Identification of genetic and epigenetic alterations in pediatric high-grade astrocytomas

Dong Anh Khuong Quang

Department of Human Genetics
Faculty of Medicine
McGill University
Montreal, Quebec, Canada
March, 2016

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

To Nha Thi

To her little brother

Table of Contents

Abstract

Brain tumors are the leading cause of cancer related mortality and morbidity in the childhood years, and until recently, most of treatments were based on protocols used in adults with limited results. Among brain tumors, astrocytomas represent the largest group. The World Health Organization (WHO) divides them into four grades based on their pathological features of malignancy. Our work is focused on pediatric grade 4 astrocytomas or glioblastomas (GBMs). These tumors are rare in children, but the prognosis is dismal with less than 20% 5-year survival rate. Many proposed therapies exist for these tumors, and one of the most promising targeted therapies tested in glioblastoma was the group of receptor tyrosine kinase (RTK) inhibitors. However, these treatments failed their expectations both in adults and children. My work has shown that RTKs have a critical role in the oncogenic process, and the inefficacy of inhibition is due to the dysregulated endosomal trafficking of RTKs in pediatric tumors, leading to kinase switching and activation. Therefore, these results support the need for a therapeutic approach targeting multiple receptors at the same time. One other critical aspect to understand GBM is to study the whole genome landscape. We and others have recently described critical somatic mutations in the gene (*H3F3A*) encoding the histone H3.3 in pediatric GBM. Interestingly, pediatric diffuse intrinsic pontine gliomas (DIPGs) also present mutations in *H3F3A*. In collaboration with Dr. Cynthia Hawkins's laboratory in Toronto, we analyzed 42 tumors for *H3F3A*, *HIST1H3B*, *IDH*, *ATRX* and *TP53* mutations, copy number alterations and clinical outcome. Our data showed that 80% of DIPGs were mutated for *H3F3A*, and that these patients have a worse prognosis. Our results emphasized the role of diagnostic biopsy for DIPG patients, and more importantly if they have an atypical presentation. In the last part of this thesis, we used an integrated approach with DNA methylation, copy number variations and gene expression data to classify adult and

pediatric GBMs into 6 groups based on their DNA methylation profile with clinico-molecular correlation. Given together, this thesis contributes to a better understanding of pediatric and adult GBMs, and this may pave the way to an individualized therapeutical approach for the patients who suffer from this debilitating tumor.

Résumé

Durant l'enfance, la principale cause de morbidité et de mortalité liées au cancer est représentée par les tumeurs cérébrales, et jusqu'à présent, la plupart des traitements sont basés sur des protocoles utilisés chez l'adulte avec peu de résultats positifs. Les astrocytomes représentent le plus large groupe de tumeurs cérébrales. L'Organisation Mondiale pour la Santé les divise en 4 grades selon des caractéristiques anatomo-pathologiques de malignité. Notre travail est centré sur les astrocytomes de grade 4 ou glioblastomes pédiatriques. Ces tumeurs sont rares chez l'enfant, mais leur pronostic est particulièrement sombre avec un taux de survie inférieur à 20% à 5 ans. Ces dernières années, une des thérapies ciblées les plus prometteuses était les inhibiteurs de récepteurs tyrosine kinase. Mais, malheureusement, les résultats des premières études conduites n'ont pas été aussi bons qu'espérés, à la fois chez l'adulte et chez l'enfant. Mon travail a confirmé que les RTK ont un rôle critique dans la pathogénicité des tumeurs pédiatriques et que leur manque d'efficacité est dû au dérèglement de la voie endosomale qui régule le trafic et la dégradation de ces RTK, entrainant la dérégulation d'autres RTK (phénomène de "kinase switching") et leur activation. Ces résultats montrent combien il est important de cibler simultanément plusieurs RTK. Un autre aspect essentiel pour comprendre la biologie des glioblastomes est d'étudier l'ensemble de leur génome. Nous avons récemment décrit, ainsi que d'autres équipes indépendantes, des mutations récurrentes du gène *H3F3A* codant pour l'histone H3.3 dans les glioblastomes pédiatriques, ainsi que dans les gliomes malins du tronc cérébral. En collaboration avec l'équipe de Cynthia Hawkins à Toronto, nous avons séquencé *H3F3A*, *HIST1H3B*, *IDH*, *ATRX* et *TP53* dans 42 gliomes malins du tronc, ainsi qu'analysé les anomalies de copie et leur évolution clinique. Nos données ont montré que 80% des DIPG présentent une mutation de *H3F3A* et que ce groupe a un plus mauvais pronostic que celui sans mutation. Ces résultats insistent sur l'importance de pratiquer une biopsie au diagnostic pour ces patients, en

particulier si leur présentation est atypique. Dans la dernière partie de cette thèse, nous avons utilisé une approche intégrative en analysant les données de méthylation de l'ADN, les variations de copies et les données d'expression génique de tumeurs pédiatriques et adultes, et ainsi nous avons pu classer ces tumeurs en 6 groupes distincts à partir de leur profil de méthylation de l'ADN avec une corrélation clinico-biologique. En conclusion, cette thèse a contribué à une meilleure connaissance des glioblastomes pédiatriques et adultes, et nous espérons que ce travail va participer à l'effort commun de la communauté médicale et scientifique pour initier la mise en place de thérapies ciblées individualisées pour ces patients avec un pronostic sombre.

Acknowledgments

This thesis represents not only the results of my research over the past few years, as part of Dr. Jabado's lab, but is also the reflection of many experiences I have encountered with remarkable people, all of whom I wish to acknowledge.

First and foremost, I wish to express my gratitude to my supervisor Dr. Nada Jabado for her constant guidance and encouragement towards making my PhD a successful and unforgettable experience. She has been a tremendous mentor, allowing me to grow as a research scientist but also as a human being. She has also been instrumental in the steps I have subsequently taken in my professional career. Her passion for science, unparalleled energy and enthusiasm, as well as her incredible work ethic are a source of admiration and she continues to be an important role model in my life. During the most difficult times of writing this thesis, she gave me the moral support and freedom I needed to progress.

I gratefully acknowledge the funding sources that made my PhD possible. During my four years in the lab, my work was generously supported by the Fonds de Recherche du Québec – Santé, the Montreal Children's Hospital Research Institute and Fondation des Etoiles, and the Association pour la Recherche sur le Cancer (France). I also received travel awards from the Human Genetics department at McGill University and from the McGill Integrated Cancer Research Training Program.

I would also like to thank my PhD committee members, Dr. Janusz Rak, Dr. Morag Park, Dr. Jacquetta Trasler. They provided invaluable feedback and scientific support during my committee meetings, as well as teaching me the importance of taking a critical approach to my work and guiding me towards excellence.

Nothing in this thesis would have been possible without invaluable collaborators from McGill University, other Canadian universities and multiple institutions overseas. They have been appropriately acknowledged individually throughout the thesis. However, I would particularly like to thank Dr. Peter Siegel from the Goodman Cancer Research Centre for his help with the tissue microarray and immunohistochemistry; Dr. Steffen Albrecht from the Pathology Department at the Montreal Children's Hospital for his enthusiasm in providing us with samples, as well as his patience with teaching me how to look at tissue slides under the microscope; and Brian Meehan who was always available at Place Toulon to troubleshot any issues that arose, and for the many insightful and useful discussions that we had about my research. I have learnt so much about high-throughput technologies at the Genome Quebec Innovation Centre, with the infinite patience of Simon Papillon-Cavanagh, François Lefebvre and Dr. Jacek Majewski - thank you for nurturing me so that I was able to develop further than the novice I started as. Both Dr. Stefan Pfister and his team from the German Cancer Research Centre and Dr. Cynthia Hawkins and her team from the Toronto Hospital for Sick Children Research Institute have made huge contributions to this thesis with their fruitful collaboration.

Many people at Place Toulon and the Department of Human Genetics helped to make me feel welcome when I arrived as a new student to Montreal. To name a few, I would like to thank Melanie Cotiangco, Marlene Aardse and Danuta Rylski from the Montreal Children's Hospital Research Institute, and Ross Mackay, Thomas Leslie, and Kandace Springer from the Department of Human Genetics. I would like to particularly emphasize the ongoing support that I have had from the Department of Human Genetics since moving overseas.

I have so many fond memories of the lab, which has become a second home. We combined hard work with great fun and, above all, excellent teamwork. To Damien Faury - I cannot tell

how much I have enjoyed working with you. Your patience and kindness, your knowledge of science and of the innermost workings of the lab are just a few of your attributes that made my time there unforgettable. To Noha Gerges - you deserve a place of honour in this acknowledgments section. You have been the best support and the best friend that I could ask for. Always there for me since we met on my first day, to listen to me and to encourage me to move forward. I cannot thank you enough for all you have done for me. To Karine Jacob - you were the first to welcome me in the lab and in Montreal, and to teach me how to work in a lab. Your friendship and generosity are priceless to me. To Margret Shirinian - you have been instrumental in my love of science. Your enthusiasm and your thoroughness are an inspiration to me. Although we now live far apart, I am blessed to count you as one of my friends. To Adam Fontebasso, Takrima Haque, Denise Bechet, Tenzin Gayden, Nisreen Ibrahim, Xiaoyang Liu, Helene Delhomelle, Djihad Hadjadj, Tiffany Lee, Hani Al-Halabi, I am very grateful for all the support we have shown to each other, helping one other to go through the difficulties of bench work with good humour.

I have written this thesis whilst training as a pediatric oncology fellow at the Royal Children's Hospital in Melbourne (Australia), and at times was desperate to finish it. I am forever grateful for the support received from the team here. I would particularly like to mention Dr. Francoise Mechinaud who has encouraged me towards a career as a clinician scientist and provided me with the support to complete this thesis, and Dr. Seong Lin Khaw for his helpful comments and whose enthusiasm to bring to the patient bedside his scientific knowledge is an example. Also, I would like to acknowledge my fellow friends who were always there to cheer me up, and to encourage and enable me to find the time to finish.

Lastly, and most importantly, nothing would have been possible without the unconditional love of my family. To be blessed with their love has given me the strength to move forward. I

cannot thank my parents enough. They have been a constant support - believing in me when I doubted myself and encouraging me when I was exhausted. I could not ask for better parents. With her constant love and patience, my sister Dong Phuong has provided me with unquestioning and invaluable support. To my partner Quan - it has been a long road, thank you for your support over all these years.

During this PhD, I have been blessed with two children, Nha Thi and her little brother who we are expecting any day now, as I write these words! They mean the world to me and are my true motivation to keep going. I dedicate this thesis to them.

Format of the thesis

This thesis encompasses five chapters and is organized in a manuscript-based format according to the guidelines and specifications outlined by the Faculty of Graduate and Postdoctoral Studies of McGill University. Chapter 1 is an introduction that provides background material about high-grade astrocytomas and diffuse intrinsic pontine gliomas for the reader to best appraise following chapters. It also presents the rationale of the study, the hypotheses and the thesis objectives. Chapters 2 to 4 are presented in manuscript format. Chapter 2 is a manuscript currently under internal review, which will shortly be submitted for publication. Chapter 3 has been published in ***Acta Neuropathologica.*** 2012 Sep;124(3):439-47. Chapter 4 has been published in ***Cancer Cell.*** 2012 Oct 16;22(4):425-37. Connecting texts are provided to link one chapter to the next. Chapter 5 includes a general discussion and a global summary of the main results of this thesis, as well as proposed future directions.

