of Sciences of the United States of America***79**, 61-65 (1982).
85 Noushmehr, H. *et al.* Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer cell***17**, 510-522, doi:10.1016/j.ccr.2010.03.017 (2010).
86 Li, Y. *et al.* The DNA methylome of human peripheral blood mononuclear cells. *PLoS biology***8**, e1000533, doi:10.1371/journal.pbio.1000533 (2010).
87 Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature***462**, 315-322, doi:10.1038/nature08514 (2009).
88 Weber, M. *et al.* Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature genetics***39**, 457-466, doi:10.1038/ng1990 (2007).
89 Weber, M. & Schubeler, D. Genomic patterns of DNA methylation: targets and function of an epigenetic mark. *Current opinion in cell biology***19**, 273-280, doi:10.1016/j.ceb.2007.04.011 (2007).
90 Bogdanovic, O. & Veenstra, G. J. DNA methylation and methyl-CpG binding proteins: developmental requirements and function. *Chromosoma***118**, 549-565, doi:10.1007/s00412-009-0221-9 (2009).
91 Baubec, T. *et al.* Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature*, doi:10.1038/nature14176 (2015).
92 Leonhardt, H., Page, A. W., Weier, H. U. & Bestor, T. H. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell***71**, 865-873 (1992).
93 Hirasawa, R. *et al.* Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. *Genes & development***22**, 1607-1616, doi:10.1101/gad.1667008 (2008).
94 Bostick, M. *et al.* UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science (New York, N.Y.)***317**, 1760-1764, doi:10.1126/science.1147939 (2007).
95 Toyota, M. *et al.* CpG island methylator phenotype in colorectal cancer. *Proceedings of the National Academy of Sciences of the United States of America***96**, 8681-8686 (1999).
96 Weisenberger, D. J. *et al.* CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nature genetics***38**, 787-793, doi:10.1038/ng1834 (2006).
97 Kim, T. Y., Zhong, S., Fields, C. R., Kim, J. H. & Robertson, K. D. Epigenomic profiling reveals novel and frequent targets of aberrant DNA methylation-mediated silencing in malignant glioma. *Cancer research***66**, 7490-7501, doi:10.1158/0008-5472.can-05-4552 (2006).
98 Martinez, R. *et al.* A microarray-based DNA methylation study of glioblastoma multiforme. *Epigenetics : official journal of the DNA Methylation Society***4**, 255-264 (2009).
99 Martinez, R., Schackert, G. & Esteller, M. Hypermethylation of the proapoptotic gene TMS1/ASC: prognostic importance in glioblastoma multiforme. *Journal of neuro-oncology***82**, 133-139, doi:10.1007/s11060-006-9264-4 (2007).
100 Nagarajan, R. P. & Costello, J. F. Epigenetic mechanisms in glioblastoma multiforme. *Seminars in cancer biology***19**, 188-197, doi:10.1016/j.semcancer.2009.02.005 (2009).

101 Ohgaki, H. & Kleihues, P. Genetic pathways to primary and secondary glioblastoma. *The American journal of pathology***170**, 1445-1453, doi:10.2353/ajpath.2007.070011 (2007).
102 Kim, Y. W. *et al.* Identification of prognostic gene signatures of glioblastoma: a study based on TCGA data analysis. *Neuro-oncology***15**, 829-839, doi:10.1093/neuonc/not024 (2013).
103 Polyak, K. & Weinberg, R. A. Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. *Nature reviews. Cancer***9**, 265-273, doi:10.1038/nrc2620 (2009).
104 Carro, M. S. *et al.* The transcriptional network for mesenchymal transformation of brain tumours. *Nature***463**, 318-325, doi:10.1038/nature08712 (2010).
105 Tso, C. L. *et al.* Primary glioblastomas express mesenchymal stem-like properties. *Molecular cancer research : MCR***4**, 607-619, doi:10.1158/1541-7786.mcr-06-0005 (2006).
106 Hockemeyer, D. *et al.* A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. *Cell stem cell***3**, 346-353, doi:10.1016/j.stem.2008.08.014 (2008).
107 Machida, Y. J., Hamlin, J. L. & Dutta, A. Right place, right time, and only once: replication initiation in metazoans. *Cell***123**, 13-24, doi:10.1016/j.cell.2005.09.019 (2005).
108 Sherr, C. J. & Roberts, J. M. Living with or without cyclins and cyclin-dependent kinases. *Genes & development***18**, 2699-2711, doi:10.1101/gad.1256504 (2004).
109 Planas-Silva, M. D. & Weinberg, R. A. The restriction point and control of cell proliferation. *Current opinion in cell biology***9**, 768-772 (1997).
