Integration of Genetic and Epigenetic Alterations in the Discovery of Molecular Drivers of Malignancy in Glioma

By

Ashley A. Smith

B.S. and B.A, Roger Williams University, Bristol RI, 02809

A dissertation submitted in partial fulfillment of the

requirements for degree of Doctor of Philosophy in

the Division of Biology and Medicine at Brown University

Providence, Rhode Island

May 2014

© Copyright 2014 by Ashley A. Smith

This dissertation by Ashley A. Smith is accepted in its present form

by the Division of Biology and Medicine as satisfying the

dissertation requirements for the degree of Doctor of Philosophy

Date _____ _____

Karl T. Kelsey, M.D., M.O.H., Advisor

Recommended to the Graduate Council

Date _____ _____

Carmen J. Marsit Ph.D. Reader (Chair)

Date _____ _____

E. Andrés Houseman, Sc.D., Reader

Date _____ _____

Yen-Tsung Huang, M.D., M.P.H., S.M., Sc.D. Reader

Date _____ _____

John K. Wiencke (Outside Reader)

Approved by the Graduate Council

Date _____ _____

Peter M. Weber, Ph.D., Dean of the Graduate School

CURRICULUM VITAE

Ashley A. Smith
Department of Pathology and Laboratory Medicine, Box G-E3,
Brown University, Providence, RI 02912
Phone: (704) 813-9269; E-mail: Ashley_Smith@brown.edu; Born: June 6, 1984 in Reading PA

EDUCATION

2013 **Ph.D.** (Pathobiology), Brown University, *Providence, RI, USA*
Qualification Exam Project: DNA methylation profiles in brain tumors associate with histology, predict outcome and reflect immune response

2006 **B.S.** (Biology), Roger Williams University, *Bristol, RI, USA*
B.A. (Chemistry), Roger Williams University, *Bristol, RI, USA*

PUBLICATIONS

1. **Smith AA,** Huang, YT, Eliot M, Houseman EA, Marsit CJ, Wiencke JK, Kelsey KT. "A novel approach to the discovery of survival biomarkers in glioma using a joint analysis of DNA methylation and gene expression". *Epigenetics.* 2013: In Review

2. Christensen BC*, **Smith AA*,** Zheng S, Koestler DC, Houseman EA, Marsit CJ, Wiemels JL, Nelson HH, Karagas MR, Wrensch MR, Kelsey KT, Wiencke JK. "DNA methylation, isocitrate dehydrogenase mutation, and survival in glioma". *JNCI.* 2010: 103(2):143-53.

3. Pietruska J, Liu X, **Smith A**, McNeil K, Weston P, Zhitkovich A, Hurt R, Kane A, " Bioavailability, intracellular mobilization of nickel, and HIF-1 activation in human lung epithelial cells exposed to metallic nickel and nickel oxide nanoparticles ". *Toxicol Sci.*, 2011: 124(1):138-48.

*** Co-authorship**

MEETING ABSTRACTS

1. **Smith AA**, Accomando WP, Wiencke JK, Houseman EA, Marsit CJ, Kelsey KT. "Utility of DMRs that characterize monocytes in non-diseased and diseased brain". In: *Society of NeuroOncology: Proceedings*; 2012 Nov. 15-18; Washington, DC, USA. Abstract ME-10.

2. Christensen BC*, **Smith AA*,** Zheng S, Koestler DC, Houseman EA, Marsit CJ, Wiemels JL, Nelson HH, Karagas MR, Wrensch MR, Kelsey KT, Wiencke JK. "*IDH* mutation defines methylation class and survival in human glioma". In: *Society of NeuroOncology* ; 2010 Nov. 19-21; Montreal, Quebec. Canada. Abstract OM-33

3. Sorger T, Ammon N, **Smith A**, Ronayne R. "The growth and adhesion of two mouse mesothelial cell lines are related to their rates of tumor formation". In: American Society for Cell Biology; 2005 Dec. 10:14; San Francisco, CA, USA. Abstract L462

- **Co-authorship**

INVITED LECTURES

1. Brown University Pathobiology Retreat, "Joint analysis of DNA methylation and gene expression on survival in glioma". August 2013. *Warren, RI, USA*

2. *Society of NeuroOncology Conference,* "*IDH* mutation defines methylation class and survival in human glioma". November 2010. *Montreal, Quebec, Canada.*

3. International Perspectives on Environmental Nanotechnology."Bioavailability and Toxicity of Nickel in Metallic Nanoparticles". October 2008. *Chicoago, IL, USA.*

TEACHING AND MENTORING EXPERIENCE

Graduate Courses

2012 **BIOL 2860 (Molecular Mechanisms of Disease)**; 1.5 lecture hours on epigenetics and cancer; Brown University Pathobiology Graduate Program, *Providence, RI*

Undergraduate Courses

2010 **BIOL 1290 (Cancer Biology)**; Teachers Assistant; Brown University, *Providence, RI, USA*

2008 **BIOL 2860 (Molecular Mechanisms of Disease)**; 1.5 lecture hours on Microscopy; Brown University Department of Pathobiology Graduate Program, *Providence, RI, USA*

2005 **BIO 325L (Cell Biology Lab)**; Teachers Assistant; Roger Williams University, *Bristol ,RI, USA*

Undergraduate Mentoring

2011 Emily Doyle (Intern, Brown University) "CpG Island methylator phenotype associated with Kras and Braf mutations in colon cancer", Brown University, *Providence, RI*

RESEARCH EXPERIENCE

Brown University, Department of Pathobiology, *Providence, RI, USA*
Project: Integration of genetic and epigenetic alterations in the discovery of molecular drives of malignancy in glioma
Doctoral Research with Advisor Dr. Karl T. Kelsey, MD, MPH (2008-2013)

Brown University, Department of Pathobiology, *Providence, RI, USA*
Project: Bioavailability, intracellular mobilization of nickel, and HIF-1 activation in human lung epithelial cells exposed to metallic nickel and nickel oxide nanoparticles
Research Assistance working under Principle Investigator Dr.Agnes Kane (2006-2008)

Roger Williams University, Department of Biology, *Bristol, RI, USA*
Project: The growth and adhesion of two mouse mesothelial cell lines are related to their rates of tumor formation
Undergraduate Research with Advisor Dr. Thomas Sorger, PhD (2003-2006)

Brown University, Rhode Island Hospital Department of Surgical Research, *Bristol, RI, US*
Project: Designed constructs for containing superior FRET proteins
Summer internship with principle investigator Dr. Minosoo Kim, PhD (2005)

Roger Williams University, Department of Marine Biology, *Bristol, RI, USA*
Project: Aquaculture
Research assistant under supervision of Brad Bourque, MS (2002-2005)

CERTIFICATIONS

Teaching Certificate I, Sheridan Teaching Center, Brown University

AWARDS/HONORS

Graduate Student and Post-doctoral Travel Award, *Brown University* (2012)
Rhoda Simper Travel Award, *Brown University, Pathobiology Program* (2011)
Graduated "magna cum laude", *Roger Williams University* (2006)
Who's Who among students (Inducted May 2006)
Tri-Beta Biological Honor Society (Inducted 2005)
Roger Williams University All Academic Team (2003-2006)
Roger Williams University Dean's List (2003-2006)
Annual Roger Williams University Achievement Scholarship (2002-2006)
Annual Unilever Bestfoods Academic Scholarship (2002-2006)

Acknowledgements

First and foremost, I would like to express my gratitude to my advisor, Dr. Karl T. Kelsey. Without your guidance, mentorship, and constant patience, I would not be the scientist I am today. Thank you for allowing me to explore my own theories even if you did not always agree, and most importantly, thank you for teaching me that it is ok to fail as long as you keep going.

I would also like to thank Dr. John K. Weincke, who I consider a second mentor. Your input and guidance is reflected throughout this thesis. Additionally, I would like to thank Dr. Carmen J. Marsit, who has acted as a mentor to me since the time at which I was applying to grad school, who continued to teach me as my professor, and who continues to support me as the chair of my committee. I would also like to thank Dr. Andrés Houseman and Dr. Yen-Tsung Huang, whose statistical advice and constant patience was critical to this thesis.

Thank you to all of my collaborators and friends in the laboratory (past and present) particularly, Graham, Haley, Billy, Scott, Rondi (lab mom), and Liz, for the unwavering support (and libation) both inside and outside of lab. In addition, I would like to express a particular thanks to Michele, who has been there for me in every aspect that one person could, I am eternally grateful.

I would also like to thank my friends (human and animal) from Windswept Farms, OSCF, and Providence Rugby, for giving me an outlet so I could maintain my sanity. Thank you to all of my friends, family, and extended family who supported me throughout this entire process including: Pop and Nonni for your support not just these past 5 years, but

throughout my life; Dad, for always helping me find the humor in unfortunate situations; my Sisterface, for constantly smacking the sense back into me #YOLO, and my Mentor and DT, both of whom had to talk me off the ledge countless times but never let me quit. Finally, I would like to thank my mother, Allison A. Smith: you are my one true inspiration and a constant reminder of why I do what I do. **ILUITU <3**

Preface

The sum of the work presented in this Ph.D. thesis has been executed by me in collaboration with internal and external investigators, who have been acknowledged appropriately in Chapters 2 and 3. My effort was critical in the planning, execution, analysis, and discussion as presented herein.

Table of Contents

ABSTRACT ……..……..…….1

CHAPTER 1: INTRODUCTION ……..…….3

THESIS OVERVIEW…...……..4

GLIOMA: PRESENTATION, DIAGNOSIS, AND TREATMENT……..5

GLIOMA: EPIDEMIOLOGY, RISK, AND SURVIVAL……10

GLIOMA: GENETICS……..11

GLIOMA: EPIGENETICS……14

GLIOMA: INTEGRATION OF GENETICS AND EPIGENETICS……17

CONCLUSION……..18

REFERENCES…….20

CHAPTER 2: DNA METHYLATION, ISOCITRATIE DEHYDROGENASE MUTATION AND SURVIVAL IN GLIOMA……..32

CONTEXT AND CAVEATS……35

ABSTRACT……...37

INTRODUCTION……39

PATIENTS, MATERIALS, AND METHODS…….....41

STATISTICAL ANALYSIS……..…46

RESULTS……...…49

DISCUSSION……54

REFERENCES……..60

CHAPTER 3: A NOVEL APPROACH TO THE DISCOVERY OF SURVIVAL BIOMARKERS IN GLIOMA USING A JOINT ANALYSIS OF DNA METHYLATION AND GENE EXPRESSION……..102

ABSTRACT……103

INTRODUCTION……...106

RESULTS……108

DISCUSSION……..111

MATERIALS AND METHODS……117

ACKNOWLEDMENTS……..122

REFERENCES……....123

CHAPTER 4: DISCUSSION ……...153

CONCLUSION……..……154

FUTURE DIRECTIONS……168

REFERENCES……....169

List of Tables

CHAPTER 2

TABLE 1: Patient demographic and tumor characteristics……..69

TABLE 2: Patient age, grade-specific glioma histology, grade, *TP53* mutation, and *EGFR* amplification stratified by *IDH* mutation status……..…70

TABLE 3: Survival analysis using multivariable Cox proportional hazards model……71

SUPPLEMENTARY TABLE 1: Primer sequences for quantitative methylation specific polymerase chain reaction (QMSP), *IDH* mutation, *TP53* mutation, and *EGFR* amplification experiments…….....73

SUPPLEMENTARY TABLE 2: Recursively partitioned mixture model methylation class by glioma histology and predicted methylation class membership for The Cancer Genome Atlas (TCGA) glioblastoma samples……..74

SUPPLEMENTARY TABLE 3: Association between GoldenGate array methylation values and quantitative methylation specific polymerase chain reaction (QMSP)……75

SUPPLEMENTARY TABLE 4: Identification numbers (ID) and RPMM methylation class membership for The Cancer Genome Atlas (TCGA) glioblastoma samples used in validation……76

SUPPLEMENTARY TABLE 5: Cellular pathways enriched among statistically significantly differentially methylated CpG loci in common among glioblastomas, astrocytomas, oligoastrocytomas, and oligodendrogliomas……..…79

SUPPLEMENTARY TABLE 6: Statistically significantly differentially hypomethylated CpG loci in human gliomas……...80

SUPPLEMENTARY TABLE 7: Statistically significantly differentially hypermethylated CpG loci in human gliomas……..….85

SUPPLEMENTARY TABLE 8: Cellular pathways enriched among statistically significantly differentially methylated CpG loci in gliomas with an *IDH* mutation compared to gliomas without *IDH* mutation…….99

SUPPLEMENTARY TABLE 9: Recursively partitioned mixture model (RPMM) methylation class membership and glioma tumor grade and histology……...100

SUPPLEMENTARY TABLE REFERENCES……...101

CHAPTER 3

TABLE 1: Patient demographic and tumor characteristics……...133

TABLE 2: Final 35 DNA methylation/gene expression pairs that are significantly associated with survival……..……134

TABLE 3: Functions of significant genes and potential mechanisms in glioma.……135

SUPPLEMENTARY TABLE 1: DNA methylation/ expression pairs that are significantly associated with survival (q-value<0.1)……...151

SUPPLEMENTARY TABLE 2: Expression-based and alternative associations of DNA methylation on gene expression and survival……152

List of Figures

CHAPTER 2

FIGURE 1: Association between glioma histologic subtypes and DNA methylation pattern……...66

FIGURE 2: Differential methylation and the ratio of hyper- to hypomethylated loci in gliomas ……..…67

FIGURE 3: Association between *IDH* mutation and methylation phenotype in gliomas …..……..…68

SUPPLEMENTARY FIGURE 1. Association between *IDH* mutation and increased *MGMT* methylation……...72

CHAPTER 3

FIGURE 1: Significant expression-based and alternative associations of DNA methylation on gene expression and survival……..131

FIGURE 2: Model for mediation analysis……...132

SUPPLEMENTARY FIGURE 1: Removal of IDH1 mutants……139

SUPPLEMENTARY FIGURE 2: Directionality of significant pairs……140

SUPPLEMENTARY FIGURE 3: Map of methylation loci locations from significant methylation/expression pairs ……141

Abstract of Integration of Genetic and Epigenetic Alterations in the Discovery of Molecular Drivers of Malignancy in Glioma, by Ashley A. Smith, Ph.D.

Brown University, May 2014

Gliomas are a family of extremely aggressive brain cancers, which, despite current treatment options, have poor prognoses. There are distinct subtypes of gliomas, and accurately identifying these is critical for diagnosis and management. Often, the pathologic diagnosis of these subtypes is difficult, and research is underway to discover novel biomarkers that aid in accurate subtype identification and prognostication. This thesis focuses on the joint analysis of DNA methylation profiles with somatic mutation and gene expression data in glioma, assessing the nature of their association with each other and, subsequently, with histology and disease outcome. The ultimate goal is to develop potential prognostic biomarkers of the disease.

DNA methylation was determined for several different grades and histologies of glioma in addition to non-brain-tumor controls. The same samples were sequenced for *IDH1/2* mutations. We, and others, discovered an *IDH* hypermethylator phenotype, showing a tight association between the occurrence of *IDH* mutation and hypermethylation. This phenotype had a higher prevalence in low-grade and secondary gliomas. Besides mutation, DNA methylation is also associated with other somatic alterations, which can alter gene expression. To better understand how DNA methylation and gene expression drive glioma, we used an integrative bioinformatics approach; our goal was to investigate DNA methylation that modulates gene expression as well as

independent DNA methylation (methylation that may exert its phenotypic effects through alternative mechanisms), assessing the nature of their association with disease survival. Our model supports the existing theory that DNA methylation can work through gene expression to influence survival outcome but also suggests that DNA methylation can work alone or through alternative mechanisms to influence glioma outcome. In addition, our approach offers an alternative method of biomarker discovery, which could potentially be used for diagnostic and therapeutic purposes. Overall, this work supports the hypothesis that somatic mutations are not solely responsible for the glioma phenotype. Epigenetics, particularly DNA methylation, is also important in both the genesis and outcome of the disease. Furthermore, our model provides an alternative approach for biomarker discovery that may also be applicable to cancers other than glioma.

Chapter 1

Thesis Overview and Introduction

Thesis Overview

Gliomas are a family of extremely aggressive brain cancers, which, despite currently available treatments, have poor prognoses, with high-grade glioblastoma multiforme (GBM) having a median survival time of 15 months. There exist many individual subtypes of glioma, which are both histologically and molecularly distinct, and accurately identifying these subtypes is critical for diagnosis, prognosis, and treatment. Often, the pathologic diagnosis of these subtypes can be difficult, and research is underway to define novel biomarkers of the disease that can assist in accurate subtype identification. There is an array of somatic alterations that can contribute to tumorigenesis, although it is now recognized that genetic alterations alone cannot explain the phenotypes of all human tumors. Currently, increasing attention is being focused on the potential for epigenetic alterations to drive these tumors. The integration of both epigenetic and genetic alterations is critical to more fully understand tumorigenesis. Using an integrated approach could be particularly valuable for studying cancers with poorly understood etiologies as well as for largely incurable cancers, such as glioma. The aims of this thesis were to focus upon the joint analysis of DNA methylation profiles with mutation and expression data in glioma, assessing their associations with histology and outcome, and evaluating their potential utility as biomarkers of the disease.

This thesis begins with a broad introduction to glioma and its histological subtypes, as well as the biology of DNA methylation alterations, gene expression changes, and mutations associated with these phenotypes. Chapter 2 provides details on the

integration of glioma DNA methylation and *IDH* mutation, resulting in the discovery of an IDH-driven hypermethylator phenotype that is associated with the survival outcome of specific glioma subtypes. Chapter 3 describes the results of a two-part bioinformatics-based analysis integrating DNA methylation and gene expression. The first part focuses on methylation-mediated changes in gene expression, which result in differential glioma survival, and the second focuses on DNA methylation mediating survival directly or through mechanisms other than direct changes in gene expression. Additionally, this analysis highlights potential biomarkers of the disease. Finally, Chapter 4 summarizes the conclusions of the previous chapters, discussing the importance of this work and provides potential future directions for this research.

Glioma: presentation, diagnosis, and treatment

Gliomas are malignant brain tumors thought to arise from glial cells or their precursors [1] and account for almost 80% of all primary malignant brain tumors[2]. Clinical presentation of the disease includes headaches, seizures, focal neurologic deficits, confusion, memory loss, and personality changes[3]. However, many patients, particularly with low-grade glioma, remain asymptomatic[4]. Patients suspected of having glioma undergo imaging for initial lesion conformation and grading[3]. Though magnetic resonance imaging (MRI) is the gold standard for investigation of suspected glioma, confirmatory diagnosis is still based on stereotactic biopsies[4,5]. New imagery methods such as diffusion and perfusion-weighted imaging, proton MR spectroscopy, and

susceptibility-weighted imaging provide even more insight into tumor grade and can influence therapeutic decisions[5].

Upon glioma conformation, a stereotactic biopsy is taken, or if placement is conducive to surgery, tumors are resected and biopsied, with the ladder method being preferable for better histological diagnosis, reduction of symptoms from mass effect, and increased efficacy of therapies[6,7]. Biopsies are classified based on guidelines set forth by the World Health Organization (WHO), which divides gliomas into several different subtypes and grades[1]. Subtypes are graded using a I-IV numerical grading system where higher numbers are associated with increased malignancy. Numerical grade is based on the presence or absence of several characteristics, including mitosis, necrosis, nuclear atypia, and endothelial cell proliferation. In addition, tumors are divided into several histological types based on their morphology and predominate cell type. The major histological types include astrocytomas, oligodendrogliomas, mixed oligoastrocytomas, and ependymomas[1]. Several subtypes can be found within each major type of glioma.

The most common subtypes of astrocytic tumors include diffuse and pylocitic astrocytomas. Diffuse astrocytomas (predominately of astrocytic origin), account for almost 80% of adult gliomas and are most frequently found in the cerebral hemispheres[1,8]. Diffuse astrocytomas (well-differentiated, anaplastic, and glioblastoma) range from grade II-IV respectively, with glioblastoma multiforme (GBM) being the most malignant of all gliomas. Pilocytic astrocytomas are generally a benign tumor with a WHO grade of I and usually arise in the cerebellum. The second major glioma type, oligodendroglioma (predominantly oligodendrocytic in origin), accounts for 5-15% of gliomas and is usually found in the cerebral hemispheres, specifically the frontal or temporal lobes.

Oligodendrogliomas are further divided into well-differentiated (grade II) and anaplastic (grade III) [1,8]. In addition, mixed oligoastrocytomas consist of a mix of both astrocytes and oligodendrocytes with both well-differentiated (grade II) and anaplastic (grade III) histologies[1]. Finally, in adults, ependymomas (predominantly of ependymal origin) are most commonly found in the spinal cord[8]. Ependymal tumors consist of 4 different subtypes subependymoma, myxopapillary, well-differentiated, and anaplastic, ranging from grade I-III[1]. Due to the heterogeneity of each of the individual subtypes and varying locations of each, glioma management and treatment can vary accordingly.

The general treatment scheme for glioma consists of resection (if applicable), radiation, and/or chemotherapy[4,9]. Due to the location and infiltrative nature of gliomas, many cannot be resected completely or remain inoperable, and tumor resection is closely associated with patient survival [9]. However, advances in surgical techniques have enhanced the ability of surgeons to preform more complete glioma resection[10]. Preoperative techniques such as MRI can work together with intraoperative techniques such as neuronavigation to aid in determining the borders of the brain lesion[10]. This technique is particularly helpful in locating small deep-seated lesions with an accuracy of about 2 mm[11]. Fluorescence-guided resection is another intraoperative imaging technique where fluorescence is used to contrast normal vs. tumor tissue, allowing for more accurate and complete resection[10]. Techniques such as functional MRI (fMRI) aid in the visualization of active parts of the brain and can be beneficial in obtaining a gross impression of the lesion preoperatively[10]. Additional techniques include CT, 3D planning, fiber tracking, and transcranial magnetic stimulation[10]. If the nature or placement of the tumor does not allow for resection, then a stereotactic biopsy is taken for diagnostic

purposes[3]. Immediately after surgery/biopsy, the main course of treatment is radiotherapy[3]. Radiotherapy is used for both low- (WHO grade II) and high-grade (WHO grade III, IV) gliomas, typically at a maximum dose of 60 Gy, as higher doses have not been associated with improved outcome and can lead to increased toxicity[4,9]. In addition to radiotherapy, chemotherapy may be used, mostly for high-grade tumors[3,9], as it is controversial whether chemotherapy should be offered to low-grade glioma patients before treatment with radiotherapy[4]. Concomitant and adjuvant temozolomide (TMZ) is the most commonly used chemotherapeutic drug for glioma treatment with advantages including oral dosing, ability to cross the blood brain barrier (BBB), preferable toxicity profile compared to other drugs, increased effectiveness, and improved quality-of-life [6,12]. Other chemotherapeutics include carmustine wafers (Gliadel) and PCV (combination of Procarbazine, CCNU, and Vincristine) [3,4,9]. Depending on the tumor grade and type and patient age, a combination of both radiotherapy and chemotherapy is often used[3,4,9]. Additionally, increased knowledge of the pathogenesis of glioma has spurred discussion and trials for targeted molecular-based[13], epigenetic-based[14] and antiangiogenic-based[12,15] therapies.

Unfortunately, the initial brain lesion is not the only concern for treatment. Another major issue with glioma patients is the management of comorbidities associated with the primary tumor. These conditions include seizures, peritumoral edema, venous thromboembolism, cognitive dysfunction, and fatigue[16]. Seizures are a common symptom of glioma, with approximately 20-62% of patients experiencing tumor-related epilepsy during the course of their disease[16]. General treatment for seizures includes a variety of antiepileptic drugs. Unfortunately, antiepileptic drugs can have unwanted interactions

with other glioma-related treatments including induction of the cytochrome P-450 system (as seen with the drug phenytoin), which increases the metabolism of many chemotherapeutic agents. For this reason, antiepileptic drugs that do not induce these enzymes (such as clonazepam) are preferred [16]. Edema is another side effect of the tumor and if not controlled can lead to serious complications and morbidity. Excess fluid build-up is caused by a disruption in the blood-brain barrier, allowing fluid into the extracellular space of the brain parenchyma. Corticosteroids are usually used to manage peritumoral edema by decreasing endothelial permeability. Unfortunately, there are several complications associated with corticosteroids, including gastrointestinal problems, steroid myopathy, and osteoporosis. Using lower doses can reduce side effects, and most subside after treatment has stopped. Venous thromboembolism (VTE) is another complication experienced by glioma patients and can be treated mechanically using elastic compression stockings as well as with anticoagulation therapies such as low molecular weight heparins. Lastly, disruption in cognitive functions and increased fatigue, though not necessarily associated with morbidity, can significantly reduce quality of life in glioma patients. Medications such as methylphenidate have been shown to improve neurobehavioral functioning, reducing fatigue and depression, while increasing cognition[16]. Finally recurrence of the primary tumor is often seen. Recurrence of low-grade glioma has been associated with increased malignancy due to transformation[17]. However, recurrence is more frequent in higher-grade tumors with a median time-to-tumor progression of ~6.9 months[18]. Unfortunately, treatment options for recurrent gliomas are limited due to difficulty of resection and drug resistance[19].

Glioma: epidemiology, risk, and survival

During the years 2005-2009 the incidence (age adjusted) of primary brain and central nervous system (CNS) tumors in the United States was approximately 20.6 per 100,000 people, with the average incidence of malignant tumors in adults (20+ years of age) ranging from 5.80-11.70 per 100,000 people[2]. Of these, gliomas accounted for 29% of all adult tumors and approximately 80% of all adult malignant tumors, with an incidence rate of 6.03 per 100,000 people. GBM and astrocytomas accounted for approximately 76% of all gliomas, with GBM having the highest incidence rate among malignant tumors. Gliomas are most commonly found in patients between the 4th and 6th decades of life, with lower grades often found at the younger end of the age range[4,7]. In addition, malignant glioma incidence is statistically significantly higher in males than in females and in caucasians compared to blacks [2].

There are few risk factors associated with glioma, with environmental/behavioral risk factors being the most attractive to study, since they are modifiable [20,21]. Of these, ionizing radiation is the only known environmental risk factor. However, it has been suggested that non-ionizing radiation could be associated with gliomagenesis. Specifically mentioned is the use of cell phones, which emit low-radiofrequency in close proximity to the head and brain. Though it is possible cell phone use could cause an increase in glioma risk, no substantial evidence for this has been provided[21]. Allergies and immunologic changes; specifically, reduced immunoglobulin E (IgE) have been inversely associated with glioma risk[22]. Genetic risk factors involved in gliomagenesis include single nucleotide polymorphisms (SNPs), which affect detoxification, DNA repair, and cell cycle regulation [3].

Low-grade pilocytic astrocytomas and ependymal tumors have the best prognosis, with an approximate 5-10 year survival rate of 91.4% and 77.6% respectively. Grade II oligodendrogliomas or astrocytomas have a survival range of 5-10 years [3] and; generally, anaplastic oligodendrogliomas (3-5 years) have a better prognosis than anaplastic astrocytomas (2-3 years) [3]. The poorest survival among gliomas is associated with GBM, where median survival is only 12-15 months[8], with a 5-year survival of only 4.7%[2]. However, recent literature has reported on the molecular complexity of these tumors in the hopes of improving survival with better diagnosis and more targeted treatments.

Glioma: genetics

The variability in the etiology, progression, and histologies of gliomas is in part due to their genetic heterogeneity, which includes somatic mutations, deletion/amplifications, copy number variation (CNV) and insertion of repetitive elements. Somatic mutations, particularly in tumor suppressor genes, were some of the first implicated in gliomagenesis. Over 65% of gliomas, predominantly low-grade and secondary GBMs, contain mutated *TP53*[3,13,23]. Mutations in the *RB1* tumor suppressor gene are observed mainly in high-grade gliomas. Additionally, p53 and RB pathways may be affected by mutations/amplifications in *MDM1/2/4/* and *CDKN2A/b* (*INK4A* and *ARF*), as well as *CDK4/6*[13,23,24]. Dysregulation of many tyrosine kinase-signaling pathways is also present in malignant glioma. For instance, *PDGFR* overexpression/amplification is ubiquitous among malignant gliomas, and *EGFR* amplification/overexpression/mutation has become a marker of high-grade glioma and

primary GBM, both of which can cause oncogenic dysregulation of PI3K-AKT-mTOR and Ras-MAPK signaling pathways[13,23]. Also associated with these pathways are mutation/deletion of *PTEN,* which is the primary negative regulator of the PI3K-AKT-mTOR signaling pathway, and mutations in *NF1*, which is the primary negative regulator of the Ras-MAPK pathway[23,24]. Loss of heterozygosity (LOH) of 1p19q is the most prevalent loss among oligodendrogliomas and a predictor of better prognosis[25]. Most recently implicated in glioma are alterations in isocitrate dehydrogenase 1/2(*IDH1/2*) [26] and telomerase reverse transcriptase (*TERT*) [27]. The metabolic enzyme IDH1/2 is mutated at high prevalence in low-grade gliomas and secondary GBMs[26]. Interestingly, patients with *IDH1* mutations tend to be younger and have a better survival outcome[26,28]. Novel mutations in the promoter region of *TERT* have also been discovered [27]; they appear to be mutually exclusive with *IDH1* mutations and demonstrate poorer outcome[29,30]. Additionally, mutations in *ATRX* (α thalassemia/mental retardation syndrome X-linked) have been observed in GBMs wild-type (WT) for TERT[31]. ATRX is involved in chromatin remodeling that is active in telomere biology[31]. Both mutations in *TERT* and *ATRX* suggest the importance of telomerase activation in the development of glioma[29]. There are several recurrent translocations reported in glioma, including the in-frame gene fusion of fibroblast growth factor receptor1/3 (*FGFR1/3*) and transforming acidic coiled-coil (*TACC*) to form *FGFR1/3-TACC3*[32] and *EGFR* fusions with septin 14 (*SEPT14*) [33]. The ladder aids in activation of the STAT3 pathway, whose dysregulation has been associated with glioma infiltration and growth[34]. Finally, genetic risk factors are also involved in glioma etiology. Extensive genome-wide association studies (GWAS) and candidate-gene studies have found associations between glioma risk and single-

nucleotide polymorphisms (SNPs) [35]. Of these, GWAS studies are the most consistently replicated, revealing 8 SNPs/near 7 different genes that are significantly associated with glioma risk: *TERT, EGFR, CCDC26, CDKN2A, PHLDB1, RTEL1,* and *TP53* [35-39].

Integration of these genetic events has allowed for increased understanding of the pathogenesis of glioma and yielded distinct genetic profiles that aid in distinguishing different subtypes for better diagnosis and treatment. Efforts put forth by Godard et al, and Nutt et al have demonstrated that gliomas can be classified based on differential gene expression[40,41], and expression-based classes correlated better with survival than histological outcome[41]. Further investigation revealed that gene expression profiles could be used to further distinguish classes within individual subtypes and aided in the discovery of prognostic markers such as *FABP7,* whose increased expression is associated with poorer outcome in GBM[42]. Further studies used gene-expression signatures to classify gliomas based on their resemblance to different stages of neurogenesis, resulting in three subclasses: proneural, proliferative, and mesenchymal[43]. These classes were further supported and refined by integrating gene-expression with copy number, and mutation data[24,44]. The integration with other genetic events resulted in the aforementioned proneural and mesenchymal classes with the addition of classical and neuronal classes[24,43,44]. Proneural classes are strongly associated with high levels of *TP53* mutation, *PDGF* amplification/mutation, *IDH1*/2 mutation, younger age, and have a trend toward increased survival. The mesenchymal subtype is defined by high expression of *CHI3L1*, *MET*, and *NF1* deletion/mutation. High levels of EGFR amplification/mutation define the classical subtype, and there is a clear difference in response to treatment observed between classical and mesenchymal subtypes[21,35,36,23]. Finally, tumors of the

neural subtype appear to be the most molecularly similar to normal brain, this group also contains the oldest patients[44].

Though genetic-based classes have aided in both prognosis and therapeutic intervention, it has become increasingly apparent that genetics alone cannot explain the phenotype of this complex disease, highlighting the need for studies focusing on not only the genome, but also the epigenome.

Glioma: epigenetics

An epigenetic trait is defined by a heritable, stable change in expression and/or cellular phenotype that does not result from change to the DNA sequence[45,46]. Epigenetic regulators include histone modifications[47], microRNA[48,49], and DNA methylation[50-52], and are critical in normal development contributing to the vast array of cellular phenotypes[52-54]. However, dysregulation of these regulators has been associated with the etiology of many human diseases[55]. Due to its assay accessibility, DNA methylation has been one of the most widely studied epigenetic events[56-59].

DNA methylation occurs on cytosines found 5' to guanines in the DNA sequence (CpG dinucleotides) [52]. Maintenance/deposition of methylation is controlled mainly by three DNA methyltransferases (DNMT1, 3A, 3B) using S-adenosyl methionine as the methyl donor[53,60]. In mammals, approximately 60-80% of CpGs are methylated[53]. CpG dinucleotides are under-represented in the genome, however, they have been found at higher than expected quantities in gene promoter regions[61], and clusters of them are referred to as CpG islands[62]. The placement of CpG islands in promoter regions of genes

allows for epigenetic regulation of transcriptional activity through structural changes in associated chromatin[53,55,60]. For instance, methylation of a CpG island in the promoter region of a gene can work together with histone modifications causing chromatin condensation and inhibition of transcriptional activity, essentially silencing expression of the gene. CpG shores (CpGs that lie ~2kb away from CpG Islands) have also been implicated in transcriptional activity as well as cell programing[63,64]. Furthermore, patterns of DNA methylation can be used to distinguish individual cell types/mixtures and tissues[52,65-67], including different regions of non-diseased brain[68]. DNA methylation is important in many normal processes besides transcriptional regulation and cell programming, including genomic imprinting, silencing of aberrant repetitive elements, and regulation of transcriptional enhancers and splice site variants[52]. Disruption of normal DNA methylation events can cause dysregulation of these processes, which has been associated with adverse health affects including diseases such as cancer [55].

One of the first epigenetic changes implicated in human cancer was a general loss of methylation in tumors compared with normal tissue[69,70]. Hypomethylation is primarily associated with aberrant expression of repetitive elements but can also lead to loss of imprinting and activation of oncogenes[69,71,72]. Furthermore, hypomethylation can promote deletions and translocations by favoring mitotic recombination[73]. Overall, hypomethylation is associated with genomic instability, which can aid in tumor progression[71,72]. Gene-specific hypermethylation is also observed in cancer and is associated with transcriptional inactivation[72-75]. Most ubiquitously observed in carcinogenesis is methylation-induced silencing of tumor suppressor genes, which can aid in tumorigenesis by altering many cancer-related pathways[74]. Patterns of methylation

can also be important prognostic and diagnostic tools in cancer. Differentially methylated regions (DMRs) are regions of the genome demonstrating variable methylation and can be used not only to distinguish different cell and tissue types; but also to aid in distinguishing normal and tumor tissue as well as individual cancer subtypes[64,76]. Genes with differential DNA methylation have become ideal candidates for biomarker selection for both the diagnosis and prognostication of disease while simultaneously highlighting potential therapeutic targets[77]. Another reason DNA methylation is so attractive to study is because, unlike genetic alterations, epigenetic alterations are potentially reversible. The reversibility of DNA methylation has been harnessed for therapeutic reasons in myelodysplastic syndromes and myelogenous leukemia, for which the Food and Drug Administration has already approved the use of drugs which prevent re-methylation (i.e. 5-azacytidine and 5-aza-2'-deoxicytidine) [78-80].

Significant advances in the field of epigenetics have led to the discovery of several epigenetically altered genes/pathways in glioma. Genome-wide hypomethylation is seen in approximately 80% of GBMs, and this loss of methylation is correlated with increased proliferation and aberrant transcriptional activity[81]. The promoter region of putative oncogene *MAGEA1* is hypomethylated in GBM and is associated with increased expression of this cancer-testis antigen [81,82]. Increased activation of MAGE proteins have been implicated in multiple cancers and are associated with T-cell recognition, p53 inhibition, and response to chemotherapy [81,82]. More commonly seen in glioma is locus-specific hypermethylation[81,83,84]. Promoter hypermethylation has been observed in many cancer-related gene pathways, including DNA repair, cell cycle progression, apoptosis, angiogenesis, and cell growth[85-89]. Disruption of any of these pathways can ultimately

lead to variable effects on survival. One example of this phenomenon is the epigenetic silencing of the DNA repair gene *MGMT*, which has become a strong predictor of glioma outcome and response to treatment[90,91]. MGMT normally functions by removing aberrant alkyl groups from the O^6 position of guanine[90,91]. In cancer treatment, *MGMT* expression can decrease the therapeutic efficacy of radiation and alkylating agents such as temozolomide by repairing therapy-induced damage to the tumor cells. DNA gene promoter methylation silencing of *MGMT* is, then, associated with significantly better survival following chemotherapeutic treatments[90,91]. Promoter methylation of *SOCS3* has been implicated in secondary and low-grade gliomagenesis via the STAT3 and MAPK-pathways[92,93]. Methylation of *SOCS3* is significantly associated with poorer survival outcome[92,93]. These examples demonstrate the impact that the epigenome can have on tumorigenesis as well as its importance for diagnosis and survival outcome and as a biomarker of the disease.

Glioma: Integration of genetics and epigenetics

The genetic landscape of glioma is fairly well studied; however, its relationship with the glioma epigenome is poorly understood. Previous literature suggests that complex somatic alterations are involved in gliomagenesis that aid not only in distinguishing glioma from other diseases but also in distinguishing different glioma subtypes. These alterations include both genetic events, such as amplifications/deletions and mutations, as well as epigenetic events such as hyper- and hypo-methylation, all of which can dysregulate cancer-related signaling pathways promoting tumorigenesis and

modulating outcome. The importance of analyses integrating the cancer genome and epigenome has been observed with the identification of a CpG island methylator phenotype in colorectal cancer[94-96]. The integration of both methylation profiles and mutation data demonstrated distinct classes of colorectal cancer, with CIMP-high tumors showing extensive promoter methylation and mutations in the *BRAF* oncogene[95]. In contrast, a CIMP-low phenotype is associated with promoter methylation of a more limited set of genes, particularly age-related genes, and is also associated with mutation in the *KRAS* oncogene[95]. CIMP-negative tumors display rare methylation as well as *TP53* mutation. The prognosis associated with these subgroups also varies, with CIMP-high tumors having the best outcome[96]. In glioma, the link between promoter methylation and gene expression has been established on a single-locus level. However, large-scale integration approaches of methylation patterns and genetic alterations in glioma have not been attempted to date.

Conclusion

This thesis aims to carefully assess the epidemiology of DNA methylation in glioma. Novel high-throughput DNA methylation arrays (Illumina), which interrogate approximately 1,500 cancer-related CpG loci, were used to identify the epigenetic determinants of methylation in glioma and how they associate with genetic alterations such as mutations. The initial results suggested the correlation of a hypermethylator phenotype and *IDH1* mutations with tumor histology and increased prognosis. To further demonstrate the importance of integrative analysis in gliomagenesis and improved

prognosis, data obtained from The Cancer Genome Atlas (TCGA) were used to determine the joint effect of DNA methylation and gene expression on survival outcome in glioma using a novel bioinformatics-based approach.

References

1. Louis DN, Ohgaki H, Wiestler OD, et al. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol.* Aug 2007;114(2):97-109.
2. Dolecek TA, Propp JM, Stroup NE, Kruchko C. CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2005-2009. *Neuro Oncol.* Nov 2012;14 Suppl 5:v1-49.
3. Wen PY, Kesari S. Malignant gliomas in adults. *N Engl J Med.* Jul 2008;359(5):492-507.
4. Sanai N, Chang S, Berger MS. Low-grade gliomas in adults. *J Neurosurg.* Nov 2011;115(5):948-965.
5. Mohammed W, Xunning H, Haibin S, Jingzhi M. Clinical applications of susceptibility-weighted imaging in detecting and grading intracranial gliomas: a review. *Cancer Imaging.* 2013;13:186-195.
6. Pouratian N, Schiff D. Management of low-grade glioma. *Curr Neurol Neurosci Rep.* May 2010;10(3):224-231.
7. Wang Y, Jiang T. Understanding high grade glioma: molecular mechanism, therapy and comprehensive management. *Cancer Lett.* May 2013;331(2):139-146.
8. V. K, A.K. A, J.C. A. *Robbins Basic Pathology*. 9 ed. Canada: Elsevier Saunders; 2013.
9. Stupp R, Tonn JC, Brada M, Pentheroudakis G, Group EGW. High-grade malignant glioma: ESMO Clinical Practice Guidelines for diagnosis, treatment

and follow-up. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO.* May 2010;21 Suppl 5:v190-193.

10. Vranic A. New developments in surgery of malignant gliomas. *Radiol Oncol.* Sep 2011;45(3):159-165.

11. Ganslandt O, Behari S, Gralla J, Fahlbusch R, Nimsky C. Neuronavigation: concept, techniques and applications. *Neurol India.* Sep 2002;50(3):244-255.

12. Bradley D, Rees J. Updates in the management of high-grade glioma. *J Neurol.* Jul 2013.

13. Masui K, Cloughesy TF, Mischel PS. Review: molecular pathology in adult high-grade gliomas: from molecular diagnostics to target therapies. *Neuropathol Appl Neurobiol.* Jun 2012;38(3):271-291.

14. Clarke J, Penas C, Pastori C, et al. Epigenetic pathways and glioblastoma treatment. *Epigenetics.* Jun 2013;8(8).

15. Plate KH, Scholz A, Dumont DJ. Tumor angiogenesis and anti-angiogenic therapy in malignant gliomas revisited. *Acta Neuropathol.* Dec 2012;124(6):763-775.

16. Wen PY, Schiff D, Kesari S, Drappatz J, Gigas DC, Doherty L. Medical management of patients with brain tumors. *Journal of neuro-oncology.* Dec 2006;80(3):313-332.

17. Jaeckle KA, Decker PA, Ballman KV, et al. Transformation of low grade glioma and correlation with outcome: an NCCTG database analysis. *Journal of neuro-oncology.* Aug 2011;104(1):253-259.

18. Stupp R, Mason W, van den Bent M, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med.* Mar 2005;352(10):987-996.

19. Haar CP, Hebbar P, Wallace GC, et al. Drug resistance in glioblastoma: a mini review. *Neurochem Res.* Jun 2012;37(6):1192-1200.

20. Fisher JL, Schwartzbaum JA, Wrensch M, Wiemels JL. Epidemiology of brain tumors. *Neurol Clin.* Nov 2007;25(4):867-890, vii.

21. Bondy ML, Scheurer ME, Malmer B, et al. Brain tumor epidemiology: consensus from the Brain Tumor Epidemiology Consortium. *Cancer.* Oct 2008;113(7 Suppl):1953-1968.

22. Wiemels JL, Wiencke JK, Patoka J, et al. Reduced immunoglobulin E and allergy among adults with glioma compared with controls. *Cancer Res.* Nov 2004;64(22):8468-8473.

23. Huse J, Holland E. Targeting brain cancer: advances in the molecular pathology of malignant glioma and medulloblastoma. *Nat Rev Cancer.* May 2010;10(5):319-331.

24. Brennan CW, Verhaak RG, McKenna A, et al. The Somatic Genomic Landscape of Glioblastoma. *Cell.* Oct 2013;155(2):462-477.

25. Aldape K, Burger PC, Perry A. Clinicopathologic aspects of 1p/19q loss and the diagnosis of oligodendroglioma. *Arch Pathol Lab Med.* Feb 2007;131(2):242-251.