Preface and contribution of authors

Understanding the mechanisms leading to pGBM tumorigenesis is a key step to finding a cure for these tumors that carry most often a fatal prognosis. Until recently, most of the research was focused on the role of RTK activation in cell proliferation and tumor formation, but findings from our lab (1) during my PhD have shifted the focus from RTK to shed light on the implication of the epigenome by identifying recurrent mutations in the histone H3.3. Our work presented in this thesis has contributed to understanding the biology of pGBM by exploring these two major biological aspects. First, within Chapter 2, we showed that the overexpression of EGFR and other RTKs was due to a disrupted endosomal pathway. By delaying EGFR degradation, SNX3, which is overexpressed in the majority of pGBM, sustains Ras pathway signalling, thus promoting cell proliferation and tumor formation. In Chapter 3, we further characterized H3.3 mutations in a specific group of high-grade gliomas, the DIPG, identifying a subgroup with H3.3-K27M mutation and worse clinical outcome. This work was published in ***Acta Neuropathologica*** (2). Based on the previous publication from our lab (1) and the results presented in Chapter 3, we aimed, in collaboration with Dr. Stephan Pfister's group, to have a comprehensive and integrative approach of the biology of GBM both in adults and pediatrics. Our most important finding was that each type of *H3F3A* mutation defines an epigenetic subgroup of GBM with a distinct global DNA methylation profile, and that they are mutually exclusive with *IDH1* mutations, which characterize a third mutation-defined subgroup. Overall, this thesis has contributed to a better understanding of the molecular biology of this dismal tumor, and has added further evidence for a personalized therapeutical approach of these patients, by identifying biomarkers and potential therapeutic targets.

Chapter 2: Sorting Nexin 3 overexpression disrupts EGFR and MET endosomal trafficking promoting cell proliferation and tumorigenicity in pediatric glioblastoma

Dong-Anh Khuong-Quang and **Takrima Haque** equally contributed to the manuscript. Both participated in the study design, and performed *in vitro* functional assays. **Dong-Anh Khuong-Quang** performed *in vivo* experiments, analyzed tissue microarray (TMA) results, performed the statistical analyses and drafted the manuscript. **Helene Delhommelle**, **Djihad Hadjadj**, **Tiffany Lee**, and **Hani Halabi** performed *in vitro* functional assays. **Steffen Albrecht** reviewed the TMA results and performed the xenograft mouse tumor histopathological analysis. **Brian Meehan** helped with the design of the *in vivo* experiments. **Zhifeng Dong** performed the immunohistochemistry of the TMA. **Damien Faury** participated to both *in vitro* and *in vivo* assays. **Peter Siegel** participated in the study design. **Stephan Pfister** provided the TMA samples. **Janusk Rak** participated in the study design. **Nada Jabado** conceived the study design, participated in data interpretation, provided leadership, managed the project and drafted the manuscript.

Chapter 3: K27M mutation in histone H3.3 defines clinically and biologically distinct subgroups of pediatric diffuse intrinsic pontine gliomas

Dong-Anh Khuong-Quang and **Pawel Buczkowicz** equally contributed to the manuscript. Both participated in the study design, collected clinical data, extracted and prepared DNA from the tumors, participated in the analysis of the sequencing data, performed the statistical analysis, and drafted the manuscript. **Dong-Anh Khuong-Quang** analysed the clinical data and constructed the survival curves. **Pawel Buczkowicz** performed the copy-number analysis. **Patricia Rakopoulos** participated in DNA preparation and in the copy-number analysis. **Xiao-Yang Liu** and **Adam M. Fontebasso** extracted DNA and participated in data interpretation. **Eric Bouffet** and **Ute Bartels** provided tumor material and clinical data. **Steffen Albrecht** provided tumor material and reviewed the histopathological diagnosis. **Jeremy Schwartzentruber** participated in sequencing analysis. **Louis Letourneau**, **Mathieu**

Bourgey, **Guillaume Bourque,** and **Alexandre Montpetit** participated in the sequencing and the copy-number analysis. **Genevieve Bourret** performed the Sanger sequencing under the supervision of **Pierre Lepage. Adam Fleming** provided tumor material and clinical data. **Peter Lichter**, **Marcel Kool**, and **Andreas von Deimling** provided tumor samples and clinical data. **Dominik Sturm** participated in the sequencing analysis and data collection. **Andrey Korshunov** provided tumor material and clinical data. **Damien Faury** extracted and prepared DNA, and participated in the data collection. **David T. Jones** participated in the sequencing analysis. **Jacek Majewski** participated in the study design and the sequencing analysis. **Stefan M. Pfister** provided tumor material and participated in the study design. **Nada Jabado and Cynthia Hawkins** are both co-senior authors and conceived the study design, participated in data interpretation, provided leadership, managed the project and drafted the manuscript. All authors approved the final manuscript.

Chapter 4: Hotspot mutations in *H3F3A* and *IDH1* define distinct epigenetic and biological subgroups of glioblastoma

Dominik Sturm, **Hendrik Witt**, **Volker Hovestadt**, and **Dong-Anh Khuong-Quang** equally contributed to the manuscript. They participated in the study design, in the results analysis, and drafted the manuscript. **Dominik Sturm**, **Hendrik Witt** and **Dong-Anh Khuong-Quang** collected and analyzed clinical data, extracted and prepared DNA from the tumors, analyzed the sequencing and DNA methylation data. **Dominik Sturm** and **Hendrik Witt** performed the gene expression and protein expression analysis. **Volker Hovestadt** performed the bioinformatics analysis, including developing an algorithm to extract copy number variant information from the DNA methylation data under the supervision of **Marc Zapatka**. **David T.W. Jones** and **Carolin Konermann** participated in DNA methylation and gene expression analysis. **Elke Pfaff**, **Martje Tönjes** and **Sebastian Bender** participated in sample preparation and data analysis. **Marcel Kool** participated in sequencing, DNA methylation and

gene expression analysis. **Martin Sill, Marc Zapatka, Natalia Becker, Manuela Zucknick,** and **Thomas Hielscher** participated in biostatistics analysis. **Xiao-Yang Liu**, **Adam M. Fontebasso**, and **Karine Jacob** participated in DNA extraction and preparation. **Marina Ryzhova** and **Steffen Albrecht**, **Marietta Wolter**, **Martin Ebinger**, **Martin U. Schuhmann, Timothy van Meter**, **Michael C. Frühwald**, **Holger Hauch**, **Arnulf Pekrun**, **Tim Niehues**, **Gregor von Komorowski**, **Jenny Madden**, **Andrew Donson**, **Nicholas K. Foreman**, **Rachid Drissi**, **Wolfram Scheurlen**, **Christel Herold-Mende**, **Christof M. Kramm**, **Jörg Felsberg**, **Benedikt Wiestler**, **Anders M. Lindroth**, **Adam Fleming**, **Magdalena Zakrzewska**, **Pawel P. Liberski**, **Krzysztof Zakrzewski**, **Peter Hauser**, **Miklos Garami**, **Almos Klekner**, **Laszlo Bognar**, **Bernhard Radlwimmer**, **Matthias Dürken**, **Maryam Fouladi**, **Andreas von Deimling**, **Camelia Monoranu**, **Wolfgang Roggendorf**, **Andreas Unterberg**, **Christian Hartmann**, **Wolfgang Wick**, **Till Milde**, **Peter van Sluis**, **Rogier Versteeg**, **Richard Volckmann**, **Tom Mikkelsen**, **Kenneth Aldape**, and **Guido Reifenberger** provided patient material and clinical information. **Andreas E. Kulozik**, **Olaf Witt**, and **V. Peter Collins** provided tumor samples and resources. **Jeremy Schwartzentruber** participated in the sequencing analysis. **Damien Faury** extracted DNA and prepared samples. **Sorana Morrissy**, **Florence Cavalli**, **Michael D. Taylor,** and **Jan Koster** participated in the bioinformatics analysis. **Jacek Majewski** coordinated the next-generation sequencing analysis, **Andrey Korshunov** contributed to the study design, the immunohistochemistry analysis of the samples, and data analysis. **Peter Lichter** provided resources and contributed to the study design and data analysis. **Christoph Plass**, **Nada Jabado**, and **Stefan M. Pfister** are co-senior authors and conceived the study design, participated in data interpretation, provided leadership, managed the project and drafted the manuscript. All authors approved the final manuscript.

List of figures, tables and supplementary material

Chapter 1: Introduction

List of abbreviations

2-HG	2-hydroxyglutarate
450HM	Infinium HumanMethylation450
5mC	5-methylcytosine
ABL1	Abelson murine leukemia viral oncogene homolog 1
aGBM	adult glioblastoma
ALT	alternative lengthening of telomeres
AML	acute myeloblastic leukaemia
ATRX	alpha thalassemia/mental retardation syndrome X-linked
BATS	Biology and Treatment Strategies
BSA	bovine serum albumin
CNA	copy number alteration
CDF	cumulative distribution function
CDKN2A	cyclin-dependent kinase Inhibitor 2A
ChIP-Seq	chromatin immunoprecipitation sequencing
CHOP	CpG hypomethylator phenotype
CIHR	Canadian Institutes of Health Research
CIMP	CpG island methylator phenotype
CLIA	Clinical Laboratory Improvement Amendments
CNS	central nervous system
CRISPR	clustered regularly interspaced short palindromic repeats
DAXX	death-associated protein 6
DIPG	diffuse intrinsic pontine glioma
DKFZ	German Cancer Research Center (Deutsches Krebsforschungszentrum)
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
ECDF	empirical cumulative distribution function
EEA1	early endosome antigen 1
EGFRvIII	epidermal growth factor receptor variant III
EGF	epidermal growth factor
EGFR	epidermal growth factor receptor
EORTC	European Organisation for Research and Treatment of Cancer