110 Blow, J. J. & Hodgson, B. Replication licensing--defining the proliferative state? *Trends in cell biology***12**, 72-78 (2002).
111 Pardee, A. B. G1 events and regulation of cell proliferation. *Science (New York, N.Y.)***246**, 603-608 (1989).
112 Holmes, W. F. *et al.* Coordinate control and selective expression of the full complement of replication-dependent histone H4 genes in normal and cancer cells. *The Journal of biological chemistry***280**, 37400-37407, doi:10.1074/jbc.M506995200 (2005).
113 Ghule, P. N. *et al.* Staged assembly of histone gene expression machinery at subnuclear foci in the abbreviated cell cycle of human embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America***105**, 16964-16969, doi:10.1073/pnas.0809273105 (2008).
114 Langst, G., Bonte, E. J., Corona, D. F. & Becker, P. B. Nucleosome movement by CHRAC and ISWI without disruption or trans-displacement of the histone octamer. *Cell***97**, 843-852 (1999).
115 Medina, R. *et al.* Epigenetic control of cell cycle-dependent histone gene expression is a principal component of the abbreviated pluripotent cell cycle. *Molecular and cellular biology***32**, 3860-3871, doi:10.1128/mcb.00736-12 (2012).
116 Marzluff, W. F. Metazoan replication-dependent histone mRNAs: a distinct set of RNA polymerase II transcripts. *Current opinion in cell biology***17**, 274-280, doi:10.1016/j.ceb.2005.04.010 (2005).
117 Groth, A., Rocha, W., Verreault, A. & Almouzni, G. Chromatin challenges during DNA replication and repair. *Cell***128**, 721-733, doi:10.1016/j.cell.2007.01.030 (2007).
118 Osley, M. A. The regulation of histone synthesis in the cell cycle. *Annual review of biochemistry***60**, 827-861, doi:10.1146/annurev.bi.60.070191.004143 (1991).
119 Mazor, T. *et al.* DNA Methylation and Somatic Mutations Converge on the Cell Cycle and Define Similar Evolutionary Histories in Brain Tumors. *Cancer cell***28**, 307-317, doi:10.1016/j.ccell.2015.07.012 (2015).

120 Mitra, P. *et al.* Identification of HiNF-P, a key activator of cell cycle-controlled histone H4 genes at the onset of S phase. *Molecular and cellular biology***23**, 8110-8123 (2003).
121 van Wijnen, A. J. *et al.* CDP/cut is the DNA-binding subunit of histone gene transcription factor HiNF-D: a mechanism for gene regulation at the G1/S phase cell cycle transition point independent of transcription factor E2F. *Proceedings of the National Academy of Sciences of the United States of America***93**, 11516-11521 (1996).
122 Stein, G. *et al.* Altered binding of human histone gene transcription factors during the shutdown of proliferation and onset of differentiation in HL-60 cells. *Proceedings of the National Academy of Sciences of the United States of America***86**, 1865-1869 (1989).
123 Ramsey-Ewing, A., Van Wijnen, A. J., Stein, G. S. & Stein, J. L. Delineation of a human histone H4 cell cycle element in vivo: the master switch for H4 gene transcription. *Proceedings of the National Academy of Sciences of the United States of America***91**, 4475-4479 (1994).
124 Braastad, C. D., Hovhannisyan, H., van Wijnen, A. J., Stein, J. L. & Stein, G. S. Functional characterization of a human histone gene cluster duplication. *Gene***342**, 35-40, doi:10.1016/j.gene.2004.07.036 (2004).
125 Stein, G. S., Stein, J. L., Van Wijnen, A. J. & Lian, J. B. Transcriptional control of cell cycle progression: the histone gene is a paradigm for the G1/S phase and proliferation/differentiation transitions. *Cell biology international***20**, 41-49, doi:10.1006/cbir.1996.0007 (1996).
126 Dominski, Z. & Marzluff, W. F. Formation of the 3' end of histone mRNA. *Gene***239**, 1-14 (1999).
127 Kadonaga, J. T. & Tjian, R. Affinity purification of sequence-specific DNA binding proteins. *Proceedings of the National Academy of Sciences of the United States of America***83**, 5889-5893 (1986).
128 Marzluff, W. F., Wagner, E. J. & Duronio, R. J. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nature reviews. Genetics***9**, 843-854, doi:10.1038/nrg2438 (2008).
129 Marzluff, W. F., Gongidi, P., Woods, K. R., Jin, J. & Maltais, L. J. The human and mouse replication-dependent histone genes. *Genomics***80**, 487-498 (2002).