26. Parsons DW, Jones S, Zhang X, et al. An integrated genomic analysis of human glioblastoma multiforme. *Science.* Sep 2008;321(5897):1807-1812.

27. Killela PJ, Reitman ZJ, Jiao Y, et al. TERT promoter mutations occur frequently in gliomas and a subset of tumors derived from cells with low rates of self-renewal. *Proceedings of the National Academy of Sciences of the United States of America.* Apr 2013;110(15):6021-6026.

28. Hartmann C, Meyer J, Balss J, et al. Type and frequency of IDH1 and IDH2 mutations are related to astrocytic and oligodendroglial differentiation and age: a study of 1,010 diffuse gliomas. *Acta Neuropathol.* Oct 2009;118(4):469-474.

29. Arita H, Narita Y, Fukushima S, et al. Upregulating mutations in the TERT promoter commonly occur in adult malignant gliomas and are strongly associated with total 1p19q loss. *Acta Neuropathol.* Aug 2013;126(2):267-276.

30. Nonoguchi N, Ohta T, Oh JE, Kim YH, Kleihues P, Ohgaki H. TERT promoter mutations in primary and secondary glioblastomas. *Acta Neuropathol.* Aug 2013.

31. Kannan K, Inagaki A, Silber J, et al. Whole-exome sequencing identifies ATRX mutation as a key molecular determinant in lower-grade glioma. *Oncotarget.* Oct 2012;3(10):1194-1203.

32. Singh D, Chan JM, Zoppoli P, et al. Transforming fusions of FGFR and TACC genes in human glioblastoma. *Science.* Sep 2012;337(6099):1231-1235.

33. Frattini V, Trifonov V, Chan JM, et al. The integrated landscape of driver genomic alterations in glioblastoma. *Nat Genet.* Aug 2013.

34. Luwor RB, Stylli SS, Kaye AH. The role of Stat3 in glioblastoma multiforme. *J Clin Neurosci.* Jul 2013;20(7):907-911.

35. Walsh KM, Anderson E, Hansen HM, et al. Analysis of 60 reported glioma risk SNPs replicates published GWAS findings but fails to replicate associations from published candidate-gene studies. *Genet Epidemiol.* Feb 2013;37(2):222-228.

36. Sanson M, Hosking FJ, Shete S, et al. Chromosome 7p11.2 (EGFR) variation influences glioma risk. *Human molecular genetics.* Jul 2011;20(14):2897-2904.

37. Shete S, Hosking FJ, Robertson LB, et al. Genome-wide association study identifies five susceptibility loci for glioma. *Nat Genet.* Aug 2009;41(8):899-904.

38. Stacey SN, Sulem P, Jonasdottir A, et al. A germline variant in the TP53 polyadenylation signal confers cancer susceptibility. *Nat Genet.* Nov 2011;43(11):1098-1103.

39. Wrensch M, Jenkins RB, Chang JS, et al. Variants in the CDKN2B and RTEL1 regions are associated with high-grade glioma susceptibility. *Nat Genet.* Aug 2009;41(8):905-908.

40. Godard S, Getz G, Delorenzi M, et al. Classification of human astrocytic gliomas on the basis of gene expression: a correlated group of genes with angiogenic activity emerges as a strong predictor of subtypes. *Cancer Res.* Oct 2003;63(20):6613-6625.

41. Nutt CL, Mani DR, Betensky RA, et al. Gene expression-based classification of malignant gliomas correlates better with survival than histological classification. *Cancer Res.* Apr 2003;63(7):1602-1607.

42. Liang Y, Diehn M, Watson N, et al. Gene expression profiling reveals molecularly and clinically distinct subtypes of glioblastoma multiforme. *Proceedings of the National Academy of Sciences of the United States of America.* Apr 2005;102(16):5814-5819.

43. Phillips HS, Kharbanda S, Chen R, et al. Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. *Cancer Cell.* Mar 2006;9(3):157-173.

44. Verhaak RG, Hoadley KA, Purdom E, et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell.* Jan 2010;17(1):98-110.

45. Berger SL, Kouzarides T, Shiekhattar R, Shilatifard A. An operational definition of epigenetics. *Genes Dev.* Apr 2009;23(7):781-783.

46. Bird A. Perceptions of epigenetics. *Nature.* May 2007;447(7143):396-398.

47. Turner BM. Cellular memory and the histone code. *Cell.* Nov 2002;111(3):285-291.

48. Chuang JC, Jones PA. Epigenetics and microRNAs. *Pediatr Res.* May 2007;61(5 Pt 2):24R-29R.

49. Costa FF. Non-coding RNAs, epigenetics and complexity. *Gene.* Feb 2008;410(1):9-17.

50. Riggs AD. X inactivation, differentiation, and DNA methylation. *Cytogenet Cell Genet.* 1975;14(1):9-25.

51. Holliday R. The inheritance of epigenetic defects. *Science.* Oct 1987;238(4824):163-170.

52. Cedar H, Bergman Y. Programming of DNA methylation patterns. *Annu Rev Biochem.* 2012;81:97-117.

53. Smith ZD, Meissner A. DNA methylation: roles in mammalian development. *Nat Rev Genet.* Mar 2013;14(3):204-220.

54. Weinhold B. Epigenetics: the science of change. *Environ Health Perspect.* Mar 2006;114(3):A160-167.

55. Portela A, Esteller M. Epigenetic modifications and human disease. *Nat Biotechnol.* Oct 2010;28(10):1057-1068.

56. Bibikova M, Barnes B, Tsan C, et al. High density DNA methylation array with single CpG site resolution. *Genomics.* Oct 2011;98(4):288-295.

57. Bibikova M, Le J, Barnes B, et al. Genome-wide DNA methylation profiling using Infinium® assay. *Epigenomics.* Oct 2009;1(1):177-200.

58. Herman JG, Graff JR, Myöhänen S, Nelkin BD, Baylin SB. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proceedings of the National Academy of Sciences of the United States of America.* Sep 1996;93(18):9821-9826.

59. Tost J, Gut IG. DNA methylation analysis by pyrosequencing. *Nat Protoc.* 2007;2(9):2265-2275.

60. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet.* Mar 2003;33 Suppl:245-254.

61. Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev.* Jan 2002;16(1):6-21.

62. Bird AP. CpG-rich islands and the function of DNA methylation. *Nature.* 1986 May 15-21 1986;321(6067):209-213.

63. Aran D, Sabato S, Hellman A. DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes. *Genome Biol.* Mar 2013;14(3):R21.

64. Doi A, Park IH, Wen B, et al. Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat Genet.* Dec 2009;41(12):1350-1353.

65. Ehrlich M, Gama-Sosa MA, Huang LH, et al. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic acids research.* Apr 1982;10(8):2709-2721.

66. Rakyan VK, Down TA, Thorne NP, et al. An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). *Genome Res.* Sep 2008;18(9):1518-1529.

67. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics.* 2012;13:86.

68. Ladd-Acosta C, Pevsner J, Sabunciyan S, et al. DNA methylation signatures within the human brain. *Am J Hum Genet.* Dec 2007;81(6):1304-1315.

69. Feinberg AP, Vogelstein B. Hypomethylation of ras oncogenes in primary human cancers. *Biochemical and biophysical research communications.* Feb 1983;111(1):47-54.

70. Gama-Sosa MA, Slagel VA, Trewyn RW, et al. The 5-methylcytosine content of DNA from human tumors. *Nucleic acids research.* Oct 1983;11(19):6883-6894.

71. Ehrlich M. DNA hypomethylation in cancer cells. *Epigenomics.* Dec 2009;1(2):239-259.

72. Baylin SB, Ohm JE. Epigenetic gene silencing in cancer - a mechanism for early oncogenic pathway addiction? *Nat Rev Cancer.* Feb 2006;6(2):107-116.

73. Esteller M. Epigenetics in cancer. *N Engl J Med.* Mar 2008;358(11):1148-1159.

74. Esteller M. CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene.* Aug 2002;21(35):5427-5440.

75. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. *Nat Rev Genet.* Jun 2002;3(6):415-428.

76. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. *Nat Rev Genet.* Aug 2011;12(8):529-541.

77. Mulero-Navarro S, Esteller M. Epigenetic biomarkers for human cancer: the time is now. *Crit Rev Oncol Hematol.* Oct 2008;68(1):1-11.

78. Issa JP, Gharibyan V, Cortes J, et al. Phase II study of low-dose decitabine in patients with chronic myelogenous leukemia resistant to imatinib mesylate.

Journal of clinical oncology : official journal of the American Society of Clinical Oncology. Jun 2005;23(17):3948-3956.

79. Kaminskas E, Farrell A, Abraham S, et al. Approval summary: azacitidine for treatment of myelodysplastic syndrome subtypes. *Clinical cancer research : an official journal of the American Association for Cancer Research.* May 2005;11(10):3604-3608.

80. Garcia-Manero G, Kantarjian HM, Sanchez-Gonzalez B, et al. Phase 1/2 study of the combination of 5-aza-2'-deoxycytidine with valproic acid in patients with leukemia. *Blood.* Nov 2006;108(10):3271-3279.

81. Nagarajan RP, Costello JF. Epigenetic mechanisms in glioblastoma multiforme. *Semin Cancer Biol.* Jun 2009;19(3):188-197.

82. Cadieux B, Ching TT, VandenBerg SR, Costello JF. Genome-wide hypomethylation in human glioblastomas associated with specific copy number alteration, methylenetetrahydrofolate reductase allele status, and increased proliferation. *Cancer Res.* Sep 2006;66(17):8469-8476.

83. Natsume A, Kondo Y, Ito M, Motomura K, Wakabayashi T, Yoshida J. Epigenetic aberrations and therapeutic implications in gliomas. *Cancer Sci.* Jun 2010;101(6):1331-1336.

84. Martinez R, Esteller M. The DNA methylome of glioblastoma multiforme. *Neurobiol Dis.* Jul 2010;39(1):40-46.

85. Amatya VJ, Naumann U, Weller M, Ohgaki H. TP53 promoter methylation in human gliomas. *Acta Neuropathol.* Aug 2005;110(2):178-184.

86. Baeza N, Weller M, Yonekawa Y, Kleihues P, Ohgaki H. PTEN methylation and expression in glioblastomas. *Acta Neuropathol.* Nov 2003;106(5):479-485.

87. Bello MJ, Rey JA. The p53/Mdm2/p14ARF cell cycle control pathway genes may be inactivated by genetic and epigenetic mechanisms in gliomas. *Cancer genetics and cytogenetics.* Jan 2006;164(2):172-173.

88. Costello JF, Berger MS, Huang HS, Cavenee WK. Silencing of p16/CDKN2 expression in human gliomas by methylation and chromatin condensation. *Cancer Res.* May 1996;56(10):2405-2410.

89. Watanabe T, Yokoo H, Yokoo M, Yonekawa Y, Kleihues P, Ohgaki H. Concurrent inactivation of RB1 and TP53 pathways in anaplastic oligodendrogliomas. *J Neuropathol Exp Neurol.* Dec 2001;60(12):1181-1189.

90. Hegi ME, Diserens AC, Gorlia T, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med.* Mar 2005;352(10):997-1003.

91. Esteller M, Garcia-Foncillas J, Andion E, et al. Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *N Engl J Med.* Nov 2000;343(19):1350-1354.

92. Martini M, Pallini R, Luongo G, Cenci T, Lucantoni C, Larocca LM. Prognostic relevance of SOCS3 hypermethylation in patients with glioblastoma multiforme. *International journal of cancer. Journal international du cancer.* Dec 2008;123(12):2955-2960.

93. Lindemann C, Hackmann O, Delic S, Schmidt N, Reifenberger G, Riemenschneider MJ. SOCS3 promoter methylation is mutually exclusive to

EGFR amplification in gliomas and promotes glioma cell invasion through STAT3 and FAK activation. *Acta Neuropathol.* Aug 2011;122(2):241-251.

94. Issa J. CpG island methylator phenotype in cancer. *Nat Rev Cancer.* Dec 2004;4(12):988-993.

95. Ogino S, Kawasaki T, Kirkner G, Loda M, Fuchs C. CpG island methylator phenotype-low (CIMP-low) in colorectal cancer: possible associations with male sex and KRAS mutations. *J Mol Diagn.* Nov 2006;8(5):582-588.

96. Shen L, Toyota M, Kondo Y, et al. Integrated genetic and epigenetic analysis identifies three different subclasses of colon cancer. *Proceedings of the National Academy of Sciences of the United States of America.* Nov 2007;104(47):18654-18659.

Chapter 2

DNA Methylation, Isocitrate Dehydrogenase Mutation, and Survival in Glioma

Brock C. Christensen, Ashley A. Smith, Shichun Zheng, Devin C. Koestler, E. Andres Houseman, Carmen J. Marsit, Joseph L. Wiemels, Heather H. Nelson, Margaret R. Karagas, Margaret R. Wrensch, Karl T. Kelsey, John K. Wiencke

BCC and AAS contributed equally to the work.

MRW, KTK, and JKW are joint lead investigators.

Journal of National Cancer Institute; January 19, 2011; 103:143-153

DNA Methylation, Isocitrate Dehydrogenase Mutation, and Survival in Glioma

Brock C. Christensen, Ashley A. Smith, Shichun Zheng, Devin C. Koestler, E. Andres Houseman, Carmen J. Marsit, Joseph L. Wiemels, Heather H. Nelson, Margaret R. Karagas, Margaret R. Wrensch, Karl T. Kelsey, John K. Wiencke

Affiliations of authors: Department of Pathology and Laboratory Medicine (BCC, AAS, CJM, KTK); Department of Community Health, Brown University, Providence, RI 02912, USA (BCC, DCK, EAH, KTK); Department of Neurological Surgery, Helen Diller Family Cancer Center (SZ, MRW, JKW); Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA (JLW); Department of Biostatistics, Harvard School of Public Health, Boston, MA (EAH); Masonic Cancer Center, Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, MN (HHN); Section of Biostatistics and Epidemiology, Department of Community and Family Medicine, Dartmouth Medical School, Lebanon, NH (MRK).

BCC and AAS contributed equally to the work.

MRW, KTK, and JKW are joint lead investigators.

Correspondence to: John K. Wiencke, PhD, Department of Neurological Surgery, Helen Diller Family Cancer Center, University of California San Francisco, San Francisco, CA 91458 (e-mail: John.Wiencke@UCSF.edu)

Funding

This study was funded by the National Institute of Health, grant numbers R01CA52689 (to MRW) and P50CA097257 (to MRW and JKW); R01CA078609, R01CA121147, R01CA126939, and R01CA100679 (to KTK); R01ES06717 and R01CA126831 (to JKW); P30CA077598 (to HHN); and Tobacco-Related Diseases Research Program, grant number 18CA-0127 (to JLW)

Notes

The funders did not have any role in the study design, collection of data, interpretation of the results, preparation of the manuscript, or the decision to submit the manuscript for publication.

Context and Caveats

Prior knowledge

Human gliomas often have mutations in the isocitrate dehydrogenase genes (*IDH1* and *IDH2*). *IDH* mutation is associated with improved survival in glioma patients. Epigenetic alterations like DNA methylation at CpG dinucleotides play an important role in gene regulation. Integration of genetic and epigenetic data is important for a better understanding of glioma development.

Study design

DNA methylation profile of CpG loci and methylation class of 131 glioma and seven non-glioma brain tissues were determined. The association between *IDH* mutation and methylation class was analyzed. Survival analysis of patients carrying *IDH* mutation vs. wild-type *IDH* was also performed.

Contribution

CpG loci showed differential methylation between glioma and non-glioma tissues. Statistically significant associations were found between DNA methylation class and histologic subtypes, and between DNA methylation class and *IDH* mutation of gliomas. Patients carrying *IDH* mutation in gliomas showed improved survival compared with patients carrying *IDH* wild-type after adjustment for age and grade-specific tumor histology.

Implications

A distinct methylation pattern in glioma tissues is associated with *IDH* mutation.

Limitations

Mutation data was not available for all tissue samples, which may have limited the statistical power of the analyses.

Abstract

Background: Although much is known about molecular and chromosomal characteristics that distinguish glioma histologic subtypes, DNA methylation patterns of gliomas and their association with other tumor features such as mutation of isocitrate dehydrogenase (*IDH*) genes, has only recently begun to be investigated.

Methods: DNA methylation of glioblastomas, astrocytomas, oligodendrogliomas, oligoastrocytomas, ependymomas, and pilocytic astrocytomas (n = 131) from the Brain Tumor Research Center at the University of California San Francisco, as well as non-tumor brain tissues (n = 7), was assessed with the Illumina GoldenGate methylation array. Methylation data were subjected to recursively partitioned mixture modeling (RPMM) to derive methylation classes. Differential DNA methylation between tumor and non-tumor was also assessed. The association between methylation class and *IDH* mutation (*IDH1* and *IDH2* isoforms) was tested using univariate and multivariable analysis for tumors with available substrate for sequencing (n = 95). Survival of glioma patients carrying mutant *IDH* (n = 56) was compared with patients carrying wild-type *IDH* (n = 39) by using a multivariable Cox proportional hazards model and Kaplan-Meier analysis. All statistical tests were two-sided.

Results: We observed a statistically significant association between RPMM methylation class and glioma histologic subtype ($P < 2.2 \times 10^{-16}$). Compared with non-tumor brain tissues, across glioma tumor histologic subtypes, the differential methylation ratios of CpG loci were statistically significantly different (Permutation $P < .0001$). Methylation class was strongly associated with *IDH* mutation in gliomas ($P = 3.0 \times 10^{-16}$). Compared with glioma patients whose tumors harbored wild-type *IDH*, patients whose tumors

harbored mutant *IDH* showed statistically significantly improved survival (HR of death = 0.27, 95% confidence interval [CI] = 0.10 to 0.72).

Conclusion: The homogeneity of methylation classes for gliomas with *IDH* mutation, despite their histologic diversity, suggests that *IDH* mutation is associated with a distinct DNA methylation phenotype and an altered metabolic profile in glioma.

Introduction

Malignant glioma is the most common form of primary malignant brain tumor and the glioma histologic subtypes include glioblastomas, grades 2 and 3 astrocytomas, grades 2 and 3 oligodendrogliomas, grades 2 and 3 oligoastrocytomas, ependymomas, and pilocytic astrocytomas (1). Presently, there are limited treatment options for glioma; glioblastoma, the most common glioma subtype, remains an incurable disease with a median survival of 15 months, even with radiation and temozolomide therapy (2).

A comprehensive appreciation of the integrated genomics and epigenomics of glioma is needed to better understand the multiple cellular pathways involved in their development, establish markers of resistance to traditional therapies, and contribute to the development of targeted therapies. Epigenetic alterations can alter gene expression, gene expression potential, or the regulation of gene function, and thereby contribute to gliomagenesis. Arguably, the most widely studied epigenetic mark is DNA methylation that occurs at cytosine residues in the context of CpG dinucleotides. Approximately half of human genes have concentrations of CpGs in their promoter regions and the methylation state of these and other gene-associated CpGs are widely regarded as critical indicators of gene regulation.

Since 2008, sequencing of gliomas has identified mutations in the isocitrate dehydrogenase 1 and 2 (*IDH1* and *IDH2*) genes (3-5). The IDH1 and IDH2 enzymes convert isocitrate to alpha (α)–ketoglutarate producing NADPH and participate in cellular metabolic processes such as glucose sensing, lipid metabolism, and oxidative respiration (reviewed in [6]). Mutations in *IDH1* are consistently found in codon 132 for arginine (R132), and mutations in *IDH2* consistently occur at the analogous amino acid R172 (3,

7). Mutations in *IDH1* and *IDH2* (*IDH* when referring to both) are unlike most cancer-associated enzyme mutations because they confer neomorphic enzyme activity rather than inactivating, or constitutively activating, the enzyme. The mutant form of IDH enzymes convert α-ketoglutarate to 2-hydroxyglutarate in an NADPH-dependent manner, and via an unknown mechanism contribute to the pathophysiology of gliomas and leukemias (5, 7, 8). *IDH* mutations occur in approximately 80% of grades 2-3 gliomas and secondary glioblastomas, but less than 10% of primary glioblastomas (4, 5). In gliomas, *IDH* mutation has been associated with genetic alterations in other genes including the tumor suppressors and oncogenes (5). *IDH* mutation also has been associated with younger age and improved survival in glioma patients (5, 9).

The somatic genetic signature of any individual tumor is critical to assessing its clinical and etiologic character. Similarly, the profile of somatic epigenetic alterations is central to forming a complete understanding of the pattern of disrupted cellular functioning responsible for the deadly behavior of gliomas. Major advances in the clinical role of epigenetics in gliomas include the findings that promoter methylation silencing of the O-6-methylguanine-DNA methyltransferase (*MGMT*) gene is associated with response to temozolomide treatment (10). Epigenetic silencing of *MGMT* gene is found in approximately 80% of gliomas with mutant *IDH1*, compared with approximately 60% of gliomas with wild-type *IDH1* (9). Other common alterations in gliomas are mutations in tumor protein p53 (*TP53*) (11) and amplification of the epidermal growth factor receptor (*EGFR*) oncogene (12). Better definitions of the somatic nature of gliomas should integrate both their genetic and epigenetic alterations. In this study, we assessed CpG methylation patterns, *IDH* mutation, *TP53* mutation, and

EGFR amplification in histologically diverse gliomas to define epigenetic subgroups of potential clinical and etiologic relevance.

Patients, Materials, and Methods

Patients and Tissue Samples

Fresh frozen tumor tissues of patients (n = 131) diagnosed with glioma between 1990 and 2003 were obtained from the University of California San Francisco (UCSF) Brain Tumor Research Center Tissue Bank. Tumors were previously reviewed by UCSF neuropathologists to assign histologic subtypes and grades according to the World Health Organization classification for patients operated on at the UCSF Medical Center (1). Tumor samples were defined as secondary glioblastoma if the patients had previous histological diagnosis of a lower-grade glioma. Non-tumor brain tissue samples were obtained from cancer-free patients (n = 7) who underwent temporal lobe resection for treatment of epilepsy at the UCSF Medical Center. Patient ages were documented at the time of initial diagnosis. Other demographic and survival data were obtained from UCSF patient records and the California Cancer Registry. The Institutional Review Board approval certification was obtained from the UCSF Committee on Human Research, and subjects provided written, informed consent for tissue collection.

Cell lines, Cell Culture, and Reagents

A431 cells (a human epidermoid cancer cell line that is known to have *EGFR* amplification and overexpression) and HT29 cells (a human colon adenocarcinoma cell line without *EGFR* amplification) were obtained from American Type Culture Collection

(ATCC, **Manassas, VA**). Cell lines were maintained in Dulbecco's Modified Eagle's Medium and RPMI 1640 medium (both from Invitrogen, Carlsbad, CA), respectively, with 10% fetal bovine serum (Hyclone, Logan UT), at 37ºC in 5% CO_2. When cultures reached 80% confluency, cells were harvested for DNA extraction.

DNA Extraction, Bisulfite Modification, and Methylation Analysis

Genomic DNA from 131 glioma tissue samples and seven non-tumor brain tissue samples was isolated from approximately 25 mg wet weight of each frozen tissue sample using QIAamp DNA mini kit (Qiagen Inc., Valencia, CA), according to the manufacturer's instructions. DNA was eluted twice in a total of 100 µl of elution buffer. The same DNA extraction method was applied to A431 and HT29 cell lines that served as *EGFR* amplification controls.

For DNA methylation analysis, 1 µg of genomic DNA was first subjected to bisulfite modification using the EZ DNA Methylation Kit (Zymo Research Corporation, Orange, CA), according to the manufacturer's instructions. Bisulfite modification converts unmethylated cytosine residues to uracil and preserves methylated cytosine residues as cytosines.

GoldenGate DNA methylation bead arrays (Illumina Inc., San Diego, CA) were used to interrogate methylation of 1505 CpG loci associated with 803 cancer-related genes, according to the manufacturer's instructions. GoldenGate methylation arrays were used to analyze bisulfite-modified DNA from 131 glioma and 7 non-tumor samples for methylation, and processed at the UCSF Institute for Human Genetics, Genomics Core Facility. The GoldenGate array methylation data were deposited in the Gene Expression Omnibus and are publicly available (accession GSE20395). The Cancer Genome Atlas

(TCGA), a public data portal, was used to obtain GoldenGate methylation array data for validation of methylation classes. Quantitative methylation-specific polymerase chain reaction (PCR) (QMSP) was used to confirm methylation data from the GoldenGate array. Candidate genes were selected based on previous studies (13-16) that reported aberrant methylation in astrocytic glioma and included *MGMT*, Ras association domain family member 1 (*RASSF1*), PYD and CARD domain containing (*PYCARD*), homeobox A9 (*HOXA9*), paternally expressed 3 (*PEG3*), and slit homolog 2 (*SLIT2*). CpGenome Universal Methylated DNA (Millipore, Billerica, MA) was bisulfite modified and used as a positive control for QMSP analysis. QMSP was performed using Applied Biosystems 7900HT Fast Real-Time PCR System (Applied Biosystems, Carlsbad, CA). The reaction plate was prepared using the Beckman Coulter automated liquid handler-Biomex 3000 (Beckman Coulter, Fullerton, CA). Each reaction contained 10.0 µL 2× Power SYBR Green PCR Master Mix (Applied Biosystems), 100-400 nM of forward and reverse primers (Supplementary Table 1, available online) and 25 ng of DNA template in a total reaction volume of 20 µL. For the amplification of *RASSF1*, 2–3% dimethyl sulfoxide (DMSO) was added to the reaction mix. PCR conditions are modified by different primer concentrations and DMSO was added to ensure that primer dimers and non-specific amplification products were not included in the threshold cycle (Ct) calculation. To confirm specificity of amplicons from QMSP, we performed dissociation curve analysis. The PCR conditions were: 95ºC for 10 minutes, and 40 cycles of 95ºC for 15 seconds, 60ºC for 30 seconds, and 72ºC for 30 seconds. SYBR Green fluorescence data was collected only during the extension reaction at 72ºC. Ct values were calculated by the 7900HT system software, and average relative quantification (RQ) values were obtained

for each sample using actin, beta (*ACTB*) amplification as the referent, where RQ = (target gene / *ACTB*) / (Universal methylation calibrator / *ACTB*). Spearman rank correlation coefficients (rho) and *P* values were calculated to assess the correlation between GoldenGate array data and QMSP results.

Mutation analysis

***IDH mutation*.** The region spanning R132 codon of *IDH1*, and the region spanning R172 codon of *IDH2* were amplified by PCR with primers designed with Primer 3 sofware (v.0.4.0) with the exception of the forward sequencing primer, which was selected from Balss et al. (4). PCR reaction mixtures contained 10–25 ng DNA, 1× buffer, 0.2 mM dNTP mix, 0.2 μM forward and reverse primers, 0.04 units of HotStarTaq, and 1 mM $MgCl_2$ (Qiagen Inc.), in a 25 μL volume. The PCR conditions were: 95ºC for 10 minutes, 40 cycles of 94ºC for 30 seconds, 60ºC for 30 seconds, and 72ºC for 1 minute. The resulting products were analyzed on a 1.5% agarose gel. DNA was purified using the QIAquick® PCR Purification Kit (Qiagen Inc.) and sent to Rhode Island Genomics and Sequencing Center at the University of Rhode Island, where it was sequenced in both directions using the BigDyeTerminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA). Sequences were analyzed with the help of Applied Biosystems Sequence Scanner Software v1.0. All primers for *IDH1* mutation analysis are listed in Supplementary Table 1, available online.

***TP53 mutation*.** For *TP53* mutation analysis, PCR–single-strand conformation polymorphism technique was used, and DNA sequencing was done as previously

described (8). Primers for PCR amplification of fragments of exons 5 to 8 of *TP53* are listed in Supplementary Table 1, available online. PCR reaction mixtures contained 50 ng DNA, 20 µmol/L dNTP, 10 mmol/L Tris-HCl (pH 9.0), 1.5 mmol/L $MgCl_2$, 0.1% Triton X-100, 10 pmol of forward and reverse primers, 1 unit Taq (Perkin-Elmer Cetus, Norwalk, CT), and 0.2 µCi [^{33}P] dCTP (DuPont New England Nuclear, Boston, MA), in a 30 µL volume. DNA with *TP53* mutation confirmed by sequencing was included as positive control. The PCR reaction was carried out using 35 cycles: 94ºC for 30 seconds, annealed for 30 seconds at 58 ºC for exons 5 and 8, and 60 ºC for exons 6 and 7 (primers listed in Supplementary Table 1, available online) and 72ºC for 1 minute. Three microliters of PCR product were mixed with 2 µL of 0.1 N NaOH and then mixed with 5 µL of gel loading buffer solution (United States Biochemical Corp. Cleveland, OH) and heated at 94ºC for 4 minutes. DNA was analyzed on 6% nondenatured polyacrylamide gel, supplemented with 10% glycerol. Electrophoresis was performed at room temperature for 20 hours and exposed to *autoradiography films* for 16 hours for detection of bands. Direct sequencing of PCR fragments for both DNA strands was done on all tumor DNAs that showed aberrant migration patterns on single-strand conformation polymorphism gel to determine the corresponding DNA sequences using dsDNA cycle sequencing system (Life Technologies, Gaithersburg, MD), as described in Wiencke et al. (17).

EGFR amplification. *EGFR* amplification was measured by a quantitative PCR method using the ABI 7900 Real-Time PCR system (Applied Biosystems) and the commonly used DNA-binding dye, SYBR Green I, which has been shown to be equivalent to

TaqMan PCR assay for the assessment of gene copy number (18). Quality control measures for the real-time SYBR green assay included running triplicate determinations for both *EGFR* and control gene, glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*). DNA from A431 and HT29 cell lines, with known copy number states for *EGFR*, served as positive and negative controls, respectively, for amplification.

Statistical Analysis

Data assembly. Methylation data were assembled with BeadStudio methylation software from Illumina. All GoldenGate methylation array data points are represented by fluorescent signals (Cy dyes) from both methylated (Cy5) and unmethylated (Cy3) alleles. The methylation level, designated as beta (β) is calculated as $\beta = (\max[Cy5, 0])/(|Cy3| + |Cy5| + 100)$, in which the average β value is derived from the approximately 30 replicate methylation measurements, because each CpG probe set is present on the array and measured in each sample approximately 30 times. Raw average β values were analyzed without normalization as recommended by Illumina. At each CpG locus, for each tissue DNA sample, the detection *P* value was used to determine sample performance; all samples had detection *P* values less than 1×10^{-5} at more than 75% of CpG loci and passed performance criteria. There were 8 CpG loci that had a median detection *P* value of greater than .05, and these 8 CpGs were excluded from the analysis. All CpG loci on the X chromosome were excluded from analysis. The final dataset contained 1413 autosomal CpG loci associated with 773 genes. For each CpG locus, the differential methylation values (delta-beta [Δβ]) were calculated by subtracting the average β value

of tumors from the mean β value of the seven non-tumor brain samples. Subsequent analyses were carried out using the R software (19). All statistical tests were two-sided.

Unsupervised Clustering, Recursively Partitioned Mixture Modeling, and Survival. Hierarchical clustering of the DNA methylation data was performed using the R function hclust with Euclidean distance metric and Ward linkage. To discern and describe the relationships between CpG methylation data and patient and tumor covariates, a modified model-based form of unsupervised clustering known as recursively partitioned mixture modeling (RPMM), was used as described in Houseman et al. (20) and as used in Christensen et al. (21). The analysis of associations between methylation class (categorical) and individual categorical covariates was performed using the Fisher exact test. To test for association between methylation class and continuous covariates, a permutation test was run with the Kruskal-Wallis test statistic, and a likelihood ratio test was used for comparing the association between methylation class and *IDH* mutation to a model including age and histology. To test for associations between *IDH* mutation and grade-specific tumor histology, and *IDH* mutation and tumor grade, Fisher's exact tests were used. To test for associations between *IDH* mutation and primary vs. secondary glioblastoma, *IDH* mutation and *TP53* mutation, and *IDH* mutation and *EGFR* amplification, Chi-square tests were used. The assumption of proportionality for Cox proportional hazards modeling was verified by calculating Pearson correlation coefficients for the corresponding set of Schoenfeld residuals with a transformation of time based on the Kaplan-Meier estimate of the survival function (22), and graphically by plotting log(survival time) vs. log(-log[survival as a function of time, t]).

Locus-by-locus analysis. To examine differential methylation between tumor and non-tumor tissues, gliomas were stratified by grade-specific histologic subtypes, and individual CpG loci were compared between subtypes of glioma and non-tumor samples using a Wilcoxon rank-sum test. Because this results in the simultaneous comparison of all CpG loci between glioma subtypes and non-tumor sample types, false discovery rate estimation and *Q*-values computed by the qvalue package in R (23) were used to adjust for multiple testing. Differentially methylated CpGs were counted as hyper- or hypomethylated if both the tumor vs. non-tumor *Q* less than .05 and the median methylation value $|\Delta\beta|$ greater than 0.2. An equivalent approach was used in the analysis of differential methylation for gliomas with mutant or wild-type *IDH*, compared with non-tumor tissues.

Pathway Analysis. A canonical pathway analysis was conducted with the use of Ingenuity Pathway Analysis software (Ingenuity Systems, Redwood City, CA). CpG gene-loci associated with the Illumina GoldenGate methylation array were used as reference and loci from differential methylation analysis, as described later in the article, were compared. The statistical significance of gene-locus enrichment within canonical pathways was measured with a Fisher's exact test.

Results

Unsupervised Clustering and Modeling of Glioma and Non-Tumor DNA Methylation Data

Histological grade and patient demographic data for the 131 gliomas and patient demographic data for the seven non-tumor brain tissues are presented in Table 1. To characterize DNA methylation of gliomas and non-tumor brain tissues, the bisulfite-modified DNA samples were hybridized to the GoldenGate DNA methylation array. Unsupervised clustering of DNA methylation data from 1413 autosomal CpG loci showed that non-tumor brain tissues cluster with each other and are distinct from tumor tissues (Figure 1, A). Furthermore, we observed that oligodendrogliomas and astrocytomas generally clustered together and demonstrated a greater number of methylated loci relative to ependymomas, pilocytic astrocytomas, as well as non-tumor brain tissues. Concomitantly, glioblastomas (also known as grade IV astrocytoma), predominantly clustered together at the bottom of the heatmap (Figure 1, A) and displayed more hypermethylated CpG loci than ependymomas.

In order to further investigate the DNA methylation patterns of gliomas and non-tumor brain tissue, we implemented an agnostic approach by applying a modified model-based form of unsupervised clustering known as recursively partitioned mixture modeling (RPMM) (20). RPMM allows for precise inference regarding the potential covariates associated with intrinsic similarities and differences in CpG methylation by generating distinct classes of DNA methylation for the modeled samples based upon the DNA methylation array data. We applied RPMM clustering to all 131 tumors, which generated 11 methylation classes (Figure 1, B). Methylation classes contain samples with DNA

methylation patterns that are most similar to each other, and samples with different DNA methylation patterns are distinguished by their membership in a different methylation class. Methylation class was statistically significantly associated with both tumor histologic subtype ($P < 2.2 \times 10^{-16}$) and grade ($P < 2.2 \times 10^{-16}$) (Supplementary Table 2, available online).

Methylation Array and Methylation Class Validation

Methylation data from GoldenGate arrays have been extensively validated by our group and others using a variety of methods (24-28). The methylation array data presented in this study were validated by correlating CpG methylation array data to quantitative methylation-specific PCR (QMSP) data for genes commonly methylated in gliomas—*MGMT, RASSF1, PYCARD, HOXA9, PEG3,* and *SLIT2* (Supplementary Table 3, available online). To determine the validity of association between histology and methylation class we utilized publicly available GoldenGate methylation array data for 71 glioblastoma samples from The Cancer Genome Atlas (TCGA). Using the RPMM classification (Figure 1, B), we predicted the methylation class for each glioblastoma sample of TCGA and confirmed that 70 of 71 (99%) TCGA glioblastoma samples were classified in RPMM methylation classes that contained glioblastoma samples (Supplementary Table 2, available online). The identification numbers and the predicted RPMM methylation classes of TCGA tumors are listed in Supplementary Table 4, available online.

Ratios of Hypermethylated to Hypomethylated CpG Loci and Tumor Histology

We examined the differential methylation (Δβ) between tumor and non-tumor brain tissues and observed a striking pattern of the number of hyper- and hypomethylated CpG loci among different tumor subtypes (Figure 2, A). Glioblastomas showed a low ratio of hyper- to hypomethylated loci (ratio = 1.3), compared with the ratio for grades 2 and 3 astrocytomas, grades 2 and 3 oligoastrocytomas, and grade 2 oligodendrogliomas (ratios = 3.7, 7.6, and 9.7, respectively). Conversely, ependymomas showed increased hypomethylation (ratio = 0.3). The ratios of hyper- to hypomethylated CpG loci were statistically significantly different across glioma tumor histologic subtypes (Permutation $P<.0001$). Histology-related hyper- and hypomethylation patterns were also evident in unsupervised hierarchical clustering of Δβ methylation values for all 1413 autosomal CpG loci (Figure 2, B).

We next assessed the cellular pathways associated with statistically significantly differentially hypomethylated and (separately) hypermethylated CpG loci that were common among glioblastomas, astrocytomas, oligoastrocytomas, and oligodendrogliomas. There were 18 CpG loci with statistically significant differential hypomethylation ($Q<.05$) and common among glioblastomas, astrocytomas, oligoastrocytomas, and oligodendrogliomas. An analysis of cellular pathways enriched among these 18 CpG loci, compared with all genes represented on the methylation array, revealed statistically significant enrichment of metabolism and biosynthesis pathways (Supplementary Table 5, available online). In addition, there were 35 statistically significantly differentially hypermethylated ($Q<0.05$) CpG loci common among glioblastomas, astrocytomas, oligoastrocytomas, and oligodendrogliomas. An analysis of

cellular pathways enriched among these 35 CpG loci showed that oxidative stress response and retinoic acid mediated apoptosis signaling pathways were statistically significantly enriched (Supplementary Table 5, available online). For each grade-specific tumor histology, all statistically significant differentially hypomethylated and hypermethylated CpG loci are detailed in Supplementary Tables 6 and 7, respectively, available online.

Glioma Methylation Classes, *IDH* Mutation, and Survival

The analysis of differentially methylated CpG loci in cellular pathways suggested that metabolic pathways as a group were commonly hypomethylated in gliomas. We hypothesized that genetic mutations in the metabolic pathways were associated with the observed DNA methylation phenotype. To test this hypothesis, we sequenced a subset of 95 tumors with available DNA for *IDH1* and *IDH2* mutations. *IDH2* mutation was detected in only two tumors, and *IDH1* mutation was detected in 55 tumors (total *IDH* mutation prevalence = 58.9%). *IDH* mutations were more common in oligoastrocytoma, oligodendroglioma, or astrocytoma histologic subtypes than in glioblastomas, pilocytic astrocytomas, or ependymomas ($P = 6.4 \times 10^{-9}$); in lower-grade than higher-grade tumors ($P = .01$); in tumors with *TP53* mutation compared with wild-type *TP53* ($P = .06$); and in younger patients (mean age = 36.6 years vs. 47.4 years, $P = .0009$) (Table 2). However, *IDH* mutation was not associated with *EGFR* amplification ($P = .10$) (Table 2). Additionally, tumors with *IDH* mutation showed statistically significantly higher *MGMT* methylation ($P = 3.6 \times 10^{-4}$) (Supplementary Figure 1, available online).

Next we investigated the number of statistically significantly differentially methylated CpG loci between tumor and non-tumor samples stratified by *IDH* mutation status. Tumors with *IDH* mutation revealed a striking contrast between the number of statistically significantly differentially hypermethylated loci, as well as the ratio of hyper- to hypomethylated loci in *IDH* mutant tumors vs. *IDH* wild-type tumors (mutant = 7.8 vs. wild-type = 0.22) (Figure 3, A). We utilized the statistically significantly differentially hypermethylated and hypomethylated CpG loci in *IDH* mutant tumors to conduct an enrichment analysis of cellular pathways. We found that cellular signaling pathways were hypermethylated, whereas metabolism and biosynthesis pathways that included starch and sucrose metabolism and pentose and glucuronate interconversion pathways, were hypomethylated in *IDH* mutant tumors (Supplementary Table 8, available online).

Methylation profiling with RPMM of the 95 gliomas with both methylation data and *IDH* mutation status resulted in nine methylation classes (Figure 3, B). Methylation classes were statistically significantly associated with patient age (Permutation $P = 3.0 \times 10^{-4}$), histology ($P<2.2 \times 10^{-16}$), and grade ($P = 6.0 \times 10^{-9}$) (Supplementary Table 9, available online). *IDH* mutation was also strongly associated with methylation class ($P = 3.0 \times 10^{-16}$) (Figure 3, C), and this association remained statistically significant when controlling for age and histology (likelihood ratio $P<.0001$). Only two methylation classes had *IDH* mutant tumors (class L and class RLLR), and greater than 98% of the tumors (all but one) in these two classes had an *IDH* mutation (Figure 3C). Furthermore, methylation classes L and RLLR were both more highly methylated than the other methylation classes (Figure 3, B).

Last, we examined the potential association between *IDH* mutation and patient survival among cases with available mutation data (n = 95) because previous studies reported increased survival among glioma patients with *IDH* mutation (3, 5). In a multivariate Cox proportional hazards model controlling for age at diagnosis, sex, and grade-specific histology, we observed that patients whose tumors harbored *IDH* mutation showed statistically significantly better survival, compared with patients (n = 39) whose tumors harbored wild-type *IDH* (HR of death = 0.27, 95% confidence interval [CI] = 0.10 to 0.72) (Figure 3, D, and Table 3).

Discussion

In this study, we demonstrate a distinct pattern of methylation across histological subtypes of glioma that is associated with genetic mutation in *IDH* gene loci. The two methylation classes associated with mutant *IDH* tumors had a homogeneous, hypermethylation-rich character compared to the methylation classes for tumors with wild-type *IDH*. Additionally, the tumors with wild-type *IDH* belonged to several distinct methylation classes. The contrast between a single homogenous hypermethylated profile and several heterogeneous hypomethylated profiles (associated with distinct histologic types) strongly suggests that *IDH* mutation "drives" the observed hypermethylated phenotype, irrespective of tumor histology. In support of this, we note that *IDH1* mutation is more robustly associated with methylation class, compared with the classical glioma tumor genetic markers like *TP53* mutation and *EGFR* amplification.