ERK	extracellular signal-regulated kinases
EZH2	enhancer of zeste homolog 2
FDA	Food and Drug Administration
FFPE	formalin-fixed, paraffin-embedded
FGFR	fibroblast growth factor receptor
FISH	fluorescence in situ hybridization
FITC	fluorescein isothiocyanate
FOXG1	forkhead box G1
G-CIMP	glioma-CpG island methylator phenotype
GCOS	GeneChip Operating Software
GEP	gene expression profiling
GFAP	glial fibrillary acidic protein
GGA3	Golgi-localized g-ear-containing Arf-binding protein 3
HE	hematoxylin and eosin
HGA	high-grade astrocytomas
HGF	hepatocyte growth factor
HGG	high-grade gliomas
HIF-1α	hypoxia inducible factor 1α
IDH1	isocitrate dehydrogenase 1
IDH2	isocitrate dehydrogenase 2
IGF1R	insulin-like growth factor 1 receptor
IHC	Immunohistochemistry
IP	immunoprecipitation
JNK	c-Jun N-terminal kinases
LAMP1	lysosomal-associated membrane protein 1
MAPK	mitogen-activated protein kinases
MDM2	louse double minute 2 homolog
MEK	mitogen-activated extracellular signal-regulated kinase
MET	hepatocyte growth factor receptor
MGMT	O^6-methylguanine–DNA methyltransferase
miRNA	micro ribonucleic acid
MLL	mixed-lineage leukemia
MRI	magnetic resonance imaging

mRNA	messenger Ribonucleic acid
NADP	nicotinamide adenine dinucleotide phosphate
NCI	National Cancer Institute
NCIC	National Cancer Institute of Canada
NF1	neurofibromin 1
NGS	next-generation sequencing
NHGRI	National Human Genome Research Institute
NKX2-2	NK2 homeobox 2
NOD/SCID	non-obese diabetic/severe combined immunodeficiency
OLIG1	oligodendrocyte transcription factor 1
OLIG2	oligodendrocyte transcription factor 2
OS	overall survival
PA	pilocytic astrocytoma
PAM	point accepted mutation
PBS	phosphate buffered saline
PCA	principal component analysis
PCR	polymerase chain reaction
PDGFRA	platelet-derived growth factor receptor alpha
PVDF	polyvinylidene difluoride
pGBM	pediatric glioblastoma
pHGG	pediatric high-grade glioma
PNET	primitive neuroectodermal tumor
PRC2	polycomb repressive complex 2
PtIns(3)P	phosphatidylinositol 3-phosphate
PVT1	Pvt1 oncogene
PX	phox homology
qRT-PCR	quantitative reverse-transcriptase polymerase chain reaction
RNA	ribonucleic acid
RT	room temperature
RTCA	real time cell analyzer
RTK	receptor tyrosine kinase
SAM	significance analysis of microarrays
SCID	severe combined immunodeficiency

SDS-PAGE	sodium dodecyl sulfate- polyacrylamide gel electrophoresis
SEM	standard error of the mean
SETD2	SET domain containing 2
SHH	sonic hedgehog
SNP	single nucleotide polymorphism
SNX3	sorting nexin 3
TALEN	transcription activator-like effector nuclease
TCGA	The Cancer Genome Atlas
TET	ten-eleven translocation
TGFα	transforming growth factor alpha
TK	tyrosine kinase
TKI	tyrosine kinase inhibitor
TMA	tissue microarrays
TP53	tumor protein p53
US	United States
VEGF	vascular endothelial growth factor
WES	whole exome sequencing
WHO	World Health Organization
Wls	wntless Wnt ligand secretion mediator
Wnt	Wingless-Type MMTV Integration Site Family
WT	wild type
YB1	Y-box binding protein 1
ZFN	zinc finger nucleases
α-KG	α-ketoglutarate

Chapter 1: Introduction

This thesis investigates genetic and epigenetic alterations in pediatric high-grade gliomas and does so through the use of high throughput assays, functional work, and next generation sequencing (NGS) techniques. This chapter will focus on the introduction of brain tumors mostly in children but will also highlight relevant information available on the adult counterpart, and will further discuss some of the main topics within this thesis.

1.1 Pediatric high-grade astrocytomas

Brain tumors are the largest group of solid neoplasias in children, accounting for approximately 20% (28.9 new cases/1,000,000 children/year) of new pediatric cancer cases in Canada (3), and are currently the leading cause of cancer-related mortality and morbidity in the pediatric years (4). Brain tumors are classified by the World Health Organization (WHO) according to their presumed cell of origin as well as their malignancy (5) (Figure 1.1). In the childhood years, the largest tumor type is the astrocytic subgroup followed by the embryonal neural tumors (Figure 1.2); the former will be discussed in detail within this chapter. Our laboratory is focused on elucidating the molecular alterations unique to high-grade pediatric astrocytomas in the hopes of discovering novel potential targets, which can be used for the treatment of patients riddled with these diseases.

In 2007, the WHO published their latest classification standard for brain tumors dividing gliomas into subtypes depending on their presumed cell of origin (5). This classification split gliomas up into these main categories: astrocytomas, ependymomas, oligoastrocytomas, and oligodendrogliomas (5). Histopathological grades were given to these tumors depending on their degree of malignancy and are the standard in tumor classification today. The assigned grades range from I to IV with each higher grade corresponding to a higher aggressivity of the

tumor (5-8). Two pediatric tumor groups will be introduced in detail in this section based on the focus of my thesis, and these are the high-grade astrocytomas (HGA) and diffuse intrinsic pontine gliomas (DIPG) tumors. In some publications reported here, high-grade gliomas (HGG) were studied together because of how rare these samples are, but it is important to remember that they are in fact distinct and encompass anaplastic astrocytoma, anaplastic oligodendroglioma, anaplastic oligoastrocytoma, anaplastic ependymoma and glioblastoma. In our laboratory, we distinguished high-grade astrocytomas from other high-grade gliomas and accordingly included only relevant tumors in our studies.

Astrocytomas comprise the most common tumors of the central nervous system (CNS) across the lifespan and account for 45% of all pediatric brain tumors (8). According to the WHO classification, they are divided into low grade (Grade I/II) and high-grade (Grade III/IV) tumors. Pilocytic astrocytomas (PA) are low-grade (Grade I) astrocytomas; they are the most common across the pediatric brain tumor cases and come with the most favourable prognosis comprising a 10-year survival rate of 96% (9). These are mainly found within the cerebellum, a relatively better location that can lend to complete surgical resection. However, PA can also arise in less favourable regions within the brain including the brainstem, thalamus, or optic pathway rendering these rare cases to be inoperable (10). Further to their location within the brain, these tumors do not typically advance to higher glioma stages. Glioblastomas (GBM) (Grade IV tumors) on the other hand have the worst prognosis amongst patients with a 5-year survival rate at 5-20% despite intensive therapeutic strategies (11-13). These tumors are rare in children representing 2.2% of all pediatric CNS tumors (14) (Figure 1.3). These are characterized by pseudopalissading necrosis, microvascular proliferation, high cellular proliferation, and nuclear atypia (5). Although pediatric GBMs (pGBM) can occur anywhere within the CNS, they most often arise in the cerebral hemispheres (6, 11). When arising in the

brainstem region, GBM and other high-grade gliomas are known as DIPGs, which are another main focus of my thesis further discussed in this section.

Treatment for pGBM stems from the typical available strategies including surgical resection if possible, followed by radio- and chemotherapies. Due to their largely infiltrative nature and, at times, difficult to reach locations, complete surgical resections of the tumor are rare and further adjuvant therapies are required (15). Treatment options are limited and it is notable that the US Food and Drug Administration (FDA) approves only two systemic chemotherapies for adult GBM: temozolomide for newly diagnosed patients and bevacizumab for patients with a recurrent tumor. To date, combined radiotherapy with temozolomide, an alkylating agent, has shown the most significant improvement leading to an increase of median survival from 12.1 to 14.6 months in the multicentre randomized trial carried out by the European Organisation for Research and Treatment of Cancer (EORTC) and the National Cancer Institute of Canada (NCIC) (16). Moreover, data from the same clinical trial showed that epigenetic silencing of the *MGMT* gene by promoter methylation is associated with favourable outcome of patients treated with concomitant radiation and temozolomide (median survival 21.7 months) compared to radiation alone (15.3 months, $p<0.001$). On the other hand, patients without methylated *MGMT* promoter did not show a significantly increased survival when treated with chemoradiotherapy versus radiotherapy alone (17). *MGMT* encodes for O^6-methylguanine DNA-methyltransferase, a protein involved in DNA repair by removing alkyl groups from the O^6 position of guanine, an important site of DNA alkylation affected by alkylating agents, including temozolomide. Bevacizumab is a humanized monoclonal antibody that binds specifically to vascular endothelial growth factor (VEGF)-A, the most angiogenic isoform of VEGF, thus preventing angiogenesis. In a relapse setting, it was granted accelerated approval by the FDA in 2009 based on results from a phase II clinical

trial where 6-month progression-free survival (PFS) was 42.6% as a monotherapy and 50.3% when combined with irinotecan, compared to 9 to 21% when patients were treated with irinotecan as single agent (18). Despite this, such strategies have only increased patient lifetime slightly without ever providing a solid treatment. Adult and pediatric GBM are histologically identical, and because of this, treatment strategies have remained the same across the lifespan probably contributing to their inefficiency (15).

DIPGs are infiltrative brainstem tumors, which represent 10-15% of all pediatric brain tumors and have a very dismal prognosis of less than 10% of patients alive at 2 years (19, 20). The mean age of diagnosis of patients is approximately 6-7 years of age with these tumors found almost exclusively in children (21). DIPGs are not categorized into one distinct grade (they vary from WHO II-IV gliomas) and instead are mainly classified based on their fatal location within the brainstem (19-22). Also, when compared to other brainstem tumors, DIPGs have been found to be radiologically distinct, thus allowing clinicians to diagnose this tumor based on clinical and radiological features only (19, 21). Treatment of DIPG, despite advances made over the last 30 years, remains ineffective leading to over 90% of cases being fatal within one year, therefore placing DIPG as one of the principal triggers of brain-related death in pediatrics (21, 22). Because the pons contains nerve nuclei critical for life-sustaining functions, surgical resection is not an option and DIPGs are rarely biopsied (20). Despite radiation therapy proving to temporarily advance neurological function in patients, its effect on overall survival is negligible with only an improvement of 5.8 months progression free survival vs. 5.0 months without treatment (21). Further to radiotherapy, several groups have investigated the effects of combining specific chemotherapeutic agents with established radiotherapeutic methods. However no improvement in response rate, event-free survival, or

overall survival was detected indicating that in DIPG patients, chemotherapy is not effective in treatment (21).

In the last decade, many clinical trials including novel targeted therapies have been conducted in adult GBMs (aGBM), and then passed down to pediatric patients with limited success. Although both pediatric and adult tumors are similar under the microscope, converging evidence showed that there are major molecular distinctions between these tumors across lifespan. The next sections of this thesis aim to highlight the most relevant and recent contributions to the field from our lab and others in response to the acute need of better understanding of the pathogenicity of the tumors leading to novel therapeutic approaches.

1.2 Molecular characterization of high-grade gliomas

1.2.1 Adult astrocytomas

Since the development of high-throughput genome analysis techniques, integrative genomic and epigenetic analyses have improved our understanding of the genetic basis of HGA. In 2008, the Cancer Genome Atlas Research Network (TCGA), a large consortium led by the National Cancer Institute (NCI) and the National Human Genome Research Institute (NHGRI) and funded by the US government, reported the first comprehensive genomic analysis of 206 predominantly primary aGBM integrating data from DNA sequence alterations of 600 genes, copy number alterations, messenger RNA (mRNA) expression, microRNA (miRNA) expression, CpG methylation (23). This study identified three critical pathways altered in aGBM: the RTK/Ras/PI3K pathway in 88% of cases, the p53 signalling in 87%, and the RB signalling in 78%. From this interim analysis, TCGA published three further studies that identified and characterized distinct molecular and epigenetic subtypes. These data have later been expanded to a bigger cohort of 543 samples, mainly with the addition of

next-generation sequencing technology (24). As expected, *EGFR* was found to be one of the most mutated genes, and at least one RTK was altered in 67.3% of GBM: *EGFR* in 57.4%, *PDGFRA* in 13.1%, *MET* in 1.6%, and *FGFR2/3* in 3.2%. Concurrent with their previous report, they confirmed the dysregulation of the three already described pathways (Figure 1.4). Because this paper was published after the results presented in this thesis, DNA methylation data will be discussed in the discussion section.