130 Girirajan, S., Campbell, C. D. & Eichler, E. E. Human copy number variation and complex genetic disease. *Annual review of genetics***45**, 203-226, doi:10.1146/annurev-genet-102209-163544 (2011).
131 Li, W. & Olivier, M. Current analysis platforms and methods for detecting copy number variation. *Physiological genomics***45**, 1-16, doi:10.1152/physiolgenomics.00082.2012 (2013).
132 Feber, A. *et al.* Using high-density DNA methylation arrays to profile copy number alterations. *Genome biology***15**, R30, doi:10.1186/gb-2014-15-2-r30 (2014).
133 Fridley, B. L. *et al.* Germline copy number variation and ovarian cancer survival. *Frontiers in genetics***3**, 142, doi:10.3389/fgene.2012.00142 (2012).
134 Haraksingh, R. R., Abyzov, A., Gerstein, M., Urban, A. E. & Snyder, M. Genome-wide mapping of copy number variation in humans: comparative analysis of high resolution array platforms. *PloS one***6**, e27859, doi:10.1371/journal.pone.0027859 (2011).
135 Toedt, G. *et al.* Molecular signatures classify astrocytic gliomas by IDH1 mutation status. *International journal of cancer. Journal international du cancer***128**, 1095-1103, doi:10.1002/ijc.25448 (2011).
136 Shete, S. *et al.* Genome-wide association study identifies five susceptibility loci for glioma. *Nature genetics***41**, 899-904, doi:10.1038/ng.407 (2009).

137 Wrensch, M. *et al.* Variants in the CDKN2B and RTEL1 regions are associated with high-grade glioma susceptibility. *Nature genetics***41**, 905-908, doi:10.1038/ng.408 (2009).
138 Arslantas, A. *et al.* The importance of genomic copy number changes in the prognosis of glioblastoma multiforme. *Neurosurgical review***27**, 58-64, doi:10.1007/s10143-003-0279-4 (2004).
139 Hodgson, J. G. *et al.* Comparative analyses of gene copy number and mRNA expression in glioblastoma multiforme tumors and xenografts. *Neuro-oncology***11**, 477-487, doi:10.1215/15228517-2008-113 (2009).
140 Lee, J. *et al.* 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, doi:10.1016/j.ccr.2006.03.030 (2006).
141 Bibikova, M. *et al.* Genome-wide DNA methylation profiling using Infinium(R) assay. *Epigenomics***1**, 177-200, doi:10.2217/epi.09.14 (2009).
142 Dedeurwaerder, S. *et al.* Evaluation of the Infinium Methylation 450K technology. *Epigenomics***3**, 771-784, doi:10.2217/epi.11.105 (2011).
143 Rakyan, V. K., Down, T. A., Balding, D. J. & Beck, S. Epigenome-wide association studies for common human diseases. *Nature reviews. Genetics***12**, 529-541, doi:10.1038/nrg3000 (2011).
144 Sandoval, J. *et al.* Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics : official journal of the DNA Methylation Society***6**, 692-702 (2011).
145 Houseman, E. A. *et al.* Model-based clustering of DNA methylation array data: a recursive-partitioning algorithm for high-dimensional data arising as a mixture of beta distributions. *BMC bioinformatics***9**, 365, doi:10.1186/1471-2105-9-365 (2008).
146 Jaffe, A. E. *et al.* Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies. *International journal of epidemiology***41**, 200-209, doi:10.1093/ije/dyr238 (2012).
147 Koestler, D. C. *et al.* Semi-supervised recursively partitioned mixture models for identifying cancer subtypes. *Bioinformatics (Oxford, England)***26**, 2578-2585, doi:10.1093/bioinformatics/btq470 (2010).
148 Zhuang, J., Widschwendter, M. & Teschendorff, A. E. A comparison of feature selection and classification methods in DNA methylation studies using the Illumina Infinium platform. *BMC bioinformatics***13**, 59, doi:10.1186/1471-2105-13-59 (2012).
149 Teschendorff, A. E. *et al.* A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics (Oxford, England)***29**, 189-196, doi:10.1093/bioinformatics/bts680 (2013).
150 Bibikova, M. *et al.* High density DNA methylation array with single CpG site resolution. *Genomics***98**, 288-295, doi:10.1016/j.ygeno.2011.07.007 (2011).
151 Wang, T. S. Eukaryotic DNA polymerases. *Annual review of biochemistry***60**, 513-552, doi:10.1146/annurev.bi.60.070191.002501 (1991).