IDH mutations are heterozygous and allow the enzyme normally responsible for conversion of isocitrate to α-ketoglutarate to convert α-ketoglutarate to 2-

hydroxyglutarate in an NADPH–dependent manner and results in accumulation of 2-hydroxyglutarate (7, 8). Despite the observed hypermethylated profile of *IDH* mutant tumors, analysis of cellular pathways showed hypomethylation of several metabolic pathways, potentially to compensate for mutation-related metabolic stress. Because the methylation profile of *IDH* mutant tumors is generally homogenous, it is possible that the hypermethylation phenotype is either selected for, or driven by, the hypomethylation of compensatory metabolic pathways, thus directly linking and temporally situating these events. The level of α-ketoglutarate has been shown to be slightly lower in *IDH1* mutant gliomas, though this decrease was not statistically significant (8). However, IDH1 localizes to the cytosol and peroxisomes, whereas IDH2 is localizes to mitochondria; and because most *IDH* mutations in gliomas are in *IDH1*, pan-cellular α-ketoglutarate levels may not represent available cytosolic α-ketoglutarate levels. Furthermore, *IDH1* R132 mutation has been shown to favor an active conformation of the enzyme, increase its affinity for NADPH, and favor reduction of α-ketoglutarate to 2-hydroxyglutarate over the conversion of isocitrate to α-ketoglutarate, which may reduce the availability of cytosolic α-ketoglutarate and NADPH (8). Hence, a potential mechanism responsible for the strong association between epigenetic profile and *IDH* mutation is related to potentially altered availability of α-ketoglutarate in these tumors. The Jumonji-domain-containing histone demethylases require α-ketoglutarate as a substrate for their enzymatic activity (29) and altered activity of these histone demethylases could lead to aberrantly remodeled chromatin, potentially resulting in epigenetic alterations at the DNA-level as well. However, studies that are beyond the scope of this manuscript would be necessary to disentangle the complex networks of chromatin remodeling enzymes, their targets, and

their responses to altered levels of enzymatic substrate. Alternatively, (or perhaps in conjunction) lower concentrations of NADPH associated with mutant *IDH1* (30) may result in a decreased capacity for reductive processes in defense against reactive oxygen species. Furthermore, α-ketoglutarate itself is a potent antioxidant (6) and its decreased availability in *IDH* mutant cells alone, or together with lower NADPH levels could drive the selection of cells with compensatory metabolic gene expression profiles mediated by altered epigenetic patterns including chromatin configuration and DNA methylation. Consistent with the suggestion that gene expression profiles are altered in association with DNA methylation related to *IDH* mutation, an analysis of glioblastoma gene expression subtypes showed that *IDH* mutation occurred almost exclusively proneural glioblastomas (31).

More broadly, and similar to the hypermethylation phenotype we describe here, hypermethylator phenotypes have previously been associated with other cancers. This phenotype was first described in colon cancer, and is commonly referred to as CpG Island Methylator Phenotype (CIMP) (32). Specifically, colorectal cancers can be divided in CIMP-high, CIMP-low, and non-CIMP based on the methylation of 5-8 specific gene promoters (33, 34). Similar to *IDH* in glioma, CIMP status in colon tumors has been associated with specific mutations; CIMP-High with *BRAF* and CIMP-Low and non-CIMP with *KRAS* (35). Recently, Noushmehr *et al.* described a CIMP in glioblastomas, termed G-CIMP, which they found to be tightly associated with *IDH1* mutation (36). In a number of lower-grade gliomas Noushmehr *et al.* performed methylation profiling of eight markers of G-CIMP and confirmed that *IDH1* mutation is associated with G-CIMP in low-grade tumors, which is consistent with our array-based findings. Furthermore,

over 83% of G-CIMP positive glioblastomas with *IDH1* mutation were of the proneural glioblastoma gene expression subtype (36), additional evidence supporting an association between distinct, *IDH*-related methylation in our data (from diverse glioma histologic subtypes), and a specific gene expression phenotype. In addition, *MGMT* methylation is often investigated in glioma since it has been associated with increased sensitivity to alkylating agents such as Temozolomide and can impact response to therapy (37). In fact, increased *MGMT* methylation can also distinguish CIMP-High and CIMP- Low from non-CIMP in colon cancer (38). Our results, consistent with previous work (9), demonstrate an association between increased *MGMT* methylation and *IDH* mutation. Finally, some studies have reported CIMP positive colon cancers to have a relatively better prognosis (39), and from both the work of Noushmehr *et al.* and ours, this appears to be consistent with the pattern of survival observed in CIMP gliomas.

The association between *IDH* mutation and a single methylation profile across several histologic subtypes suggests that genetic and epigenetic alterations are not independent. This observation also has profound implications for the development of new therapies for glioma. Although pharmacological inhibition of 2-hydroxyglutarate has been suggested as a possible approach to treating *IDH* mutant gliomas (40) such drugs do not yet exist. However, DNA methylation is a modifiable therapeutic target; DNA methyltransferase inhibitors and histone deacetylase inhibitors are in clinical trials and showing some promise for the treatment of hematopoietic malignancies (41-43). Our work suggests that a simple diagnostic test for DNA methylation (or mutation) can identify a class of tumors for which the modification of DNA methylation may have therapeutic efficacy. This class of tumors is not discernable by any of the classic

histopathologic or tumor markers for glioma. The recognition that *IDH* mutation has value as a clinical prognostic marker and is associated with a broad DNA methylation phenotype suggests that glioma therapeutic protocols that reverse DNA methylation should be pursued.

Our study has a few limitations. Although we studied 131 histologically diverse tumors, we did not have *IDH* mutation, *TP53* mutation, and *EGFR* amplification data on all subjects and had somewhat limited statistical power to explore the relationships between *IDH* mutation and these alterations. Future investigations that include larger numbers of histologically diverse samples and higher-resolution methylation array techniques, along with measurements of other somatic alterations (*IDH* mutation, mRNA expression, and copy number) will afford a more comprehensive understanding of the molecular and chromosomal characteristics that distinguish glioma subtypes. Understanding whether these glioma molecular and chromosomal subtypes are differentially associated with glioma risk loci (44) also will help to understand the etiology and possibly outcomes of this often-catastrophic disease.

In summary, our work demonstrates a clear relationship between genetic and epigenetic events in human gliomas by associating *IDH* mutations with a homogenous methylation profile, and demonstrates that profiles of methylation differ by histologic subtype of disease. Additionally, and consistent with previous work, we also showed that patients with *IDH* mutation have ~~a~~ significantly improved survival. Advances in therapy for glioma may be realized by targeting DNA methylation. Much attention has recently been given to the utility of *MGMT* methylation in predicting response to therapy, and our

data further suggest that other DNA methylation markers may improve clinical assessment, guide therapies, and potentially uncover novel therapeutic avenues altogether.

References

1. Kleihues P, Burger PC, Scheithauer BW. The new WHO classification of brain tumours. Brain Pathol 1993;3(3):255-68.

2. Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N Engl J Med 2005;352(10):987-96.

3. Parsons DW, Jones S, Zhang X, Lin JC, Leary RJ, Angenendt P, et al. An integrated genomic analysis of human glioblastoma multiforme. Science 2008;321(5897):1807-12.

4. Balss J, Meyer J, Mueller W, Korshunov A, Hartmann C, von Deimling A. Analysis of the IDH1 codon 132 mutation in brain tumors. Acta Neuropathol 2008;116(6):597-602.

5. Yan H, Parsons DW, Jin G, McLendon R, Rasheed BA, Yuan W, et al. IDH1 and IDH2 mutations in gliomas. N Engl J Med 2009;360(8):765-73.

6. Reitman ZJ, Yan H. Isocitrate dehydrogenase 1 and 2 mutations in cancer: alterations at a crossroads of cellular metabolism. J Natl Cancer Inst 2010;102(13):932-41.

7. Ward PS, Patel J, Wise DR, Abdel-Wahab O, Bennett BD, Coller HA, et al. The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate. Cancer Cell 2010;17(3):225-34.

8. Dang L, White DW, Gross S, Bennett BD, Bittinger MA, Driggers EM, et al. Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. Nature 2009;462(7274):739-44.

9. Sanson M, Marie Y, Paris S, Idbaih A, Laffaire J, Ducray F, et al. Isocitrate dehydrogenase 1 codon 132 mutation is an important prognostic biomarker in gliomas. J Clin Oncol 2009;27(25):4150-4.

10. Hegi ME, Diserens AC, Gorlia T, Hamou MF, de Tribolet N, Weller M, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. N Engl J Med 2005;352(10):997-1003.

11. Mashiyama S, Murakami Y, Yoshimoto T, Sekiya T, Hayashi K. Detection of p53 gene mutations in human brain tumors by single-strand conformation polymorphism analysis of polymerase chain reaction products. Oncogene 1991;6(8):1313-8.

12. Wong AJ, Bigner SH, Bigner DD, Kinzler KW, Hamilton SR, Vogelstein B. Increased expression of the epidermal growth factor receptor gene in malignant gliomas is invariably associated with gene amplification. Proc Natl Acad Sci U S A 1987;84(19):6899-903.

13. Dallol A, Krex D, Hesson L, Eng C, Maher ER, Latif F. Frequent epigenetic inactivation of the SLIT2 gene in gliomas. Oncogene 2003;22(29):4611-6.

14. Maegawa S, Itaba N, Otsuka S, Kamitani H, Watanabe T, Tahimic CG, et al. Coordinate downregulation of a novel imprinted transcript ITUP1 with PEG3 in glioma cell lines. DNA Res 2004;11(1):37-49.

15. Stone AR, Bobo W, Brat DJ, Devi NS, Van Meir EG, Vertino PM. Aberrant methylation and down-regulation of TMS1/ASC in human glioblastoma. Am J Pathol 2004;165(4):1151-61.

16. Yu J, Zhang H, Gu J, Lin S, Li J, Lu W, et al. Methylation profiles of thirty four promoter-CpG islands and concordant methylation behaviours of sixteen genes that may contribute to carcinogenesis of astrocytoma. BMC Cancer 2004;4:65.

17. Wiencke J, Aldape K, McMillan A, Wiemels J, Moghadassi M, Miike R, et al. Molecular features of adult glioma associated with patient race/ethnicity, age, and a polymorphism in O6-methylguanine-DNA-methyltransferase. Cancer Epidemiol Biomarkers Prev 2005;14(7):1774-83.

18. De Preter K, Speleman F, Combaret V, Lunec J, Laureys G, Eussen B, et al. Quantification of MYCN, DDX1, and NAG gene copy number in neuroblastoma using a real-time quantitative PCR assay. Mod Pathol 2002;15(2):159-66.

19. R Development CT. R: A Language and Environment for Statistical Computing. In. Vienna, Austria: R Foundation for Statistical Computing; 2007.

20. Houseman EA, Christensen BC, Marsit CJ, Karagas MR, Wrensch MR, Yeh RF, 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 2008;9(365).

21. Christensen BC, Houseman EA, Godleski JJ, Marsit CJ, Longacker JL, Roelofs CR, et al. Epigenetic profiles distinguish pleural mesothelioma from normal pleura and predict lung asbestos burden and clinical outcome. Cancer Res 2009;69(1):227-34.

22. Grambsch P, Therneau T. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika 1994;81(3):515-26.

23. Storey J, Taylor J, Siegmund D. Strong control, conservative point estimation, and simultaneous conservative consistency of false discovery rates: A unified approach. J Royal Stat Soc 2004;Series B(66):187-205.

24. Bibikova M, Lin Z, Zhou L, Chudin E, Garcia EW, Wu B, et al. High-throughput DNA methylation profiling using universal bead arrays. Genome Res 2006;16(3):383-93.

25. Byun HM, Siegmund KD, Pan F, Weisenberger DJ, Kanel G, Laird PW, et al. Epigenetic profiling of somatic tissues from human autopsy specimens identifies tissue- and individual-specific DNA methylation patterns. Hum Mol Genet 2009;18(24):4808-17.

26. Christensen BC, Houseman EA, Marsit CJ, Zheng S, Wrensch MR, Wiemels JL, et al. Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. PLoS Genet 2009;5(8):e1000602.

27. Irizarry RA, Ladd-Acosta C, Carvalho B, Wu H, Brandenburg SA, Jeddeloh JA, et al. Comprehensive high-throughput arrays for relative methylation (CHARM). Genome Res 2008;18(5):780-90.

28. Ladd-Acosta C, Pevsner J, Sabunciyan S, Yolken RH, Webster MJ, Dinkins T, et al. DNA methylation signatures within the human brain. Am J Hum Genet 2007;81(6):1304-15.

29. Krieg AJ, Rankin EB, Chan D, Razorenova O, Fernandez S, Giaccia AJ. Regulation of the histone demethylase JMJD1A by hypoxia-inducible factor 1 alpha enhances hypoxic gene expression and tumor growth. Mol Cell Biol 2010;30(1):344-53.

30. Winkler BS, DeSantis N, Solomon F. Multiple NADPH-producing pathways control glutathione (GSH) content in retina. Exp Eye Res 1986;43(5):829-47.

31. Verhaak RGW, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, et al. Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell 2010;17(1):98-110.

32. Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB, Issa JP. CpG island methylator phenotype in colorectal cancer. Proc Natl Acad Sci U S A 1999;96(15):8681-6.

33. Ogino S, Nosho K, Kirkner G, Kawasaki T, Chan A, Schernhammer E, et al. A cohort study of tumoral LINE-1 hypomethylation and prognosis in colon cancer. J Natl Cancer Inst 2008;100(23):1734-8.

34. Issa J. CpG island methylator phenotype in cancer. Nat Rev Cancer 2004;4(12):988-93.

35. Ogino S, Kawasaki T, Kirkner G, Loda M, Fuchs C. CpG island methylator phenotype-low (CIMP-low) in colorectal cancer: possible associations with male sex and KRAS mutations. J Mol Diagn 2006;8(5):582-8.

36. Noushmehr H, Weisenberger DJ, Diefes K, Phillips HS, Pujara K, Berman BP, et al. Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. Cancer Cell 2010;17(5):510-22.

37. Chin L, Meyerson M, Aldape K, Bigner D, Mikkelsen T, VandenBerg S, et al. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature 2008;455(7216):1061-1068.

38. Ogino S, Kawasaki T, Kirkner G, Suemoto Y, Meyerhardt J, Fuchs C. Molecular correlates with MGMT promoter methylation and silencing support CpG island methylator phenotype-low (CIMP-low) in colorectal cancer. Gut 2007;56(11):1564-71.

39. Ogino S, Goel A. Molecular classification and correlates in colorectal cancer. J Mol Diagn 2008;10(1):13-27.

40. Frezza C, Tennant DA, Gottlieb E. IDH1 mutations in gliomas: when an enzyme loses its grip. Cancer Cell 2010;17(1):7-9.

41. Borthakur G, Huang X, Kantarjian H, Faderl S, Ravandi F, Ferrajoli A, et al. Report of a phase 1/2 study of a combination of azacitidine and cytarabine in acute myelogenous leukemia and high-risk myelodysplastic syndromes. Leuk Lymphoma;51(1):73-8.

42. Santos FP, Kantarjian H, Garcia-Manero G, Issa JP, Ravandi F. Decitabine in the treatment of myelodysplastic syndromes. Expert Rev Anticancer Ther;10(1):9-22.

43. Mercurio C, Minucci S, Pelicci PG. Histone deacetylases and epigenetic therapies of hematological malignancies. Pharmacol Res.

44. Wrensch M, Jenkins RB, Chang JS, Yeh RF, Xiao Y, Decker PA, et al. Variants in the CDKN2B and RTEL1 regions are associated with high-grade glioma susceptibility. Nat Genet 2009;41(8):905-8.

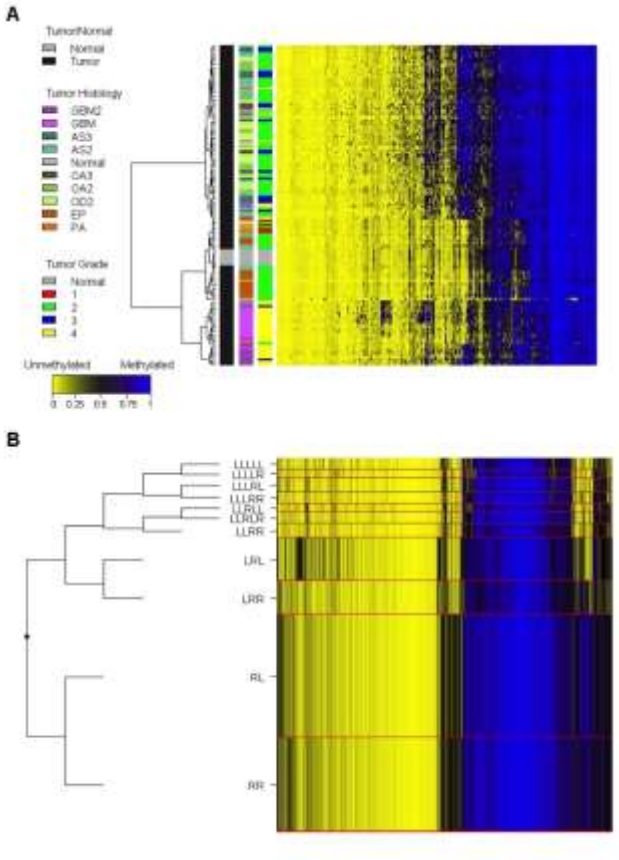


Figure 1. Association between glioma histologic subtypes and DNA methylation pattern. A) The average methylation beta (β) values of both gliomas (n = 131) and non-tumor tissue samples (n =7) were subjected to unsupervised hierarchical clustering based on Euclidean distance metric and Ward linkage and are shown in the heatmap. Each row represents a sample and each column represents a CpG locus (all 1413 autosomal loci). The scale bar at the bottom shows the range of β values (0 to 1). Tissue histology and grade are defined in color keys next to the heatmap, on the left. GBM2 = secondary glioblastoma multiforme; GBM = primary glioblastoma multiforme; AS3 = grade 3 astrocytoma; AS2 = grade 2 astrocytoma; OA3 = grade 3 oligoastrocytoma; OA2 = grade 2 oligoastrocytoma; OD2 = grade 2 oligodendroglioma; EP = ependymoma; PA = pilocytic astrocytoma. B) Recursively partitioned mixture model (RPMM) of glioma and non-tumor brain tissue samples (n = 138). Methylation profile classes are stacked in rows separated by red lines and class height corresponds to the number of samples in each class. Class methylation at each CpG locus (columns) is the mean methylation for all samples in a class. To the left of the RPMM is the clustering dendrogram. In the heatmap and RPMM, blue designates methylated CpG loci (average β = 1), and yellow designates unmethylated CpG loci (average β = 0).

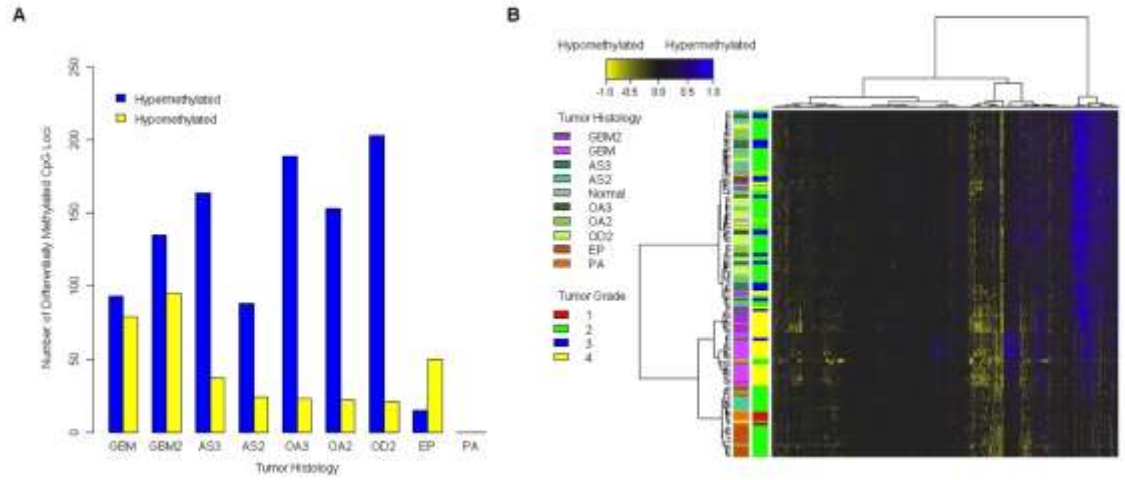


Figure 2. Differential methylation and the ratio of hyper- to hypomethylated loci in gliomas. Differential methylation values (Δβ) were calculated by subtracting tumor average β value from the mean value of the non-tumor brain samples (n = 7) for each CpG locus. A) The number of statistically significantly differentially hyper- and hypomethylated loci ($Q<0.05$ and $|\Delta\beta|>0.2$), are plotted by grade-specific glioma histology. GBM = primary glioblastoma multiforme; GBM2 = secondary glioblastoma multiforme; AS3 = grade 3 astrocytoma; AS2 = grade 2 astrocytoma; OA3 = grade 3 oligoastrocytoma; OA2 = grade 2 oligoastrocytoma; OD2 = grade 2 oligodendroglioma; EP = ependymoma; PA = pilocytic astrocytoma. B) Δβ values for all tumors (n = 131) were subjected to unsupervised hierarchical clustering based on Euclidean distance metric and Ward linkage. Each row represents a sample and each column represents a CpG locus (all 1413 autosomal loci). The scale bar at the top shows the range of Δβ values (-1 to 1). Tissue histology and grade are defined in color keys next to the heatmap on the left. In the heatmap blue designates differentially hypermethylated CpG loci in tumors ($\Delta\beta = 1$), and yellow designates differentially hypomethylated CpG loci in tumors ($\Delta\beta = -1$).

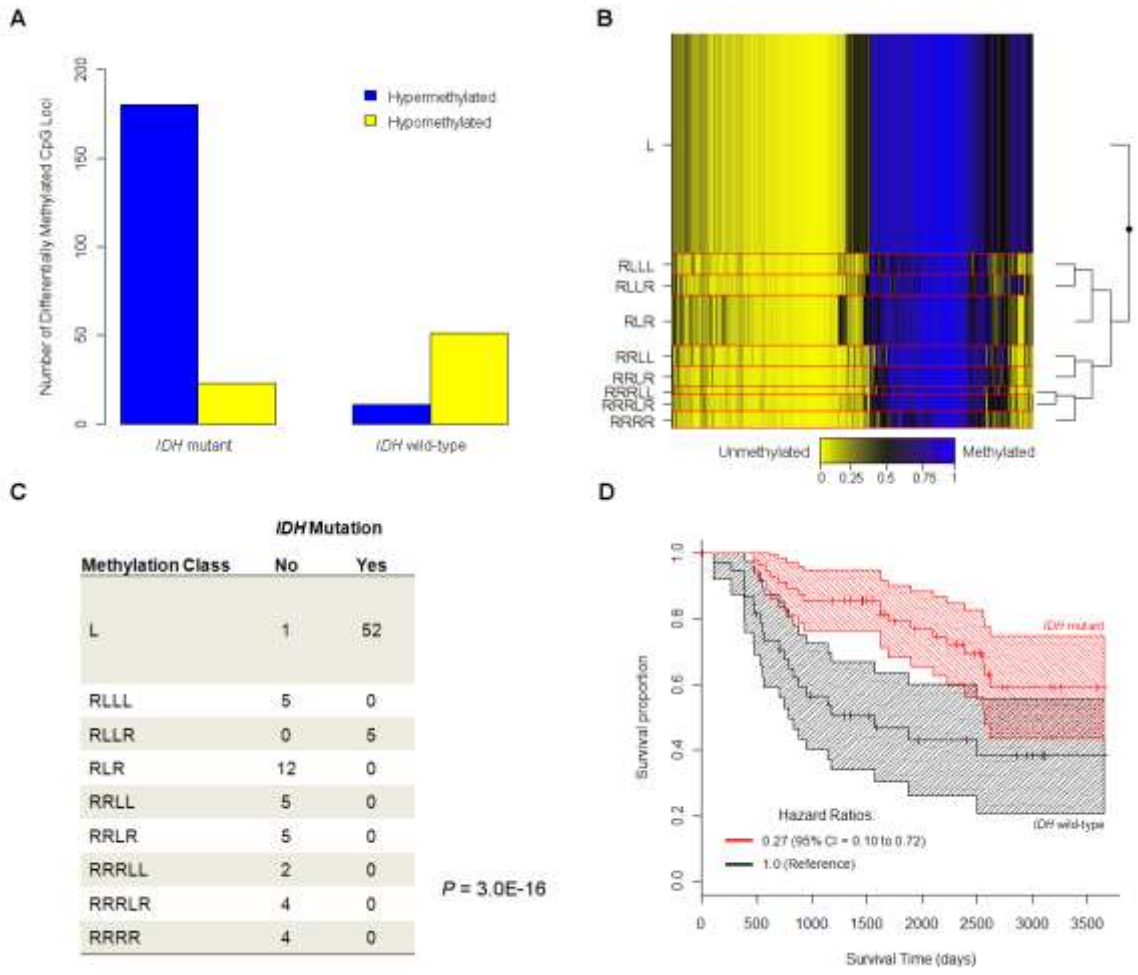


Figure 3. Association between *IDH* mutation and methylation phenotype in gliomas. A) The number of statistically significantly differentially hyper- and hypomethylated loci (Q<0.05 and |Δβ|>0.2), are plotted by tumor *IDH* mutation status. B) Recursively partitioned mixture model (RPMM) of glioma samples with both methylation and mutation data (n = 95). Methylation profile classes are stacked in rows separated by red lines, class height corresponds to the number of samples in each class. Class methylation at each CpG locus (columns) is the mean methylation for all samples in a class where blue designates methylated CpG loci (average β = 1), and yellow designates unmethylated CpG loci (average β = 0). To the right of the RPMM is the clustering dendrogram. C) Methylation-class-specific *IDH* mutation status (Fisher's *P* = 3.0E-16). D) Kaplan-Meier survival probability strata for *IDH* mutant (red, n = 56) and *IDH* wild-type (black, n = 39) tumors, tick marks are censored observations and banding patterns represent 95% confidence intervals (CIs).

Table 1. Patient demographic and tumor characteristics*

Characteristic	Non-tumor brain tissue (n=7)	Tumor histology and grade of glioma tissues (n=131)								
		Primary Glioblastoma multiforme (n=20)	Secondary Glioblastoma multiforme (n=12)	Grade 3 Aastrocytoma (n=9)	Grade 2 Astrocytoma (n=20)	Grade 3 Oligoastrocytoma (n=9)	Grade 2 Oligoastrocytoma (n=22)	Grade 2 Oligodendroglioma (n=20)	Eppendymoma (n=15)	Pilocytic Astrocytoma (n=4)
Age at diagnosis, y										
Median	33	55	34.5	40	40	40	33	35.5	41	28.5
Range	23 - 42	21 - 78	18 - 49	23 - 57	21 - 64	26 - 52	19 - 48	20 - 59	19 - 70	22 - 39
Sex, No. (%)										
Female	3 (43)	7 (35)	4 (33)	6 (67)	10 (50)	4 (44)	9 (41)	10 (50)	4 (27)	2 (50)
Male	4 (57)	13 (65)	8 (67)	3 (33)	10 (50)	5 (56)	13 (59)	10 (50)	11 (73)	2 (50)
Race, No. (%)										
White	-	18 (90)	11 (92)	5 (56)	17 (85)	7 (78)	16 (73)	18 (90)	12 (80)	4 (100)
Hispanic	-	1 (5)	1 (8)	2 (22)	1 (5)	0	1 (4)	1 (5)	2 (13)	0
American Indian	-	0	0	0	1 (5)	0	0	0	0	0
Asian	-	0	0	1 (11)	1 (5)	0	2 (9)	1 (5)	0	0
Unknown	7 (100)	1 (5)	0	1 (11)	0	2 (22)	3 (14)	0	1 (7)	0
Survival, d										
Median	NA	759	1244	1933	1584	3007	2937	2532	2498	2789
Range	NA	108-2477	466-4973	516-4494	305-4043	603-6459	612-5843	4-5988	478-5983	948-3279

*Non-tumor brain tissues (n=7) were obtained from cancer-free patients who underwent temporal lobe resection for treatment of epilepsy at the UCSF Medical Center. Glioma tissues (n=131) were obtained between 1990 and 2003 from the University of California San Francisco Brain Tumor Research Center Tissue Bank.

Table 2. Patient age, grade-specific glioma histology, grade, *TP53* mutation, and *EGFR* amplification stratified by *IDH* mutation status*

Patient age and tumor characteristics	*IDH* Mutation†	
	No	**Yes**
Age at diagnosis, y		*P* = 9.0E-04‡
Median age (range)	49 (17–78)	35 (20–59)
Mean age (SD)	47.4 (17.5)	36.6 (8.7)
Tumor histology§, No. (%)		*P* = 6.4E-09‖
Grade 2 Astrocytoma	5 (26)	14 (74)
Grade 3 Astrocytoma	0 (0)	4 (100)
Ependymoma	14 (100)	0 (0)
Primary Glioblastoma	15 (79)	4 (21)
Secondary Glioblastoma (*P*=.005)¶	1 (14)	6 (86)
Grade 2 Oligoastrocytoma	2 (13)	13 (87)
Grade 2 Oligodendroglioma	1 (6)	15 (94)
Tumor grade, No. (%)		*P* = .01#
1	-	-
2	22 (34)	42 (66)
3	0 (0)	5 (100)
4	16 (62)	10 (38)
TP53 mutation, No. (%)		*P* = .06**
No	27 (63)	16 (37)
Yes	5 (31)	11 (69)
EGFR amplification, No. (%)		*P* = .10††
No	28 (51)	27 (49)
Yes	5 (100)	0 (0)

* Analysis of patient age and tumor characteristics vs isocitrate dehydrogenase (*IDH*) gene mutation status. *TP53* = tumor protein 53. *EGFR* = epidermal growth factor receptor.

† *IDH* gene mutation was assessed by sequencing tumor DNA.

‡ Association between age and *IDH* mutation was assessed using two-sided Student's t-test.
§ Tumors were previously reviewed by neuropathologists at the University of California San Francisco to assign histologic subtypes and grades according to the World Health Organization classification.
‖ Association between grade-specific histology and *IDH* mutation was assessed using two-sided Fisher's exact test.
¶ Association between primary vs. secondary glioblastoma and *IDH* mutation was assessed using two-sided χ^2 test.
# Association between tumor grade and *IDH* mutation was assessed using two-sided Fisher's exact test

** Association between *TP53* mutation and *IDH* mutation was assessed using two-sided χ^2 test.
†† Association between *EGFR* amplification and *IDH* mutation was assessed using two-sided χ^2 test.

Table 3. Survival analysis using multivariable Cox proportional hazards model*

Variable	HR† (95% CI)
Age	1.03 (1.00 to 1.06)
Sex	
Female	1.0 (Referent)
Male	0.73 (0.34 to 1.55)
IDH mutation‡	
No	1.0 (Referent)
Yes	0.27 (0.10 to 0.72)
Histology§	
Grade 2 astrocytoma	1.0 (Referent)
Grade 3 astrocytoma	1.79 (0.35 to 9.13)
Ependymoma	0.25 (0.06 to 1.06)
Primary glioblastoma	1.77 (0.60 to 5.22)
Secondary glioblastoma	3.94 (1.20 to 12.9)
Grade 2 oligoastrocytoma	2.8 (0.06 to 1.39)
Grade 3 oligoastrocytoma‖	-
Grade 2 oligodendroglioma	0.75 (0.21 to 2.69)

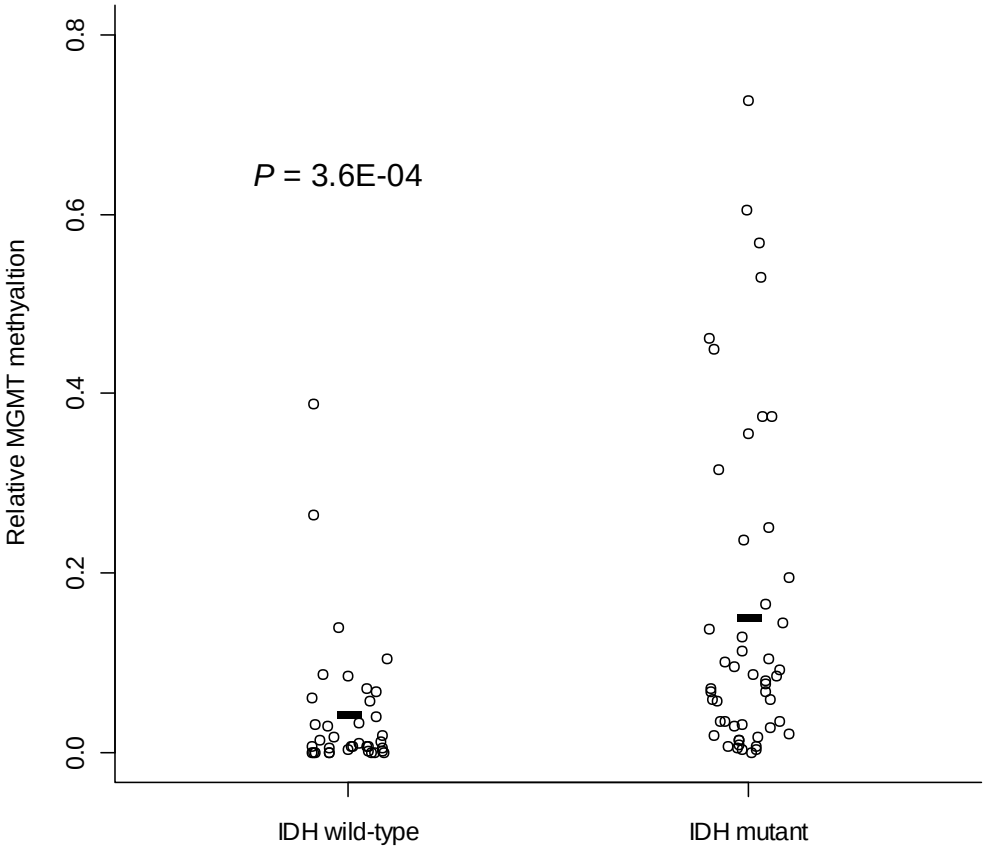
* Cox proportional hazards model of survival included age, sex, *IDH* mutation, and grade-specific histology. HR = hazards ratio, CI = confidence interval, *IDH* = isocitrate dehydrogenase gene.

† Adjusted HR values.

‡ *IDH* gene mutation was assessed by sequencing tumor DNA.

§Tumors were previously reviewed by neuropathologists at the University of California San Francisco to assign histologic subtypes and grades according to the World Health Organization classification.

‖ n =1, HR = 1.4E-07, standard error = 4,910, confidence interval indeterminable.



Supplementary Figure 1. Association between *IDH* mutation and increased *MGMT* methylation. *IDH* mutation status vs. relative *MGMT* methylation from quantitative methylation specific PCR demonstrates statistically significantly increased *MGMT* methylation among tumors with *IDH* mutation ($P = 3.6 \times 10^{-4}$). Black bars indicate mean relative *MGMT* methylation in *IDH* wild-type (0.04) and *IDH* mutant tumors (0.17).

Supplementary Table 1. Primer sequences for quantitative methylation specific polymerase chain reaction (QMSP), *IDH* mutation, *TP53* mutation, and *EGFR* amplification experiments*

Experiment	Forward 5'-3'	Reverse 5'-3'	Amplicon Size	Reference
QMSP				
RASSF1A-M	GTGTTAACGCGTTGCGTATC	AACCCCGCGAACTAAAAACG	94	Yu et al. 2004 (1)
MGMT3-M	GATTTGGTGAGTGTTTGGGTC	ACCACTCGAAACTACCACCG	79	This study
HOXA9-M	GAATTTAAGGGTTGTTCGGGC	GACCGCTCAAAAAATACCGCG	81	This study
PYCARD-M	GGTTGTAGCGGGGTGAGC	CGACGATCAAATTCTCCAACG	96	Stone et al. 2004 (2)
PEG3-M	TCGTCGTATTTGTCGTTAATTAATTC	GCAAACGCTATCCTAATTAATTAAACG	123	Maegawa et al. 2004 (3)
SLIT2-M	TTTAGGTTGCGGCGGAGTC	CAACGAACCCGTAACAAAACG	147	Dallol et al. 2003 (4)
ACTB	TGGTGATGGAGGAGGTTTAGTAAGT	AACCAATAAAACCTACTCCTCCCTTAA	133	Harden et al. 2003 (5)
IDH1				
Amplification	ATATTCTGGGTGGCACGGTCTT	CCTTGCTTAATGGGTGTAGATACCA	227	This study
Sequencing F	CGGTCTTCAGAGAAGCCATT			This study
Sequencing R	CATGCAAAATCACATTATTGCCAAC			This study
IDH2				
Amplification	TTCTGGTTGAAAGATGGCG	CAGGTCAGTGGATCCCCTC	251	This study
Sequencing F	ATGGCGGCTGCAGTGGG			This study
Sequencing R	CAGGTCAGTGGATCCCCTC			This study
TP53				
Exon 5	GTTCACTTGTGCCCTGA	AGCCCTGTCGTCTCT		Wiencke et al. 2005 (6)
Exon 6	CTCTGATTCCTCACTG	CCAGAGACCCCAGTTGCAAACC		Wiencke et al. 2005 (6)
Exon 7	TGCTTGCCACAGGTCT	ACAGCAGGCCAGTGT		Wiencke et al. 2005 (6)
Exon 8	AGGACCTGATTTCCTTAC	TCTGAGGCATAACTGC		Wiencke et al. 2005 (6)
Gene Amplification				
EGFR	CCGCATTAGCTCTTAGACCCA	GAATGCAACTTCCCAAAATGTGC	98	This study
GAPDH	CTCCCCACACACATGCACTTA	CCTAGTCCCAGGGCTTTGATT	99	This study

* *RASSF1*=Ras association domain family member 1, *MGMT*=O-6-methylguanine-DNA methyltransferase, *HOXA9*=homeobox A9, *PYCARD*=PYD and CARD domain containing, *PEG3*=paternally expressed 3, *SLIT2*=slit homolog 2, *ACTB*=actin, beta, *IDH*=isocitrate dehydrogenase, *TP53*=tumor protein 53, *EGFR*=epidermal growth factor receptor, *GAPDH*=glyceraldehyde-3-phosphate dehydrogenase.

Supplementary Table 2. Recursively partitioned mixture model methylation class by glioma histology and predicted methylation class membership for The Cancer Genome Atlas (TCGA) glioblastoma samples*

Methylation Class	AS2	AS3	EP	GBM	GBM2	OA2	OA3	OD2	PA	Predicted TCGA GBM Class
LLLLL	0	0	3	0	0	1	0	0	0	0
LLLLR	0	0	3	0	0	0	0	0	0	0
LLLRL	0	0	5	0	0	0	0	0	0	0
LLLRR	1	0	3	0	0	0	0	0	0	0
LLRLL	1	0	0	0	0	0	0	1	1	0
LLRLR	0	0	0	0	0	1	0	0	3	0
LLRR	4	0	0	0	0	1	0	0	0	1
LRL	0	1	0	14	0	0	0	0	0	56
LRR	0	0	1	3	8	0	0	0	0	10
RL	12	8	0	2	4	10	4	3	0	4
RR	2	0	0	1	0	9	5	16	0	0

* AS2=grade 2 Astrocytoma, AS3=grade 3 astrocytoma, EP=ependymoma, GBM=primary glioblastoma multiforme, GBM2=secondary glioblastoma multiforme, OA2=grade 2 oligoastrocytoma, OA3=grade 3 oligoastrocytoma, OD2=grade 2 oligodendroglioma, PA=pilocytic astrocytoma, TCGA=The Cancer Genome Atlas. Tumors were previously reviewed by UCSF neuropathologists to assign histologic subtypes and grades according to the World Health Organization classification.

Supplementary Table 3. Association between GoldenGate array methylation values and quantitative methylation specific polymerase chain reaction (QMSP)*

***GENE*_CpG †**	**No. ‡**	**Spearman (rho) §**	***P* ‖**
*PEG3*_E496	110	0.32	5.90E-04
*HOXA9*_E252	117	0.52	1.50E-09
*HOXA9*_E252	117	0.53	6.90E-10
*MGMT*_P272	110	0.45	7.40E-07
*MGMT*_P281	110	0.47	2.30E-07
*PYCARD*_E87	107	0.81	< 2.2E-16
*PYCARD*_P150	107	0.26	6.70E-03
*PYCARD*_P393	107	0.43	5.00E-06
*RASSF1A*_E116	118	0.7	< 2.2E-16
*RASSF1A*_P244	118	0.59	3.50E-12
*SLIT2*_E111	106	0.4	2.30E-05
*SLIT2*_P208	106	0.4	2.40E-05

* *PEG3* = paternally expressed 3, *HOXA9* = homeobox A9, *MGMT* = O-6-methylguanine-DNA methyltransferase, *PYCARD* = PYD and CARD domain containing, *RASSF1* = Ras association domain family member 1, *SLIT2* = slit homolog 2.
† This column lists the Illumina GoldenGate methylation array annotation for CpGs where the gene name is listed first in all capital letters and italics followed by an E for exon or P for promoter to indicate the location of the CpG relative to the transcription start site, and the number indicates the distance of the CpG from the transcription start site.
‡ Number of samples with both GoldenGate array and QMSP methylation data.
§ Spearman correlation coefficient (rho)
|| Two-sided Spearman's rank correlation test for association between GoldenGate array methylation value and QMSP methylation value.

Supplementary Table 4. Identification numbers (ID) and RPMM methylation class membership for The Cancer Genome Atlas (TCGA) glioblastoma samples used in validation.

TCGA ID	RPMM methylation class
TCGA-02-0001-01C-01D-0186-05	LRR
TCGA-02-0002-01A-01D-0186-05	LRL
TCGA-02-0003-01A-01D-0186-05	LRL
TCGA-02-0006-01B-01D-0186-05	LRL
TCGA-02-0007-01A-01D-0186-05	LRL
TCGA-02-0009-01A-01D-0186-05	LRL
TCGA-02-0010-01A-01D-0186-05	LRR
TCGA-02-0011-01B-01D-0186-05	LRR
TCGA-02-0014-01A-01D-0186-05	LRR
TCGA-02-0021-01A-01D-0186-05	LRL
TCGA-02-0024-01B-01D-0186-05	LRR
TCGA-02-0027-01A-01D-0186-05	LRL
TCGA-02-0028-01A-01D-0186-05	RL
TCGA-02-0033-01A-01D-0186-05	LRL
TCGA-02-0034-01A-01D-0186-05	LRL
TCGA-02-0037-01A-01D-0186-05	LRL
TCGA-02-0038-01A-01D-0186-05	LRL
TCGA-02-0043-01A-01D-0186-05	LRL
TCGA-02-0046-01A-01D-0186-05	LRL
TCGA-02-0047-01A-01D-0186-05	LRR
TCGA-02-0052-01A-01D-0186-05	LRL
TCGA-02-0054-01A-01D-0186-05	LRL
TCGA-02-0055-01A-01D-0186-05	LRL
TCGA-02-0057-01A-01D-0186-05	LRL
TCGA-02-0058-01A-01D-0186-05	RL
TCGA-02-0060-01A-01D-0186-05	LRL
TCGA-02-0064-01A-01D-0199-05	LRL
TCGA-02-0069-01A-01D-0199-05	LRR
TCGA-02-0071-01A-01D-0199-05	LRL
TCGA-02-0074-01A-01D-0199-05	LRL
TCGA-02-0075-01A-01D-0199-05	LRL

TCGA-02-0080-01A-01D-0199-05	RL
TCGA-02-0083-01A-01D-0199-05	LRL
TCGA-02-0085-01A-01D-0199-05	LRL
TCGA-02-0086-01A-01D-0199-05	LRL
TCGA-02-0089-01A-01D-0199-05	LRL
TCGA-02-0099-01A-01D-0199-05	LRL
TCGA-02-0102-01A-01D-0199-05	LRL
TCGA-02-0107-01A-01D-0199-05	LRL
TCGA-02-0113-01A-01D-0199-05	LRL
TCGA-02-0114-01A-01D-0199-05	LRR
TCGA-02-0115-01A-01D-0199-05	LRL
TCGA-02-0116-01A-01D-0199-05	LRL
TCGA-06-0119-01A-08D-0218-05	LRL
TCGA-06-0121-01A-04D-0218-05	LRL
TCGA-06-0122-01A-01D-0218-05	LRL
TCGA-06-0124-01A-01D-0218-05	LRL
TCGA-06-0125-01A-01D-0218-05	LRL
TCGA-06-0126-01A-01D-0218-05	LRL
TCGA-06-0128-01A-01D-0218-05	RL
TCGA-06-0129-01A-01D-0218-05	LRR
TCGA-06-0130-01A-01D-0218-05	LRL
TCGA-06-0133-01A-02D-0218-05	LRL
TCGA-06-0137-01A-01D-0218-05	LRL
TCGA-06-0137-01A-02D-0218-05	LRL
TCGA-06-0137-01A-03D-0218-05	LRL
TCGA-06-0137-01B-02D-0218-05	LRL
TCGA-06-0139-01A-01D-0218-05	LLRR
TCGA-06-0140-01A-01D-0218-05	LRL
TCGA-06-0141-01A-01D-0218-05	LRR
TCGA-06-0142-01A-01D-0218-05	LRL
TCGA-06-0143-01A-01D-0218-05	LRL
TCGA-06-0145-01A-01D-0218-05	LRL
TCGA-06-0145-01A-02D-0218-05	LRL
TCGA-06-0145-01A-03D-0218-05	LRL
TCGA-06-0145-01A-04D-0218-05	LRL
TCGA-06-0145-01A-05D-0218-05	LRL

TCGA-06-0145-01A-06D-0218-05	LRL
TCGA-06-0147-01A-01D-0218-05	LRL
TCGA-06-0148-01A-01D-0218-05	LRL
TCGA-06-0169-01A-01D-0218-05	LRL

Supplementary Table 5. Cellular pathways enriched among statistically significantly differentially methylated CpG loci in common among glioblastomas, astrocytomas, oligoastrocytomas, and oligodendrogliomas*.