In 2010, Verhaak *et al.* reported a gene expression-based molecular classification of aGBM (25), which confirmed and expanded a previous study by Phillips *et al.* (26). Consensus clustering identified four subtypes with specific expression of gene signatures: Classical, Mesenchymal, Proneural, and Neural (Figure 1.5). In the classical subtype, GBMs were characterized by chromosome 7 amplification paired with chromosome 10 loss, high rate of high-level *EGFR* amplification, focal 9p21.3 homozygous deletion targeting *CDKN2A*, and lack of *TP53* mutations. Focal deletions at 17q11.2, containing *NF1*, lower *NF1* expression level, and *NF1* mutations were the most common alterations occurring in the mesenchymal subtype along with expression of mesenchymal markers. The major features in the proneural class were alterations in *PDGFRA* and point mutations in *IDH1*. Tumors also showed high expression of oligodendrocytic development genes (*PDGFRA*, *NKX2-2*, *OLIG2*) and proneural development genes. The neural subtype was unified by the expression of neuron markers. In this study, the proneural group displayed characteristics associated with secondary GBMs including *IDH1* and *TP53* mutations (27-29) and younger age (30) with a trend toward longer survival. Later that year, Noushmehr *at al.* (31) used DNA methylation arrays (Illumina® Infinium HumanMethylation27) to profile 272 TCGA aGBM. They identified a distinct methylation subgroup within the proneural group characterized by a glioma-CpG island methylator phenotype (G-CIMP) similar to what has been described in colon cancer as

cancer-specific CpG island hypermethylation of a subset of genes (32) (Figure 1.6). This phenotype was associated with secondary GBM and with *IDH1* mutation. Patients in this subgroup had a better outcome. In their independent validation set composed of gliomas across all grades, *IDH1* mutated samples presented this methylation phenotype whatever their malignancy.

Although high-density arrays have largely contributed to the advances of knowledge in both pediatric and adult HGGs, the advent of NGS, which performs cost-effective massive parallel sequencing, has changed our approach to the investigation of human disorders, whether they are Mendelian or complex (33). In the field of cancer genomics, NGS has led to a dramatic expansion of the catalogue of somatic mutations, providing an unprecedented understanding of cancer biology. In 2008, Parsons *et al.* (28) published the first application of NGS to aGBM. They sequenced 20,661 protein-coding genes in 22 aGBM and found a novel recurrent mutation in *IDH1* in 12% of samples. Five of the 6 samples mutated were secondary GBMs. All the samples harboured the same heterozygous point mutation, a change of a guanine with an adenine at position 395 of the transcript (G395A) leading to the replacement of an arginine with a histidine at amino acid residue 132 of the protein (R132H). Patients with this *IDH1* mutation were younger and had a significant increased survival. A larger screening of *IDH* mutation in 445 CNS tumors and 494 non-CNS tumors revealed that *IDH1* mutations at residue R132 affected more that 70% of diffuse astrocytomas, anaplastic astrocytomas, and secondary GBM, as well as oligodendrogliomas and anaplastic oligoastrocytomas, but no mutation was found in pilocytic astrocytomas, ependymomas, medulloblastomas, or non CNS tumors, and they were rare in primary GBM (5%) (27). Nine of the grade II or III gliomas negative for *IDH1* presented mutations in *IDH2* at residue 172. As reported previously (28), these patients were younger (32 years vs. 59 years; $p<0.001$) with an increased overall

survival (31 months vs. 15 months; p=0.002). Further NGS of human cancers found *IDH1/2* genes to be also mutated in acute myeloblastic leukemia (AML) (34).

IDH1 encodes for $NADP^+$-dependent isocitrate dehydrogenase 1, a critical enzyme that participates in the Krebs cycle, converting isocitrate to α-ketoglutarate (α-KG). Mutant IDH1 reduces the activity of the wild-type enzyme and produces the oncometabolite 2-hydroxyglutarate (2-HG). These mutations are considered as gain-of-function mutations (35, 36). Interestingly, 2-HG is known to accumulate in L-2-hydroxyglutaric aciduria caused by the deficiency of 2-hydroxyglutarate dehydrogenase. Patients affected by this autosomal recessive disorder develop leukoencephalopathy and have an increased risk of developing brain tumors (37). *IDH1*-mutant gliomas have been shown to have elevated levels of hypoxia inducible factor 1α (HIF-1α), a major transcription factor involved in cell response to hypoxia and participating in key pathways in oncogenesis such as angiogenesis, glucose metabolism, or cell invasion (36).

Following the identification of the subgroup of tumors with a G-CIMP profile and harbouring *IDH1* mutations tumors within the proneural group as described above (31), several studies shed light on the pathogenicity of *IDH1/2* mutations. Turcan *et al.* (38) showed that this unique mutation is sufficient to establish this specific DNA profile associated with extensive, coordinated hypermethylation at specific loci. Working on primary human astrocytes, they were able to induce the remodelling of the methylome by introducing *IDH1* mutation into the cells. This occurred via alteration of histone marks (increased levels of H3K9me2, H3K27me3 and H3k36me3) and DNA hypermethylation. As mentioned earlier, *IDH* mutation leads to 2-HG accumulation, and therefore blocks several α-KG-dependent dioxygenases, including histone demethylases and the TET family of 5-methylcytosine (5mC) hydroxylases, through competitive inhibition (39). TET proteins are a group of 5-methylcytosine

hydroxylases that convert 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), using α-KG as a cosubstrate (40). Thus, the TET CIMP pattern of these cells is explained by decreased levels of 5hmC due to the inhibition of TET, key player in the DNA methylation pathway (39). Interestingly, no mutation in TET family members have been found in gliomas as it has been in myeloid malignancies (41). Another pathway to tumorigenicity was demonstrated by Craig Thompson's group and was published in the same issue of *Nature* than the study from Turcan *et al.* (38). He showed *in vitro* that *IDH* mutation impaired histone demethylation and induced a block to cell differentiation. A significant increase in repressive histone methylation marks was associated to this.

1.2.2 Pediatric GBM and DIPG

While considerable information is available for adult GBMs, substantially fewer genetic/molecular data exist in children due to limited clinical material. Based on similar histology, current treatments in children are driven from adult studies and similarly show little therapeutic success (15). This is despite the presence of numerous clinical trials associating conventional therapies (surgery, radiation and chemotherapy mainly temozolomide) in association to adjuvant therapies including anti-angiogenic and/or anti-receptor kinase agents in adults. Today, GBM is still the most treatment-refractory CNS tumor in pediatrics and adults with a 5-year overall survival rate ranging from 5 to 20% in children (4, 15, 42).

Despite sharing the same histopathologic characteristics, pGBMs are molecularly distinct from adult primary and secondary GBM (15). In children, GBMs develop *de novo* in the vast majority of cases. *TP53* mutations are observed in one third of patients (43) while overexpression of the protein regardless of its mutational status is considered to correlate with poor outcome in contrast to adult GBM where this marker is of no prognostic value (44). EGFR is frequently overexpressed, in up to 80% of samples, however only 7% have gene

amplification in children (45) indicating a different mechanism for the up-regulation of this RTK in pGBM. EGFR can be also mutated and the most common mutant found in about 30% of aGBM (46), EGFRvIII, is quite rarely identified in pGBM (2%) (47). EGFRvIII is a receptor with a deletion of amino acid residues 6–273 in its extracellular domain, resulting in a truncated receptor with a non-functional ligand binding pocket and absence of the dimerization arm (48). In spite of its inability to bind any ligand, the receptor is constitutively active, and is able to activate Ras-Erk1/2 (49) and PI3K-Akt pathways (50).

In 2007, our laboratory was the first to publish gene-expression profiling of pGBM (51). Principal component analysis (PCA) of a cohort of 32 pediatric and 7 adult tumors (Figure 1.7) revealed that aGBM clustered separately from pGBM and that pGBM could be divided in two subgroups, one with Ras/Akt activity and a poor prognosis compared to the second one with non-active Ras/Akt. We further identified potential candidate genes uniquely involved in pGBM oncogenesis, including *YB1* and *SNX3*. Studying SNX3 is one of the aims of this thesis and will be discussed later in this introduction. These findings were reproduced in an independent study using formalin fixed-paraffin embedded (FFPE) pGBM samples, highlighting the utility of archival FFPE tissues as they are more widely available and often accessible with clinical and follow-up data (52).

Further to our gene-expression profiling (GEP) study, copy number profiling studies confirmed that pGBMs harboured a distinct spectrum of genomic alterations compared to aGBMs (53-55). The main study from Paugh *et al.* (55) integrated the results of single-nucleotide polymorphism (SNP) microarray and gene expression microarray analyses from 78 *de novo* pediatric HGGs (pHGG) including 7 DIPGs and 10 radiation-induced pHGGs. Although the majority of samples showed multiple genomic imbalances, 19% lacked large-scale copy number changes. This was also seen in another study published shortly after (54).

Dr. Suzanne Baker's group found that *PDGFRA* was the predominant focal amplification (12%) within a dysregulated PDGFRα pathway, even in many tumors lacking this gene amplification. Because of that, the PDGFRα pathway seemed critical to pHGG genesis and was likely activated by multiple mechanisms. The most significant focal homozygous deletion encompassed *CDKN2A/CDKN2B* in 19% of tumors. Compared to aGBM, pGBMs showed frequent gain of chromosome 1q (30%) and less frequent chromosome 7 gain (13%) and 10q loss (35%). None of the pGBM harboured *IDH1* mutation (55).

Aside from pGBM, the severe problem in terms of rarity of tissue samples for research purposes is even more critical in DIPGs, which are rarely biopsied at diagnosis. Recently, several copy-number and gene expression studies have elucidated candidate oncogenic alterations in DIPGs. One of the main findings reported by Zarghooni *et al.* (56) was a recurrent amplification of *PDGFRA* in 36% of DIPGs, as well as other receptor tyrosine kinases to a lesser extent. This was later confirmed by two other studies (57, 58).