152 Greshock, J. *et al.* A comparison of DNA copy number profiling platforms. *Cancer research***67**, 10173-10180, doi:10.1158/0008-5472.can-07-2102 (2007).
153 Perry, G. H. *et al.* The fine-scale and complex architecture of human copy-number variation. *American journal of human genetics***82**, 685-695, doi:10.1016/j.ajhg.2007.12.010 (2008).

154 Morris, T. J. *et al.* ChAMP: 450k Chip Analysis Methylation Pipeline. *Bioinformatics (Oxford, England)***30**, 428-430, doi:10.1093/bioinformatics/btt684 (2014).
155 Fernandez-Jimenez, N. *et al.* Accuracy in copy number calling by qPCR and PRT: a matter of DNA. *PloS one***6**, e28910, doi:10.1371/journal.pone.0028910 (2011).
156 McCaughan, F. *et al.* Microdissection molecular copy-number counting (microMCC)--unlocking cancer archives with digital PCR. *The Journal of pathology***216**, 307-316, doi:10.1002/path.2413 (2008).
157 Ceulemans, S., van der Ven, K. & Del-Favero, J. Targeted screening and validation of copy number variations. *Methods in molecular biology (Clifton, N.J.)***838**, 311-328, doi:10.1007/978-1-61779-507-7_15 (2012).
158 van den Boom, J., Wolter, M., Blaschke, B., Knobbe, C. B. & Reifenberger, G. Identification of novel genes associated with astrocytoma progression using suppression subtractive hybridization and real-time reverse transcription-polymerase chain reaction. *International Journal of Cancer***119**, 2330-2338, doi:10.1002/ijc.22108 (2006).
159 Reitman, Z. J. & Yan, H. Isocitrate dehydrogenase 1 and 2 mutations in cancer: alterations at a crossroads of cellular metabolism. *Journal of the National Cancer Institute***102**, 932-941, doi:10.1093/jnci/djq187 (2010).
160 Koivunen, P. *et al.* Transformation by the (R)-enantiomer of 2-hydroxyglutarate linked to EGLN activation. *Nature***483**, 484-488, doi:10.1038/nature10898 (2012).
161 Xu, W. *et al.* Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. *Cancer cell***19**, 17-30, doi:10.1016/j.ccr.2010.12.014 (2011).
162 Johnson, B. E. *et al.* Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. *Science (New York, N.Y.)***343**, 189-193, doi:10.1126/science.1239947 (2014).
163 Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *The New England journal of medicine***366**, 883-892, doi:10.1056/NEJMoa1113205 (2012).
164 Turcan, S. *et al.* IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature***483**, 479-483, doi:10.1038/nature10866 (2012).
165 Merkle, F. T. & Alvarez-Buylla, A. Neural stem cells in mammalian development. *Current opinion in cell biology***18**, 704-709, doi:10.1016/j.ceb.2006.09.008 (2006).
166 Horbinski, C., Miller, C. R. & Perry, A. Gone FISHing: clinical lessons learned in brain tumor molecular diagnostics over the last decade. *Brain pathology (Zurich, Switzerland)***21**, 57-73, doi:10.1111/j.1750-3639.2010.00453.x (2011).
167 Bengtsson, M., Stahlberg, A., Rorsman, P. & Kubista, M. Gene expression profiling in single cells from the pancreatic islets of Langerhans reveals lognormal distribution of mRNA levels. *Genome research***15**, 1388-1392, doi:10.1101/gr.3820805 (2005).
168 Dalerba, P. *et al.* Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nature biotechnology***29**, 1120-1127, doi:10.1038/nbt.2038 (2011).
169 Kalisky, T., Blainey, P. & Quake, S. R. Genomic analysis at the single-cell level. *Annual review of genetics***45**, 431-445, doi:10.1146/annurev-genet-102209-163607 (2011).
170 Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y. & Tyagi, S. Stochastic mRNA synthesis in mammalian cells. *PLoS biology***4**, e309, doi:10.1371/journal.pbio.0040309 (2006).

171 Sherry, M. M., Reeves, A., Wu, J. K. & Cochran, B. H. STAT3 is required for proliferation and maintenance of multipotency in glioblastoma stem cells. *Stem cells (Dayton, Ohio)***27**, 2383-2392, doi:10.1002/stem.185 (2009).
172 Di Stefano, A. L. *et al.* Detection, Characterization, and Inhibition of FGFR-TACC Fusions in IDH Wild-type Glioma. *Clinical cancer research : an official journal of the American Association for Cancer Research***21**, 3307-3317, doi:10.1158/1078-0432.ccr-14-2199 (2015).