Cellular Pathway	*P*†
Hypomethylated	
Methane Metabolism	.02
Stilbene, Coumarine and Lignin Biosynthesis	.02
Metabolism of Xenobiotics by Cytochrome P450	.02
PXR/RXR Activation	.02
Retinol Metabolism	.04
TREM1 Signaling	.04
Phenylalanine Metabolism	.05
Hypermethylated	
Retinoic acid Mediated Apoptosis Signaling	.005
Primary Immunodeficiency Signaling	.01
RAN Signaling	.02
NRF2-mediated Oxidative Stress Response	.03
EGF Signaling	.05

* CpG loci with statistically significantly differential methylation ($Q<0.05$ and $|\Delta\beta|>0.2$) between tumor and non-tumor tissue were examined for cellular pathway enrichment with Ingenuity pathways analysis software. PXR=nuclear receptor subfamily 1, group I, member 2; RXR=retinoid X receptor, gamma; TREM1=triggering receptor expressed on myeloid cells 1; RAN=RAN, member RAS oncogene family; NRF2=nuclear factor (erythroid-derived 2)-like 2; EGF=epidermal growth factor.

† Two-sided Fisher's exact test for enrichment of genes whose CpG loci are represented among the genes in the listed pathways.

Supplementary Table 6. Statistically significantly differentially hypomethylated CpG loci in human gliomas.

GENE_CpG*	*Q*-value	Median Δβ Value
Primary glioblastoma		
CASP10_P334_F	0.002	-0.587
CD82_P557_R	0.002	-0.289
CDK2_P330_R	0.002	-0.234
DDR1_P332_R	0.002	-0.299
DSG1_P159_R	0.002	-0.249
GFAP_P1214_F	0.002	-0.293
GSTM2_P109_R	0.002	-0.261
LEFTY2_P561_F	0.002	-0.452
MPO_E302_R	0.002	-0.227
MPO_P883_R	0.002	-0.637
PSCA_P135_F	0.002	-0.205
PTHR1_P258_F	0.002	-0.509
TRIP6_P1274_R	0.002	-0.500
TRPM5_P721_F	0.002	-0.265
UGT1A1_P315_R	0.002	-0.276
IFNG_E293_F	0.002	-0.274
MKRN3_P108_F	0.002	-0.355
NOTCH4_E4_F	0.002	-0.561
WNT8B_E487_F	0.002	-0.226
IFNG_P188_F	0.002	-0.205
IL8_E118_R	0.002	-0.666
TMPRSS4_P552_F	0.002	-0.356
TRIP6_P1090_F	0.002	-0.504
GFAP_P56_R	0.002	-0.375
SERPINE1_E189_R	0.002	-0.251
CCL3_E53_R	0.002	-0.302
PADI4_P1011_R	0.002	-0.361
CASP10_P186_F	0.002	-0.630
SH3BP2_E18_F	0.002	-0.267
ACVR1_P983_F	0.002	-0.321
BLK_P14_F	0.002	-0.273
SLC14A1_P369_R	0.002	-0.318
SPP1_E140_R	0.002	-0.382
MMP10_E136_R	0.003	-0.290
PRSS1_P1249_R	0.003	-0.393
TNK1_P221_F	0.003	-0.207
MKRN3_E144_F	0.003	-0.230
MPL_P657_F	0.003	-0.376

GENE_CpG*	*Q*-value	Median Δβ Value
Grade 3 Astrocytoma		
ACVR1_P983_F	0.007	-0.363
CASP10_P334_F	0.007	-0.412
CD82_P557_R	0.007	-0.289
DDR1_P332_R	0.007	-0.349
GFAP_P1214_F	0.007	-0.324
GSTM2_P109_R	0.007	-0.255
IL16_P93_R	0.007	-0.394
IL8_E118_R	0.007	-0.578
LEFTY2_P561_F	0.007	-0.426
MKRN3_E144_F	0.007	-0.480
MKRN3_P108_F	0.007	-0.509
MPO_P883_R	0.007	-0.514
PADI4_P1011_R	0.007	-0.246
PTHR1_P258_F	0.007	-0.446
TRIP6_P1090_F	0.007	-0.497
TRIP6_P1274_R	0.007	-0.489
UGT1A1_P315_R	0.007	-0.335
CCL3_E53_R	0.008	-0.261
CDK2_P330_R	0.008	-0.235
HBII_52_E142_F	0.008	-0.253
SERPINE1_P519_F	0.010	-0.454
P2RX7_P597_F	0.012	-0.491
PLA2G2A_P528_F	0.012	-0.210
ACVR1_E328_R	0.015	-0.353
CXCL9_E268_R	0.015	-0.205
JAK3_P1075_R	0.015	-0.455
MMP2_P303_R	0.015	-0.291
PDGFRA_E125_F	0.015	-0.266
GFAP_P56_R	0.019	-0.312
EPM2A_P113_F	0.023	-0.205
PRSS1_E45_R	0.023	-0.329
PRSS1_P1249_R	0.023	-0.389
TJP2_P518_F	0.023	-0.335
TRPM5_P721_F	0.026	-0.213
BLK_P14_F	0.033	-0.217
PTK7_E317_F	0.033	-0.219
NOTCH4_E4_F	0.049	-0.304
Grade 2 Astrocytoma		

TNFSF10_E53_F	0.003	-0.458
S100A2_E36_R	0.003	-0.228
SERPINE1_P519_F	0.003	-0.507
TJP2_P518_F	0.003	-0.347
JAK3_P1075_R	0.004	-0.493
KLK11_P103_R	0.004	-0.383
ACVR1_E328_R	0.005	-0.304
HBII_52_E142_F	0.005	-0.261
IL16_P93_R	0.005	-0.312
S100A2_P1186_F	0.005	-0.454
ZNFN1A1_E102_F	0.005	-0.252
IL8_P83_F	0.005	-0.585
NAT2_P11_F	0.005	-0.203
PI3_P1394_R	0.005	-0.367
SHB_P691_R	0.005	-0.254
EMR3_P39_R	0.006	-0.481
KLK10_P268_R	0.006	-0.208
CD86_P3_F	0.008	-0.436
CSF3R_P8_F	0.009	-0.370
GSTM2_P453_R	0.010	-0.263
NOS2A_P288_R	0.010	-0.264
MBD2_P233_F	0.011	-0.254
MMP2_P303_R	0.011	-0.332
MMP9_P189_F	0.011	-0.413
PDGFRA_E125_F	0.011	-0.250
FGFR2_P460_R	0.013	-0.327
PRSS1_E45_R	0.013	-0.304
ALPL_P433_F	0.016	-0.227
FGF1_E5_F	0.016	-0.213
PADI4_P1158_R	0.016	-0.243
VAV1_P317_F	0.016	-0.214
MC2R_P1025_F	0.018	-0.243
CASP10_E139_F	0.022	-0.305
NGFR_P355_F	0.022	-0.263
PIK3R1_P307_F	0.022	-0.234
STAT5A_E42_F	0.024	-0.298
C4B_E171_F	0.027	-0.205
CPA4_E20_F	0.027	-0.258
DDR2_E331_F	0.032	-0.246
HDAC1_P414_R	0.032	-0.522
TNFSF10_P2_R	0.032	-0.369
Secondary Glioblastoma		
CD82_P557_R	0.012	-0.298
DDR1_P332_R	0.012	-0.310
GFAP_P1214_F	0.012	-0.371
LEFTY2_P561_F	0.012	-0.402
TRIP6_P1090_F	0.012	-0.458
TRIP6_P1274_R	0.012	-0.476
UGT1A1_P315_R	0.012	-0.316
GFAP_P56_R	0.013	-0.450
GSTM2_P109_R	0.013	-0.260
MKRN3_E144_F	0.013	-0.347
PTHR1_P258_F	0.014	-0.233
CDK2_P330_R	0.017	-0.232
MPO_P883_R	0.018	-0.246
MKRN3_P108_F	0.019	-0.499
ACVR1_E328_R	0.019	-0.321
IL12B_P392_R	0.021	-0.301
ACVR1_P983_F	0.025	-0.356
IL16_P93_R	0.025	-0.258
IL8_E118_R	0.026	-0.416
CASP10_P334_F	0.032	-0.269
SLC14A1_P369_R	0.037	-0.274
P2RX7_P597_F	0.047	-0.230
Grade 3 Oligoastrocytoma		
ACVR1_P983_F	0.003	-0.428
CD82_P557_R	0.003	-0.246
DDR1_P332_R	0.003	-0.349
GFAP_P1214_F	0.003	-0.377
GSTM2_P109_R	0.003	-0.255
LEFTY2_P561_F	0.003	-0.427
MKRN3_E144_F	0.003	-0.384
MKRN3_P108_F	0.003	-0.615
MPO_P883_R	0.003	-0.346
TRIP6_P1090_F	0.003	-0.540
TRIP6_P1274_R	0.003	-0.528
UGT1A1_P315_R	0.003	-0.451
CASP10_P334_F	0.004	-0.313
CDK2_P330_R	0.004	-0.247
ACVR1_E328_R	0.005	-0.332
IL16_P93_R	0.005	-0.381
GFAP_P56_R	0.007	-0.452
PDGFRA_E125_F	0.007	-0.236
PWCR1_P357_F	0.007	-0.204

CD82_P557_R	0.002	-0.270	IL8_E118_R	0.009	-0.521
CXCL9_E268_R	0.002	-0.441	TRPM5_P721_F	0.014	-0.211
DSG1_P159_R	0.002	-0.406	BLK_P14_F	0.040	-0.205
EMR3_P1297_R	0.002	-0.406	RUNX1T1_E145_R	0.040	-0.229
			Grade 2 Oligoastrocytoma		
GABRA5_P1016_F	0.002	-0.339			
GFAP_P1214_F	0.002	-0.408	ACVR1_E328_R	0.006	-0.403
GSTM2_P109_R	0.002	-0.251	CDK2_P330_R	0.006	-0.251
IFNG_E293_F	0.002	-0.326	DDR1_P332_R	0.006	-0.334
IFNG_P188_F	0.002	-0.447	GFAP_P1214_F	0.006	-0.275
IL8_P83_F	0.002	-0.615	GSTM2_P109_R	0.006	-0.237
ITK_P114_F	0.002	-0.312	IL16_P93_R	0.006	-0.360
JAK3_P1075_R	0.002	-0.562	IL8_E118_R	0.006	-0.497
KLK10_P268_R	0.002	-0.211	LEFTY2_P561_F	0.006	-0.369
KLK11_P103_R	0.002	-0.355	MKRN3_E144_F	0.006	-0.267
KRT1_P798_R	0.002	-0.297	TRIP6_P1090_F	0.006	-0.456
LEFTY2_P561_F	0.002	-0.489	TRIP6_P1274_R	0.006	-0.509
MPO_E302_R	0.002	-0.427	UGT1A1_P315_R	0.006	-0.385
MPO_P883_R	0.002	-0.633	MKRN3_P108_F	0.007	-0.488
PADI4_P1011_R	0.002	-0.647	CD82_P557_R	0.007	-0.295
PI3_P1394_R	0.002	-0.330	GFAP_P56_R	0.007	-0.444
PRSS1_E45_R	0.002	-0.498	CASP10_P334_F	0.007	-0.277
PRSS1_P1249_R	0.002	-0.664	PDGFRA_E125_F	0.007	-0.250
PSCA_P135_F	0.002	-0.468	ACVR1_P983_F	0.007	-0.341
PTHR1_P258_F	0.002	-0.560	P2RX7_P597_F	0.007	-0.284
PWCR1_P357_F	0.002	-0.238	MPO_P883_R	0.008	-0.258
SPI1_P929_F	0.002	-0.206	HTR2A_E10_R	0.010	-0.248
SPP1_P647_F	0.002	-0.291	PTHR1_P258_F	0.024	-0.235
			Grade 2 Oligodendroglioma		
TMPRSS4_P552_F	0.002	-0.520			
TRIP6_P1274_R	0.002	-0.539	ACVR1_E328_R	0.002	-0.420
UGT1A1_P315_R	0.002	-0.419	DDR1_P332_R	0.002	-0.302
WNT8B_E487_F	0.002	-0.369	GFAP_P1214_F	0.002	-0.243
CDK2_P330_R	0.002	-0.241	GSTM2_P109_R	0.002	-0.254
CSF3R_P8_F	0.002	-0.498	IL16_P93_R	0.002	-0.401
DDR1_P332_R	0.002	-0.328	MKRN3_E144_F	0.002	-0.207
IL16_P93_R	0.002	-0.421	MKRN3_P108_F	0.002	-0.515
MKRN3_P108_F	0.002	-0.496	TRIP6_P1274_R	0.002	-0.511
P2RX7_P597_F	0.002	-0.562	UGT1A1_P315_R	0.002	-0.426
PSCA_E359_F	0.002	-0.434	MPO_P883_R	0.002	-0.204
TNK1_P221_F	0.002	-0.254	CASP10_P334_F	0.003	-0.316
TRPM5_P721_F	0.002	-0.448	TRIP6_P1090_F	0.003	-0.429
HBII_52_E142_F	0.002	-0.468	LEFTY2_P561_F	0.004	-0.308
NOTCH4_E4_F	0.002	-0.604	ACVR1_P983_F	0.004	-0.259

CASP10_P334_F	0.003	-0.554	PDGFRA_E125_F	0.004	-0.256
HLA_DQA2_E93_F	0.003	-0.226	CDK2_P330_R	0.005	-0.233
ACVR1_E328_R	0.003	-0.370	SPP1_E140_R	0.005	-0.292
ACVR1_P983_F	0.003	-0.442	CD82_P557_R	0.007	-0.235
BLK_P14_F	0.003	-0.360	IL8_E118_R	0.008	-0.354
GFAP_P56_R	0.003	-0.471	PEG3_E496_F	0.008	-0.276
GLI2_E90_F	0.003	-0.320	GFAP_P56_R	0.009	-0.374
MKRN3_E144_F	0.003	-0.531	**Ependymoma**		
MPL_P657_F	0.003	-0.449	ACVR1_E328_R	0.007	-0.387
PLA2G2A_P528_F	0.003	-0.399	ACVR1_P983_F	0.007	-0.525
ZNFN1A1_E102_F	0.003	-0.267	CD82_P557_R	0.007	-0.291
PLA2G2A_E268_F	0.004	-0.277	CDK2_P330_R	0.007	-0.236
TNFSF8_P184_F	0.004	-0.219	DDR1_P332_R	0.007	-0.335
CCR5_P630_R	0.004	-0.271	FGF1_E5_F	0.007	-0.393
EMR3_P39_R	0.004	-0.288	FGF1_P357_R	0.007	-0.288
FGF7_P44_F	0.004	-0.308	FGFR2_P460_R	0.007	-0.413
CCL3_E53_R	0.005	-0.589	GFAP_P1214_F	0.007	-0.456
CD1A_P6_F	0.005	-0.215	GSTM2_P109_R	0.007	-0.263
EMR3_E61_F	0.005	-0.458	GSTM2_P453_R	0.007	-0.399
PTPRH_E173_F	0.005	-0.489	LEFTY2_P561_F	0.007	-0.466
SERPINE1_P519_F	0.005	-0.492	MMP14_P13_F	0.007	-0.561
TEK_P526_F	0.005	-0.228	MPL_P657_F	0.007	-0.453
TGFB1_P833_R	0.005	-0.334	RIPK1_P744_R	0.007	-0.446
TRIP6_P1090_F	0.005	-0.521	RIPK1_P868_F	0.007	-0.414
IL8_E118_R	0.006	-0.563	SERPINE1_P519_F	0.007	-0.504
ALPL_P433_F	0.007	-0.278	SH3BP2_E18_F	0.007	-0.330
HLA_DPA1_P28_R	0.007	-0.225	SPARC_P195_F	0.007	-0.209
NOS2A_P288_R	0.007	-0.501	TNK1_P221_F	0.007	-0.291
PDGFRA_E125_F	0.007	-0.288	TRIP6_P1274_R	0.007	-0.511
SH3BP2_E18_F	0.007	-0.300	UGT1A1_P315_R	0.007	-0.421
CSF3_P309_R	0.010	-0.302	FASTK_P598_R	0.008	-0.292
PEG3_E496_F	0.010	-0.245	MATK_P190_R	0.008	-0.294
CASP10_P186_F	0.012	-0.548	SERPINE1_E189_R	0.008	-0.262
DDR2_E331_F	0.012	-0.410	SHB_P691_R	0.008	-0.299
HLA_DPA1_P205_R	0.012	-0.351	SLC14A1_P369_R	0.008	-0.487
SLC14A1_P369_R	0.012	-0.361	FAS_P65_F	0.009	-0.237
IL6_P213_R	0.017	-0.306	CHI3L2_E10_F	0.013	-0.253
KIAA0125_E29_F	0.017	-0.245	GFAP_P56_R	0.013	-0.508
MMP2_P303_R	0.017	-0.329	HPN_P374_R	0.013	-0.527
SIN3B_P514_R	0.017	-0.201	IL1RN_P93_R	0.013	-0.335
SPP1_E140_R	0.017	-0.356	NAT2_P11_F	0.013	-0.335
AOC3_P890_R	0.021	-0.237	HPN_P823_F	0.015	-0.573
LEFTY2_P719_F	0.021	-0.268	ALPL_P433_F	0.020	-0.334

SLC22A18_P216_R	0.021	-0.201	TGFBI_P173_F	0.020	-0.295
TJP2_P518_F	0.021	-0.347	TJP2_P518_F	0.020	-0.375
VAMP8_P114_F	0.021	-0.337	TRIP6_P1090_F	0.020	-0.523
NAT2_P11_F	0.024	-0.322	IL8_E118_R	0.022	-0.560
PADI4_P1158_R	0.033	-0.364	MUC1_P191_F	0.022	-0.262
PDGFB_P719_F	0.033	-0.318	ZNFN1A1_E102_F	0.022	-0.203
CCKAR_P270_F	0.038	-0.453	IGF2AS_P203_F	0.026	-0.241
ZNF264_P397_F	0.038	-0.272	HTR2A_E10_R	0.029	-0.292
MBD2_P233_F	0.043	-0.300	RBP1_P426_R	0.029	-0.400
SHB_P691_R	0.043	-0.284	S100A2_P1186_F	0.029	-0.255
			MMP10_E136_R	0.032	-0.367
			TNFSF10_E53_F	0.032	-0.342
			EPHA2_P203_F	0.035	-0.318
			LCN2_P86_R	0.035	-0.332
			SEPT5_P464_R	0.035	-0.243

* This column lists the Illumina GoldenGate methylation array annotation for CpGs where the gene name is listed first in all capital letters followed by an E for exon or P for promoter to indicate the location of the CpG relative to the transcription start site, and the number indicates the distance of the CpG from the transcription start site, and F indicates forward strand and R indicates reverse strand.

Supplementary Table 7. Statistically significantly differentially hypermethylated CpG loci in human gliomas.

GENE_ CpG*	*Q*-value	Median Δβ Value	GENE_ CpG*	*Q*-value	Median Δβ Value
Primary Glioblastoma			**Grade 3 Oligoastrocytoma**		
AATK_P519_R	0.002	0.214	AATK_P519_R	0.003	0.217
CD40_P372_R	0.002	0.310	ABCG2_P310_R	0.003	0.748
FZD9_E458_F	0.002	0.771	ALOX12_E85_R	0.003	0.647
IRAK3_E130_F	0.002	0.277	ALOX12_P223_R	0.003	0.359
IRAK3_P185_F	0.002	0.538	ATP10A_P147_F	0.003	0.603
MEST_E150_F	0.002	0.521	BMP4_P123_R	0.003	0.668
MEST_P4_F	0.002	0.477	BMP4_P199_R	0.003	0.231
SLIT2_P208_F	0.002	0.281	CCKAR_E79_F	0.003	0.377
TES_P182_F	0.002	0.675	CD40_E58_R	0.003	0.541
TNFRSF10A_P91_F	0.002	0.706	CD40_P372_R	0.003	0.314
TP73_P945_F	0.002	0.291	CD81_P272_R	0.003	0.599
CD81_P272_R	0.002	0.603	CD9_P504_F	0.003	0.582
GFI1_P45_R	0.002	0.455	CDH3_E100_R	0.003	0.720
HOXA9_P1141_R	0.002	0.748	CDH3_P87_R	0.003	0.764
MEST_P62_R	0.002	0.464	CDKN1B_P1161_F	0.003	0.226
TAL1_E122_F	0.002	0.250	CFTR_P115_F	0.003	0.238
TNFRSF10D_E27_F	0.002	0.664	COL18A1_P494_R	0.003	0.802
HTR1B_P222_F	0.002	0.570	CRIP1_P874_R	0.003	0.597
RAB32_P493_R	0.002	0.370	CTNNA1_P382_R	0.003	0.809
TNFRSF10A_P171_F	0.002	0.550	DDIT3_P1313_R	0.003	0.444
DIO3_P674_F	0.002	0.512	DES_E228_R	0.003	0.752
FLT3_E326_R	0.002	0.669	DNAJC15_P65_F	0.003	0.343
FLT4_E206_F	0.002	0.278	DSC2_E90_F	0.003	0.829
HOXA11_E35_F	0.002	0.472	EIF2AK2_P313_F	0.003	0.604
F2R_P839_F	0.002	0.416	ELK3_P514_F	0.003	0.585
HOXA5_P1324_F	0.002	0.329	EPHA2_P203_F	0.003	0.282
IRAK3_P13_F	0.002	0.252	EPHA2_P340_R	0.003	0.476
KIAA1804_P689_R	0.002	0.370	ERBB3_E331_F	0.003	0.814
MT1A_P600_F	0.002	0.245	ERCC6_P698_R	0.003	0.435
CD40_E58_R	0.002	0.318	ERN1_P809_R	0.003	0.252
PRKCDBP_E206_F	0.002	0.634	ESR2_E66_F	0.003	0.790
DNAJC15_P65_F	0.003	0.324	ESR2_P162_F	0.003	0.796
DSC2_E90_F	0.003	0.596	EYA4_P508_F	0.003	0.441
MAPK10_E26_F	0.003	0.204	EYA4_P794_F	0.003	0.369
ALOX12_P223_R	0.003	0.296	F2R_P839_F	0.003	0.438
DES_E228_R	0.003	0.221	FABP3_E113_F	0.003	0.592
HOXA9_E252_R	0.003	0.676	FAS_P65_F	0.003	0.618

ISL1_E87_R	0.003	0.236	FES_E34_R	0.003	0.842
MOS_E60_R	0.003	0.595	FES_P223_R	0.003	0.770
ALOX12_E85_R	0.004	0.587	FGFR3_P1152_R	0.003	0.365
BMP4_P199_R	0.004	0.222	FRZB_E186_R	0.003	0.805
HOXA11_P698_F	0.004	0.777	FZD9_E458_F	0.003	0.762
HTR1B_E232_R	0.004	0.412	GFI1_P45_R	0.003	0.535
RARRES1_P426_R	0.004	0.340	GLI3_P453_R	0.003	0.788
GATA6_P21_R	0.005	0.579	GNMT_E126_F	0.003	0.806
GFI1_E136_F	0.005	0.425	GNMT_P197_F	0.003	0.682
LY6G6E_P45_R	0.005	0.230	GP1BB_P278_R	0.003	0.261
SLC22A3_P528_F	0.005	0.203	GRB7_P160_R	0.003	0.389
TES_E172_F	0.005	0.412	GUCY2D_P48_R	0.003	0.507
ZNF215_P129_R	0.005	0.233	HFE_E273_R	0.003	0.796
HOXB2_P488_R	0.006	0.351	HHIP_P578_R	0.003	0.273
RARA_P176_R	0.006	0.349	HIC_1_seq_48_S103_R	0.003	0.228
TUSC3_E29_R	0.006	0.676	HIC2_P498_F	0.003	0.432
DIO3_P90_F	0.006	0.286	HOXA9_E252_R	0.003	0.280
HIC_1_seq_48_S103_R	0.006	0.210	HRASLS_P353_R	0.003	0.359
HOXA5_P479_F	0.006	0.213	HS3ST2_E145_R	0.003	0.596
TAL1_P594_F	0.006	0.794	ICA1_P61_F	0.003	0.330
GATA6_P726_F	0.008	0.565	ICA1_P72_R	0.003	0.759
GJB2_P931_R	0.008	0.249	IFNGR2_P377_R	0.003	0.217
TDGF1_P428_R	0.008	0.259	IGF1_E394_F	0.003	0.225
ISL1_P379_F	0.009	0.222	IGFBP2_P306_F	0.003	0.574
DIO3_E230_R	0.010	0.471	IL17RB_E164_R	0.003	0.813
RAN_P581_R	0.010	0.327	INSR_P1063_R	0.003	0.428
ZP3_E90_F	0.010	0.424	IRAK3_E130_F	0.003	0.260
DDIT3_P1313_R	0.011	0.370	IRAK3_P13_F	0.003	0.279
TWIST1_E117_R	0.011	0.265	IRF5_E101_F	0.003	0.435
MYOD1_E156_F	0.014	0.317	ITPR2_P804_F	0.003	0.754
NEFL_P209_R	0.014	0.616	JAK3_E64_F	0.003	0.625
IPF1_P750_F	0.016	0.439	JAK3_P156_R	0.003	0.425
EYA4_P794_F	0.018	0.214	KIAA1804_P689_R	0.003	0.745
GSTM1_P266_F	0.020	0.316	KIT_P367_R	0.003	0.488
BMP4_P123_R	0.022	0.420	KLK10_P268_R	0.003	0.492
HOXA9_P303_F	0.022	0.394	LOX_P313_R	0.003	0.781
PALM2_AKAP2_P420_R	0.022	0.252	LY6G6E_P45_R	0.003	0.313
TFAP2C_P765_F	0.022	0.368	LYN_P241_F	0.003	0.786
JAK3_P156_R	0.024	0.340	MAP3K1_E81_F	0.003	0.822
SFN_P248_F	0.024	0.344	MAP3K1_P7_F	0.003	0.796
IPF1_P234_F	0.030	0.389	MATK_P190_R	0.003	0.532

MAP3K1_P7_F	0.030	0.450	MEST_E150_F	0.003	0.606
TNFRSF10C_P7_F	0.030	0.398	MEST_P4_F	0.003	0.495
BCR_P346_F	0.032	0.284	MEST_P62_R	0.003	0.525
SCGB3A1_E55_R	0.032	0.287	MET_E333_F	0.003	0.794
CAV2_E33_R	0.036	0.363	MGMT_P272_R	0.003	0.240
FZD9_P175_F	0.036	0.391	MGMT_P281_F	0.003	0.202
HOXA5_E187_F	0.036	0.259	MMP14_P13_F	0.003	0.323
ITPR2_P804_F	0.040	0.203	MOS_E60_R	0.003	0.694
TAL1_P817_F	0.040	0.254	MST1R_P392_F	0.003	0.225
TSP50_P137_F	0.040	0.259	MT1A_E13_R	0.003	0.567
CRIP1_P874_R	0.044	0.234	MT1A_P49_R	0.003	0.397
IGFBP1_P12_R	0.044	0.403	MYCL1_P502_R	0.003	0.753
HOXB2_P99_F	0.049	0.243	MYLK_P469_R	0.003	0.376
PYCARD_E87_F	0.049	0.308	NOTCH3_P198_R	0.003	0.733
WRN_P969_F	0.049	0.287	PAX6_P1121_F	0.003	0.518
Secondary Glioblastoma			POMC_P400_R	0.003	0.404
AATK_P519_R	0.002	0.215	PRKCDBP_E206_F	0.003	0.732
ABCG2_P310_R	0.002	0.475	PYCARD_E87_F	0.003	0.710
ALOX12_E85_R	0.002	0.633	PYCARD_P393_F	0.003	0.276
ALOX12_P223_R	0.002	0.354	RAB32_E314_R	0.003	0.542
BMP4_P199_R	0.002	0.234	RAB32_P493_R	0.003	0.476
CD81_P272_R	0.002	0.651	RAN_P581_R	0.003	0.410
CFTR_P115_F	0.002	0.453	RARRES1_P426_R	0.003	0.549
CFTR_P372_R	0.002	0.631	RASSF1_E116_F	0.003	0.894
CRIP1_P274_F	0.002	0.465	RASSF1_P244_F	0.003	0.514
CRIP1_P874_R	0.002	0.522	RBP1_E158_F	0.003	0.889
DES_E228_R	0.002	0.662	RBP1_P150_F	0.003	0.797
DSC2_E90_F	0.002	0.696	RBP1_P426_R	0.003	0.504
EYA4_P794_F	0.002	0.337	SCGB3A1_E55_R	0.003	0.795
FZD9_E458_F	0.002	0.708	SEPT9_P374_F	0.003	0.223
GFI1_E136_F	0.002	0.729	SLC22A3_P528_F	0.003	0.276
GFI1_P45_R	0.002	0.545	TES_P182_F	0.003	0.666
HIC_1_seq_48_S103_R	0.002	0.222	TGFB2_E226_R	0.003	0.830
HOXA9_E252_R	0.002	0.656	THBS1_E207_R	0.003	0.837
IRAK3_E130_F	0.002	0.612	TIMP3_seq_7_S38_F	0.003	0.419
IRF5_E101_F	0.002	0.231	TJP1_P390_F	0.003	0.812
JAK3_P156_R	0.002	0.369	TNFRSF10A_P171_F	0.003	0.566
KIAA1804_P689_R	0.002	0.761	TNFRSF10A_P91_F	0.003	0.727
MAPK10_E26_F	0.002	0.209	TNFRSF10D_E27_F	0.003	0.716
MEST_E150_F	0.002	0.563	TNFRSF10D_P70_F	0.003	0.238
MEST_P4_F	0.002	0.491	TRIP6_E33_F	0.003	0.752

MEST_P62_R	0.002	0.500	VAV2_P1182_F	0.003	0.360
MGMT_P272_R	0.002	0.225	ZMYND10_E77_R	0.003	0.696
MGMT_P281_F	0.002	0.394	ZMYND10_P329_F	0.003	0.653
MST1R_P392_F	0.002	0.222	ZP3_E90_F	0.003	0.745
MT1A_E13_R	0.002	0.531	CAV2_E33_R	0.004	0.817
MT1A_P49_R	0.002	0.517	CFTR_P372_R	0.004	0.698
MT1A_P600_F	0.002	0.245	CRIP1_P274_F	0.004	0.459
PRKCDBP_E206_F	0.002	0.697	GSTM1_P266_F	0.004	0.466
RAB32_E314_R	0.002	0.520	MAPK10_E26_F	0.004	0.233
RAB32_P493_R	0.002	0.457	MMP14_P208_R	0.004	0.424
RASSF1_E116_F	0.002	0.888	NGFR_P355_F	0.004	0.207
RASSF1_P244_F	0.002	0.463	PGF_P320_F	0.004	0.661
RBP1_E158_F	0.002	0.894	TAL1_P594_F	0.004	0.574
SYK_E372_F	0.002	0.348	WT1_P853_F	0.004	0.384
TAL1_P594_F	0.002	0.656	CCNA1_E7_F	0.005	0.626
TNFRSF10A_P171_F	0.002	0.558	CTNNB1_P757_F	0.005	0.308
TNFRSF10A_P91_F	0.002	0.716	E2F5_P516_R	0.005	0.361
TNFRSF10C_E109_F	0.002	0.381	ENC1_P484_R	0.005	0.409
TNFRSF10D_E27_F	0.002	0.713	EVI1_P30_R	0.005	0.719
TNFRSF10D_P70_F	0.002	0.240	GRB7_E71_R	0.005	0.226
COL18A1_P494_R	0.002	0.485	HTR1B_E232_R	0.005	0.580
CTNNA1_P382_R	0.002	0.315	IGFBP2_P353_R	0.005	0.227
ERBB3_E331_F	0.002	0.606	IRAK3_P185_F	0.005	0.383
ESR2_P162_F	0.002	0.750	TDGF1_P428_R	0.005	0.257
ZMYND10_E77_R	0.002	0.683	TNFRSF1B_P167_F	0.005	0.318
ERG_E28_F	0.002	0.687	BCR_P346_F	0.007	0.586
RARRES1_P426_R	0.002	0.524	GFI1_E136_F	0.007	0.650
TES_P182_F	0.002	0.608	IGFBP1_P12_R	0.007	0.476
THBS1_E207_R	0.002	0.825	SH3BP2_P771_R	0.007	0.447
CCNA1_E7_F	0.003	0.752	ZNF215_P129_R	0.007	0.674
DNAJC15_P65_F	0.003	0.367	CPA4_E20_F	0.009	0.249
ERN1_P809_R	0.003	0.216	FAS_P322_R	0.009	0.252
GNMT_E126_F	0.003	0.797	GRB10_P260_F	0.009	0.723
GP1BB_P278_R	0.003	0.241	IGFBP1_E48_R	0.009	0.537
HIC2_P498_F	0.003	0.262	TNFRSF10C_P7_F	0.009	0.404
HOXB2_P488_R	0.003	0.365	CDKN1C_P626_F	0.011	0.489
HS3ST2_E145_R	0.003	0.709	FGFR2_P266_R	0.011	0.324
ICA1_P61_F	0.003	0.319	GJB2_P931_R	0.011	0.373
TDGF1_P428_R	0.003	0.229	HDAC1_P414_R	0.011	0.242
GJB2_P931_R	0.003	0.313	HOXB2_P488_R	0.011	0.288
ERCC6_P698_R	0.004	0.289	ITGA6_P298_R	0.014	0.465
RBP1_P150_F	0.004	0.758	NR2F6_E375_R	0.014	0.239
SLC22A3_P528_F	0.004	0.284	PTCH2_E173_F	0.014	0.368

FRZB_E186_R	0.004	0.806	STAT5A_E42_F	0.014	0.218
HHIP_P578_R	0.004	0.464	TGFBI_P173_F	0.014	0.352
HOXA9_P303_F	0.004	0.304	CRK_P721_F	0.017	0.344
ICA1_P72_R	0.004	0.498	EVI1_E47_R	0.017	0.555
SCGB3A1_E55_R	0.004	0.783	JUNB_P1149_R	0.017	0.316
SFN_P248_F	0.004	0.288	MC2R_P1025_F	0.017	0.209
TJP1_P390_F	0.004	0.616	MCM2_P260_F	0.017	0.213
AREG_P217_R	0.005	0.411	MOS_P27_R	0.017	0.353
BMP4_P123_R	0.005	0.619	MT1A_P600_F	0.017	0.301
CD40_P372_R	0.005	0.318	PLAUR_E123_F	0.017	0.627
F2R_P839_F	0.005	0.441	PLSCR3_P751_R	0.017	0.377
GNMT_P197_F	0.005	0.666	TMEFF1_P626_R	0.017	0.235
GRB7_E71_R	0.005	0.399	AREG_P217_R	0.021	0.475
GRB7_P160_R	0.005	0.227	EPHB4_E476_R	0.021	0.224
GSTM1_P266_F	0.005	0.453	FGFR2_P460_R	0.021	0.279
HFE_E273_R	0.005	0.819	LAMC1_P808_F	0.021	0.372
HOXA9_P1141_R	0.005	0.733	PTK2_P735_R	0.021	0.311
HOXB2_P99_F	0.005	0.267	TFAP2C_P765_F	0.021	0.265
MAP3K1_E81_F	0.005	0.712	WRN_P969_F	0.026	0.309
MAP3K1_P7_F	0.005	0.691	COL1A2_P407_R	0.032	0.222
PYCARD_E87_F	0.005	0.693	HDAC5_E298_F	0.032	0.286
RAN_P581_R	0.005	0.383	RARRES1_P57_R	0.032	0.346
RBP1_P426_R	0.005	0.501	KCNK4_P171_R	0.040	0.303
TNFRSF10C_P7_F	0.005	0.454	FGF1_P357_R	0.048	0.208
ZMYND10_P329_F	0.005	0.615	HCK_P858_F	0.048	0.314
ITPR2_P804_F	0.006	0.631	MMP9_P237_R	0.048	0.221
RARRES1_P57_R	0.006	0.354	PADI4_P1158_R	0.048	0.208
ST6GAL1_P164_R	0.006	0.629	**Grade 2 Oligoastrocytoma**		
FES_E34_R	0.007	0.743	ALOX12_E85_R	0.006	0.614
ZP3_E90_F	0.007	0.385	CRIP1_P274_F	0.006	0.420
EPHA2_P340_R	0.009	0.330	FZD9_E458_F	0.006	0.669
TAL1_E122_F	0.009	0.285	LOX_P313_R	0.006	0.769
CD40_E58_R	0.010	0.458	MT1A_P600_F	0.006	0.275
CDH3_E100_R	0.010	0.648	RBP1_E158_F	0.006	0.725
CTNNB1_P757_F	0.010	0.207	SEPT9_P374_F	0.006	0.208
ELL_P693_F	0.010	0.362	CD40_E58_R	0.007	0.467
FRZB_P406_F	0.010	0.290	MEST_P4_F	0.007	0.473
HTR1B_E232_R	0.010	0.371	PRKCDBP_E206_F	0.007	0.676
JAK3_E64_F	0.010	0.432	AATK_P519_R	0.007	0.210
PAX6_P1121_F	0.010	0.258	CAV2_E33_R	0.007	0.580
WNT10B_P823_R	0.010	0.274	ERBB3_E331_F	0.007	0.698
HOXA11_P698_F	0.015	0.713	FAS_P65_F	0.007	0.593

CDH3_P87_R	0.017	0.589
EYA4_P508_F	0.017	0.373
FES_P223_R	0.017	0.667
FGFR3_P1152_R	0.017	0.256
HCK_P858_F	0.017	0.403
MOS_E60_R	0.017	0.624
DIO3_E230_R	0.021	0.384
LYN_P241_F	0.024	0.626
MET_E333_F	0.024	0.284
NOTCH3_P198_R	0.024	0.292
POMC_P400_R	0.024	0.341
TGFB2_E226_R	0.024	0.777
ATP10A_P147_F	0.028	0.314
CCKAR_E79_F	0.028	0.315
FAS_P65_F	0.028	0.589
HOXA5_P479_F	0.028	0.273
HRASLS_P353_R	0.028	0.312
IMPACT_P234_R	0.028	0.341
TIMP3_seq_7_S38_F	0.028	0.308
EVI1_E47_R	0.038	0.237
IRAK3_P185_F	0.038	0.595
PGF_P320_F	0.049	0.271
SGCE_E149_F	0.049	0.266
ST6GAL1_P528_F	0.049	0.647
TAL1_P817_F	0.049	0.201
Grade 3 Astrocytoma		
AATK_P519_R	0.007	0.211
ABCG2_P310_R	0.007	0.728
ALOX12_E85_R	0.007	0.616
ALOX12_P223_R	0.007	0.325
CD40_P372_R	0.007	0.282
CD81_P272_R	0.007	0.462
CFTR_P372_R	0.007	0.692
COL18A1_P494_R	0.007	0.635
CRIP1_P874_R	0.007	0.473
CTNNA1_P382_R	0.007	0.726
DES_E228_R	0.007	0.517
DSC2_E90_F	0.007	0.804
EPHA2_P340_R	0.007	0.433
ERBB3_E331_F	0.007	0.866
ERN1_P809_R	0.007	0.217
ESR2_P162_F	0.007	0.790
GFI1_P45_R	0.007	0.297
GLI3_P453_R	0.007	0.592
ICA1_P61_F	0.007	0.380
IGFBP1_P12_R	0.007	0.539
MATK_P190_R	0.007	0.433
MT1A_E13_R	0.007	0.410
TAL1_P594_F	0.007	0.225
ZP3_E90_F	0.007	0.371
ALOX12_P223_R	0.007	0.341
BMP4_P199_R	0.007	0.233
CCNA1_E7_F	0.007	0.639
COL18A1_P494_R	0.007	0.600
DDIT3_P1313_R	0.007	0.434
ESR2_P162_F	0.007	0.677
MAP3K1_E81_F	0.007	0.810
NOTCH3_P198_R	0.007	0.672
RASSF1_E116_F	0.007	0.760
RBP1_P150_F	0.007	0.679
BMP4_P123_R	0.007	0.645
CD81_P272_R	0.007	0.539
DSC2_E90_F	0.007	0.701
ENC1_P484_R	0.007	0.250
EYA4_P794_F	0.007	0.353
F2R_P839_F	0.007	0.411
FES_P223_R	0.007	0.709
GRB10_P260_F	0.007	0.553
GRB7_P160_R	0.007	0.362
HIC_1_seq_48_S103_R	0.007	0.204
JAK3_P156_R	0.007	0.380
MYCL1_P502_R	0.007	0.646
PAX6_P1121_F	0.007	0.322
POMC_P400_R	0.007	0.376
RAB32_E314_R	0.007	0.413
RASSF1_P244_F	0.007	0.436
TES_P182_F	0.007	0.277
TIMP3_seq_7_S38_F	0.007	0.441
TJP1_P390_F	0.007	0.756
TNFRSF10D_E27_F	0.007	0.696
TNFRSF10D_P70_F	0.007	0.226
ZMYND10_P329_F	0.007	0.654
ABCG2_P310_R	0.007	0.689
AREG_P217_R	0.007	0.446