Using next-generation sequencing, we reported for the first time in humans recurrent mutations in the replication-independent histone 3 variant 3 (*H3F3A*) in pGBM (1). These mutations lead to amino-acid changes in two key residues (K27 and G34), encoding p.Lys27Met and p.Gly34Arg or p.Gly34Val alterations. These residues are known to be involved in histone post-transcriptional modifications, and the mutations were specific to pediatric and young adult GBM. They significantly overlapped with mutations in *TP53* and in alpha thalassemia/mental retardation syndrome X-linked gene (*ATRX*). The protein encoded by *ATRX* is a SWI/SNF-like chromatin remodelling protein, and mutated tumors display abnormal telomeres characteristic of alternative lengthening of telomeres (59) (Figure 1.8). At the same time, Dr. Suzanne Baker's group from the Saint Jude Children's Research Hospital reported histone 3 K27M mutations in DIPG in *H3F3A* and *HIST1H3B* (encoding H3.1)

genes thus confirming our findings regarding this critical mutation (60). This finding unravelled a pivotal role of epigenetic dysregulation in pGBM. Since then, further mutations in histone 3 have been reported (61) in bone and cartilage tumors: 95% of chondroblastomas presented either *H3F3B* p.Lys36Met mutation or *H3F3A* p.Lys36Met mutation, 92% of giant cell tumors of bone presented either a *H3F3A* p.Gly34Trp or p.Gly34Leu. *H3F3A* or *H3F3B* mutations were also reported in clear cell chondrosarcoma, osteosarcoma and conventional sarcoma (61). Interestingly, histone H3.3 variant is expressed throughout the cell cycle as well as in quiescent cells, whereas the "DNA synthesis-coupled" histone H3.1 is produced during the S phase and incorporated *de novo* in newly replicated chromatin and during DNA repair. Therefore, unsurprisingly, high levels of nucleosomal H3.3 can be found in post-mitotic cells such as neurons (62). Histone H3 variants show distinct patterns of localization though the genome and participe in chromatin dynamics through the timing and sites of their incorporation. This is promoted by dedicated histone chaperones. H3.3 deposition occurs on DNA sequences that are transiently nucleosome free, during transcription and DNA repair, but also on heterochromatic subtelomeric and pericentric regions (63). Several studies have shed light on the critical role played by H3.3 during many developmental processes such as gametogenesis and fertilization (63). Overall, our lab has identified critical mutations in the regulatory histone H3.3 and in the complex DAXX/ATRX providing an example of interplay of genetic and epigenetic events in driving cancer. Further studies are needed to better understand how these mutations lead to tumor formation and to identify targets that could improve the prognosis of this dismal cancer.

1.3 Receptor tyrosine kinase in pediatric gliomas

1.3.1 Overview of receptor tyrosine kinases

Receptor Tyrosine Kinases (RTKs) play a prominent role in the control of a variety of intracellular processes including many metabolic and physiological processes in the brain, such as proliferation and differentiation, cell survival and metabolism, cell migration, and cell-cycle control (64). So far, 58 RTKs have been described in humans and are organized into 20 families. They all have a similar molecular architecture, with ligand-binding domains in the extracellular region, a single transmembrane helix, and a cytoplasmic region that contains the protein tyrosine kinase (TK) domain plus additional carboxy (C-) terminal and juxtamembrane regulatory regions. Their structure, mechanism of activation and signalling pathways triggered are highly conserved in evolution from the nematode *Caenorhabditis elegans* to humans consistent with the key regulatory roles that they play. Since their discovery in the 50s, many diseases have been described to result from RTK alteration, including genetic changes and abnormal activation or regulation. Mutations in RTKs and aberrant activation of their intracellular signalling pathways have been causally linked to cancers as well as diabetes, inflammation, severe bone disorders, and arteriosclerosis (65). In physiological conditions, RTKs are activated by growth factor binding to the extradomain of the protein, inducing their dimerization or oligodimerization. This conformational change leads to the activation of the intracellular tyrosine kinase domain, enabling the recruitment of signalling proteins to initiate a downstream signalling cascade. In the case of EGFR, its activation will stimulate proliferation through the Ras pathway but also cell survival through the Akt pathway, among other functions. Moreover, these pathways are interconnected in a complex and dynamic network where RTKs function as key regulatory nodes (Figure 1.9) (65). In parallel, this process is balanced by the simultaneous recruitment of endocytic adaptor

proteins, which enhance the internalization of RTKs into the endocytic trafficking network, thereby allowing for their removal from the cell surface (Figure 1.10) (66).

Dysregulated activation of many RTKs through mechanisms including mutation, overexpression, structural rearrangements, disruption of autocrine/paracrine loops, and inactivation of regulatory constraints is implicated in multiple human neoplasias (67). More precisely, most chromosomal translocations result in RTK variants that are located in the cytosol and are activated independently of ligand binding. These variants fail to enter the endocytic pathway and thus escape downregulation through lysosomal degradation (68). In aGBM, somatically acquired mutations of EGFR confer to this RTK the ability to evade ubiquitin-mediated sorting to degradation (69). In addition to mutations/translocations that prevent RTKs from an efficient downregulation, there is growing evidence to support a role for disruption of the endocytic pathway in oncogenesis. In fact, the major mechanism for downregulation of EGFR is the endocytic pathway (70). Receptor trafficking plays a role in the regulation of receptor signalling by controlling the magnitude or the specificity of the response (71, 72), thus the targeting to lysosome is important to prevent a sustained activation that can lead to cell transformation (67). Moreover, RTKs can still signal from endosomes after internalisation, and this signalling will end when these RTKs are dephosphorylated and degraded (68). Therefore, aberrant endocytosis that could delay the internalization, trafficking and/or degradation of RTKs may promote oncogenesis through sustained signalling. However, we do not have yet a clear molecular understanding of how each signalling pathway impacts on RTK signalling and stability and how this contributes to cancer progression.

Since the discovery of epidermal growth factor (EGF) (73), its receptor (EGFR) and other receptor tysoine kinases in the 70s, their role in CNS development has been extensively studied. It is now well known that RTKs play a critical role in the development and

subsequent maintenance of all cell types in the CNS as well as their precursors. EGFR and its ligands are expressed from embryogenesis, throughout brain development and into adulthood. *In vitro* and *in vivo* studies have shown them to be involved in proliferation, migration, differentiation and survival in all CNS cell types.

In cancer, somatically acquired mutations in RTKs are widespread; several studies on pediatric and adult GBMs highlighted their importance in gliomagenesis (15, 74). As previously discussed in this introduction, TCGA project on adult glioblastoma has shown that 90% of them have an alteration on RTKs or downstream pathways (24), the most commonly affected being EGFR (57%) and PDGFRα (10%).

EGFR is the best characterized RTK in aGBM. It is mostly expressed in primary GBM, with about 40% of tumors with gene amplification. As mentioned earlier, the most common genetic rearrangement is the deletion of exons 2-7 leading to EGFRvIII (75). Interestingly, almost half of the tumors harbouring *PDGRFA* alteration also harboured concurrent *EGFR* alterations, as well as the majority of the tumors with *MET* alterations (24). Based on these observations, as well as numerous previous studies that have highlighted *in vitro* and *in vivo* the role of EGFR and EGFRvIII in GBM oncogenesis, clinical trials including tyrosine kinase inhibitors (TKI) have been conducted since the early 2000 with limited efficacy (76). Interestingly, a recent study has shown that EGFR was dephosphorylated after gefitinib treatment but without any changes in the phosphorylation of downstream targets, thus emphasizing the fact that most likely there were other sources of activation of these targets in recurrent aGBM (77).

In pediatric high-grade astrocytomas, it is known that EGFR is overexpressed since the late 90s, although genetic alterations are not frequent (45). In addition to EGFR overexpression,

one study uncovered an overexpression of additional RTKs in pGBM including PDGFRα (45%) and PDGFRβ (32%) in a cohort of 42 patients using immunohistochemistry (IHC) (47). Overexpression of these RTKs did not confer a prognostic factor to the patients. MET overexpression has not been investigated to date in pGBM, however in adults, its overexpression was detected in 29% of 62 patients and was associated with poor prognosis (78). In a recent integrated analysis of pGBM, 19% of 58 tumors showed *PDGFRA* amplification, 1.7% *MET* amplification, and no tumor showed *EGFR* amplification (55). Even though *PDGFRA* amplification rate is high in this cohort and was more frequent in relapsed GBM, the authors reported that many tumors that lacked amplification of *PDGFRA* still showed PDGFRα overexpression (55), confirming that overexpression of RTK in pGBM is not only due to gene amplification.

1.3.2 Endosomal trafficking of RTKs

As previously discussed, dysregulation of RTK signalling plays a critical role in cell proliferation and tumor formation in aGBM. Moreover, molecular analysis has shed light on a potential role for this signalling in pediatric gliomagenesis. In normal cells, one major player of signal regulation is the intracellular trafficking machinery, which contributes to spacial and temporal control of RTK signalling. The three major points of regulation are endocytic internalization of the activated receptor, its lysosomal degradation and/or its recycling back to the plasma membrane (Figure 1.10A). The lysosomal degradation of the receptor will result in the termination of the signalling process, whereas being recycled back to the cell membrane will result in a sustained signalling if the ligand is still available. Furthermore, increasing evidence has shown that activated RTKs can still signal after internalization into intracellular compartments (79).

After ligand stimulation, the RTK changes its conformation, leading to its dimerization or oligodimerization, and phosphorylation. This initiates downstream signalling cascades. The receptor enters then the endocytic system either via the clathrin-mediated pathway or clathrin-independent pathway (80). The classic clathrin-dependent endocytic pathway is characterized by the recruitment of soluble clathrin from the cytoplasm to the plasma membrane, forming coating pits then giving rise to clathrin-coated vesicles. After endocytosis, these vesicles are uncoated and fuse with early endosomes which are highly enriched in phosphatidylinositol-3-phosphate (PtIns(3)P) (81). For a long time, this pathway was thought to be the major pathway for internalization of EGFR. However, emerging data have identified and characterized novel clathrin-independent pathways involved in EGFR trafficking, including caveolar endocytosis (82, 83). Several studies support the concept that a cell can exploit a specific internalization route for a given receptor depending on the amount of ligand, therefore controlling cellular response (84, 85).

From the early endosomes, RTKs are sorted either to late endosomes/multivesicular bodies and lysosomes for degradation, or recycled back to the plasma membrane via recycling endosomes (Figure 1.10A). Modification of the receptors during endocytic trafficking, including dephosphorylation, ubiquitylation, and dissociation from the ligand, contributes to signal attenuation.

The concept of signalling endosomes originates initially from studies in neurons (86) where clathrin-coated vesicles and early endosomes were shown to carry activated TrkA receptors interacting with signalling molecules of ERK1/2 and PI3K/Akt pathways (87). Furthermore, functional importance of signalling from early endosomes was provided by studies using a system where reversibly inhibited EGFRs were subsequently activated in endosomes. These activated receptors were able to recruit signalling molecules and induce cell proliferation and

migration (88). Although late endosomes are generally the place of signal attenuation, they still can contain activated receptors and their downstream effectors (Figure 1.10B) (89).

RTKs control many fundamental cell behaviours, and their implication in a majority of cancers highlights their importance in cell decision-making. They are tightly controlled by several modes of regulation. The best-known ones are their production both at a transcriptional and post-transcriptional level, and ligand availability. More specifically, in cancer, aberrant RTK activation is often caused by gene amplification, receptor overexpression, autocrine activation or gain-of-function mutations. However, dysregulation of their endocytosis is clearly an emerging mechanism of oncogenic activation in some human cancers. One interesting example is EGFRvIII, the most frequent mutant in aGBM. Its oncogenic role is not only due to its constitutive activation but also because of its insufficient internalisation and ubiquitination, and therefore delayed degradation (90). Dysregulation of sorting proteins is also a main actor in altered trafficking. Recently, GGA3 (Golgi-localized g-ear-containing Arf-binding protein 3) has been described as a key factor of Met recycling independently of its ubiquitination status. *In vitro*, GGA3 depletion has been shown to increase MET degradation upon HGF stimulation, impairing MAPK activation and cell migration (91).