EYA4_P794_F	0.007	0.312
F2R_P839_F	0.007	0.403
FRZB_P406_F	0.007	0.588
FZD9_E458_F	0.007	0.701
GFI1_E136_F	0.007	0.706
GFI1_P45_R	0.007	0.573
GSTM1_P266_F	0.007	0.518
HHIP_P578_R	0.007	0.379
HIC_1_seq_48_S103_R	0.007	0.216
HOXA9_E252_R	0.007	0.549
ICA1_P61_F	0.007	0.336
ICA1_P72_R	0.007	0.695
IFNGR2_P377_R	0.007	0.220
IL6_P611_F	0.007	0.336
IRF5_E101_F	0.007	0.202
ITPR2_P804_F	0.007	0.648
JAK3_P156_R	0.007	0.383
KIAA1804_P689_R	0.007	0.700
MAP3K1_E81_F	0.007	0.791
MAP3K1_P7_F	0.007	0.735
MEST_E150_F	0.007	0.466
MOS_E60_R	0.007	0.646
MT1A_P600_F	0.007	0.249
MYLK_P469_R	0.007	0.355
PGF_P320_F	0.007	0.535
PRKCDBP_E206_F	0.007	0.617
PTPN6_E171_R	0.007	0.248
RAB32_E314_R	0.007	0.552
RAB32_P493_R	0.007	0.422
RAN_P581_R	0.007	0.378
RARA_E128_R	0.007	0.313
RARA_P176_R	0.007	0.570
RARRES1_P426_R	0.007	0.514
RASSF1_E116_F	0.007	0.826
RBP1_E158_F	0.007	0.859
RBP1_P426_R	0.007	0.495
SCGB3A1_E55_R	0.007	0.733
SYK_E372_F	0.007	0.258
TAL1_E122_F	0.007	0.547
TES_P182_F	0.007	0.518
TNFRSF10A_P171_F	0.007	0.542
TNFRSF10A_P91_F	0.007	0.709
CDKN1B_P1161_F	0.007	0.227
CRIP1_P874_R	0.007	0.462
CTNNA1_P382_R	0.007	0.666
DES_E228_R	0.007	0.320
IGFBP2_P306_F	0.007	0.441
IL17RB_E164_R	0.007	0.716
MEST_E150_F	0.007	0.464
MET_E333_F	0.007	0.691
EPHA2_P340_R	0.008	0.438
FES_E34_R	0.008	0.690
FGFR3_P1152_R	0.008	0.336
GFI1_E136_F	0.008	0.614
GUCY2D_P48_R	0.008	0.515
IRF5_E101_F	0.008	0.222
KIAA1804_P689_R	0.008	0.609
LY6G6E_P45_R	0.008	0.262
MEST_P62_R	0.008	0.475
RARRES1_P426_R	0.008	0.489
TNFRSF10A_P91_F	0.008	0.554
TNFRSF10C_P7_F	0.008	0.394
CD9_P504_F	0.008	0.489
CTNNB1_P757_F	0.008	0.264
ERN1_P809_R	0.008	0.216
ESR2_E66_F	0.008	0.411
GSTM1_P266_F	0.008	0.375
IRAK3_P185_F	0.008	0.472
TNFRSF10A_P171_F	0.008	0.526
BMPR2_P1271_F	0.009	0.371
CCKAR_E79_F	0.009	0.407
ERCC6_P698_R	0.009	0.403
EVI1_E47_R	0.009	0.416
HHIP_P578_R	0.009	0.226
HRASLS_P353_R	0.009	0.368
TGFB2_E226_R	0.009	0.731
FABP3_E113_F	0.010	0.516
RAB32_P493_R	0.010	0.424
RAN_P581_R	0.010	0.386
RUNX1T1_P103_F	0.010	0.405
SH3BP2_P771_R	0.010	0.306
ZNF215_P129_R	0.010	0.462
ATP10A_P147_F	0.011	0.524
CD40_P372_R	0.011	0.294

TNFRSF10D_E27_F	0.007	0.712	HTR1B_E232_R	0.011	0.270
TNFRSF10D_P70_F	0.007	0.223	ICA1_P72_R	0.011	0.640
ZP3_E90_F	0.007	0.557	ITPR2_P804_F	0.011	0.579
BMP4_P199_R	0.008	0.228	MAP3K1_P7_F	0.011	0.760
CD40_E58_R	0.008	0.447	THBS1_E207_R	0.011	0.674
EPHA1_P119_R	0.008	0.509	CDH3_P87_R	0.012	0.515
ESR2_E66_F	0.008	0.767	FRZB_E186_R	0.012	0.684
HOXA11_P698_F	0.008	0.683	GNMT_E126_F	0.012	0.715
HOXA9_P1141_R	0.008	0.536	GNMT_P197_F	0.012	0.645
HS3ST2_E145_R	0.008	0.636	PYCARD_E87_F	0.012	0.652
MAPK10_E26_F	0.008	0.206	RBP1_P426_R	0.012	0.492
MT1A_P49_R	0.008	0.711	SCGB3A1_E55_R	0.012	0.706
POMC_P400_R	0.008	0.362	TRIP6_E33_F	0.012	0.641
PYCARD_E87_F	0.008	0.649	DNAJC15_P65_F	0.013	0.311
RASSF1_P244_F	0.008	0.420	ELK3_P514_F	0.013	0.523
TJP1_P390_F	0.008	0.861	EPHA2_P203_F	0.013	0.258
ZMYND10_E77_R	0.008	0.634	EYA4_P508_F	0.013	0.361
ZMYND10_P329_F	0.008	0.643	KLK10_P268_R	0.013	0.482
ZNF215_P129_R	0.008	0.606	SEMA3B_E96_F	0.013	0.223
CCNA1_E7_F	0.010	0.839	CFTR_P372_R	0.015	0.588
DDIT3_P1313_R	0.010	0.459	PGF_P320_F	0.015	0.413
ERCC6_P698_R	0.010	0.381	CPA4_E20_F	0.016	0.231
ERG_E28_F	0.010	0.743	EVI1_P30_R	0.016	0.269
FABP3_E113_F	0.010	0.547	GP1BB_P278_R	0.016	0.209
FES_E34_R	0.010	0.731	IGF1_E394_F	0.016	0.250
HDAC5_E298_F	0.010	0.222	MMP14_P13_F	0.016	0.222
MYCL1_P502_R	0.010	0.659	ERG_E28_F	0.018	0.394
SOX17_P303_F	0.010	0.235	GJB2_P931_R	0.018	0.243
TAL1_P594_F	0.010	0.624	JAK3_E64_F	0.018	0.564
TIMP3_seq_7_S38_F	0.010	0.543	LYN_P241_F	0.018	0.563
CFTR_P115_F	0.012	0.279	PTCH2_E173_F	0.018	0.297
DNAJC15_P65_F	0.012	0.261	VAV2_P1182_F	0.018	0.328
MEST_P62_R	0.012	0.444	JUNB_P1149_R	0.020	0.252
MOS_P27_R	0.012	0.209	ST6GAL1_P164_R	0.020	0.254
SEPT9_P374_F	0.012	0.205	BCR_P346_F	0.022	0.248
SOX17_P287_R	0.012	0.351	MYLK_P469_R	0.022	0.285
TNFRSF10C_P7_F	0.012	0.472	EIF2AK2_P313_F	0.024	0.562
BMPR2_P1271_F	0.015	0.277	HFE_E273_R	0.024	0.607
CRIP1_P274_F	0.015	0.417	HIC2_P498_F	0.024	0.228
CTNNB1_P757_F	0.015	0.318	IL6_P611_F	0.024	0.384
CTSL_P264_R	0.015	0.445	KIT_P367_R	0.024	0.313
GNMT_E126_F	0.015	0.751	PLAUR_E123_F	0.024	0.527
KIT_P367_R	0.015	0.374	ZNF215_P71_R	0.024	0.261

NOTCH3_P198_R	0.015	0.643
PLSCR3_P751_R	0.015	0.400
ACVR1C_P363_F	0.019	0.423
CDKN1B_P1161_F	0.019	0.366
EYA4_P508_F	0.019	0.510
GLI3_P453_R	0.019	0.668
HOXA9_P303_F	0.019	0.313
HRASLS_P353_R	0.019	0.342
MEST_P4_F	0.019	0.472
PTCH2_E173_F	0.019	0.282
RBP1_P150_F	0.019	0.695
SH3BP2_P771_R	0.019	0.378
TP73_P945_F	0.019	0.348
CAV2_E33_R	0.023	0.273
E2F5_P516_R	0.023	0.271
FES_P223_R	0.023	0.705
GJB2_P931_R	0.023	0.314
GUCY2D_P48_R	0.023	0.261
LYN_P241_F	0.023	0.508
NRG1_P558_R	0.023	0.324
PAX6_P1121_F	0.023	0.307
TNFRSF10C_E109_F	0.023	0.270
WNT10B_P823_R	0.023	0.295
BMP4_P123_R	0.026	0.609
CCKAR_E79_F	0.026	0.350
CD9_P504_F	0.026	0.545
CDH3_E100_R	0.026	0.717
CDH3_P87_R	0.026	0.705
ELK3_P514_F	0.026	0.339
EPHA2_P203_F	0.026	0.304
FAS_P65_F	0.026	0.623
FGFR3_P1152_R	0.026	0.356
FRZB_E186_R	0.026	0.782
GNMT_P197_F	0.026	0.674
GRB7_E71_R	0.026	0.708
GRB7_P160_R	0.026	0.388
HCK_P858_F	0.026	0.377
HFE_E273_R	0.026	0.776
HIC2_P498_F	0.026	0.354
HLA_F_E402_F	0.026	0.366
IRAK3_P185_F	0.026	0.490
JAK3_E64_F	0.026	0.669
ITGA6_P298_R	0.026	0.252
MOS_E60_R	0.026	0.570
TGFBI_P173_F	0.026	0.226
HCK_P858_F	0.028	0.306
MAPK10_E26_F	0.032	0.208
CDH3_E100_R	0.039	0.541
TNFRSF1B_P167_F	0.039	0.252
ACVR1C_P363_F	0.043	0.263
TEK_E75_F	0.043	0.200
MMP14_P208_R	0.046	0.246
MMP7_E59_F	0.046	0.208
TMEFF1_P626_R	0.050	0.216
Grade 2 Oligodendroglioma		
AATK_P519_R	0.002	0.208
ABCG2_P310_R	0.002	0.678
ALOX12_E85_R	0.002	0.617
ALOX12_P223_R	0.002	0.341
ATP10A_P147_F	0.002	0.581
BMP4_P123_R	0.002	0.646
BMP4_P199_R	0.002	0.230
BMPR2_P1271_F	0.002	0.431
CAV2_E33_R	0.002	0.763
CCKAR_E79_F	0.002	0.381
CCNA1_E7_F	0.002	0.662
CD40_E58_R	0.002	0.378
CD40_P372_R	0.002	0.307
CD81_P272_R	0.002	0.610
CD9_P504_F	0.002	0.527
CDH3_P87_R	0.002	0.706
CDKN1B_P1161_F	0.002	0.419
CFTR_P372_R	0.002	0.500
COL18A1_P494_R	0.002	0.725
CRIP1_P274_F	0.002	0.389
CRIP1_P874_R	0.002	0.449
CTNNA1_P382_R	0.002	0.675
CTNNB1_P757_F	0.002	0.319
DDIT3_P1313_R	0.002	0.427
DES_E228_R	0.002	0.551
DSC2_E90_F	0.002	0.707
ELK3_P514_F	0.002	0.553
EPHA2_P203_F	0.002	0.223
EPHB6_E342_F	0.002	0.213

LOX_P313_R	0.026	0.632	ERBB3_E331_F	0.002	0.714
MATK_P190_R	0.026	0.362	ERCC6_P698_R	0.002	0.444
MET_E333_F	0.026	0.684	ERN1_P809_R	0.002	0.226
MT1A_E13_R	0.026	0.534	ESR2_E66_F	0.002	0.733
TERT_P360_R	0.026	0.256	ESR2_P162_F	0.002	0.768
TGFB2_E226_R	0.026	0.838	EYA4_P508_F	0.002	0.478
THBS1_E207_R	0.026	0.782	EYA4_P794_F	0.002	0.361
TRIP6_E33_F	0.026	0.701	F2R_P839_F	0.002	0.349
ATP10A_P147_F	0.033	0.528	FES_E34_R	0.002	0.735
BCR_P346_F	0.033	0.273	FES_P223_R	0.002	0.700
IMPACT_P234_R	0.033	0.636	FGFR3_P1152_R	0.002	0.353
RUNX1T1_P103_F	0.033	0.497	FZD9_E458_F	0.002	0.697
AREG_P217_R	0.040	0.491	GFI1_P45_R	0.002	0.314
IGFBP1_P12_R	0.040	0.535	GLI3_P453_R	0.002	0.635
PLAUR_E123_F	0.040	0.391	GNMT_E126_F	0.002	0.739
RARRES1_P57_R	0.040	0.281	GNMT_P197_F	0.002	0.662
CEBPA_P1163_R	0.049	0.453	GP1BB_P278_R	0.002	0.212
COL1A2_E299_F	0.049	0.298	GRB7_E71_R	0.002	0.584
COL1A2_P407_R	0.049	0.245	HHIP_P578_R	0.002	0.312
COL1A2_P48_R	0.049	0.233	HRASLS_P353_R	0.002	0.382
GRB10_P260_F	0.049	0.614	ICA1_P61_F	0.002	0.398
Grade 2 Astrocytoma			ICA1_P72_R	0.002	0.717
CD81_P272_R	0.012	0.551	IGFBP2_P306_F	0.002	0.437
ERBB3_E331_F	0.012	0.634	IL17RB_E164_R	0.002	0.794
FZD9_E458_F	0.012	0.641	INSR_P1063_R	0.002	0.507
TNFRSF10A_P91_F	0.012	0.655	IRAK3_E130_F	0.002	0.409
COL18A1_P494_R	0.013	0.625	IRAK3_P185_F	0.002	0.530
MEST_P4_F	0.013	0.447	IRF5_E101_F	0.002	0.301
ESR2_P162_F	0.014	0.778	JAK3_E64_F	0.002	0.512
DSC2_E90_F	0.015	0.716	JAK3_P156_R	0.002	0.379
GFI1_P45_R	0.015	0.384	KIAA1804_P689_R	0.002	0.691
ERCC6_P698_R	0.017	0.203	KIT_P367_R	0.002	0.450
MEST_E150_F	0.017	0.401	KLK10_P268_R	0.002	0.466
MT1A_P600_F	0.017	0.246	LOX_P313_R	0.002	0.775
TGFB2_E226_R	0.017	0.703	LYN_P241_F	0.002	0.634
TJP1_P390_F	0.017	0.726	MAP3K1_E81_F	0.002	0.809
ALOX12_E85_R	0.018	0.607	MAP3K1_P7_F	0.002	0.745
CFTR_P372_R	0.018	0.496	MATK_P190_R	0.002	0.478
CRIP1_P274_F	0.018	0.412	MEST_E150_F	0.002	0.485
TNFRSF10A_P171_F	0.018	0.523	MEST_P4_F	0.002	0.468
DES_E228_R	0.019	0.329	MEST_P62_R	0.002	0.473
ERN1_P809_R	0.019	0.212	MET_E333_F	0.002	0.764

RARRES1_P426_R	0.019	0.488	MT1A_P600_F	0.002	0.272
TES_P182_F	0.019	0.478	NGFR_P355_F	0.002	0.261
CAV2_E33_R	0.019	0.385	NOTCH3_P198_R	0.002	0.684
CCNA1_E7_F	0.019	0.617	NTSR1_P318_F	0.002	0.292
FABP3_E113_F	0.019	0.448	PAX6_P1121_F	0.002	0.388
HHIP_P578_R	0.019	0.233	PGF_P320_F	0.002	0.586
JAK3_P156_R	0.019	0.385	PLAUR_E123_F	0.002	0.448
TNFRSF10D_E27_F	0.019	0.689	POMC_P400_R	0.002	0.397
TNFRSF10D_P70_F	0.019	0.201	PRKCDBP_E206_F	0.002	0.669
ZMYND10_P329_F	0.019	0.639	PYCARD_E87_F	0.002	0.658
DDIT3_P1313_R	0.021	0.445	RAB32_E314_R	0.002	0.645
EYA4_P794_F	0.021	0.315	RAN_P581_R	0.002	0.359
PRKCDBP_E206_F	0.021	0.611	RARRES1_P426_R	0.002	0.478
SH3BP2_P771_R	0.021	0.415	RASSF1_E116_F	0.002	0.786
ABCG2_P310_R	0.023	0.686	RASSF1_P244_F	0.002	0.460
CTNNA1_P382_R	0.023	0.609	RBP1_E158_F	0.002	0.792
ICA1_P72_R	0.023	0.527	RBP1_P150_F	0.002	0.668
MATK_P190_R	0.023	0.438	RBP1_P426_R	0.002	0.488
RAB32_E314_R	0.023	0.420	SCGB3A1_E55_R	0.002	0.752
BMP4_P199_R	0.025	0.206	SEPT9_P374_F	0.002	0.225
HIC_1_seq_48_S103_R	0.025	0.201	SYK_E372_F	0.002	0.227
IGF1_E394_F	0.025	0.254	TES_P182_F	0.002	0.358
LOX_P313_R	0.025	0.771	TGFB2_E226_R	0.002	0.774
MEST_P62_R	0.025	0.426	THBS1_E207_R	0.002	0.757
PTCH2_E173_F	0.025	0.244	TIMP3_seq_7_S38_F	0.002	0.622
RBP1_E158_F	0.025	0.736	TJP1_P390_F	0.002	0.853
RBP1_P150_F	0.025	0.635	TNFRSF10A_P171_F	0.002	0.542
RUNX1T1_P103_F	0.025	0.209	TNFRSF10A_P91_F	0.002	0.633
F2R_P839_F	0.026	0.377	TNFRSF10C_P7_F	0.002	0.441
GLI3_P453_R	0.026	0.508	TNFRSF10D_E27_F	0.002	0.711
LY6G6E_P45_R	0.026	0.296	TNFRSF10D_P70_F	0.002	0.229
MYCL1_P502_R	0.026	0.570	TRIP6_E33_F	0.002	0.673
THBS1_E207_R	0.026	0.719	WT1_P853_F	0.002	0.354
DNAJC15_P65_F	0.028	0.274	ZMYND10_E77_R	0.002	0.553
PGF_P320_F	0.028	0.467	ZMYND10_P329_F	0.002	0.664
SCGB3A1_E55_R	0.028	0.617	ZNF215_P129_R	0.002	0.688
ALOX12_P223_R	0.029	0.329	ZP3_E90_F	0.002	0.601
BMP4_P123_R	0.029	0.588	CPA4_E20_F	0.002	0.255
CD40_E58_R	0.029	0.372	FABP3_E113_F	0.002	0.409
GSTM1_P266_F	0.029	0.426	HIC_1_seq_48_S103_R	0.002	0.209
HRASLS_P353_R	0.029	0.310	IGFBP2_P353_R	0.002	0.213

NOTCH3_P198_R	0.029	0.647	ITPR2_P804_F	0.002	0.544
POMC_P400_R	0.029	0.353	MOS_E60_R	0.002	0.665
RAB32_P493_R	0.029	0.371	PODXL_P1341_R	0.002	0.375
ZMYND10_E77_R	0.029	0.446	RAB32_P493_R	0.002	0.436
CTNNB1_P757_F	0.032	0.221	TGFBI_P173_F	0.002	0.279
GRB7_P160_R	0.032	0.311	CDH3_E100_R	0.003	0.720
RBP1_P426_R	0.032	0.477	GRB10_P260_F	0.003	0.390
CCKAR_E79_F	0.035	0.335	GUCY2D_P48_R	0.003	0.389
EPHA2_P340_R	0.035	0.345	HTR1B_E232_R	0.003	0.418
ATP10A_P147_F	0.037	0.453	LY6G6E_P45_R	0.003	0.318
RAN_P581_R	0.037	0.361	MT1A_P49_R	0.003	0.276
EYA4_P508_F	0.040	0.468	TNFRSF10C_E109_F	0.003	0.292
JAK3_E64_F	0.040	0.556	CTSD_P726_F	0.003	0.321
KIAA1804_P689_R	0.040	0.552	DNAJC15_P65_F	0.003	0.335
MAP3K1_E81_F	0.040	0.769	MYCL1_P502_R	0.003	0.592
RARA_P176_R	0.040	0.503	PADI4_P1158_R	0.003	0.262
EPHA2_P203_F	0.044	0.295	SH3BP2_P771_R	0.003	0.375
GNMT_E126_F	0.044	0.660	TEK_E75_F	0.003	0.222
ITPR2_P804_F	0.044	0.486	TMEFF1_P626_R	0.003	0.288
MYLK_P469_R	0.044	0.249	IFNGR2_P377_R	0.003	0.200
PLAUR_E123_F	0.044	0.442	MMP14_P13_F	0.003	0.280
PYCARD_E87_F	0.044	0.646	SLC22A3_P528_F	0.003	0.231
RASSF1_E116_F	0.044	0.612	CASP6_P201_F	0.004	0.247
TNFRSF10C_P7_F	0.044	0.384	EIF2AK2_P313_F	0.004	0.722
AREG_P217_R	0.047	0.391	GFI1_E136_F	0.004	0.532
CRIP1_P874_R	0.047	0.415	HS3ST2_E145_R	0.004	0.305
IGFBP1_P12_R	0.047	0.515	PTCH2_E173_F	0.004	0.265
LYN_P241_F	0.047	0.410	AREG_P217_R	0.004	0.494
MAP3K1_P7_F	0.047	0.719	EPHA2_P340_R	0.004	0.289
RASSF1_P244_F	0.047	0.393	EVI1_P30_R	0.004	0.427
Ependymoma			FAS_P65_F	0.004	0.591
EVI2A_P94_R	0.009	0.237	HFE_E273_R	0.004	0.627
RASSF1_E116_F	0.013	0.574	MMP2_P303_R	0.004	0.383
ERN1_P809_R	0.017	0.210	MYLK_P469_R	0.004	0.461
IFNGR2_P377_R	0.017	0.208	RUNX1T1_P103_F	0.004	0.247
TDGF1_P428_R	0.017	0.314	STAT5A_E42_F	0.004	0.258
PTPRO_P371_F	0.026	0.259	FRZB_E186_R	0.005	0.719
RAB32_P493_R	0.029	0.345	GRB7_P160_R	0.005	0.396
SPP1_E140_R	0.029	0.215	HIC2_P498_F	0.005	0.306
FZD9_E458_F	0.032	0.201	IGFBP1_E48_R	0.005	0.490
KLK10_P268_R	0.032	0.349	IMPACT_P234_R	0.005	0.466
RASSF1_P244_F	0.032	0.406	PYCARD_P393_F	0.005	0.218

TES_P182_F	0.032	0.234
HOXA11_P698_F	0.041	0.252
NFKB1_P496_F	0.041	0.236
CTSD_P726_F	0.045	0.296

VAV2_P1182_F	0.005	0.399
AHR_P166_R	0.005	0.660
CDKN1C_P626_F	0.005	0.252
FAS_P322_R	0.005	0.371
HCK_P858_F	0.005	0.323
IGFBP7_P297_F	0.005	0.229
LAMC1_P808_F	0.005	0.363
COL1A2_P407_R	0.006	0.256
HOXA11_P698_F	0.006	0.499
IGF2R_P396_R	0.006	0.272
ZNF215_P71_R	0.006	0.349
CD81_P211_F	0.007	0.360
GSTM1_P266_F	0.007	0.247
MMP14_P208_R	0.007	0.374
PLSCR3_P751_R	0.007	0.218
RARRES1_P57_R	0.007	0.260
MMP2_P197_F	0.008	0.272
TNF_P158_F	0.008	0.235
IGF1_E394_F	0.009	0.310
MLH3_P25_F	0.009	0.250
MT1A_E13_R	0.009	0.489
SEMA3B_E96_F	0.009	0.237
ERG_E28_F	0.010	0.338
EVI1_E47_R	0.012	0.638
JUNB_P1149_R	0.012	0.361
GJB2_P931_R	0.013	0.288
IGFBP1_P12_R	0.013	0.414
TNFRSF1B_P167_F	0.013	0.385
WRN_P969_F	0.013	0.294
ACVR1C_P363_F	0.015	0.378
BCR_P346_F	0.015	0.267
COL1A2_P48_R	0.015	0.275
FGFR2_P460_R	0.015	0.363
ITGB4_P517_F	0.015	0.410
FGFR2_P266_R	0.017	0.375
TNFRSF1B_E5_F	0.017	0.201
E2F5_P516_R	0.019	0.303
FGF1_P357_R	0.019	0.238
HOXA5_E187_F	0.019	0.321
PTPN6_E171_R	0.019	0.238
IL6_P611_F	0.021	0.319
PTK2_P735_R	0.027	0.467
CALCA_E174_R	0.030	0.225

TFAP2C_P765_F	0.037	0.250
CREB1_P819_F	0.041	0.248
CRK_P721_F	0.046	0.266
MAPK12_P416_F	0.046	0.214
SH3BP2_E18_F	0.046	0.216
SHB_P691_R	0.046	0.223

* This column lists the Illumina GoldenGate methylation array annotation for CpGs where the gene name is listed first in all capital letters followed by an E for exon or P for promoter to indicate the location of the CpG relative to the transcription start site, the number indicates the distance of the CpG from the transcription start site, and F indicates forward strand and R indicates reverse strand.

Supplementary Table 8. Cellular pathways enriched among statistically significantly differentially methylated CpG loci in gliomas with an *IDH* mutation compared to gliomas without *IDH* mutation*.

Pathways enriched in *IDH* mutant gliomas	***P*†**
Hypermethylated	
Protein Kinase A Signaling	.05
Angiopoietin Signaling	.06
RAN Signaling	.10
Hypomethylated	
Methane Metabolism	.03
Stilbene, Coumarine and Lignin Biosynthesis	.03
Metabolism of Xenobiotics by Cytochrome P450	.03
PXR/RXR Activation	.04
Retinol Metabolism	.05
Phenylalanine Metabolism	.06
Starch and Sucrose Metabolism	.09
Pentose and Glucuronate Interconversions	.09
Androgen and Estrogen Metabolism	.10

* CpG loci with statistically significantly differential methylation ($Q<0.05$ and $|\Delta\beta|>0.2$) between *IDH* wild-type and *IDH* mutant gliomas were examined for cellular pathway enrichment with Ingenuity pathways analysis software. RAN=RAN, member RAS oncogene family; PXR=nuclear receptor subfamily 1, group I, member 2; RXR=retinoid X receptor, gamma.

† Two-sided Fisher's exact test *P* value for enrichment of genes whose CpG loci are represented in among those in the listed pathways.

Supplementary Table 9. Recursively partitioned mixture model (RPMM) methylation class membership and glioma tumor grade and histology*.

	IDH Mutation		Tumor Grade			Tumor histology†							
Methylation Class	**No**	**Yes**	**2**	**3**	**4**	**AS2**	**AS3**	**EP**	**GBM**	**GBM2**	**OA2**	**OA3**	**OD2**
L	1	52	43	5	5	14	4	0	3	2	13	1	16
RLLL	5	0	1	0	4	0	0	1	3	1	0	0	0
RLLR	0	5	0	0	5	0	0	0	1	4	0	0	0
RLR	12	0	0	0	12	0	0	0	12	0	0	0	0
RRLL	5	0	5	0	0	0	0	5	0	0	0	0	0
RRLR	5	0	5	0	0	0	0	5	0	0	0	0	0
RRRLL	2	0	2	0	0	1	0	1	0	0	0	0	0
RRRLR	4	0	4	0	0	4	0	0	0	0	0	0	0
RRRR	4	0	4	0	0	0	0	2	0	0	2	0	0
	$P = 3.0x10^{-16}$‡		**$P < 2.2x10^{-16}$§**										**$P < 2.2x10^{-16}$‖**

* Methylation classes from recursively partitioned mixture model (RPMM) of gliomas with *IDH* mutation data stratified by *IDH* mutation status, tumor grade, and grade-specific tumor histology, all statistical tests are two-sided.

† AS2=grade 2 Astrocytoma, AS3=grade 3 astrocytoma, EP=ependymoma, GBM=primary glioblastoma multiforme, GBM2=secondary glioblastoma multiforme, OA2=grade 2 oligoastrocytoma, OA3=grade 3 oligoastrocytoma, OD2=grade 2 oligodendroglioma. Tumors were previously reviewed by UCSF neuropathologists to assign histologic subtypes and grades according to the World Health Organization classification.

‡ Fisher's exact test *P* value for association between RPMM methylation class and *IDH* mutation status.

§ Fisher's exact test *P* value for association between RPMM methylation class and tumor grade.

‖ Fisher's exact test *P* value for association between RPMM methylation class and grade-specific tumor histology.

References

1. Yu J, Zhang H, Gu J, Lin S, Li J, Lu W, et al. Methylation profiles of thirty four promoter-CpG islands and concordant methylation behaviours of sixteen genes that may contribute to carcinogenesis of astrocytoma. BMC Cancer 2004;4:65.

2. Stone AR, Bobo W, Brat DJ, Devi NS, Van Meir EG, Vertino PM. Aberrant methylation and down-regulation of TMS1/ASC in human glioblastoma. Am J Pathol 2004;165(4):1151-61.

3. Maegawa S, Itaba N, Otsuka S, Kamitani H, Watanabe T, Tahimic CG, et al. Coordinate downregulation of a novel imprinted transcript ITUP1 with PEG3 in glioma cell lines. DNA Res 2004;11(1):37-49.

4. Dallol A, Krex D, Hesson L, Eng C, Maher ER, Latif F. Frequent epigenetic inactivation of the SLIT2 gene in gliomas. Oncogene 2003;22(29):4611-6.

5. Harden SV, Tokumaru Y, Westra WH, Goodman S, Ahrendt SA, Yang SC, et al. Gene promoter hypermethylation in tumors and lymph nodes of stage I lung cancer patients. Clin Cancer Res 2003;9(4):1370-5.

6. Wiencke JK, Aldape K, McMillan A, Wiemels J, Moghadassi M, Miike R, et al. Molecular features of adult glioma associated with patient race/ethnicity, age, and a polymorphism in O6-methylguanine-DNA-methyltransferase. Cancer Epidemiol Biomarkers Prev 2005;14(7):1774-83.

Chapter 3

A novel approach to the discovery of survival biomarkers in glioma using a joint analysis of DNA methylation and gene expression

Ashley A. Smith, Yen-Tsung Huang, Melissa Eliot, E. Andres Houseman,

Carmen J. Marsit, John K. Wiencke, and Karl T. Kelsey

In Review: Epigenetics; November 18, 2013

A novel approach to the discovery of survival biomarkers in glioblastoma using a joint analysis of DNA methylation and gene expression

Ashley A. Smith[1], Yen-Tsung Huang[2], Melissa Eliot[2], E. Andres Houseman[3], Carmen J. Marsit[4,5], John K. Wiencke[6], and Karl T. Kelsey[1, 2]

[1] Department of Pathology and Laboratory Medicine, Brown University, Providence, Rhode Island, 02912
[2] Department of Epidemiology, Brown University, Providence, Rhode Island, 02903
[3] Department of Public Health, Oregon State University, Corvallis, Oregon, 97331
[4] Department of Pharmacology and Toxicology, Geisel School of Medicine at Dartmouth, Hanover, New Hampshire, 03755
[5] Department of Community and Family Medicine and Section of Biostatistics and Epidemiology, Geisel School of Medicine at Dartmouth, Dartmouth, New Hampshire, 03755
[6] Department of Neurological Surgery, University of California San Francisco, San Francisco, California, 94143

Corresponding Author: Karl T. Kelsey,
Department of Pathology and Laboratory Medicine, Brown University,
Box G-E3, Providence, RI 02912.
Phone: 401-863-6420
Fax: 401-863-9008
E-mail: Karl_Kelsey@brown.edu

Running Title: Methylation and expression predict glioma survival

Conflict of interest: The authors have no potential conflicts of interest to report.

Funding: This work was funded by grants from the National Cancer Institute [CA100679, K.T.K.]; and the National Institutes of Health [CA126831, J.K.W]

Abbreviations: GBM: glioblastoma multiforme; CNV: copy number variation; G-CIMP: Glioma CpG Island Methylator Phenotype; AFT: accelerated failure time; iBag: integrative Bayesian analysis; BH: Benjamini-Hochberg; FDR: false discovery rate; DF: degree-of-freedom

Abstract

Glioblastoma multiforme (GBM) is the most aggressive of all brain tumors with a median survival under 1.5 years. Recently, epigenetic alterations have been found to play key roles in both glioma genesis and clinical outcome, demonstrating the need to integrate genetic and epigenetic data into predictive models. To enhance current models through discovery of novel predictive biomarkers, we employed a genome wide, agnostic strategy to specifically capture both expression-based (methylation-directed changes in gene expression) and alternative associations of DNA methylation with disease survival in glioma. Human GBM-associated DNA methylation, gene expression, *IDH1* mutation status, and survival data were obtained from The Cancer Genome Atlas. DNA methylation loci and expression probes were paired by gene, and their subsequent association with survival was determined by applying an accelerated failure time model to previously published alternative and expression-based association equations. Significant associations were seen in 27 unique methylation/expression pairs with expression-based, alternative, and combinatorial associations observed (10, 13, and 4 pairs, respectively). The majority of the DNA methylation loci that were predictive were located within CpG islands, and all but three of the locus pairs showed negative correlations, suggesting that for most loci, the methylation/expression pairs were inversely related, consistent with methylation-associated gene regulatory action. Our results indicate that changes in DNA methylation are associated with altered survival outcome through both coordinate changes in gene expression and alternative mechanisms. Furthermore, our approach offers an alternative method of biomarker discovery using a

priori gene pairing and precise targeting to identify novel sites for loci-specific therapeutic intervention.

Keywords: glioma, DNA methylation, gene expression, biomarker, mediation analysis

Introduction

Glioblastoma multiforme (GBM) is the most aggressive of all brain tumors, and accounts for approximately 70% of all malignant gliomas.[1] Despite current treatments, patients with GBMs have a median survival of only 12-15 months.[1] This disease is thought to result from the outgrowth of clonal populations that harbor a combination of somatic gene alterations that are likely complex.[1] Genetic alterations include dysregulation of many angiogenic and proliferative pathways including amplification of EGFR and overexpression of VEGF.[1] In addition, dysregulation in many members of the PI(3)K /Akt/RAS signaling pathway have also been implicated in the disease.[1] In 2006, Phillips et al used these genetic alterations, as well as copy number variations (CNV), to distinguish subclasses of GBM, which had prognostic implications.[2] These analyses were further supported by several studies that assessed known, prevalent mutations in GBMs (*EGFR*, *PTEN*, *IDH1*, *TP53*, and *NF1*), copy number alterations, and expression changes in an integrative approach in order to more precisely define of GBM subtypes important for survival prediction. These data and approaches strongly support the hypothesis that GBMs harbor a complex combination of somatic alterations that determine their phenotype. [3, 4]

Recently, Frattini et al (2013), used a novel statistical approach in an attempt to identify drivers of gliomagenesis through integration of somatic mutations and CNV.[5] They classified three types of GBM: 1) GBM having deletions at sites containing mutations, 2) GBM having amplifications at sites containing mutations, and 3) GBM with recurrent mutations and no alteration in CNV.[5] They also identified fusion products involving the EGFR-SEPT14 loci. Their integrative analysis further added to the genetic understanding

of GBM pathogenesis as well as marked specific targets for possible therapeutic intervention. [5]

Epigenetics (particularly DNA methylation) also plays an important role in gliomagenesis and glioma survival. Gene promoter DNA methylation has long been associated with gene silencing and research has now identified a role for methylation in selecting alternate transcripts and gene promoters, giving rise to somatic events that can impact disease survival. [6-10] [11, 12] Our group and others have reported an association between isocitrate dehydrogenase 1 and 2 (*IDH1/2)* mutations and a hypermethylator phenotype in gliomas that is associated with early age of onset and increased patient survival, specifically in lower grade gliomas and secondary GBM.[6, 15] Our data, which looked at a TCGA independent population, also demonstrated and association between TP53 and G-CIMP and a lack of association between EGFR and G-CIMP, and an overall increase in methylation genome-wide. [16]

DNA methylation does not act solely through the mediation of gene expression (the mechanism that we designate as an expression-based association). DNA methylation has also been found to associate with chromosomal instability, the induction of splice variants, alterations in enhancer regions, changes in microRNA binding regions and expression control regions, and mutations. These somatic changes (which we designate as an alternative association) could also greatly impact survival, but are much less well studied.[6-10]

These reports have highlighted the crosstalk between various types of carcinogenic somatic alterations and the need for a better understanding of the complex nature of the

pattern of somatic gene inactivation, involving genetic and epigenetic alterations that impact upon both the genesis of and survival from glioma. Although there has been a call for these integrative biomarkers that can sharpen predictive tools, most research has focused on the integration of genetic alterations (e.g. mutations) and their association with survival.[5, 18, 19] Here, we have made use of The Cancer Genome Atlas (TCGA) data sets to test our bioinformatics-based approach for identifying novel biomarkers of phenotypically important relationships among DNA methylation, gene expression, and survival in GBM.

Results

DNA methylation and gene expression are significantly associated in GBM samples

After removal of all *IDH1* mutant samples and replicates to prevent survival bias, the final phase 1 and phase 2 datasets contained n=73 and n=168 samples, respectively. Patient demographic data for all 241 GBM samples are presented in Table 1. Expression and methylation loci were paired by gene symbol for all 241 samples, resulting in a total of 66,202 unique methylation and expression pairs, which were used for the following analysis. In order to ensure functionality of methylation loci in the following analysis an initial screen was conducted to determine the association of methylation and expression with in the same gene. To identify the methylation loci that regulate gene expression level, a linear model, as specified in Equation 2 (see Materials and Methods), was performed using the combined phase 1 and phase 2 datasets (n=241). Pairs were chosen as significant if they had a q-value <0.05. Out of all 66,202 corresponding loci for both expression and methylation, 9821 were found to be significantly associated with each

other (84.3% negatively correlated, 15.7% positively correlated). Samples were then separated back into the original phase 1 (n=73) and phase 2 (n=168) sets for survival analysis.

DNA methylation and gene expression pairs are significantly associated with patient survival in GBM samples

To determine DNA methylation and gene expression pairs that are not only significantly associated with each other, but also significantly associated with survival, a Cox proportional hazards model was run on phase 1, phase 2, and pooled datasets. We used the Cox model to investigate the effect of gene expression, DNA methylation, and their interaction term on survival, adjusting for age, gender, and study. 'Study' was included in as a model variable as a precautionary measure due to the inherent difference in how the presence of *IDH1* mutation was determined for each of the two datasets. As previously mentioned, tumors with a G-CIMP phenotype or IDH mutation were removed from this analysis due to their association with increased survival in GBM patients. Analysis of the phase 1 data set (n=73) yielded 878 pairs (from the original 9821) that were significantly associated with survival ($p<0.05$). Those 878 pairs were re-run using the phase 2 data set (n=168) using the same model, which reveals 100 pairs with $p<0.05$. Finally, we assessed effects of the 100 pairs on overall survival using the pooled dataset (n=241) (Supplementary Material, Table S1). Pairs significantly correlated with survival were chosen based on the q-value (BH) of the pooled model (cutoff: $q<0.10$). A total of 36 unique methylation/expression pairs from 29 genes were significantly associated with

survival. Of these 36 unique pairs, CpG locus cg23134520 was found to contain a SNP (rs6032566) and was removed from further analysis. This yielded 35 unique methylation/expression pairs from 28 different genes, which were used for the final mediation analysis (Table 2).

Association of methylated loci with survival can be decomposed into i) those whose action is mediated through expression and ii) those whose association with survival is not mediated in this fashion.

We first estimated the association of DNA methylation with survival mediated through presumptive effect on gene expression (expression-based association) and then assessed the association not directly mediated through gene expression (alternative association). The expression-based and alternative associations of paired loci with survival were estimated for the top 35 unique methylation/expression pairs (chosen from the linear model and Cox proportional hazards model) by using an accelerated failure time (AFT) model (see Supplementary Material, Table S2). This yielded 10 unique methylation/expression pairs where expression-mediated methylation was associated with survival outcome (or significant expression-based associations) (Fig. 1A), 13 methylation/expression pairs where methylation did not work through expression of the same gene to effect survival (significant alternative association) (Fig. 1B), and 4 methylation/expression pairs where methylation exerted its effect on survival outcome directly and through gene expression (both significant alternative and expression-based associations) (Fig. 1C). Of the 27 significant methylation and expression pairs, 22 DNA

methylation loci were located within a CpG Island and, in general, pairs within the same gene had similar effects on survival (Fig. 1 A-C). In addition, all but three of the locus pairs (associated with *CACNB1*, *RFXANK*, and *RAB21*), had negative correlations, suggesting that the majority of the methylation/expression pairs were inversely related (see Supplementary Material, Fig. S2). Additionally, exon locations of methylation loci from significant pairs can be seen in supplementary material, Fig. S3

Discussion

The association of alterations in DNA methylation and gene expression in GBM with disease survival has been a major focus of recent studies, as it is apparent that outcome is not solely driven by somatic mutation. These previous studies generally identified loci whose methylation was inversely correlated with expression and examined that impact of those loci on patient outcome. Uniquely, in our study, we focused upon methylation and attempted to classify the effects of methylation on survival into those mediated by expression and those not mediated by expression, thereby expanding the potential biomarker pool.

In 2013, Wang et al used an integrative Bayesian analysis (iBAG) approach to analyze the association of DNA methylation with changes in gene expression and subsequently evaluated the association of changes in gene expression on GBM survival.[21] This linear approach was able to identify several genes with significant associations of gene expression modulated by methylation. Consistent with this data, several genes that we identified to be significantly modulated by DNA methylation, including *OSMR*, *STEAP1*,

and *GRB10,* were also reported by Wang et al in their findings. [21] However, methylation not only exerts its effects on survival through expression of its associated gene, but also can operate through a variety of other mechanisms, including chromosomal fragility/instability, splicing variants, enhancer regions, and dysregulation of microRNA. [6-10] Etcheverry et al (2010) investigated the impact of DNA methylation on gene expression and outcome in GBM.[22] Their analysis focused on the relationship between DNA methylation and gene expression and the association of methylation on survival. They identified 421 CpG sites that were significantly inversely correlated between methylation and expression, 291 of these CpG sites matched what we found to be correlated in our analysis. They also identified 13 genes, that appeared to have consistent differential methylation and expression (between GBM and control brain) but were negatively correlated, suggesting that the regulation of these genes may be epigenetically modulated.[22] However, Wang et al did not consider the joint effect of methylation and expression on outcome. In addition, *IDH1* mutant-associated samples were removed from our study to ensure that the final results would not reflect a bias toward the *IDH1* hypermethylator phenotype due to its association with increased survival.[6]

Our final model focuses not only on how methylation acts through expression to associate with survival; but also assesses how methylation can associate with survival directly or as a proxy for alternative mechanisms (Fig. 2). The final 27 significant methylation/expression pairs (contain genes associated with invasion, angiogenesis, and metabolism, and many have been previously linked to brain/glioma (Table 3). Of the 20 genes that contained the significant pairs, to our knowledge none are associated with common amplifications and deletions found in GBM.[23] Ten pairs (from seven genes) had

a significant expression-based association with survival, suggesting that DNA methylation in these genes affects survival outcome via gene expression of the associated gene. Interestingly, two genes contained multiple significant methylation/expression pairs. One of these genes, oncostatin M receptor (*OSMR*), contained two significant pairs, both with the same gene expression probe, but paired with different DNA methylation loci. The DNA methylation loci for these pairs fall in a CpG island within 550 bp of the transcription start site of the *OSMR* gene and the pairs showed a negative correlation, suggesting that methylation of these loci could inhibit gene expression. The locus pairs (cg03138091_A_24_P388860 and cg26475085_A_24_P388860) were associated with a significant expression-based association for each CpG. It is known that OSMR beta associates with Interleukin 31 Receptor alpha (*IL31RA*) to form the Interleukin 31 receptor (IL31) complex which activates signal transducer and activator of transcription 3 (*STAT3*).[24] Priester et al (2013) recently demonstrated that silencing of *STAT3* inhibits glioma single cell infiltration and tumor growth, suggesting that *STAT3* plays an important role in the invasiveness of gliomas.[25] If *OSMR* is silenced via DNA methylation of its promoter, this could lead to a decrease in *OSMR* gene expression and its association with *IL31RA*, inhibiting the subsequent activation of *STAT3*. Without activated *STAT3,* GBM growth and infiltration could be attenuated, potentially causing an increase in survival. This proposed mechanism supports the expression-based association of *OSMR* methylation on survival in the present study.

In addition to the 10 pairs with significant expression-based associations, there were also 14 methylation/expression pairs (in 12 genes) with significant alternative associations. This suggests that in these genes, DNA methylation is associated with survival either

directly or through mechanisms other than direct changes in gene expression. For instance, aquaporin 1 (*AQP1*) contained one methylation/expression pair, which is located in a CpG island within 300 bp of the transcription start site of the *AQP1* gene, and the pair showed a negative correlation, suggesting that methylation of this locus could inhibit gene expression. The major function of aquaporins (AQPs) is transportation of water across cell membranes, the disruption of which has been shown to disturb the blood-brain barrier and lead to cerebral edema.[26-28] *AQP1* and *AQP4* are most abundantly expressed in the nervous system, and though *AQP4* has been more heavily studied, the expression of both has been observed in GBM and found to correlate with malignancy, specifically with cytotoxic cerebral edema, angiogenesis, and migration/invasion.[26, 29, 30] Recently, it has been shown that both *AQP1* and *AQP4* are direct targets of microRNA 320a (miR-320a) and that increased miR-320a is associated with a reduction in *AQP1/4* expression.[31] Therefore, a possible mechanistic explanation for the alternative association we observe involves methylation of the microRNA target region on *AQP1* inhibiting the binding of miR-320a and ultimately allowing transcription of *AQP1*.

Interestingly, there were four methylation/expression pairs (three genes) that had both significant alternative and expression-based associations. Of interest is the gene growth factor receptor-bound protein 10 (*GRB10*), which contained two significant pairs, both with the same DNA methylation locus but paired with different gene expression probes. The DNA methylation locus for these pairs fall in a CpG island of the *GRB10* gene, and the pairs showed a negative correlation. The loci pairs (cg24302095_A_24_P235266 and cg24302095_A_24_P235268) have significant alternative associations that suggest that with a 5% increase in methylation, a decrease in survival may be observed; but the pairs

also have significant expression-based associations. *GRB10* is an imprinted gene that is differentially expressed from two promoters. In the brain, it is paternally expressed.[32] *GRB10* interacts with receptor tyrosine kinases and signaling molecules, most commonly insulin receptors and insulin-like growth factor receptors.[32, 33] In addition, monoallelic expression appears to be limited to fetal brain, skeletal muscle, and, most recently, placenta.[32, 33] Not only is expression of *GRB10* tissue specific, but it is also isoform specific.[32] Currently, 13 different splice variants of *GRB10* have been identified, with all but one being expressed in the brain.[33] Overexpression of some isoforms has been shown to suppress growth.[32] Yonghao et al (2011) found decreased expression of *GRB10* in many human tumor types, including gliomas, compared to corresponding normal tissue.[34] These tumor samples demonstrated a negative correlation between *GRB10* and *PTEN* expression. Furthermore, in a murine cell line, stabilization of Grb10 due to mTORC1-mediated phosphorylation resulted in inhibition of PI3K and ERK-MAPK pathways, suggesting a role for Grb10 as a tumor suppressor. [34] Conversely, Nord et al (2009), using a 32K bacterial artificial chromosomes array, found human *GRB10* to be a putative novel oncogene in glioblastoma.[35] Mechanistic differences might be attributed to inherent imprinting differences in *GRB10* between mice and humans. Nonetheless, DNA methylation of this CpG locus has the potential to cause alternative splice sites and may be responsible for the different isoforms of *GRB10*. Therefore, it is plausible that both the alternative and expression-based associations of this gene have a significant outcome on survival. Further potential mechanisms for genes that contained significant pairs can be found in Table 3.

There were several limitations to our work. First, we relied upon publically available data, which did not have complete mutation and survival data. We used a previously validated approach to control for this, but this remains a limitation [9 7]. To address the issue of missing survival data we used an accelerated failure time model to predict the survival time of censored values. In order to ensure functionality of methylation loci in our analysis, an initial screen was conducted, and only methylation and expression pairs that were significantly correlated within the same gene were used. Due to limited patient data, our study consisted only of primary GBM; however, promoter methylation of many GBM associated genes is more common in secondary GBM (ie. 11% promoter methylation for *MGMT* [36]), which may explain the lack of detection of previously described genes associated with promoter methylation in glioma. Additionally, there was one pediatric patient out of the 241 samples (age 10) that was not removed from the study prior to analyses.

Our approach focuses on methylations that regulate expression of the same gene, as mentioned above, and would miss methylation loci that do not regulate gene expression and are associated with survival through the alternative mechanism. To establish no association with gene expression, difficulties such as distinguishing null findings due to severe multiple comparisons from those with true biology will be an issue. Our approach circumvents the difficulty and is driven by biology: methylation that regulates gene expression is more likely to be functional and thus affects cancer survival.

Overall, our findings are consistent with the well-accepted concept that DNA methylation can associate with survival outcome via alterations in gene expression (e.g., *OSMR*). Our findings also suggest that methylation can associate with survival outcome through

mechanisms other than dysregulation of gene transcription, including disruption of microRNA function, as possible in the case of *AQP1*. Additionally, some methylation/expression pairs have both significant alternative and expression-based associations, suggesting that different tumors are using discrete mechanisms, yielding different survival outcomes, as described for the proposed alternative and expression-based associations of GRB10. It should be noted that promoter methylation of *MGMT*, which is frequent in low-grade and secondary GBM[11, 12],was observed to be significantly correlated with *MGMT* gene expression (data not shown), but was not observed in our final list of significant pairs. This may be attributable to the data quality (e.g. treatment data), or the relatively large number of subjects required to detect an interaction between treatment and methylation at this locus.

Importantly, our data suggest that this approach might profitably be applied to cancers other than GBM. Our method also brings to light pathways for future study as potential mechanisms in the pathogenesis of glioma. Though additional validation is needed, our work supports the concept that DNA methylation can function both through gene expression, and more directly or through alternative mechanisms, to modulate survival outcomes among glioblastoma patients.

Materials and Methods:

External Data Sets

Methylation, expression, and mutation data for glioblastoma multiforme (GBM) were downloaded from The Cancer Genome Atlas (TCGA) for two different sample sets.

Level 1 HumanMethylation27 (Illumina) DNA methylation data and level 2 AgilentG4502A_07_1 and 2 gene expression data were downloaded for all available GBM batches. GBM batches 1, 2, 3, and 10 were used as the phase 1 set and GBM batches 16, 20, 26, 38, and 62 were used as the phase 2 data set. Patient samples lacking covariate data were removed; samples were further restricted to patients diagnosed with glioblastoma who were alive 30 days after their date of diagnosis. Data sets were not combined in further analyses due to the fact that phase 2 data did not have definitive *IDH* mutation status. Since *IDH* mutations are associated with survival we were hesitant to combine the two datasets as mis-identification of IDH mutations could grossly affect findings.

Recursively partitioned mixture model to determine IDH1 mutation status

Patient survival, DNA methylation, gene expression, and *IDH1* mutation data (phase 1 set only), was obtained for primary glioblastoma multiforme (GBM) samples. It has been widely acknowledged that *IDH1* mutants are almost exclusively associated with a hypermethylater (G-CIMP) phenotype, and this phenotype is associated with increased survival in glioma.[10,11] Therefore, we wanted to remove IDH mutant samples from our study so results would not be biased due to increased survival associated with this mutation. Since *IDH* mutation data was not available for the phase 2 sample set, we employed a recursively partitioned mixture model (RPMM) as described by Houseman et al [20] and used in Christensen and Smith et al.[6] The RPMM successfully divided the phase 1 set into seven classes (see Supplementary Material, Fig. S1), and the samples in the top two most highly methylated classes, along with the samples having IDH mutations in the

phase 1 set, were removed (TCGA.14.1458, TCGA.16.1460, TCGA.19.1788, TCGA.14.1456, TCGA.28.1756, TCGA.14.4157, TCGA.32.4208).

Methylation Data

Methylation beta values were extracted from raw idat files using GenomeStudio software (Illumina), which calculates beta values using M/(M+U+100), where M is the methylated signal, U is the unmethylated signal, and 100 is an arbitrary offset. Replicates that did not correlate were removed (TCGA.06.0137, TCGA.06.0145). For methylation loci, all loci that contained a detection p-value > 0.05 for any sample were removed from further analysis. Since approximately 25% of the survival data is censored, censored survival times were estimated using an accelerated failure time (AFT) model based on the equation below.

Equation 1. $\log(T) = b_0 + b_1 Age + b_2 Gender + b_3 Study + b_4(Age * Study) + b_5(Study * Gender) + \mu\epsilon$

Where T follows a Weibull distribution [37] (μ is a scale parameter and ϵ follows an extreme value distribution). Next, methylation values were normalized for bead chip to control potential batch effect using the ComBat method [38] with adjustment of age, gender, survival, censored data, and survival-censored interaction.

Expression Data

TCGA expression and methylation subject identification numbers were matched; all non-matching samples were removed from the datasets. Replicates in expression samples were either averaged or chosen based on the closest mean and standard deviation to the

methylation distribution across all samples. The final data sets consist of a phase 1 dataset (n=73) and a phase 2 dataset (n=168) that contain complete data on overall survival, DNA methylation and gene expression with samples considered G-CIMP removed.

Final methylation/expression locus pairs

Methylation and expression loci were merged based on gene of origin. Annotation files for both platforms (HumanMethylation27 and AgilentG4502A_07_1 and 2) were downloaded from TCGA and matched by gene symbol, (using the manufacturer's annotation) yielding 66,202 methylation/expression pairs. It should be noted that there are usually several methylation loci and/or expression probes found within each gene, so while each pair is unique upon merging, an individual methylation or expression locus may be repeated among several pairs.

Statistical Analysis

To choose statistically significant methylation and expression pairs, expression was regressed on methylation in the pooled (n=241) dataset. The associated p-values were adjusted for false discovery rate (FDR) using the Benjamini-Hochberg (BH) procedure.[39] All methylation/expression pairs that had a q-value <0.05 were identified as being significantly associated with each other (n=9821 pairs).

To further siphon out statistically significant pairings, pairs were then assessed using a Cox proportional hazards model for the effect of expression, methylation, and their interaction on survival, controlling for age, gender, and study (when applicable). A three degree-of-freedom (DF) Chi-square test was performed to test for significance of

expression, methylation, and their cross-product interaction. The three-DF models were repeated for both phase 1 (n=73) and phase 2 datasets (n=168) separately and the pooled dataset (n=241). In order to reduce false positives, final statistically significant pairs were selected for having p-values <0.05 in both phase 1 and phase 2 datasets and q-values of <0.1 in the pooled dataset.

The associations of methylation and expression on survival were determined by a mediation analysis adopted from VanderWeele [37] using the following equations for the expression-based and alternative associations of methylation on survival:

Equation 2. $\boldsymbol{E}[E|M,c] = \beta_0 + \beta_1 M + \beta_2 c$

Equation 3. $\log(T) = \theta_0 + \theta_1 M + \theta_2 E + \theta_3 EM + \theta_4 c + \nu\varepsilon \log(T) = \theta_0 + \theta_1 M + \theta_2 E + \theta_3 EM + \theta_4 c + \nu\varepsilon$

Equation 4. $\Delta_{M\to E\to T} = (\theta_2\beta_1 + \theta_3\beta_1 m)(m - m^*)$

Equation 5. $\Delta_{M\to T} = \{\theta_1 + \theta_3(\beta_0 + \beta_1 m^* + \beta_2 c + \theta_2\sigma^2))(m - m^*) + 0.5\theta_3^2\sigma^2(m^2 - m^{*2})$,

where T is survival time, E is expression, M is methylation, c is study, σ^2 is the variance of the error term in Equation 2, ε is a random error in Equation 3 following the extreme value distribution, and ν is a scale parameter. For our purposes, m* is median methylation and (m-m*) is the change in methylation we are interested in observing. For example, we would set m-m* to 0.05 if we wanted to look at the change in survival for a 5% increase in methylation. Equation 2 represents the linear model for the association between expression and methylation, and Equation 3 represents the accelerated failure

time model with interaction between methylation and expression. β_0-β_2 are the regression parameters for the linear model, and θ_0-θ_4 are the regression parameters for the accelerated failure time model. We used a stepwise mediation analysis that considers the relationships between methylation and expression (Equation 2) and their joint effect on survival (Equation 3). In our case, an alternative association is the effect that methylation alone (or as a proxy for alternative mechanisms) has on survival, and expression-based association is the effect of methylation on survival mediated through gene expression. Equation 4 represents the expression-based association, and Equation 5 represents the alternative association of methylation on survival, [37] both of which can be estimated by fitting the models in Equations 2 and 3. We used bootstrap to find the variances and confidence intervals of the expression-based and alternative associations.

To determine directionality of the association of methylation on expression, we looked at the coefficient in the linear model regressing expression on methylation (Equation 2). A negative coefficient suggests that methylation and expression are inversely related (i.e., increased methylation is associated with decreased expression and vise versa). A positive correlation demonstrates that methylation and expression are directly related (i.e., increased methylation is associated with increased expression).

Acknowledgments

The authors would like to thank all individuals involved in the TCGA, particularly the patients who donated samples for use in this research.

References

1. Wen PY, Kesari S. Malignant gliomas in adults. N Engl J Med 2008; 359:492-507.

2. Phillips HS, Kharbanda S, Chen R, Forrest WF, Soriano RH, Wu TD, Misra A, Nigro JM, Colman H, Soroceanu L, et al. Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. Cancer Cell 2006; 9:157-73.

3. TCGA. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature 2008; 455:1061-8.

4. Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, Miller CR, Ding L, Golub T, Mesirov JP, et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell 2010; 17:98-110.

5. Frattini V, Trifonov V, Chan JM, Castano A, Lia M, Abate F, Keir ST, Ji AX, Zoppoli P, Niola F, et al. The integrated landscape of driver genomic alterations in glioblastoma. Nat Genet 2013.

6. Christensen BC, Smith AA, Zheng S, Koestler DC, Houseman EA, Marsit CJ, Wiemels JL, Nelson HH, Karagas MR, Wrensch MR, et al. DNA methylation, isocitrate dehydrogenase mutation, and survival in glioma. Journal of the National Cancer Institute 2011; 103:143-53.

7. Ehrlich M. DNA hypomethylation in cancer cells. Epigenomics 2009; 1:239-59.

8. Lopez-Serra P, Esteller M. DNA methylation-associated silencing of tumor-suppressor microRNAs in cancer. Oncogene 2012; 31:1609-22.

9. Wan J, Oliver VF, Zhu H, Zack DJ, Qian J, Merbs SL. Integrative analysis of tissue-specific methylation and alternative splicing identifies conserved transcription factor binding motifs. Nucleic Acids Res 2013.

10. Aran D, Hellman A. DNA methylation of transcriptional enhancers and cancer predisposition. Cell 2013; 154:11-3.

11. Hegi M, Diserens A, Gorlia T, Hamou M, de Tribolet N, Weller M, Kros J, Hainfellner J, Mason W, Mariani L, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. N Engl J Med 2005; 352:997-1003.

12. Esteller M, Garcia-Foncillas J, Andion E, Goodman SN, Hidalgo OF, Vanaclocha V, Baylin SB, Herman JG. Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. N Engl J Med 2000; 343:1350-4.

13. Killela PJ, Reitman ZJ, Jiao Y, Bettegowda C, Agrawal N, Diaz LA, Friedman AH, Friedman H, Gallia GL, Giovanella BC, et al. TERT promoter mutations occur frequently in gliomas and a subset of tumors derived from cells with low rates of self-renewal. Proc Natl Acad Sci U S A 2013; 110:6021-6.

14. Castelo-Branco P, Choufani S, Mack S, Gallagher D, Zhang C, Lipman T, Zhukova N, Walker EJ, Martin D, Merino D, et al. Methylation of the TERT promoter and risk stratification of childhood brain tumours: an integrative genomic and molecular study. Lancet Oncol 2013; 14:534-42.

15. Noushmehr H, Weisenberger D, Diefes K, Phillips H, Pujara K, Berman B, Pan F, Pelloski C, Sulman E, Bhat K, et al. Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. Cancer Cell 2010; 17:510-22.

16. Christensen BC, Smith AA, Zheng S, Koestler DC, Houseman EA, Marsit CJ, Wiemels JL, Nelson HH, Karagas MR, Wrensch MR, et al. DNA methylation, isocitrate dehydrogenase mutation, and survival in glioma. J Natl Cancer Inst 2011; 103:143-53.

17. Godard S, Getz G, Delorenzi M, Farmer P, Kobayashi H, Desbaillets I, Nozaki M, Diserens AC, Hamou MF, Dietrich PY, et al. Classification of human astrocytic gliomas on the basis of gene expression: a correlated group of genes with angiogenic activity emerges as a strong predictor of subtypes. Cancer Res 2003; 63:6613-25.

18. Rivenbark AG, Coleman WB. Dissecting the molecular mechanisms of cancer through bioinformatics-based experimental approaches. J Cell Biochem 2007; 101:1074-86.

19. Ferté C, Trister AD, Huang E, Bot BM, Guinney J, Commo F, Sieberts S, André F, Besse B, Soria JC, et al. Impact of bioinformatic procedures in the development and translation of high-throughput molecular classifiers in oncology. Clin Cancer Res 2013; 19:4315-25.

20. Houseman E, Christensen B, Yeh R, Marsit C, Karagas M, Wrensch M, Nelson H, Wiemels J, Zheng S, Wiencke J, 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 2008; 9:365.

21. Wang W, Baladandayuthapani V, Morris JS, Broom BM, Manyam G, Do KA. iBAG: integrative Bayesian analysis of high-dimensional multiplatform genomics data. Bioinformatics 2013; 29:149-59.

22. Etcheverry A, Aubry M, de Tayrac M, Vauleon E, Boniface R, Guenot F, Saikali S, Hamlat A, Riffaud L, Menei P, et al. DNA methylation in glioblastoma: impact on gene expression and clinical outcome. BMC Genomics 2010; 11:701.

23. Rao SK, Edwards J, Joshi AD, Siu IM, Riggins GJ. A survey of glioblastoma genomic amplifications and deletions. J Neurooncol 2010; 96:169-79.

24. Chattopadhyay S, Tracy E, Liang P, Robledo O, Rose-John S, Baumann H. Interleukin-31 and oncostatin-M mediate distinct signaling reactions and response patterns in lung epithelial cells. J Biol Chem 2007; 282:3014-26.

25. Priester M, Copanaki E, Vafaizadeh V, Hensel S, Bernreuther C, Glatzel M, Seifert V, Groner B, Kögel D, Weissenberger J. STAT3 silencing inhibits glioma single cell infiltration and tumor growth. Neuro Oncol 2013; 15:840-52.

26. Papadopoulos MC, Verkman AS. Aquaporin water channels in the nervous system. Nat Rev Neurosci 2013; 14:265-77.

27. Bonomini F, Francesca B, Rezzani R. Aquaporin and blood brain barrier. Curr Neuropharmacol 2010; 8:92-6.

28. Wolburg H, Noell S, Fallier-Becker P, Mack AF, Wolburg-Buchholz K. The disturbed blood-brain barrier in human glioblastoma. Mol Aspects Med 2012; 33:579-89.

29. El Hindy N, Bankfalvi A, Herring A, Adamzik M, Lambertz N, Zhu Y, Siffert W, Sure U, Sandalcioglu IE. Correlation of aquaporin-1 water channel protein

expression with tumor angiogenesis in human astrocytoma. Anticancer Res 2013; 33:609-13.

30. Saadoun S, Papadopoulos MC, Hara-Chikuma M, Verkman AS. Impairment of angiogenesis and cell migration by targeted aquaporin-1 gene disruption. Nature 2005; 434:786-92.

31. Sepramaniam S, Armugam A, Lim KY, Karolina DS, Swaminathan P, Tan JR, Jeyaseelan K. MicroRNA 320a functions as a novel endogenous modulator of aquaporins 1 and 4 as well as a potential therapeutic target in cerebral ischemia. J Biol Chem 2010; 285:29223-30.

32. Blagitko N, Mergenthaler S, Schulz U, Wollmann HA, Craigen W, Eggermann T, Ropers HH, Kalscheuer VM. Human GRB10 is imprinted and expressed from the paternal and maternal allele in a highly tissue- and isoform-specific fashion. Hum Mol Genet 2000; 9:1587-95.

33. Monk D, Arnaud P, Frost J, Hills FA, Stanier P, Feil R, Moore GE. Reciprocal imprinting of human GRB10 in placental trophoblast and brain: evolutionary conservation of reversed allelic expression. Hum Mol Genet 2009; 18:3066-74.

34. Yu Y, Yoon SO, Poulogiannis G, Yang Q, Ma XM, Villén J, Kubica N, Hoffman GR, Cantley LC, Gygi SP, et al. Phosphoproteomic analysis identifies Grb10 as an mTORC1 substrate that negatively regulates insulin signaling. Science 2011; 332:1322-6.

35. Nord H, Hartmann C, Andersson R, Menzel U, Pfeifer S, Piotrowski A, Bogdan A, Kloc W, Sandgren J, Olofsson T, et al. Characterization of novel and complex

genomic aberrations in glioblastoma using a 32K BAC array. Neuro Oncol 2009; 11:803-18.

36. Cecener G, Tunca B, Egeli U, Bekar A, Tezcan G, Erturk E, Bayram N, Tolunay S. The promoter hypermethylation status of GATA6, MGMT, and FHIT in glioblastoma. Cell Mol Neurobiol 2012; 32:237-44.

37. VanderWeele TJ. Causal mediation analysis with survival data. Epidemiology 2011; 22:582-5.

38. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 2007; 8:118-27.

39. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Jornal of the Royal Statistical Society Series B (Methodological) 1995; 57:289-300.

40. Ruan B, Pong K, Jow F, Bowlby M, Crozier RA, Liu D, Liang S, Chen Y, Mercado ML, Feng X, et al. Binding of rapamycin analogs to calcium channels and FKBP52 contributes to their neuroprotective activities. Proc Natl Acad Sci U S A 2008; 105:33-8.

41. Hanissian SH, Teng B, Akbar U, Janjetovic Z, Zhou Q, Duntsch C, Robertson JH. Regulation of myeloid leukemia factor-1 interacting protein (MLF1IP) expression in glioblastoma. Brain Res 2005; 1047:56-64.

42. Zagzag D, Salnikow K, Chiriboga L, Yee H, Lan L, Ali MA, Garcia R, Demaria S, Newcomb EW. Downregulation of major histocompatibility complex antigens in invading glioma cells: stealth invasion of the brain. Lab Invest 2005; 85:328-41.

43. Halestrap AP. The SLC16 gene family - structure, role and regulation in health and disease. Mol Aspects Med 2013; 34:337-49.

44. Miranda-Gonçalves V, Honavar M, Pinheiro C, Martinho O, Pires MM, Cordeiro M, Bebiano G, Costa P, Palmeirim I, Reis RM, et al. Monocarboxylate transporters (MCTs) in gliomas: expression and exploitation as therapeutic targets. Neuro Oncol 2013; 15:172-88.

45. Colen CB, Shen Y, Ghoddoussi F, Yu P, Francis TB, Koch BJ, Monterey MD, Galloway MP, Sloan AE, Mathupala SP. Metabolic targeting of lactate efflux by malignant glioma inhibits invasiveness and induces necrosis: an in vivo study. Neoplasia 2011; 13:620-32.

46. Atkinson GP, Nozell SE, Benveniste ET. NF-kappaB and STAT3 signaling in glioma: targets for future therapies. Expert Rev Neurother 2010; 10:575-86.

47. Smith SJ, Tilly H, Ward JH, Macarthur DC, Lowe J, Coyle B, Grundy RG. CD105 (Endoglin) exerts prognostic effects via its role in the microvascular niche of paediatric high grade glioma. Acta Neuropathol 2012; 124:99-110.

48. Akhtar S, McIntosh P, Bryan-Sisneros A, Barratt L, Robertson B, Dolly JO. A functional spliced-variant of beta 2 subunit of Kv1 channels in C6 glioma cells and reactive astrocytes from rat lesioned cerebellum. Biochemistry 1999; 38:16984-92.

49. Yang X, Zhang Y, Li S, Liu C, Jin Z, Wang Y, Ren F, Chang Z. Rab21 attenuates EGF-mediated MAPK signaling through enhancing EGFR internalization and degradation. Biochemical and biophysical research communications 2012; 421:651-7.

50. Totong R, Schell T, Lescroart F, Ryckebüsch L, Lin YF, Zygmunt T, Herwig L, Krudewig A, Gershoony D, Belting HG, et al. The novel transmembrane protein Tmem2 is essential for coordination of myocardial and endocardial morphogenesis. Development 2011; 138:4199-205.

51. Smith KA, Lagendijk AK, Courtney AD, Chen H, Paterson S, Hogan BM, Wicking C, Bakkers J. Transmembrane protein 2 (Tmem2) is required to regionally restrict atrioventricular canal boundary and endocardial cushion development. Development 2011; 138:4193-8.

52. Halestrap AP, Meredith D. The SLC16 gene family-from monocarboxylate transporters (MCTs) to aromatic amino acid transporters and beyond. Pflugers Arch 2004; 447:619-28.

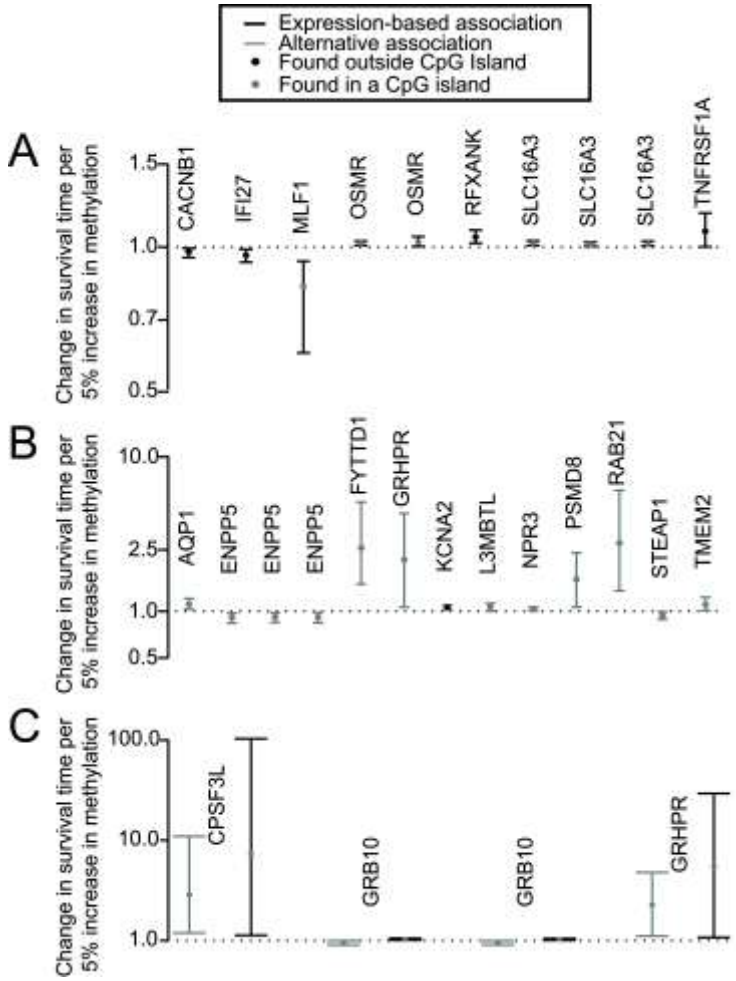


Figure 1. Significant expression-based and alternative associations of DNA methylation on gene expression and survival. The 35 unique DNA methylation/gene expression pairs were subjected to an Accelerated Failure Time (AFT) survival model and applied to alternative and expression-based equations (2-5 in methods). This yielded a total of 27 significant methylation/expression pairs, 10 had significant expression-based associations (A), 14 had significant alternative associations (B), and 4 had both significant expression-based and alternative associations (C). Grey lines indicate alternative associations, black lines indicate expression-based associations, grey circles indicate that the methylation locus for that gene pair was found in CpG Island, and black circles indicate that the methylation locus for that gene pair was not found in a CpG island. The y-axis indicates the change in survival time per 5% increase in methylation; therefore, effects that fall above the line are associated with an increase in survival and effects that fall below the line are associated with a decrease in survival.

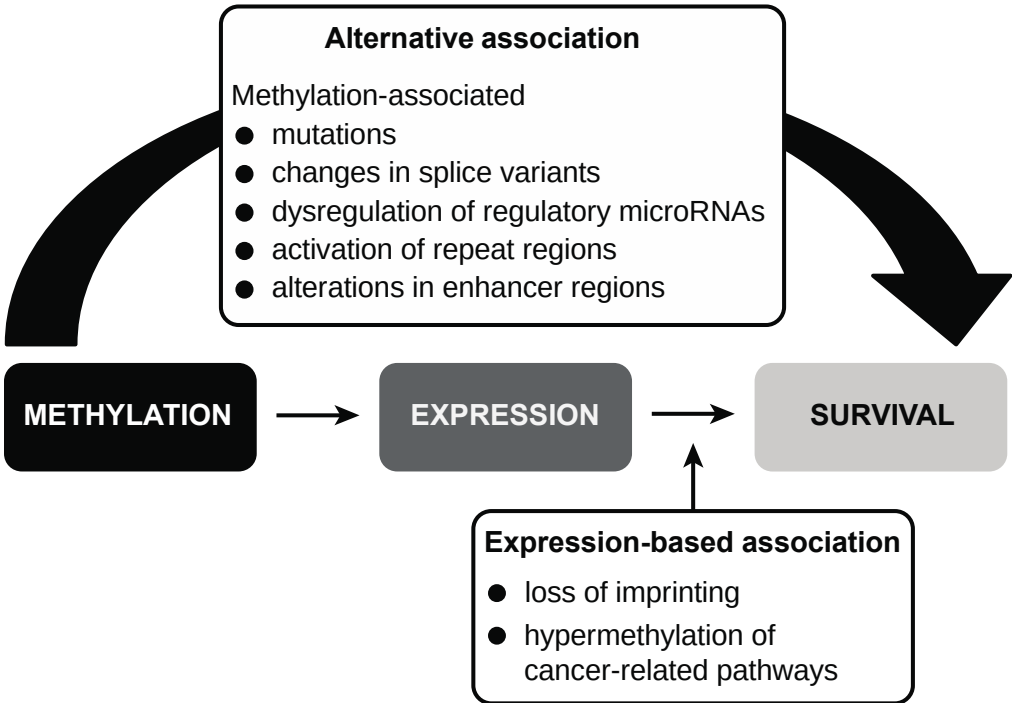


Figure 2. Model for mediation analysis. First a linear model adjusted for study was used to determine significantly correlated methylation/expression pairs. Next, a Cox proportional hazards model was used to find significant association between survival and expression, methylation, and their interaction term (adjusting for age, gender, and study). An accelerated failure time model was used to estimate the association between survival and expression, methylation, and their interaction term (adjusting for age, gender, and study), and a mediation analysis was performed to estimate the alternative and expression-based associations on glioma survival.

Table 1. Patient demographic and tumor* characteristics

	Data Sets		
Characteristic	**Training Set (n=73)**	**Testing Set (n=168)**	**Pooled (n=241)**
Age, years			
Median	56	60	59
Range	18-86	10-86	10-86
Sex, n(%)			
Female	31 (42.5)	69 (41.1)	100 (41.5)
Male	42 (57.5)	99 (58.9)	141(58.5)
****Survival (months)**			
Median	12.58	10.6	11.3
Range	1.37-60.0	1.08-60.0	1.08-60.0

***All tumor data obtained from The Cancer Genome Atlas (TCGA)**

****Censored at 60 months (5 years)**

Table. 2. Final 35 DNA methylation/gene expression pairs that are significantly associated with survival

TargetID_Reporter.REF	SYMBOL
cg17942096_A_23_P165180	RFXANK
cg18345635_A_23_P147345	SLC16A3
cg23943801_A_23_P128166	RAB21
cg27626424_A_23_P34449	LOR
cg05743054_A_23_P419947	MLF1
cg18345635_A_23_P158725	SLC16A3
cg18345635_A_23_P147349	SLC16A3
cg11558474_A_23_P94552	TMEM2
cg01781266_NM_018222_2_3793	PARVA
cg05845503_A_24_P141275	GRHPR
cg05845503_A_23_P60225	GRHPR
cg04551925_A_23_P19894	AQP1
cg00973286_A_23_P139715	TNFRSF1A
cg16773028_A_32_P40593	KCNA2
cg03138091_A_24_P388860	OSMR
cg26475085_A_24_P388860	OSMR
cg24812523_A_23_P14346	AKAP6
cg24302095_A_24_P235266	GRB10
cg24302095_A_24_P235268	GRB10
cg22166290_A_24_P402580	BCL11A
cg03764161_A_23_P203330	FAM111A
cg17726022_A_24_P261734	SLC38A1
cg17726022_A_23_P326510	SLC38A1
cg07663789_A_23_P327451	NPR3
cg04006554_A_23_P214244	ENPP5
cg04006554_A_23_P214240	ENPP5
cg04006554_NM_021572_2_2378	ENPP5
cg05788437_A_23_P80826	FYTTD1
cg06038049_A_23_P35029	CPSF3L
cg20089715_A_23_P405754	CACNB1
cg24219058_A_23_P310921	PCDH7
cg20091959_A_23_P210445	L3MBTL
cg18138552_A_23_P67464	PSMD8
cg20161089_A_24_P270460	IFI27
cg18320336_A_24_P406335	STEAP1

Table 3. Functions of significant genes and potential mechanisms in glioma

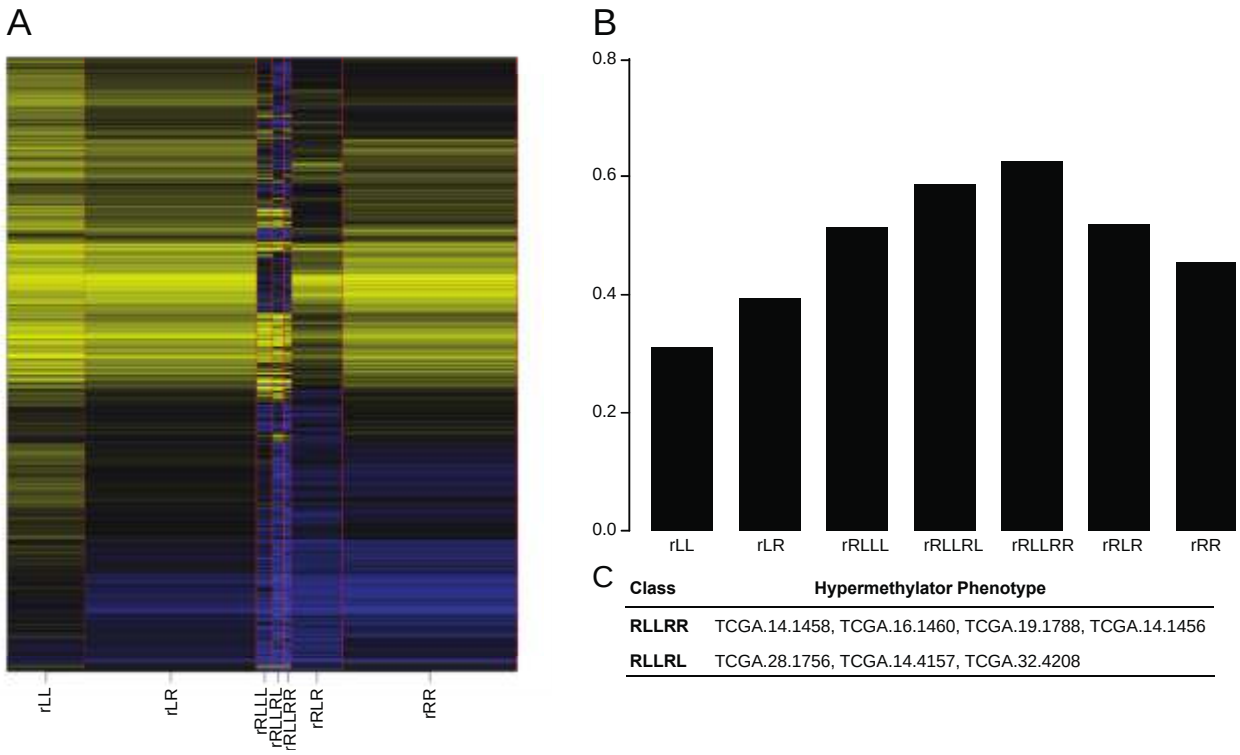
SYMBOL	NAME	FUNCTION (GeneCards®)	Potential expression-based role in glioma survival	Potential alternative role in glioma survival	Ref. #
CACNB1	calcium channel, voltage-dependent, beta 1 subunit	Involved in modulating G protein inhibition	It has been proposed that CACNB1 can protect neurons from Ca(2+)-induced cell death by modulating Ca(2+) channels; therefore, methylation-induced inhibition of CACNB1 could lead to loss of their neuroprotective activities (Ruan B et al 2008).		37
IFI27	interferon, alpha-inducible protein 27	Promotes cell death through mediation of IFN-alpha	**?**		
MLF1	myeloid leukemia factor 1	Oncoprotein that may be involved in lineage commitment	MLF1 and MLF1-like protein were found to co-localize and be over expressed in GBM tumors suggesting they play a role in glioma pathogenesis and survival. (Hanissian SH et al 2005). Dysregulation in expression of MLH1 via methylation could lead to differential survival outcomes.		38
OSMR	oncostatin M receptor	Member of the type 1 cytokine receptor family which heterodimerizes with interleukin 31, which as a complex can induce signaling events	Dysregulation of STAT3 activation via epigenetic induced silencing (Chattopadhyay et al 2007; Priester et al 2013).		22-23
RFXANK	regulatory factor X-associated ankyrin-containing protein	Forms a complex with regulatory factor X-associated protein and regulatory factor 5, which can then bind X box motif regions of some major histocompatibility (MHC)	Methylation-induced decrease in RFXANK could inhibit MHC class II activation, which is associated with glioma tumor invasion (Zagzag D et al, 2005).		39

		class II molecules, leading to activation			
SLC16A3	solute carrier family 16, member 3 (monocarboxylic acid transporter 4)	Part of a family of monocarboxylate transporters that catalyze lactic acid and pyruvate transport across plasma membranes	Differential SLC16A3 expression causing dysregulation of glycolytic metabolism via MCTs (Halestrap AP et al 2004 and 2013; Miranda-Gonçalves V et al 2013; Colen CB et al 2011).		40-42,49
TNFRSF1A	tumor necrosis factor receptor superfamily, member 1A	This receptor can activate NF-kappaB, mediate apoptosis, and function as a regulator of inflammation	Methylation induced changes in gene expression can dysregulate NF-kappaB pathway, which has been previously associated with glioma tumorigenesis and could be a possible therapeutic target of this disease (Atkinson GP et al 2010).		43
AQP1	aquaporin 1 (Colton blood group)	Molecular water channel protein		Methylation-mediated dysregulation of microRNA mir-320a binding region (Papadopoulos MC et al 2013; Bonomini F et al 2010; Wolburg H et al 2012; El Hindy Ner et al 2013; Saadoun S et al 2005; Sepramaniam S et al 2010).	24-29
ENPP5	ectonucleotide pyrophosphatase/ phosphodiesterase 5 (putative)	It may play a role in neuronal cell communication		Possible dysregulation in angiogenic signaling (Smith SJ et al 2012).	44

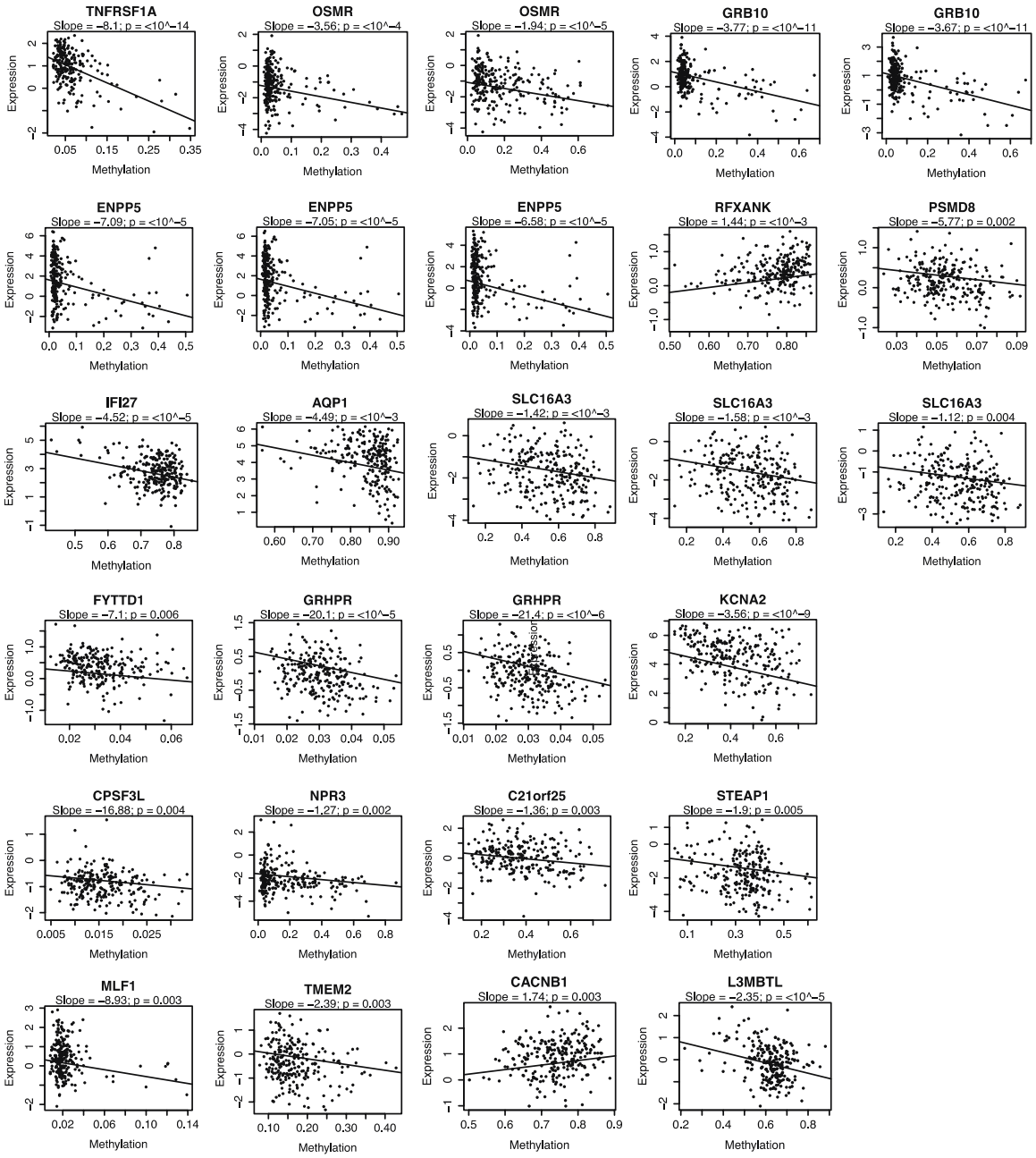
FYTTD1	forty-two-three domain containing 1	Required for mRNA export from the nucleus to the cytoplasm	**?**	
KCNA2	potassium voltage-gated channel, shaker-related subfamily, member 2	Voltage-gated ion channel that has a multitude of different functions ranging from regulation of neurotransmitter release, heart rate, insulin secretion, and neuronal excitability	Contains an alternatively spliced product in glioma cells which could contribute to the inactivation rate of the k(+) current Akhtar S et al 1999)	45
L3MBTL	Lethal (3) Malignant Brain Tumor-like 1 (Drosophila)	Polycomb group gene which functions to regulate gene activity via chromatin modifications	**?**	
NPR3	natriuretic peptide receptor C/guanylate cyclase C (atrionatriuretic peptide receptor C)	Natriuretic peptide receptor that regulates blood volume/pressure, pulmonary hypertension, cardiac function and some metabolic/growth processes	**?**	
PSMD8	proteasome (prosome, macropain) 26S subunit, non-ATPase, 8	Regulatory subunit of the 26S multicatalytic proteinase complex, which is involved in the ATP-dependent degradation of ubiquitinated proteins	**?**	
RAB21	RAB21, member RAS oncogene family	GTP-binding protein involved in integrin internalization and recycling	Rab21 expression has been found to attenuate Epidermal growth factor (EFG)-mediated mitogen-activated protein kinase (MAPK) by inducing EGF-receptor degradation (Yang X et al 2012).	46

STEAP1	six transmembrane epithelial antigen of the prostate 1	Found to be upregulated in multiple cancer cells lines and may be a potential metalloreductase		**?**	
TMEM2	transmembrane protein 2	Involved in coordination of myocardial and endocardial morphogenesis (Totong R et al 2011, Smith KA et al 2011)		**?**	47-48
CPSF3L	cleavage and polyadenylation specific factor 3-like	Catalytic subunit of the integrator complex, which mediates the 3-prime end processing of small nuclear RNAs U1 and U3	**?**	**?**	
GRB10	growth factor receptor-bound protein 10	Growth receptor binding protein that interacts with insulin and insulin-like growth-factor receptors	Methylation induced loss of imprinting (Blagitko N et al 2009; Monk D et al 2009; Yu Y et al 2011 ;Nord H et al 2009).	Methylation changes in splice variants, leading to expression of alternatively functioning isoforms (Blagitko N et al 2009; Monk D et al 2009; Yu Y et al 2011 ;Nord H et al 2009).	30-33
GRHPR	glyoxylate reductase/hydroxypyruvate reductase	Enzyme that plays a role in metabolism and reduces hydroxypyruvate to D-glycerate and glyoxylate to glycolate and oxidizes D-glycerate to hydroxypyruvate	**?**	**?**	

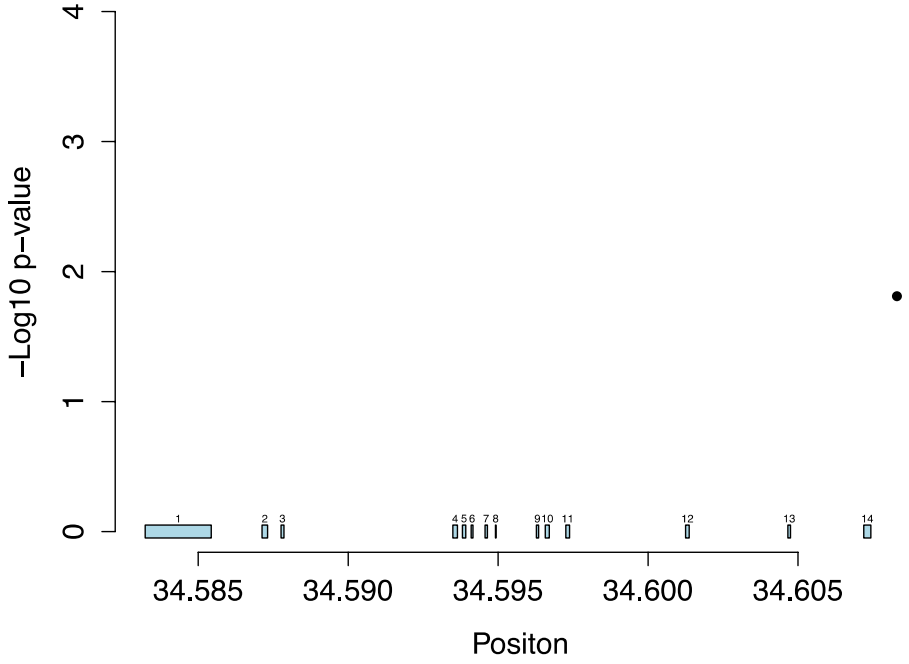
? - Possible mechanisms relating to glioma and significant expression-based or alternative association are unknown.