1.3.3 Sorting nexin family

Sorting nexins (SNXs) are a family of hydrophilic proteins involved in the regulation of cellular trafficking. They are unified by the presence of a phosphoinositide-binding motif, known as the SNX phox homology (PX) domain. This domain aids in the targeting of the SNX protein to phosphoinositide-enriched membranes such as the sorting endosomal compartment (92). Sorting nexins function in diverse processes, including endocytosis, endosomal sorting and endosomal signalling (93, 94). But precise functions of many SNXs

are not yet known and their cellular regulation has not been addressed (94). However some sorting nexins are involved in cancer progression (95-97). SNX1 was the first mammalian SNX to be characterized and its overexpression has been shown to decrease activated EGFR on the cell surface (98). SNX1 also interacts with other members of the receptor tyrosine kinase family including PDGFR (99). More recently, SNX1 was described to be down-regulated in 75% of colon cancer by immunohistochemical staining (71) and this was mediated by the direct control of microRNA-95 (97). In addition, *in vitro*, colon cancer cell lines with a down-regulation of SNX1 show increased proliferation, decreased apoptosis, which promote tumorigenicity (95). In contrast, SNX5 overexpression inhibits EGFR degradation. SNX5 interacts with SNX1 but its role in EGFR degradation is independent of this interaction. However, SNX1 overexpression is able to attenuate the effect of SNX5, therefore these two proteins may play antagonistic roles in regulating trafficking of EGFR (100). Overexpression of SNX13 inhibits EGFR degradation *in vitro* (101) and SNX13-null mice are embryonic lethal with growth retardation and defects in neural tube closure (102). In B-cell acute lymphoblastic leukemia, *SNX2* was recently described as a novel fusion partner of *ABL1* (103).

Phosphoinositides that specifically bind SNXs are considered as regulators of signalling pathways (104). PtIns(3)P are produced by class III phosphatidylinositol-3-kinase (PI3K) or Vps34 and are involved in many signalling events including regulation of endosomal trafficking (105). Vps34 is located mainly on intracellular membranes (106). Frequent mutation or deletion of components of the Vps34 complexes have been found in breast, ovarian and colon cancer (107) but their role in oncogenesis remains unclear (107).

Our group has published the first study of gene expression profiling in pGBM (51) and has confirmed that, even though they share common aberrations of major signal transduction and

cell cycle pathways, pediatric tumors are molecularly distinct from their adult counterparts. Transcriptional profiling analyses distinguished two genetically and prognostically different subsets of pGBM based on their differing association with active Ras/Akt pathways (51, 52).

We identified Sorting Nexin 3 (SNX3) to be specifically overexpressed in pGBM associated with Ras/Akt pathway activation (51). SNX3 belongs to the sorting nexin family and is involved in the regulation of endosomal trafficking of membrane receptors, including RTKs (99, 100, 108-112). SNX3 is a small protein of 162 amino acids and its PX domain binds directly and selectively to phosphatidylinositol-3-phosphates (PtIns(3)P) on early endosomes (108, 113). Moreover, the generation of PtIns(3)P is required for SNX3 membrane recruitment (108, 114). Overexpression of SNX3 leads to the expansion of tubular-vesicular structures that are characteristic of early, recycling, and late endosomes and delays EGFR degradation with internalized complexes being retained in the swollen structures rather than being targeted to the lysosome for degradation (108, 110). Up to now, no link has been shown between SNX3 and cancer.

1.3.4 Other mechanisms of regulation

Over the last several years, targeted therapies were developed mainly towards RTK inhibitors, and have proven to be inefficient in adult and pediatric GBMs, in spite of supporting evidence of the role of RTKs in their oncogenesis (115). We believe that several mechanisms may lead to tumor resistance to RTK inhibitors, including kinase switching mediated by SNX3, and/or chromatin remodelling alterations in gliomagenesis.

1.4 Rationale

As described throughout this introduction, recent advances in our knowledge of both pediatric and adult glioblastomas have shed light on their molecular differences (116). RTKs, mostly

EGFR, play a major role in GBM formation, as shown with cumulative evidence in *in vitro* and mouse models. However, development of targeted therapies against RTKs has proven to be insufficient to improve the prognosis of affected adult or pediatric patients. In the specific case of pGBM, one major difference is that, although they are often overexpressed, most of RTKs present limited genetic alteration. Thus, understanding what dysregulates RTKs in pediatric tumors is critical to propose adequate therapies overcoming them. One critical mechanism of regulation is the endocytic pathway. Evidence of RTK disrupted trafficking has been shown to promote tumorigenesis in many types of tumors, including GBMs. From the gene expression profiling we generated (51), we identified Sorting Nexin 3 as a potential protein involved in pGBM formation where it was specifically overexpressed in the subset of tumors with an activation of Ras/Akt pathways. Given the functions of SNX3 in endosomal trafficking and regulation of membrane receptors including EGFR (108) and its specific overexpression in pGBM with an activated Ras/Akt pathway, it may play a major role in the regulation of RTK trafficking in pGBM. Therefore, SNX3 may explain why RTKs are overexpressed in pGBM without gene amplification but also resistance to targeted therapies against RTK through kinase switching.

Using next-generation sequencing, we and others (1, 60) have described recurrent histone variant H3.3 point mutations (K27M and G34R/V aminoacid changes) in one third of pediatric glioblastomas and almost 80% of DIPGs. Interestingly, histone mutations were not found in aGBM, adding further evidence that adult and pediatric GBM have distinct molecular profiles and mechanisms of progression. Interestingly, mutations in *H3F3A* and *IDH1/2* are key drivers that alter chromatin-remodelling regulation. As previously discussed in the introduction of this thesis, recent publications have described the role of *IDH* in histone mark regulation, with *IDH* mutations impairing histone demethylation, resulting in a block to

cell differentiation (117). *IDH1* mutations also induce genome-wide alterations in DNA methylation, described as the CpG island methylator phenotype (38), characteristic of proneural glioblastomas (31). Based on this evidence, we believe that *H3F3A* mutations are the pediatric counterpart of *IDH* mutations in young adult GBM. Because of the crucial role of DNA methylation in the regulation of gene expression and chromatin organization and because of the availability of genome-scale DNA methylation screening technologies, we used DNA methylation analysis as the first step to study epigenetic alterations in our tumors of interest. Even though not fully understood, there is an interdependent relationship between DNA methylation and histone marks including the one at lysine 27 of histone H3 (118-120).

1.5 Thesis hypotheses and objectives

Pediatric and adult GBM are molecularly distinct and this thesis aims to better characterize pediatric GBMs in order to identify potential targets. We focused on the two main themes that have been identified to play a driver role in pediatric tumorigenesis: RTK signalling and histone mutations. Therefore, this thesis will focus on investigating the trafficking of RTK in pGBM and better characterize tumors with *H3F3A* mutations. Our working hypotheses are:

1. SNX3 overexpression delays the degradation of multiple RTKs in pediatric GBMs, leading to a sustained activation of downstream signalling pathways, thus promoting cell proliferation and tumor formation. This includes EGFR, which can be trapped within early endosomes where it potentially continues to signal, leading to sustained activation of intracellular pathways (Ras pathway, etc.). Other RTKs, including PDGFR which is known to be overexpressed in high-grade pediatric astrocytomas, may additionally be affected by the same mechanism; this may also hold true for RTKs which are yet to be investigated, such as MET (47, 121). The sustained coactivation of multiple RTKs could decrease the efficiency of

targeted therapy against a single receptor like EGFR (122-124), and SNX3 may play a role at a critical point in RTK regulation. We hypothesize that its overexpression supports the concept of kinase switching.

2. *H3F3A* mutations are a hallmark alteration in pHGG including DIPG, distinguishing them from their adult counterparts. These mutations are specific of the tumor location within the brain. Tumors can be subgrouped based on molecular characteristics.

The objectives of this thesis were as follows:

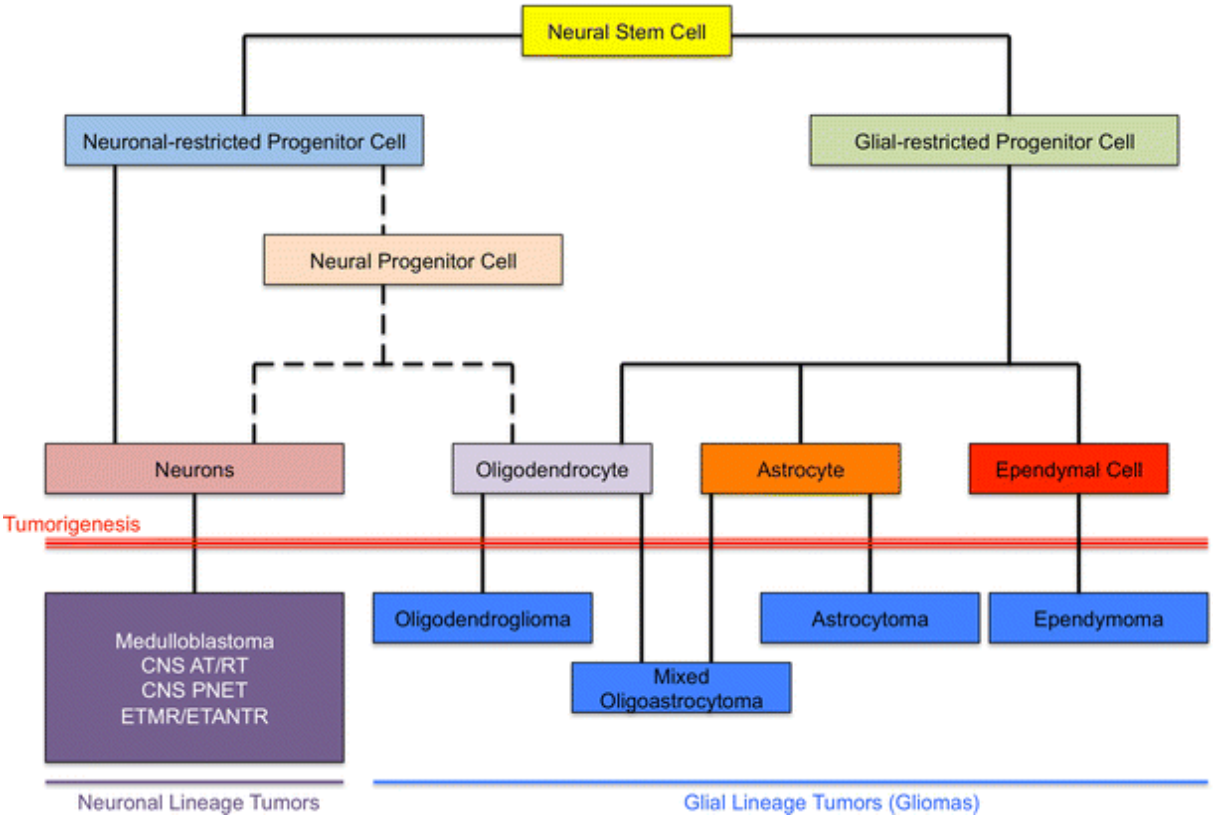
1. To determine the role of SNX3 overexpression in the disruption of RTK trafficking, promoting cell proliferation and tumor formation by delaying RTK degradation, hence sustaining their signalling, using both pediatric and adult GBM cell lines. *In vivo* studies in immunocompromised mice will further confirm the role of SNX3 in tumor formation.