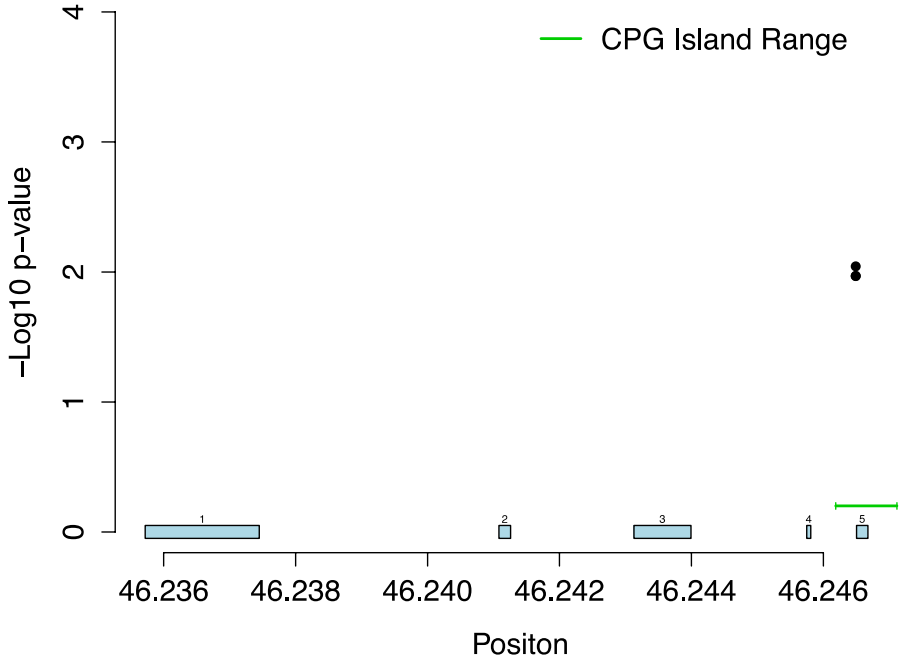


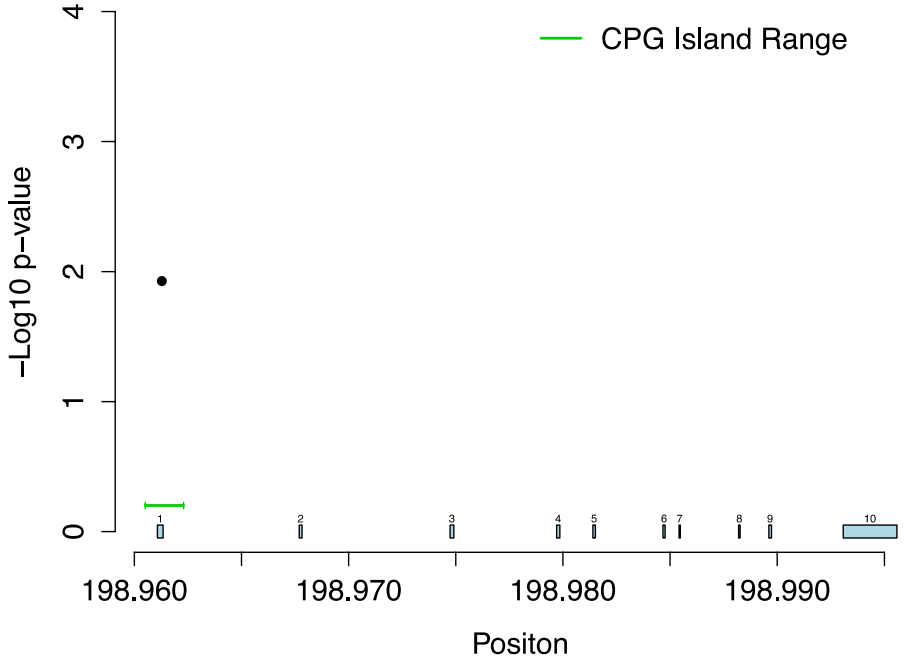
Supplementary Figure 1. Removal of IDH1 mutants. A Recursively partitioned mixture model (RPMM) was run on the top 5000 most variable CpG loci from the HumanMethylation27 (Illumina) DNA methylation array testing set (n=190) in order to determine hypermethylated classes, which have been previously associated with IDH1 mutation and increased survival in glioma(1A). The average methylation of each class was plotted (1B) and tumors in the top two most hypermethylated classes (RLLRR and RLLRL) were removed from the analysis as possible IDH1 mutant containing samples (1C).

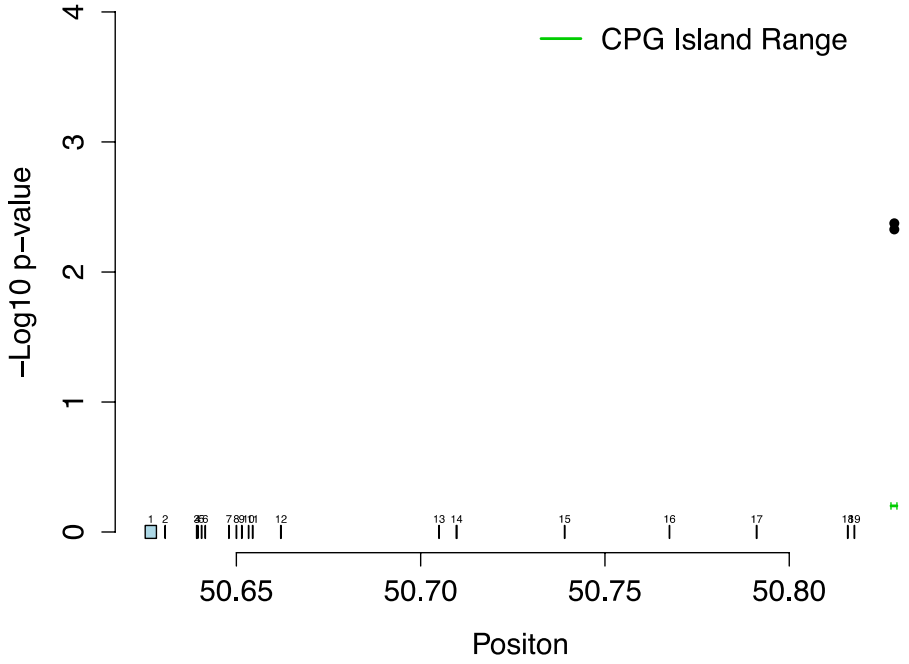


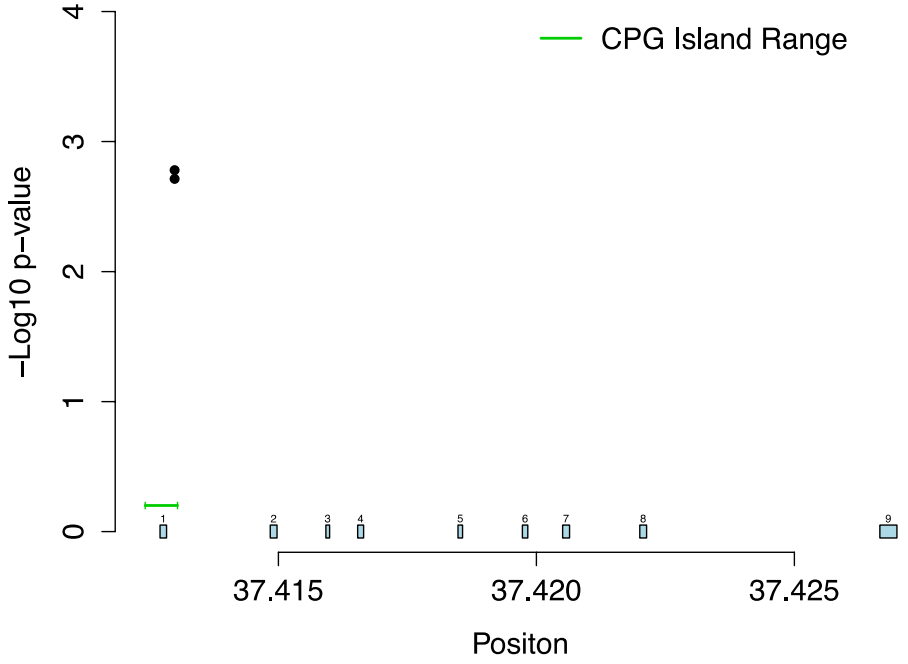
Supplementary Figure 2. Directionality of significant pairs. Gene expression values were plotted against DNA methylation values. A negative correlation demonstrates that methylation and expression are going in opposite directions (i.e. an increase in methylation is associated with a decrease in expression) and a positive correlation demonstrates that methylation and expression are going in the same directions (i.e. an increase in methylation is associated with an increase in expression).

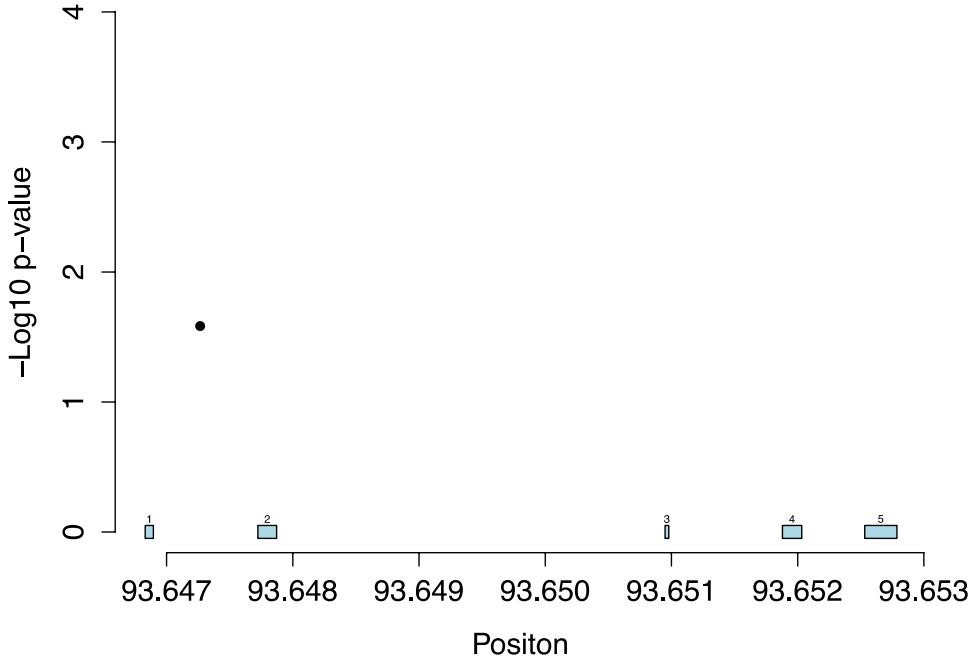

−Log10 p−value
0
1
2
3
4
1
2
3
4
5
6
7
8
9
10
11
12
13
14
34.585
34.590
34.595
34.600
34.605
Positon

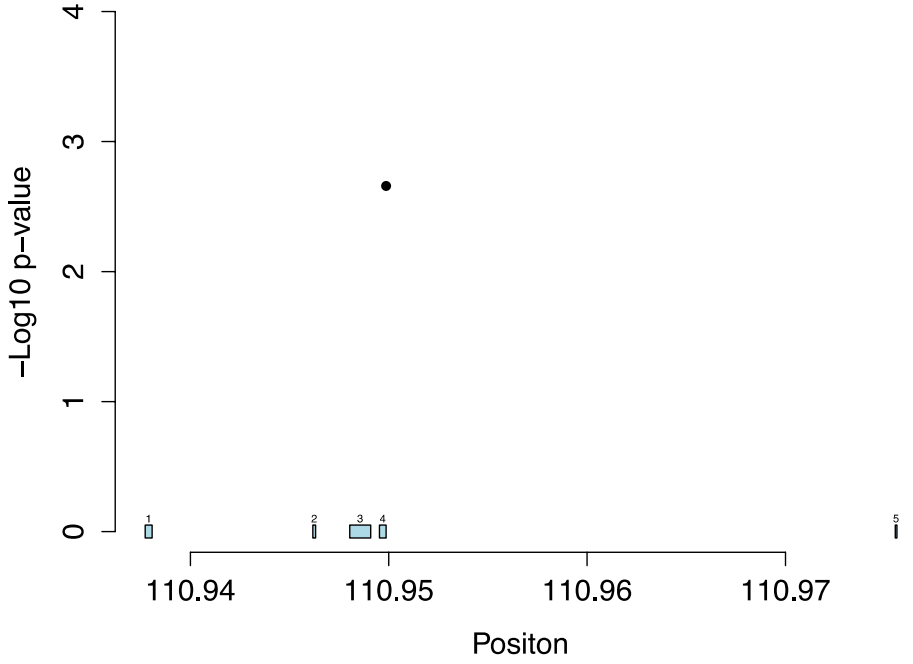

CPG Island Range
−Log10 p−value
0
1
2
3
4
1
2
3
4
5
46.236
46.238
46.240
46.242
46.244
46.246
Positon

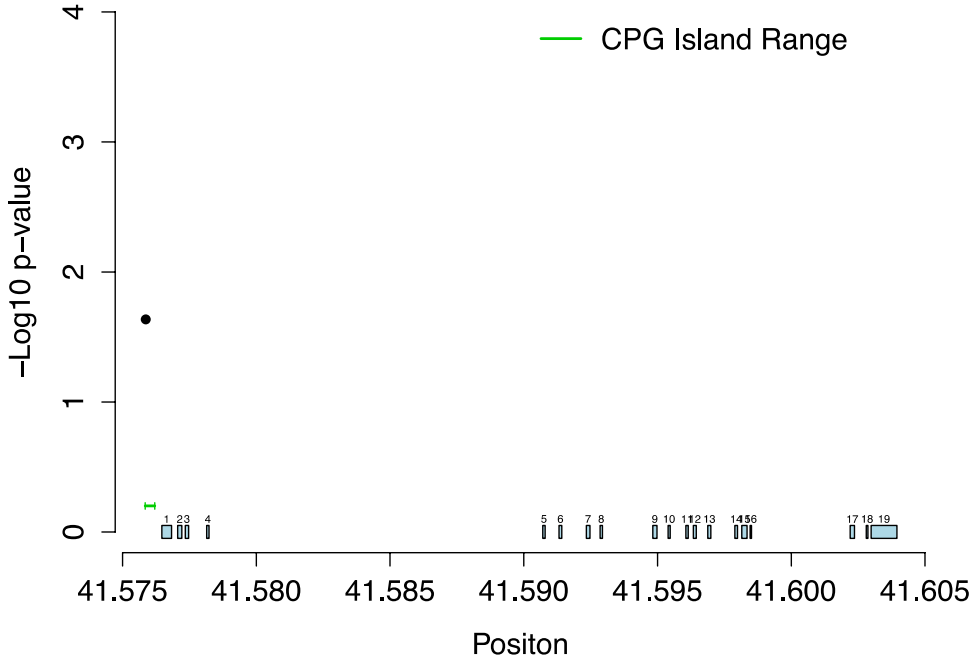

CPG Island Range
−Log10 p−value
0
1
2
3
4
198.960
198.970
198.980
198.990
Positon

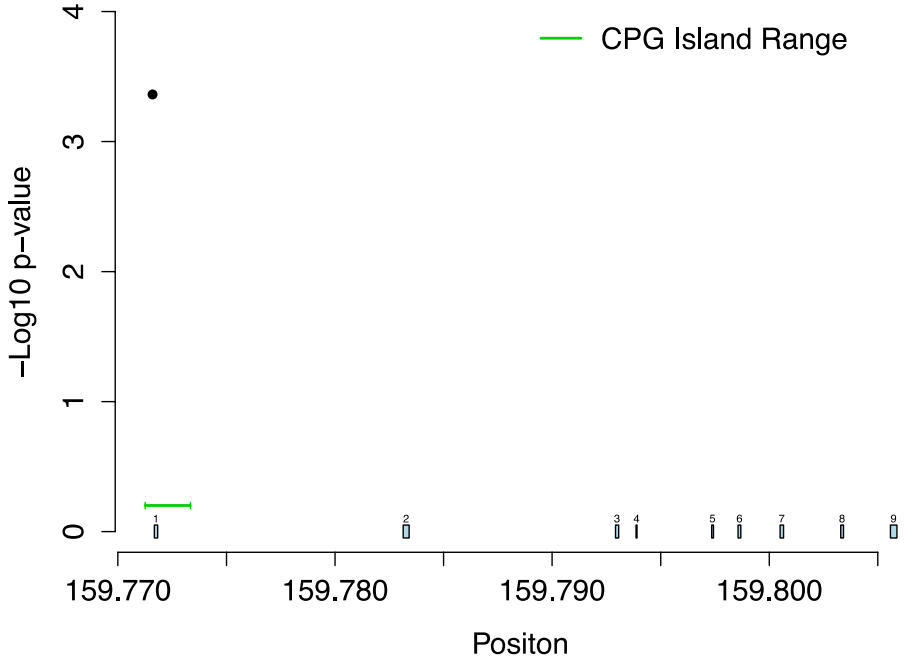

CPG Island Range
−Log10 p−value
0
1
2
3
4
50.65
50.70
50.75
50.80
Positon

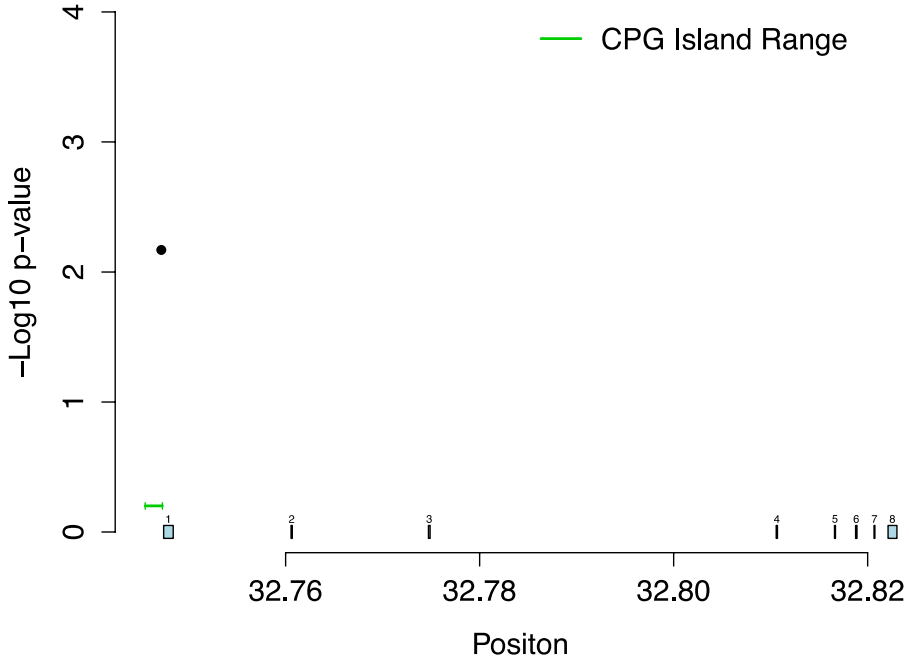

CPG Island Range
−Log10 p−value
0
1
2
3
4
1
2
3
4
5
6
7
8
9
37.415
37.420
37.425
Positon

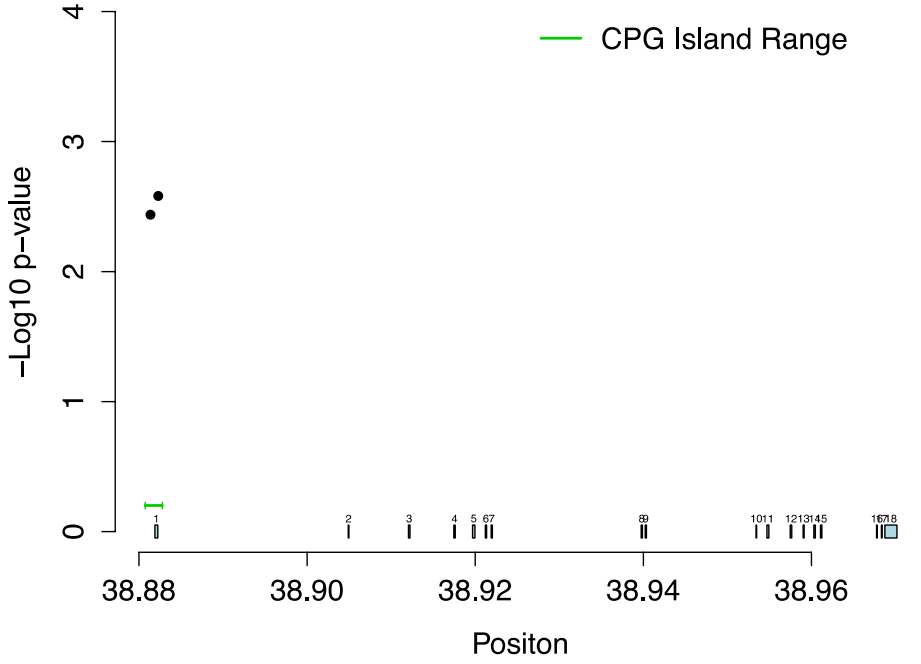

−Log10 p−value
0
1
2
3
4
1
2
3
4
5
93.647
93.648
93.649
93.650
93.651
93.652
93.653
Positon

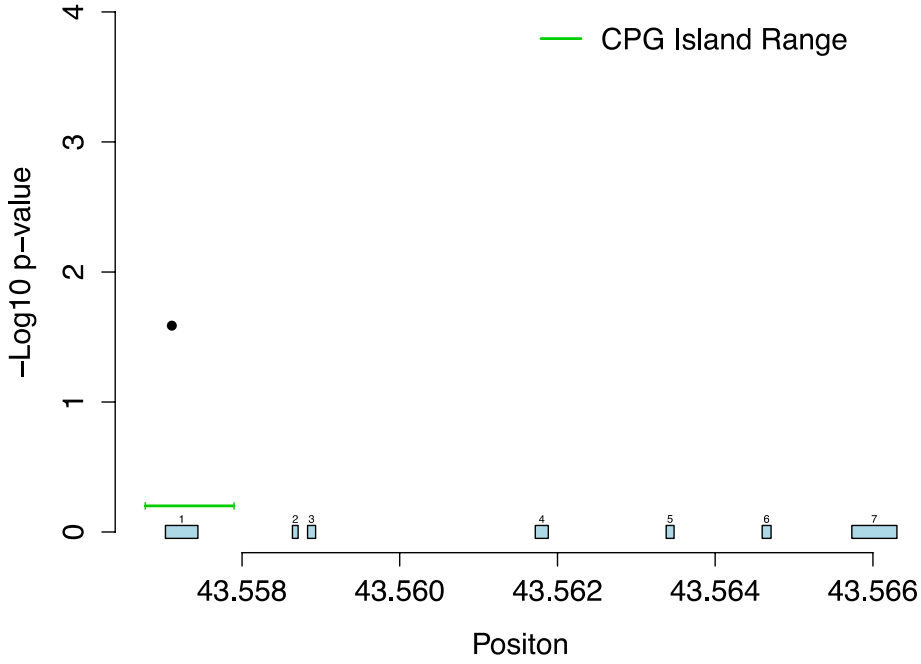

−Log10 p−value
0
1
2
3
4
110.94
110.95
110.96
110.97
Positon

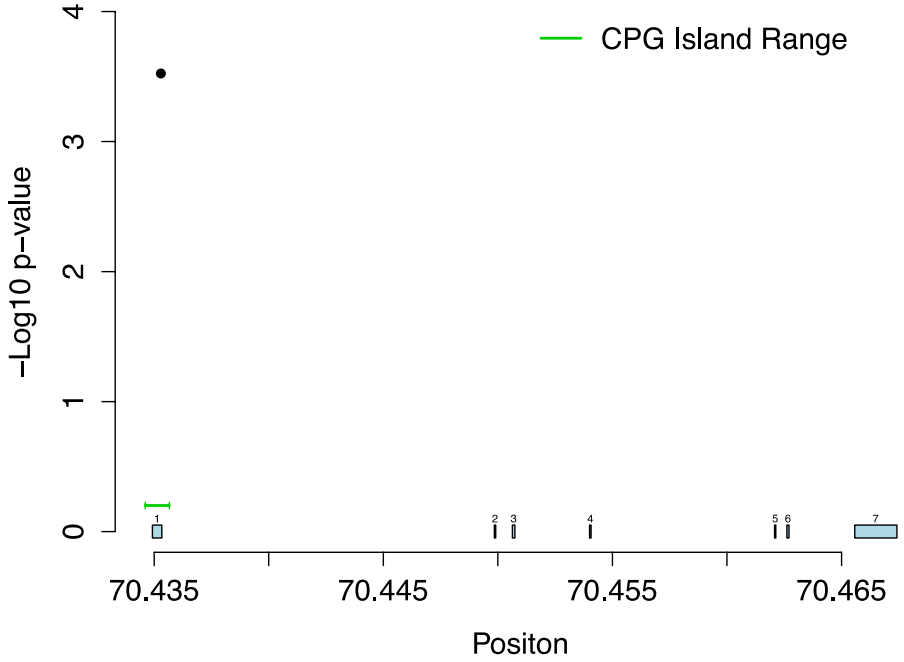

CPG Island Range
−Log10 p−value
0
1
2
3
4
41.575
41.580
41.585
41.590
41.595
41.600
41.605
Positon

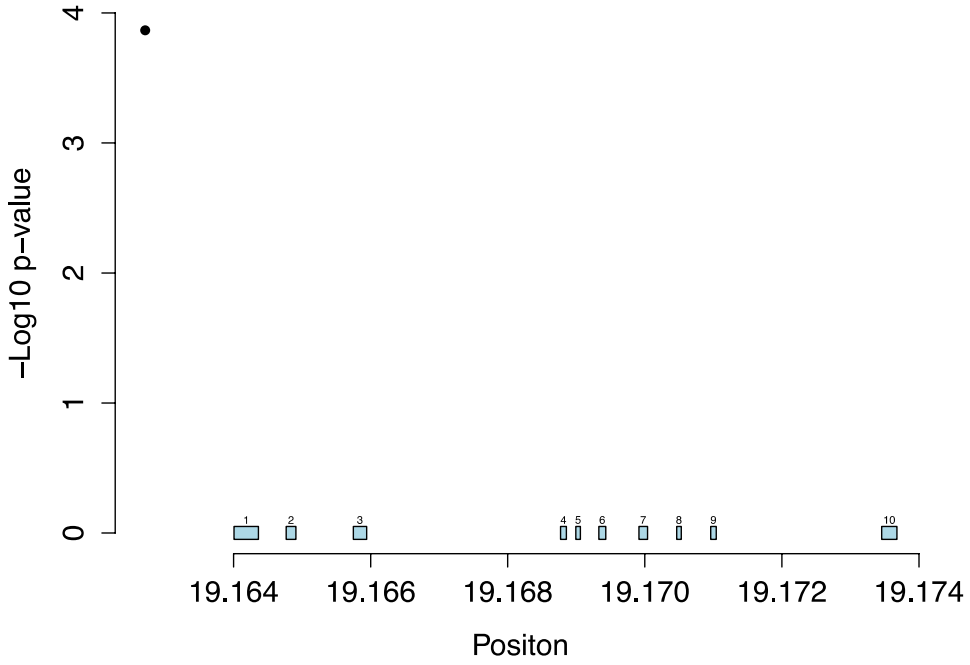

CPG Island Range
−Log10 p−value
0
1
2
3
4
159.770
159.780
159.790
159.800
Positon

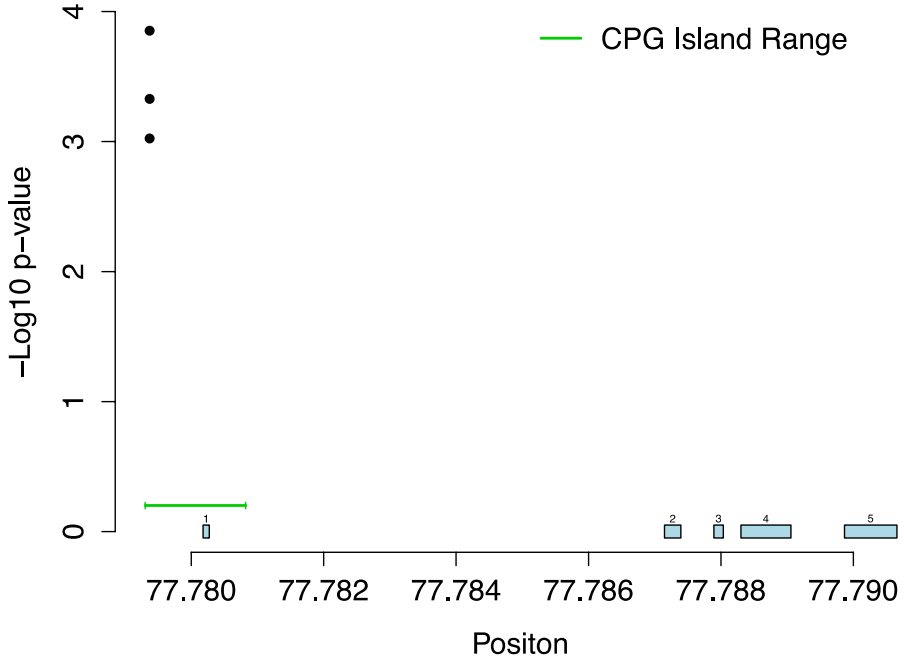

CPG Island Range
−Log10 p−value
0
1
2
3
4
32.76
32.78
32.80
32.82
Positon

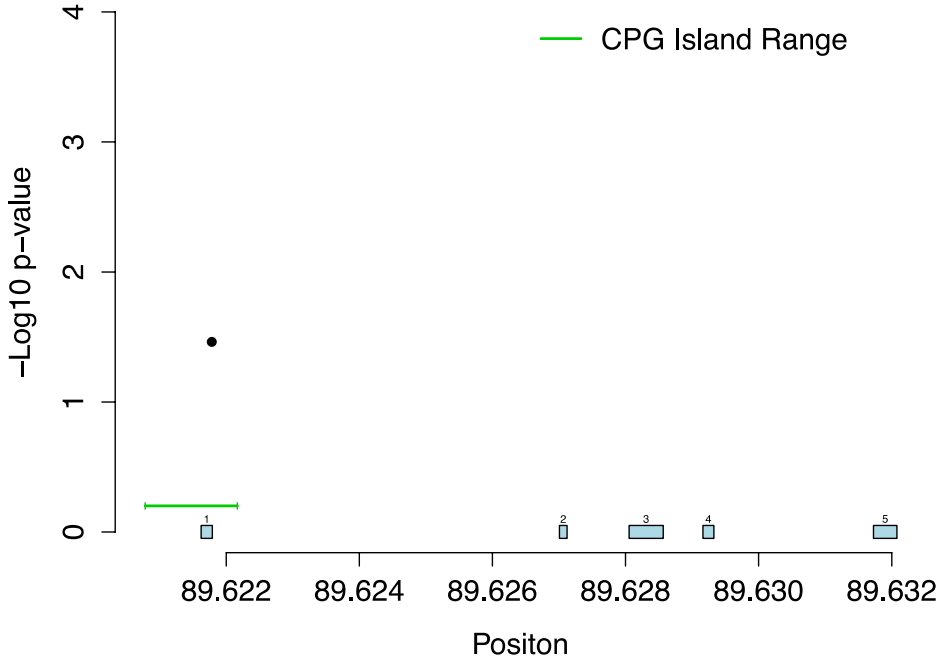

CPG Island Range
−Log10 p−value
0
1
2
3
4
38.88
38.90
38.92
38.94
38.96
Positon

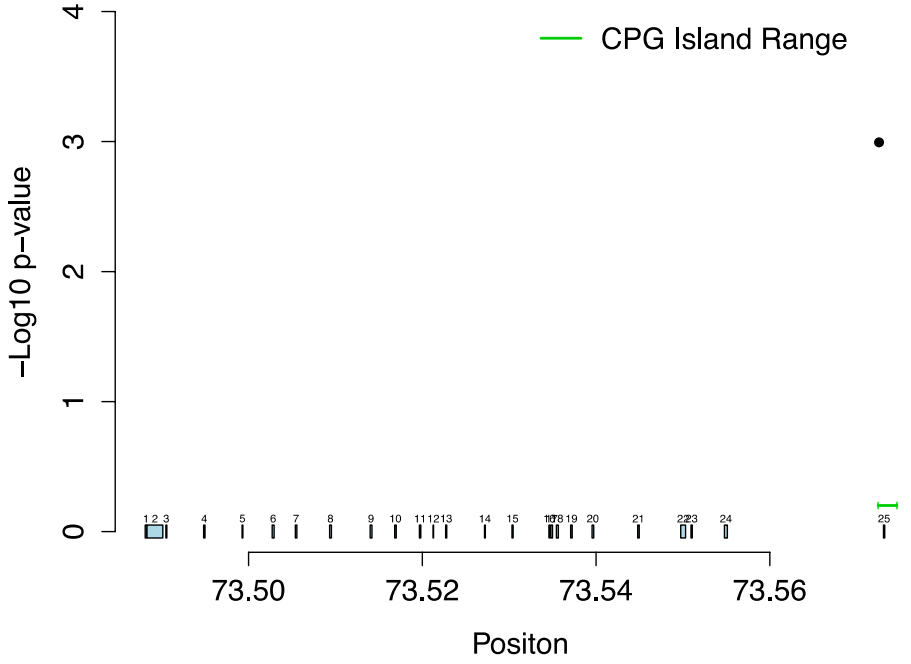

CPG Island Range
−Log10 p−value
0
1
2
3
4
43.558
43.560
43.562
43.564
43.566
Positon

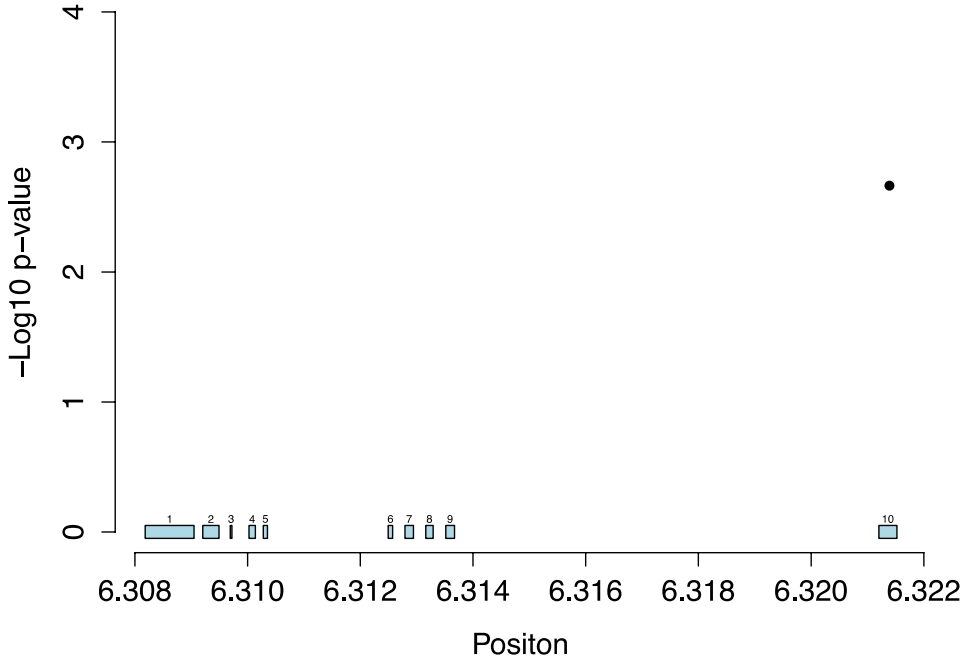

CPG Island Range
−Log10 p−value
0
1
2
3
4
1
2
3
4
5
6
7
70.435
70.445
70.455
70.465
Positon

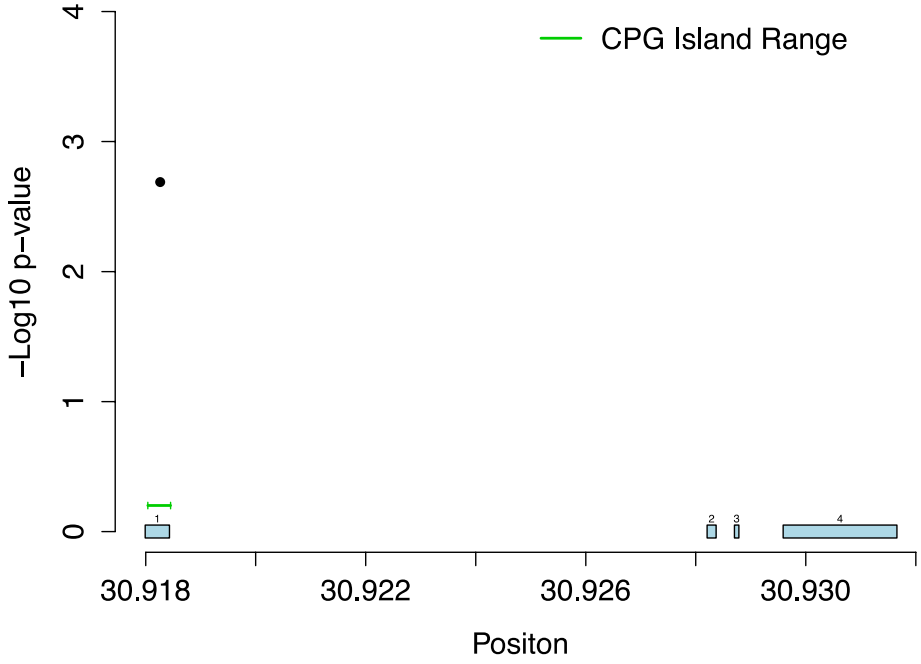

−Log10 p−value
0
1
2
3
4
1
2
3
4
5
6
7
8
9
10
19.164
19.166
19.168
19.170
19.172
19.174
Positon


CPG Island Range
−Log10 p−value
0
1
2
3
4
1
2
3
4
5
77.780
77.782
77.784
77.786
77.788
77.790
Positon


CPG Island Range
−Log10 p−value
0
1
2
3
4
1
2
3
4
5
89.622
89.624
89.626
89.628
89.630
89.632
Positon


CPG Island Range
−Log10 p−value
0
1
2
3
4
73.50
73.52
73.54
73.56
Positon


−Log10 p−value
0
1
2
3
4
6.308
6.310
6.312
6.314
6.316
6.318
6.320
6.322
Positon



Supplementary Fig. 3 Map of methylation loci locations from significant methylation/expression pairs. Each methylation locus obtained from significant methylation and expression pairs was plotted according to its genome location (Illumina annotation file). Exon locations were obtained from Genome Browser, with variants chosen based on the highest number of exons for which methylation loci fell within an exon as opposed to an intron. If the methylation locus was found within a CpG Island, that CpG island range was plotted in green (Illumina annotation file). CPSF3L is not plotted due to the fact that the accession number for this gene (NM_032179.1) was not available on genome browser.