2. To better characterize genetic alterations in DIPG by investigating the frequency of *H3F3A* and *HIST1H3B* mutations as well as *ATRX* mutations. Furthermore, we aim to assess whether *ATRX* mutations overlap with histone H3.3 and/or *TP53* mutations similar to our findings in supratentorial GBM (1) and to investigate the clinical and biologic features of DIPG subgroups based on histone H3.3 mutation status.

3. To investigate the heterogeneity of glioblastoma across the entire age spectrum, and to elucidate the impact of *H3F3A* mutations on the GBM epigenome using genome-scale DNA methylation arrays.

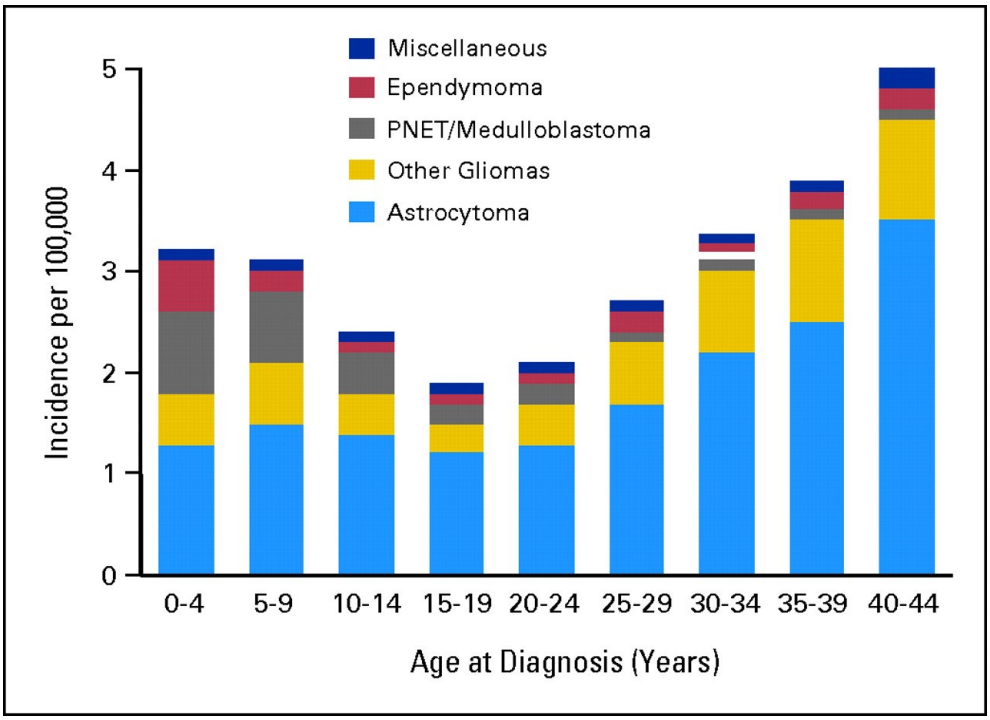
1.6 Figures

Figure 1.1 – Central nervous system cellular development and tumorigenesis



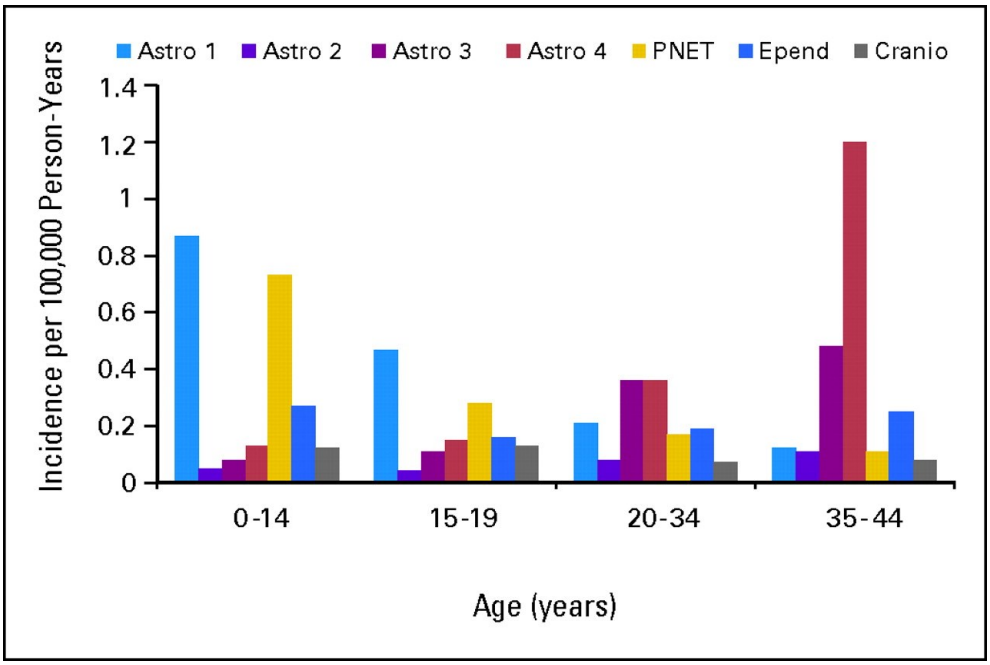
Graphic depiction showing differentiation of neural stem cells into neuronal and glial differentiation pathways and subsequent tumorigenesis from presumed cells of origin. *CNS* Central nervous system, *PNET* primitive neuroectodermal tumor, *AT/RT* atypical teratoid/rhabdoid tumor, *ETMR* embryonal tumor with multilayered rosettes, *ETANTR* embryonal tumor with abundant neuropil and true rosettes. Adapted with permission from (125).

Figure 1.2 – Gliomas, including astrocytomas, are the largest subgroup of brain tumor across the lifespan



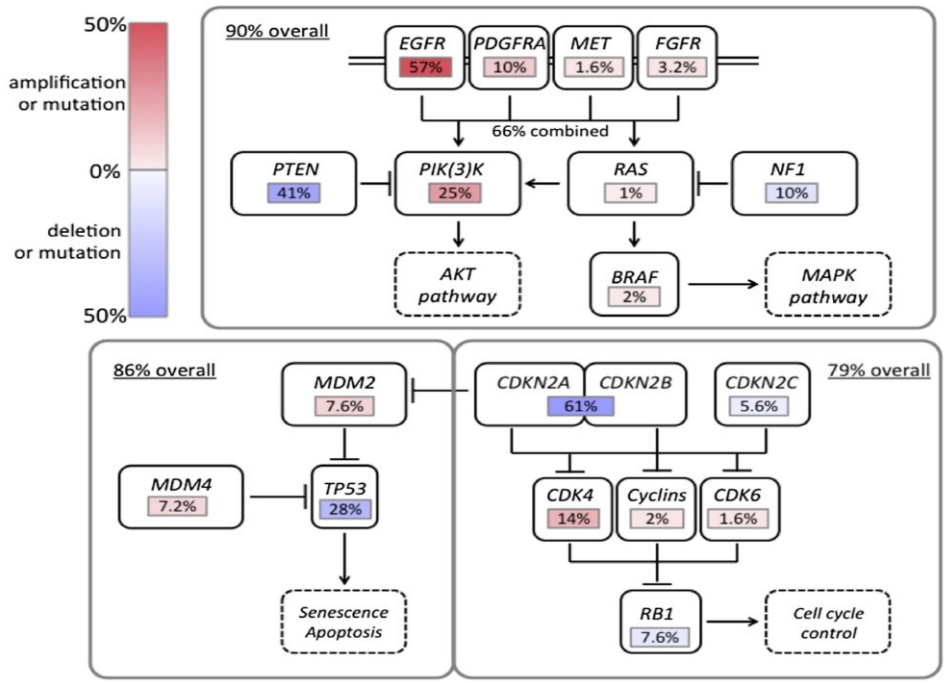
US cancer incidence from Surveillance, Epidemiology, and End Results (SEER), 1975 to 1998, by CNS tumor type. Broadly, glial tumors include astrocytomas, ependymomas, and other (which includes oligodendroglioma). Neural tumors include medulloblastoma, CNS primitive neuroectodermal tumors (PNETs), and pineoblastoma. Miscellaneous category includes germ cell tumors, craniopharyngioma, and choroid plexus tumors. Adapted with permission from (6).

Figure 1.3 – Distribution of astrocytomas by subtype across the age spectrum



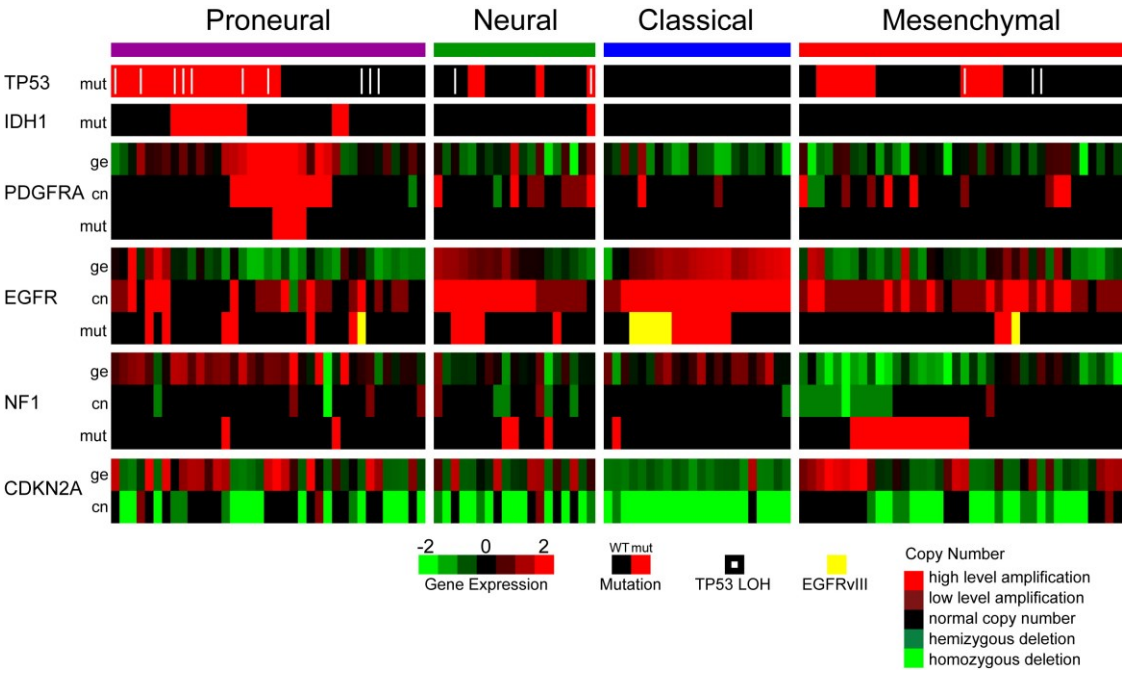
Incidence of astrocytomas (Astro; grade I, II, III, and IV), medulloblastoma/CNS primitive neuroectodermal tumors (PNETs), ependymoma (Epend), and craniopharyngioma (Cranio), from the Central Brain Tumor Registry of the United States, 1997 to 2001. The majority of CNS tumors decrease steadily through infancy, childhood, adolescence, and into adulthood. Grade II, III, and IV astrocytomas are uncommon in children and increase in incidence steadily into old age. Adapted with permission from (6).