Supplementary Table 1. DNA methylation/ expression pairs that are significantly associated with survival (q-value<0.1)

Methylation Loci_Expression Probe	Gene Symbol		p-value		q-value**
cg17942096_A_23_P165180	RFXANK	0.0007	0.0124	0.0001	0.007
cg18345635_A_23_P147345	SLC16A3	0.0123	0.0075	0.0001	0.007
cg23943801_A_23_P128166	RAB21	0.0133	0.0063	0.0003	0.0078
cg27626424_A_23_P34449	LOR	0.0238	0.0078	0.0004	0.0078
cg05743054_A_23_P419947	MLF1	0.0205	0.0001	0.0004	0.0078
cg18345635_A_23_P158725	SLC16A3	0.0082	0.0336	0.0005	0.0078
cg18345635_A_23_P147349	SLC16A3	0.0091	0.0179	0.0009	0.0127
cg11558474_A_23_P94552	TMEM2	0.0192	0.0171	0.001	0.0127
cg01781266_NM_018222_2_3793	PARVA	0.0204	0.0052	0.0015	0.0157
cg05845503_A_24_P141275	GRHPR	0.0399	0.0191	0.0017	0.0157
cg05845503_A_23_P60225	GRHPR	0.0305	0.0220	0.0019	0.0157
cg04551925_A_23_P19894	AQP1	0.0486	0.0302	0.002	0.0157
cg00973286_A_23_P139715	TNFRSF1A	0.0451	0.0231	0.0022	0.0157
cg16773028_A_32_P40593	KCNA2	0.0474	0.0032	0.0022	0.0157
cg03138091_A_24_P388860	OSMR	0.0070	0.0261	0.0026	0.0174
cg26475085_A_24_P388860	OSMR	0.0227	0.0104	0.0037	0.0228
cg24812523_A_23_P14346	AKAP6	0.0049	0.0239	0.0039	0.0229
cg24302095_A_24_P235266	GRB10	0.0244	0.0006	0.0042	0.0234
cg24302095_A_24_P235268	GRB10	0.0238	0.0005	0.0047	0.0244
cg22166290_A_24_P402580	BCL11A	0.0179	0.0449	0.0053	0.0244
cg03764161_A_23_P203330	FAM111A	0.0425	0.0082	0.0053	0.0244
cg17726022_A_24_P261734	SLC38A1	0.0346	0.0388	0.0054	0.0244
cg17726022_A_23_P326510	SLC38A1	0.0389	0.0335	0.0063	0.0273
cg07663789_A_23_P327451	NPR3	0.0217	0.0261	0.0068	0.0282
cg04006554_A_23_P214244	ENPP5	0.0113	0.0140	0.0091	0.0362
cg04006554_A_23_P214240	ENPP5	0.0068	0.0107	0.0107	0.0398
cg04006554_NM_021572_2_2378	ENPP5	0.0071	0.0115	0.0107	0.0398
cg05788437_A_23_P80826	FYTTD1	0.0232	0.0000	0.0118	0.0422
cg06038049_A_23_P35029	CPSF3L	0.0031	0.0322	0.0154	0.0515
cg20089715_A_23_P405754	CACNB1	0.0044	0.0123	0.0154	0.0515
cg24219058_A_23_P310921	PCDH7	0.0078	0.0085	0.0224	0.0723
cg20091959_A_23_P210445	L3MBTL	0.0286	0.0136	0.0232	0.0725
cg18138552_A_23_P67464	PSMD8	0.0310	0.0205	0.0259	0.0758
cg20161089_A_24_P270460	IFI27	0.0462	0.0093	0.0261	0.0758
cg23134520_A_23_P143218	ACOT8*	0.0463	0.0374	0.0265	0.0758
cg18320336_A_24_P406335	STEAP1	0.0374	0.0277	0.0345	0.0958

ACOT8 was removed from further analysis as SNP rs6032566 lies in the CpG locus of interest (cg23144520)q-value based on n=878**

Supplementary Table 2. Expression-based and alternative associations of DNA methylation on gene expression and survival.

Methylation Loci_Expression Probe	SYMBOL	Alternative Association	Confidence Interval (lower bound)	Confidence Interval (upper bound)	Expression-Based Association	Confidence Interval (lower bound)	Confidence Interval (upper bound)
cg24812523_A_23_P14346	AKAP6	0.999	0.96	1.04	1.007	0.994	1.018
cg04551925_A_23_P19894	**AQP1**	**1.111**	**1.03**	**1.21**	0.982	0.952	1.012
cg22166290_A_24_P402580	BCL11A	0.998	0.92	1.11	0.989	0.964	1.011
cg20089715_A_23_P405754	**CACNB1**	1.029	0.97	1.10	**0.977**	**0.950**	**0.994**
cg06038049_A_23_P35029	**CPSF3L**	**2.863**	**1.19**	**10.97**	**7.237**	**1.129**	**103.291**
cg04006554_A_23_P214240	**ENPP5**	**0.915**	**0.84**	**0.97**	0.997	0.980	1.015
cg04006554_A_23_P214244	**ENPP5**	**0.920**	**0.85**	**0.98**	0.994	0.977	1.011
cg04006554_NM_021572_2_2378	**ENPP5**	**0.917**	**0.84**	**0.98**	0.997	0.979	1.014
cg03764161_A_23_P203330	FAM111A	1.172	0.70	1.80	0.970	0.717	1.441
cg05788437_A_23_P80826	**FYTTD1**	**2.577**	**1.50**	**5.09**	0.768	0.406	1.287
cg24302095_A_24_P235266	**GRB10**	**0.961**	**0.90**	**1.00**	**1.029**	**1.012**	**1.053**
cg24302095_A_24_P235268	**GRB10**	**0.958**	**0.90**	**0.99**	**1.028**	**1.012**	**1.049**
cg05845503_A_23_P60225	**GRHPR**	**2.156**	**1.06**	**4.30**	4.370	0.953	21.125
cg05845503_A_24_P141275	**GRHPR**	**2.277**	**1.11**	**4.79**	**5.474**	**1.074**	**29.351**
cg20161089_A_24_P270460	**IFI27**	0.941	0.85	1.05	**0.961**	**0.928**	**0.989**
cg16773028_A_32_P40593	**KCNA2**	**1.061**	**1.03**	**1.10**	0.996	0.980	1.010
cg20091959_A_23_P210445	**L3MBTL**	**1.064**	**1.00**	**1.12**	1.013	0.997	1.034
cg27626424_A_23_P34449	LOR	1.036	1.00	1.07	1.001	0.994	1.009
cg05743054_A_23_P419947	**MLF1**	1.124	0.73	2.38	**0.825**	**0.596**	**0.933**
cg07663789_A_23_P327451	**NPR3**	**1.033**	**1.00**	**1.07**	0.998	0.991	1.003
cg03138091_A_24_P388860	**OSMR**	0.984	0.95	1.02	**1.018**	**1.008**	**1.030**
cg26475085_A_24_P388860	**OSMR**	1.041	0.95	1.15	**1.025**	**1.003**	**1.052**
cg01781266_NM_018222_2_3793	PARVA	1.059	0.95	1.15	0.990	0.963	1.017
cg24219058_A_23_P310921	PCDH7	1.043	0.78	1.76	0.939	0.810	1.005
cg18138552_A_23_P67464	**PSMD8**	**1.610**	**1.07**	**2.39**	0.850	0.625	1.080
cg23943801_A_23_P128166	**RAB21**	**2.764**	**1.36**	**6.06**	1.240	0.729	2.441
cg17942096_A_23_P165180	**RFXANK**	0.922	0.84	1.01	**1.050**	**1.019**	**1.086**
cg18345635_A_23_P147345	**SLC16A3**	1.009	0.98	1.06	**1.017**	**1.007**	**1.031**
cg18345635_A_23_P147349	**SLC16A3**	1.010	0.98	1.05	**1.011**	**1.003**	**1.024**
cg18345635_A_23_P158725	**SLC16A3**	1.011	0.98	1.06	**1.015**	**1.006**	**1.029**
cg17726022_A_23_P326510	SLC38A1	1.040	0.72	1.51	0.875	0.702	1.032
cg17726022_A_24_P261734	SLC38A1	1.052	0.73	1.55	0.868	0.698	1.021
cg18320336_A_24_P406335	**STEAP1**	**0.939**	**0.89**	**0.99**	0.995	0.981	1.004
cg11558474_A_23_P94552	**TMEM2**	**1.107**	**1.01**	**1.23**	1.017	0.994	1.042
cg00973286_A_23_P139715	**TNFRSF1A**	1.096	0.91	1.28	**1.080**	**1.001**	**1.181**

***Associations in bold type are significant**

Chapter 4

Discussion

Conclusion

Malignant adult gliomas are the most common type of brain cancer[1]. In the past decade, advances in diagnosis and treatment, particularly the use of the alkylating agent temozolomide, have only led to minimal improvement in patient survival [2,3]. Glioma survival outcome has been found to be associated with age, adjuvant treatments, giant-cell subtype and oligodendroglia differentiation[2]. In addition, advances in imaging techniques have allowed for better diagnosis[4,5] and more complete resection of malignant tumors, which has also been correlated with improved patient survival[5,6]. However, advances in the classification of glioma based on its molecular landscape are the most clinically relevant[7-10]. The classification of glioma types/subtypes using both genetic and epigenetic profiles has not only enhanced our knowledge of gliomagenesis but has also highlighted both molecular predictors of survival and possible therapeutic targets of glioma[11].

Genetic markers of glioma such as somatic alterations in the p53[12-15], Rb[15], EGFR[16], PI3K, and VEGF signaling pathways have now been well established[2,17], allowing for treatments targeting specific genes and proteins. More successful targeted therapies include anti-angiogenic drugs, including the commonly used bevacizumab[18]. Bevacizumab is a monoclonal antibody against VEGF, which, upon binding, inhibits VEGF activity[19]. In 2009 bevacizumab was granted accelerated approval by the FDA as a single agent for the treatment of recurrent GBM[19]. However, some studies have shown that anti-angiogenic drugs can enhance invasion and metastasis[20,21]. It has been suggested

that anti-angiogenic drugs should be used in combination with drugs that inhibit progression, invasion, and/or metastasis to increase overall survival[22,23].

EGFR is a tyrosine kinase activated growth factor that is involved in the activation of many signaling pathways, including RAS-MEK-ERK and PI3K-AKT[24], and is strongly dysregulated in glioma and consistently amplified or mutated in GBM[16]. Many targeted therapies, including monoclonal antibodies, vaccines, tyrosine-kinase inhibitors, and RNA-based agents have been under review[25]. Though research on these drugs seems promising, drug resistance is a common endpoint, re-emphasizing the need for novel drug targets[25]. Recently, greater attention has been given to epigenetics-based targets, both for prognostic and therapeutic purposes.

The role of epigenetics in cancer biology has only recently started to come into focus. Epigenetics encompasses events such as histone modifications[26], DNA methylation[27,28], and the targeting of genes by microRNA[29,30], all of which are capable of changing an individual's gene expression and/or cellular phenotype without directly changing the DNA sequence[31,32]. One of the most intensely studied areas of epigenetics is DNA methylation, which entails the addition of methyl groups to CpG dinucleotides[33]. DNA methylation is catalyzed by DNA methyltransferases and causes condensation of chromatin structure, which can lead to dysregulation of gene transcription[33,34]. DNA methylation of gene promoters is strongly implicated in a variety of cancers, including gliomas, and has been associated variable prognosis[35]. The best-known example of this association in glioma is the promoter methylation of methyl guanine methyl transferase (*MGMT*), which is associated with increased survival after treatment[36,37]. The fact that epigenetics does not involve actual alterations in the DNA sequence makes it more

appealing to study because unlike genetic alterations, epigenetic alterations are potentially reversible. Agents inhibiting re-methylation, such as 5-azacytidine, have already been approved for the treatment of hematopoietic cancers[38-40]. However, the use of such drugs on solid tumors has proven less effective[41]. Additionally, the lack of specificity of DNA-methylating agents is a prevalent concern, as it can lead to global demethylation and consequent expression of oncogenes and transposable elements, ultimately causing genomic instability[41].

DNA methylation has become a reliable source of biomarkers, as methylation profiles can distinguish cell lineages[42], tissues[43], and disease subtypes, and contribute to improvements in diagnosis, prognosis, and treatment outcome[35]. On a single locus level, DNA methylation has aided in the treatment and survival of glioma, as seen with the aforementioned promoter methylation of *MGMT*[44]. Loss of methylation on a global level has become a defining tumor characteristic[45,46].

Progress in the understanding of both the genetic and epigenetic landscapes of glioma has led to advances in both the diagnosis and treatment of the disease. Despite these advancements, disease survival remains low. Some researchers theorize that patients would benefit from targeting the molecular landscape as a whole, not just specific somatic alterations[47]. This theory would rely heavily on the molecular classification of tumors and how specific profiles or molecular characteristics associate with both treatment and survival. The molecular classification of glioma began with the integration of genetic alterations. At the forefront of this research was Phillips et al, who defined 3 classes of glioma based on the integration of copy number variation (CNV), gene expression, and activation of cell signaling[8]. The classes, proneural, proliferative,

and mesenchymal, each resemble different stages of neurogenesis and each was differentially associated with outcome, with the proliferative and mesenchymal classes demonstrating the poorest survival[8]. These analyses were further supported by the addition of mutation data, which helped refine associations seen between classes and survival outcome and led to the identification of novel drivers of gliomagenesis[10,48,49]. However, as these studies only exhibit the genetic diversity of glioma and its subtypes, there is a need for further analyses that integrate other important aspects of gliomagenesis, such as epigenetics.

The integration of genetics and epigenetics has greatly enhanced our knowledge of cancer biology, as seen with CpG island methylator phenotypes (CIMP), which can distinguish different tumor subtypes and are significantly associated with outcome. In glioma, the link between promoter methylation and gene expression has been established on a single-locus level. However, no large-scale attempts integrating methylation patterns and genetic alterations in glioma have been made to date. The goal of this thesis was to integrate genetic and epigenetic profiles to obtain molecular drivers of malignancy and survival in glioma.

In chapter 2, we discuss the relationship between the common glioma mutant isocitrate dehydrogenase (IDH) and its association with DNA methylation. First DNA methylation signatures of GBM, astrocytomas (AS), oligodendrogliomas (OD), oligoastrocytomas (OA), ependymomas (EP), and pilocytic astrocytomas (PA) (n=131) and those of non-glioma brain tissues (n=7) were obtained using the Infinium GoldenGate array, which interrogates CpG methylation loci in ~1500 cancer related genes. Tumors and non-glioma tissues were then clustered based on their methylation

status (β-value), with samples having the most similar methylation patterns clustering together. As expected, gliomas clustered separately from non-glioma brain tissue. Furthermore, a pathway analysis of differentially methylated loci (based on Δβ between glioma and non-glioma brain tissue) demonstrated that as a whole, metabolic pathways were commonly hypomethylated. Interestingly, the metabolite IDH1/2 has recently been found to be mutated in approximately 80% of low-grade gliomas and secondary GBM, and in <10% of primary GBM[50,51]. This prompted us to investigate the association between *IDH* and DNA methylation in glioma. Recursively partitioned mixture modeling was used to cluster only glioma samples, resulting in nine classes that were significantly associated with age, histology, and grade. Not surprisingly, *IDH* mutants were associated with histological subtypes, with an increased number of mutants found in low-grade gliomas and secondary GBM compared with primary GBM. In addition, *IDH* mutants were exclusively associated with the two homogenous hypermethylated classes, where non-mutants were heterogeneously distributed among the remaining seven classes. This novel finding suggested *IDH* as a potential driver of a hypermethylator phenotype. In fact, associations between methylation class and both *TP53* and *EGFR* were less robust than that of mutant *IDH*, further supporting the role of *IDH* as a driver of the observed hypermethylator phenotype. A Cox proportional hazards model showed that patients whose tumors harbored *IDH* mutants had significantly improved outcome compared with patients whose tumors harbored non-mutant *IDH*, suggesting *IDH* mutation in association with hypermethylation as a potential prognostic biomarker of glioma. The prognostic value of the hypermethylator phenotype, or CIMP, was first observed in colorectal cancer. CIMP classes are determined based on mutations in *BRAF* and/or *KRAS* and promoter

methylation levels[52], and CIMP subtypes are associated with differential prognosis. Recently, Noushmehr et al confirmed the relationship of mutant *IDH* with promoter methylation of CIMP-associated loci and successfully defined a glioma-CIMP (G-CIMP) [7]. They found that G-CIMP-positive tumors are frequently found in younger patients with low-grade gliomas, and these patients often show better survival outcomes, which supports our findings[7]. Additionally, gene expression data revealed that G-CIMP tumors are enriched in a portion of the previously identified proneural subtype[7], which has also been associated with a better prognosis[8,10].

Mechanistically, the link between IDH mutants and DNA methylation is still under debate. Mutant IDH is a neomorphic enzyme that, instead of catalyzing the oxidative decarboxylation of isocitrate into α-ketoglutarate (α –KG), actually converts α –KG into oncometabolite 2-hydroxyglutare (2-HG) in an NADPH-dependent manner[53]. Accumulation of 2-HG has been seen in diseases such as 2-hydroxyglutaric aciduria, which has been associated with increased risk of glioma[53]. In 2011, Xu et al found 2-HG to be a weak antagonist of α–KG, which, at high concentrations, can inhibit α–KG-dependent dioxygenases such as histone demethylases and TET2 5-methylcytosine hydroxylases (5mC). Inhibition of histone demethylases can limit the removal of histone-associated methyl groups causing an increase in normal histone methylation. TET2 normally catalyzes the conversion of 5mC to 5-hydroxymethylcytosine (5hmC), which can lead to demethylation of DNA. Therefore, inhibition of TET2 can lead to an increase in DNA methylation. Consequently, the increased production of 2-HG from mutated IDH can cause dysregulation of the normal methylome[54]. In 2012, Turcan et al lent further support to the connection between IDH and the G-CIMP[55]. The group found enriched

methylation of the histone marks H3K9 and H3K27 in cells expressing mutant *IDH*, which have been shown to promote DNA methylation through recruitment of DMNTs [56]. Furthermore, expression of *TET2* was inhibited in *IDH* mutant samples, decreasing production of 5hmC and further supporting previous findings of a possible mechanistic link between *IDH* mutations and G-CIMP[55].

The success of the integration of genetic and epigenetic alterations in defining a prognostically relevant G-CIMP class further demonstrates the need for analyses that can aid in the discovery of other drivers and potential biomarkers of gliomagenesis. Of particular interest are those gliomas that fall outside of the IDH-driven methylator phenotype.

In chapter 3, we employed a genome-wide, agnostic strategy for the discovery of novel predictive biomarkers related to the prognosis of glioma. In the previous chapter, we focused on the associations of *IDH* mutant gliomas and methylation. Interestingly, though *IDH* mutants were exclusive to the two hypermethylated classes, wild-type *IDH* was homogenously distributed among the lower methylated classes, suggesting alternative mechanisms of glioma pathogenesis in these patients. Uniquely, in our study, we focused on *IDH* wild-type samples and the role methylation plays alone or in conjunction with gene expression in the pathogenesis and survival of primary GBM. Not surprisingly, the 27 genes found to be significantly associated with survival in our study are involved in invasion, angiogenesis, and metastasis, and many were previously found to be associated with brain/glioma. We found 10 methylation/expression pairs that had a significant expression-based association with survival, suggesting that DNA methylation in these genes affects survival outcome via expression of the associated gene, supporting

the commonly accepted paradigm that methylation effects survival through gene expression. Interestingly, of these 10 methylation/expression pairs, two were found within the *OSMR* gene. As mentioned in chapter 3, OSMR is associated with STAT3 activation via the JAK/STAT signaling pathway. In a glioma cell line, inhibition of STAT3 activation was associated with reduced cell migration and invasion and mice with *STAT3* knockdown tumors exhibited increased survival compared to controls, suggesting that inhibition of *STAT3* is important in gliomagenesis and survival. Therefore, methylation induced silencing of *OSMR* could inhibit the activation of STAT3, thereby attenuating glioma cell migration and invasion. Evidence of methylation-induced silencing of *OSMR* has already been shown in colorectal (CR) cancer. Methylation-induced silencing of *OSMR* expression was associated with increased growth due to inhibition of the *OSMR* substrate OSM[57,58]. Furthermore, CR tumors with increased *OSMR* promoter methylation were associated with a non-invasive phenotype[58], suggesting that *OSMR* could predict a class of tumors that are associated with improved survival. This data supports a possible link between promoter methylation, gene expression, and survival outcome, and suggests methylation of *OSMR* as a potential biomarker of a novel prognostic phenotype.

In our work, 14 methylation/expression pairs were found to have significant alternative associations, suggesting that methylation can function through alternative mechanisms other than expression, to effect survival. Increased expression of the water channel *AQP1*, which had a significant alternative association in our analysis, has recently been observed in GBM. Interestingly, *AQP1* has been shown to contain targets for regulatory microRNAs. Osmotically regulated microRNAs miR-708 and miR-666 were found to

inhibit *AQP1* expression in BDL endothelial cells, and low AQP1 levels were associated with reduced angiogenesis and fibrosis in a mouse model of liver cirrhosis[59]. Additionally, hypoxically activated miR-214 was correlated with decreased AQP1 expression in HUVEC cells[60], and miR-320a has been found to directly target *AQP1* and is associated with decreased mRNA and protein expression of AQP1 during cerebral ischemia. Importantly, AQP1 is associated with cytotoxic cerebral edema, angiogenesis, and invasion in GBM, suggesting that suppression of AQP1 expression could increase survival outcome[61]. Therefore, a possible mechanistic explanation for the alternative association we observed in our work involves methylation of the microRNA target region on *AQP1*, which would inhibit the binding of miR-320a and ultimately result in increased expression of *AQP1*. This model could explain the alternative mechanism associated with glioma survival in this instance[61,62]. Unexpectedly, there were four methylation-expression pairs that had both significant alternative and expression-based associations, suggesting that methylation can function simultaneously through both expression-based and alternative mechanisms to significantly impact survival. This phenomenon was observed with locus pairs found within the imprinted *GRB10* gene. *GRB10* has recently been implicated as both a putative oncogene in glioma [63] and potential tumor suppressor[64]. The *GRB10* gene is of particular interest because it has been found to contain 13 different splice variants, expression of all but one of which ($\gamma 2$) has been found in the brain[65]. GRB10 has been shown to have both an inhibitory and stimulatory effect on IGF-1-related proliferation, though not specifically in brain tissue[66]. Though the reason behind its conflicting effects is not yet understood, one theory is that different GRB10 isoforms have different regulatory functions but compete for similar substrates. DNA methylation

has the ability to regulate differential isoform production via alternative splicing[67]. Uniquely, imprinting of *GRB10* is tissue-dependent. Monoallelic expression is seen in skeletal muscle and placenta (maternally expressed in humans) and in the brain (paternally expressed in humans) [65]. Disruption of maternal imprinting in mice leads to overgrowth and insulin sensitivity throughout life, while in adult mice, deletion of *GRB10* is associated with increased total body mass and up-regulation of cancer associated genes[68]. Unfortunately, ablation of imprinting in the paternal allele has not been shown to affect growth[68]. However, it is important to note that mice have been found to only have maternally imprinted *GRB10* as opposed to humans, who show biallelic imprinting. Overall, this suggests that associations with survival can occur due to both loss of imprinting (expression-based association) and through expression of differentially functioning alternative isoforms (alternative association). It is plausible that these effects could be seen simultaneously in genes within the same tumor and work synergistically, or these effects could occur separately within different tumors, allowing the gene and its associated effect (alternative or expression-based) to be used as possible markers of tumor type. Overall, these findings corroborate the common idea that methylation operates through expression to affect survival outcome, but they also suggest that methylation can associate with survival outcome through mechanisms other than dysregulation of gene transcription. Though additional validation studies are needed, our method may lead to the identification of novel putative genetic and epigenetic biomarkers of glioma that could potentially be useful as therapeutic targets Importantly, this approach could be applicable to cancers other than glioma, and the model can be adjusted to include other variables of interest. For instance, instead of focusing on DNA

methylation and gene expression, one could focus on DNA methylation and microRNA. Thus, our analysis provides a conceivable method of biomarker discovery that may be broadly clinically applicable.

Unfortunately, there were several limitations to this study. First, we had to rely on publically available data, which did not have complete mutation and survival data. For the missing mutation data, we used an RPMM to remove the hypermethylated classes that our group[9] and others[7] have previously found to exclusively contain *IDH* mutants. To address issue of missing survival data we used an accelerated failure time model to predict the survival time of censored values in order to control for survival in a combat model. In order to ensure functionality of methylation loci in our analysis, an initial screen was conducted, and only methylation and expression pairs that were significantly correlated within the same gene were used. Unfortunately, this approach entailed the exclusion of loci that may affect the expression of genes from distant locations, as seen with the methylation of enhancer regions. Additionally, the Infinium HumanMethylation27 array that was used to determine methylation status in these samples is strongly biased towards proximal promoter regions[69]. Therefore, in future studies it may be beneficial to look at the correlation of methylation loci and all gene expression probes, without focusing on pairs within a single gene. In 2013, Aran et al explored how DNA methylation of distal regulatory sites in normal and malignant cell lines associates with gene expression levels across the genome[70]. First, they developed a model, which, at a score of less than or equal to 0.85, successfully determined genes that undergo promoter methylation-dependent expression in variable methylation sites (VMS) of malignant cell lines with 2.63% sensitivity and a 12.8% false discovery rate. This

model was then applied to VMS +/- 1 megabase of the transcription start site of over 17,000 human genes, excluding sites that fell within 5 KB of promoters/alternative promoters of the associated genes. This yielded 1,911 pairs (486 genes), 1,041 of which were distal regulatory sites largely within promoter and enhancer regions. Further analyses suggested that high-scoring pairs were enriched in transcriptional enhancers and bound transcription factors in a methylation-dependent manner. Furthermore, analysis of the 1,911 distal methylation sites in normal vs. malignant mammary epithelial cells revealed a methylation-dependent association between high-score enhancer regions and the expression of their associated genes and suggested that methylation levels of these enhancers associate better with transcriptional regulation than promoter methylation. Moreover, both hypomethylation and hypermethylation of enhancers was observed in different malignant cell types, suggesting a role for differential enhancer methylation in cancer[70]. In addition differential enhancer methylation may be useful in differentiating different glioma subtypes. Due to limited patient data, our study consisted only of primary GBM; however, using our method to look at several different histologies could support pre-existing or aid in the discovery of new subclasses of glioma. This is further supported by our observation of methylation and expression pairs that were significant for both expression-based and alternative associations, demonstrating that the pathogenesis of these tumors involves discrete mechanisms that have differential effects on survival outcome. Of further interest is the prognostic signature of gliomas both before and after treatment. Shukla et al used both treated (radio therapy and concomitant temozolomide) patient samples and a series of cell culture experiments (using 5-Aza-2'-deoxycytidine treatment) to identify a methylation-based prognostic signature in high-

grade glioma comprising nine genes[71]. Using a methylation-based risk-score, the methylation statuses of these nine genes could identify patients as either low-risk or high-risk, with the latter having significantly lower survival. Unsurprisingly, the low-risk group contained a high percentage of IDH1 mutants, proneural associated genes, and G-CIMP tumors. Additionally, the high-risk group was associated with activated NF-kB signaling. Further studies demonstrated that inhibition of NF-kB lead to enhanced sensitivity to chemotherapeutic agents, Not only does this explain the decreased survival observed in high-risk groups, but it also suggests NF-kB as a probable therapeutic target in cases where normal therapy is not successful[71].

Our research has aided in the discovery of putative glioma biomarkers, as observed in chapter 2 with the association of IDH and a hypermethylator phenotype with increased survival. Though more validation is required, we have shown the importance of analyses integrating multiple somatic alterations and their associations with outcome, as shown in chapter 3 with the integration of DNA methylation and gene expression. These analyses supported the common idea that DNA methylation works through gene expression to affect survival. Our analysis also demonstrated a unique method of biomarker discovery that can easily lend itself to diseases other than glioma. Most importantly, our analysis demonstrated significant alternative associations, suggesting that DNA methylation can also operate through alternative or combined mechanisms to affect outcome. An alternative association, as defined in chapter 3, is when DNA methylation affects survival without directly influencing gene expression. In this case, DNA methylation does not directly alter gene transcription via promoter methylation, but may change gene expression and survival via dysregulation of microRNAs and enhancer

regions. In addition, DNA methylation could also affect survival by promoting/hindering genomic fragility and instability. Survival in patients with tumors having a hypermethylator phenotype, as seen with the *IDH*-associated hypermethylator phenotype discussed in chapter 2, has often been associated with promoter methylation-induced silencing of tumor suppresser genes. However, with increased knowledge of the role epigenetics plays in cancer and survival, promoter methylation may not be the only relevant epigenetic mechanism at play in the methylator phenotype. As previously discussed, alternative mechanisms such as methylation inhibition of microRNAs and their target regions and methylation of distal sites associated with enhancer or polycomb regions could impact gene expression and survival. Additionally, tumors with a hypermethylator phenotype are generally associated with increased survival. This could also be explained by the enhanced genomic stability observed in these tumors. In cancer, genomic instability is associated with hypomethylation, which can cause increased expression of aberrant transposons and a potential subsequent decrease in survival. Theoretically, the increased hypermethylation seen in CIMP-positive tumors could manifest in methylation of transposons/repeat elements, thereby increasing stability relative to CIMP-negative tumors. This is another possible explanation for the increased survival associated with CIMP-positive tumors vs CIMP-negative tumors. Collectively, this research has enhanced our knowledge of gliomagenesis and has further demonstrated the molecular complexity of glioma.

Future Directions

Though our method was successful in demonstrating the importance of integration of molecular phenotypes in the classification of molecular drivers of malignancy in glioma, it will be important to properly validate these analyses with a separate set of GBMs. Additionally, our methods could benefit from the use of new high-throughput DNA methylation arrays, such as the HumanMethylation450 BeadChip array (Illumina), which allows for increased coverage of the genome compared to the HumanMethylation27 array, including the interrogation of CpG shores, whose differential methylation patterns have become increasingly recognized as important biomarkers of disease. This increased coverage would allow us to look at the correlation between DNA methylation and gene expression beyond the transcription start site. Following the example of Aran et al, who looked at DNA methylation of distal regulatory sites[70], we could tailor our integrated analysis to focus on the association of CpG shore methylation with both gene expression and survival. It has been shown that altered methylation at distal regulatory sites found within enhancer regions correlates with altered gene expression more strongly than altered methylation at promoter regions[70]. With this knowledge, it would be expected that our analysis would demonstrate increased expression-based associations with survival.

References

1. Wen PY, Kesari S. Malignant gliomas in adults. *N Engl J Med.* Jul 2008;359(5):492-507.

2. Masui K, Cloughesy TF, Mischel PS. Review: molecular pathology in adult high-grade gliomas: from molecular diagnostics to target therapies. *Neuropathol Appl Neurobiol.* Jun 2012;38(3):271-291.

3. Weller M, Cloughesy T, Perry JR, Wick W. Standards of care for treatment of recurrent glioblastoma--are we there yet? *Neuro Oncol.* Jan 2013;15(1):4-27.

4. Mohammed W, Xunning H, Haibin S, Jingzhi M. Clinical applications of susceptibility-weighted imaging in detecting and grading intracranial gliomas: a review. *Cancer Imaging.* 2013;13:186-195.

5. Clarke JL, Chang SM. Neuroimaging: diagnosis and response assessment in glioblastoma. *Cancer J.* 2012 Jan-Feb 2012;18(1):26-31.

6. Sanai N, Berger MS. Glioma extent of resection and its impact on patient outcome. *Neurosurgery.* Apr 2008;62(4):753-764; discussion 264-756.

7. Noushmehr H, Weisenberger D, Diefes K, et al. Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell.* May 2010;17(5):510-522.

8. Phillips HS, Kharbanda S, Chen R, et al. Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. *Cancer Cell.* Mar 2006;9(3):157-173.

9. Christensen BC, Smith AA, Zheng S, et al. DNA methylation, isocitrate

dehydrogenase mutation, and survival in glioma. *Journal of the National Cancer Institute.* Jan 19 2011;103(2):143-153.

10. Verhaak RG, Hoadley KA, Purdom E, et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell.* Jan 2010;17(1):98-110.

11. von Deimling A, Korshunov A, Hartmann C. The next generation of glioma biomarkers: MGMT methylation, BRAF fusions and IDH1 mutations. *Brain Pathol.* Jan 2011;21(1):74-87.

12. Watanabe T, Yokoo H, Yokoo M, Yonekawa Y, Kleihues P, Ohgaki H. Concurrent inactivation of RB1 and TP53 pathways in anaplastic oligodendrogliomas. *J Neuropathol Exp Neurol.* Dec 2001;60(12):1181-1189.

13. Bello MJ, Rey JA. The p53/Mdm2/p14ARF cell cycle control pathway genes may be inactivated by genetic and epigenetic mechanisms in gliomas. *Cancer genetics and cytogenetics.* Jan 2006;164(2):172-173.

14. Amatya VJ, Naumann U, Weller M, Ohgaki H. TP53 promoter methylation in human gliomas. *Acta Neuropathol.* Aug 2005;110(2):178-184.

15. Costello JF, Berger MS, Huang HS, Cavenee WK. Silencing of p16/CDKN2 expression in human gliomas by methylation and chromatin condensation. *Cancer Res.* May 1996;56(10):2405-2410.

16. Heimberger AB, Suki D, Yang D, Shi W, Aldape K. The natural history of EGFR and EGFRvIII in glioblastoma patients. *J Transl Med.* Oct 2005;3:38.

17. Huse J, Holland E. Targeting brain cancer: advances in the molecular

pathology of malignant glioma and medulloblastoma. *Nat Rev Cancer.* May 2010;10(5):319-331.

18. Norden AD, Drappatz J, Wen PY. Antiangiogenic therapies for high-grade glioma. *Nat Rev Neurol.* Nov 2009;5(11):610-620.

19. Cohen MH, Shen YL, Keegan P, Pazdur R. FDA drug approval summary: bevacizumab (Avastin) as treatment of recurrent glioblastoma multiforme. *Oncologist.* Nov 2009;14(11):1131-1138.

20. Ebos JM, Lee CR, Cruz-Munoz W, Bjarnason GA, Christensen JG, Kerbel RS. Accelerated metastasis after short-term treatment with a potent inhibitor of tumor angiogenesis. *Cancer Cell.* Mar 2009;15(3):232-239.

21. Pàez-Ribes M, Allen E, Hudock J, et al. Antiangiogenic therapy elicits malignant progression of tumors to increased local invasion and distant metastasis. *Cancer Cell.* Mar 2009;15(3):220-231.

22. You WK, Sennino B, Williamson CW, et al. VEGF and c-Met blockade amplify angiogenesis inhibition in pancreatic islet cancer. *Cancer Res.* Jul 2011;71(14):4758-4768.

23. Narayana A, Kelly P, Golfinos J, et al. Antiangiogenic therapy using bevacizumab in recurrent high-grade glioma: impact on local control and patient survival. *J Neurosurg.* Jan 2009;110(1):173-180.

24. Oda K, Matsuoka Y, Funahashi A, Kitano H. A comprehensive pathway map of epidermal growth factor receptor signaling. *Mol Syst Biol.* 2005;1:2005.0010.

25. Taylor TE, Furnari FB, Cavenee WK. Targeting EGFR for treatment of glioblastoma: molecular basis to overcome resistance. *Curr Cancer Drug*

Targets. Mar 2012;12(3):197-209.

26. Turner BM. Cellular memory and the histone code. *Cell.* Nov 2002;111(3):285-291.

27. Riggs AD. X inactivation, differentiation, and DNA methylation. *Cytogenet Cell Genet.* 1975;14(1):9-25.

28. Holliday R. The inheritance of epigenetic defects. *Science.* Oct 1987;238(4824):163-170.

29. Costa FF. Non-coding RNAs, epigenetics and complexity. *Gene.* Feb 2008;410(1):9-17.

30. Chuang JC, Jones PA. Epigenetics and microRNAs. *Pediatr Res.* May 2007;61(5 Pt 2):24R-29R.

31. Berger SL, Kouzarides T, Shiekhattar R, Shilatifard A. An operational definition of epigenetics. *Genes Dev.* Apr 2009;23(7):781-783.

32. Bird A. Perceptions of epigenetics. *Nature.* May 2007;447(7143):396-398.

33. Smith ZD, Meissner A. DNA methylation: roles in mammalian development. *Nat Rev Genet.* Mar 2013;14(3):204-220.

34. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet.* Mar 2003;33 Suppl:245-254.

35. Portela A, Esteller M. Epigenetic modifications and human disease. *Nat Biotechnol.* Oct 2010;28(10):1057-1068.

36. Esteller M, Garcia-Foncillas J, Andion E, et al. Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *N Engl J*

Med. Nov 2000;343(19):1350-1354.

37. Hegi ME, Diserens AC, Gorlia T, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med.* Mar 2005;352(10):997-1003.

38. Garcia-Manero G, Kantarjian HM, Sanchez-Gonzalez B, et al. Phase 1/2 study of the combination of 5-aza-2'-deoxycytidine with valproic acid in patients with leukemia. *Blood.* Nov 2006;108(10):3271-3279.

39. Kaminskas E, Farrell A, Abraham S, et al. Approval summary: azacitidine for treatment of myelodysplastic syndrome subtypes. *Clinical cancer research : an official journal of the American Association for Cancer Research.* May 2005;11(10):3604-3608.

40. Issa JP, Gharibyan V, Cortes J, et al. Phase II study of low-dose decitabine in patients with chronic myelogenous leukemia resistant to imatinib mesylate. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* Jun 2005;23(17):3948-3956.

41. PM H, Z L, H K. Demethylation Agents in the Treatment of Cancer. *Pharmaceuticals.* July 2 2010 2010;3:2022-2044.

42. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics.* 2012;13:86.

43. Rakyan VK, Down TA, Thorne NP, et al. An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). *Genome Res.* Sep 2008;18(9):1518-1529.

44. Hegi M, Diserens A, Gorlia T, et al. MGMT gene silencing and benefit from

temozolomide in glioblastoma. *N Engl J Med.* Mar 2005;352(10):997-1003.

45. Feinberg AP, Vogelstein B. Hypomethylation of ras oncogenes in primary human cancers. *Biochemical and biophysical research communications.* Feb 1983;111(1):47-54.

46. Gama-Sosa MA, Slagel VA, Trewyn RW, et al. The 5-methylcytosine content of DNA from human tumors. *Nucleic acids research.* Oct 1983;11(19):6883-6894.

47. Olar A, Aldape KD. Using the Molecular Classification of Glioblastoma to Inform Personalized Treatment. *J Pathol.* Oct 2013.

48. Network CGAR. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature.* Oct 2008;455(7216):1061-1068.

49. Frattini V, Trifonov V, Chan JM, et al. The integrated landscape of driver genomic alterations in glioblastoma. *Nat Genet.* Aug 2013.

50. Parsons DW, Jones S, Zhang X, et al. An integrated genomic analysis of human glioblastoma multiforme. *Science.* Sep 2008;321(5897):1807-1812.

51. Hartmann C, Meyer J, Balss J, et al. Type and frequency of IDH1 and IDH2 mutations are related to astrocytic and oligodendroglial differentiation and age: a study of 1,010 diffuse gliomas. *Acta Neuropathol.* Oct 2009;118(4):469-474.

52. Ogino S, Kawasaki T, Kirkner G, Loda M, Fuchs C. CpG island methylator phenotype-low (CIMP-low) in colorectal cancer: possible associations with male sex and KRAS mutations. *J Mol Diagn.* Nov 2006;8(5):582-588.

53. Dang L, White D, Gross S, et al. Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature.* Dec 2009;462(7274):739-744.

54. Xu W, Yang H, Liu Y, et al. Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. *Cancer Cell.* Jan 2011;19(1):17-30.

55. Turcan S, Rohle D, Goenka A, et al. IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature.* Mar 2012;483(7390):479-483.

56. Jin B, Li Y, Robertson KD. DNA methylation: superior or subordinate in the epigenetic hierarchy? *Genes & cancer.* Jun 2011;2(6):607-617.

57. Friedrich M, Höss N, Stögbauer F, et al. Complete inhibition of in vivo glioma growth by oncostatin M. *J Neurochem.* Mar 2001;76(5):1589-1592.

58. Hibi K, Goto T, Sakuraba K, et al. Methylation of OSMR gene is frequently observed in non-invasive colorectal cancer. *Anticancer Res.* Apr 2011;31(4):1293-1295.

59. Huebert RC, Jagavelu K, Hendrickson HI, et al. Aquaporin-1 promotes angiogenesis, fibrosis, and portal hypertension through mechanisms dependent on osmotically sensitive microRNAs. *Am J Pathol.* Oct 2011;179(4):1851-1860.

60. Rutkovskiy A, Valen G, Vaage J. Cardiac aquaporins. *Basic Res Cardiol.* Nov 2013;108(6):393.

61. El Hindy N, Bankfalvi A, Herring A, et al. Correlation of aquaporin-1 water channel protein expression with tumor angiogenesis in human astrocytoma. *Anticancer Res.* Feb 2013;33(2):609-613.

62. Sepramaniam S, Armugam A, Lim KY, et al. MicroRNA 320a functions as a novel endogenous modulator of aquaporins 1 and 4 as well as a potential therapeutic target in cerebral ischemia. *The Journal of biological chemistry.* Sep 2010;285(38):29223-29230.

63. Nord H, Hartmann C, Andersson R, et al. Characterization of novel and complex genomic aberrations in glioblastoma using a 32K BAC array. *Neuro Oncol.* Dec 2009;11(6):803-818.

64. Yu Y, Yoon SO, Poulogiannis G, et al. Phosphoproteomic analysis identifies Grb10 as an mTORC1 substrate that negatively regulates insulin signaling. *Science.* Jun 2011;332(6035):1322-1326.

65. Monk D, Arnaud P, Frost J, et al. Reciprocal imprinting of human GRB10 in placental trophoblast and brain: evolutionary conservation of reversed allelic expression. *Human molecular genetics.* Aug 2009;18(16):3066-3074.

66. Vecchione A, Marchese A, Henry P, Rotin D, Morrione A. The Grb10/Nedd4 complex regulates ligand-induced ubiquitination and stability of the insulin-like growth factor I receptor. *Mol Cell Biol.* May 2003;23(9):3363-3372.

67. Maunakea AK, Chepelev I, Cui K, Zhao K. Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell Res.* Nov 2013;23(11):1256-1269.

68. Desbuquois B, Carré N, Burnol AF. Regulation of insulin and type 1 insulin-like growth factor signaling and action by the Grb10/14 and SH2B1/B2 adaptor proteins. *The FEBS journal.* Feb 2013;280(3):794-816.

69. Bibikova M, Le J, Barnes B, et al. Genome-wide DNA methylation profiling

using Infinium® assay. *Epigenomics.* Oct 2009;1(1):177-200.

70. Aran D, Sabato S, Hellman A. DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes. *Genome Biol.* Mar 2013;14(3):R21.

71. Shukla S, Pia Patric IR, Thinagararjan S, et al. A DNA Methylation Prognostic Signature of Glioblastoma: Identification of NPTX2-PTEN-NF-κB Nexus. *Cancer Res.* Nov 2013;73(22):6563-6573.