Figure 1.4 – Landscape of pathway alterations in GBM



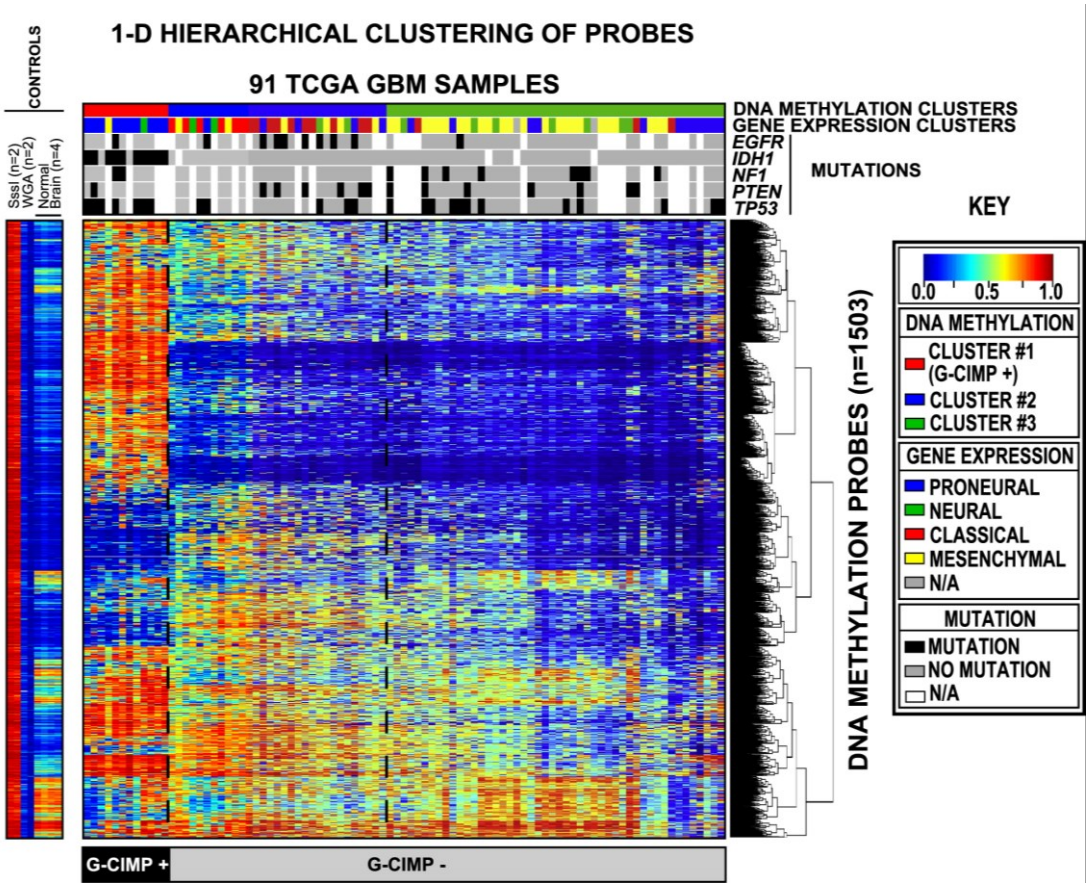
Overall alteration rate is summarized for canonical PI3K/MAPK, p53 and Rb regulatory pathways. Alterations affecting canonical signal transduction and tumor suppressor pathways are summarized for 251 GBM with both exome sequencing and DNA copy-number data. Rearrangements are underestimated in this summary since RNA-seq data were available for only a subset of cases with exome sequencing data (153/291, 61%). Adapted with permission from (24).

Figure 1.5 – Integrated view of gene expression and genomic alterations across glioblastoma subtypes



Gene expression data (ge) were standardized (mean equal to zero, standard deviation equal to 1) across the 202 data set; data are shown for the 116 samples with both mutation and copy number data. Mutations (mut) are indicated by a red cell, a white pipe indicates loss of heterozygosity, and a yellow cell indicates the presence of an EGFRvIII mutation. Copy number events (cn) are illustrated by bright green for homozygous deletions, green for hemizygous deletions, black for copy number neutral, red for low-level amplification, and bright red for high-level amplifications. A black cell indicates no detected alteration. Adapted with permission from (25).

Figure 1.6 – Clustering of TCGA aGBM tumors and control samples identifies a CpG Island Methylator Phenotype

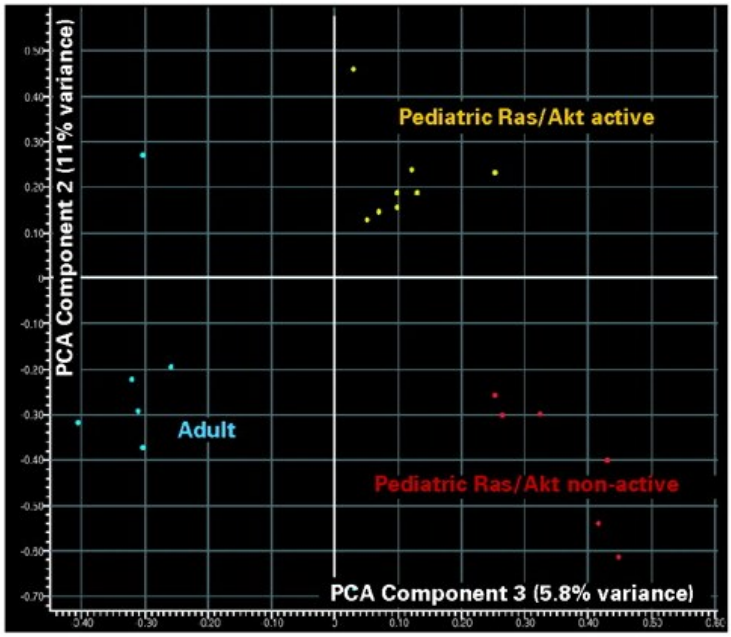


Unsupervised consensus clustering was performed with the 1503 Infinium DNA methylation probes whose DNA methylation beta values varied the most across the 91 TCGA GBM samples. DNA methylation clusters are distinguished with a color code at the top of the panel: red, consensus cluster 1 (n = 12 tumors); blue, consensus cluster 2 (n = 31 tumors); and green, consensus cluster 3 (n = 48 samples). Each sample within each DNA Methylation cluster are color labeled as described in the key for its gene expression cluster membership (proneural, neural, classical, and mesenchymal). The somatic mutation status of five genes (*EGFR*, *IDH1*, *NF1*, *PTEN*, and *TP53*) are indicated by the black squares, the gray squares indicate the absence of mutations in the sample, and the white squares indicate that the gene was not

screened in the specific sample. G-CIMP-positive samples are labeled at the bottom of the matrix.

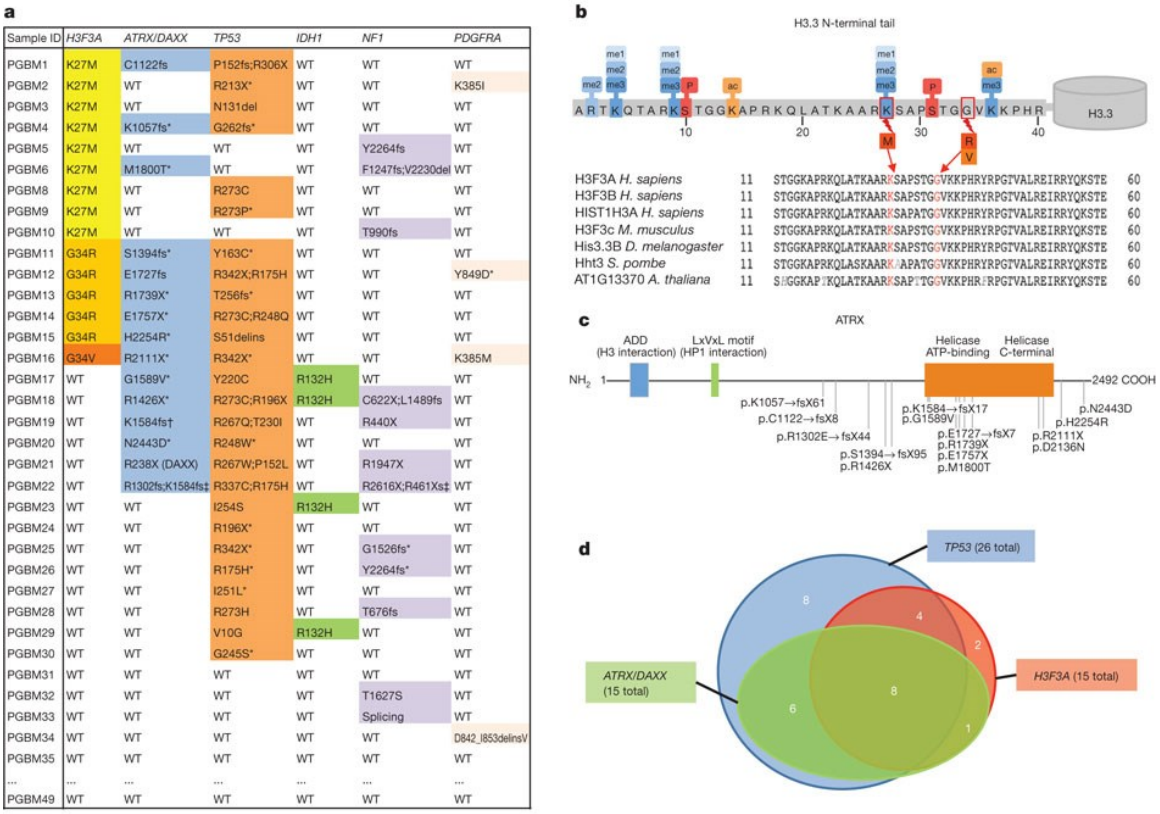
Each row represents a probe; each column represents a sample. The level of DNA methylation (beta value) for each probe, in each sample, is represented with a color scale as shown in the legend; white indicates missing data. M.SssI-treated DNA (n = 2), WGA-DNA (n = 2), and normal brain (n = 4) samples are included in the heatmap but did not contribute to the unsupervised clustering. The probes in the eight control samples are listed in the same order as the y axis of the GBM sample heatmap. Adapted with permission from (31).

Figure 1.7 – Molecular profiling identifies two subsets of pediatric GBMs distinct from adult GBMs



The 14 pediatric (p-) and seven adult (a-) glioblastoma (GBM) samples were subjected to a principal components analysis (PCA) based on the expression profile measured on 15,068 individual probes. The figure is a two-dimensional plot of PCA components 2 and 3, which resulted in a clear differentiation between aGBM and pGBM and can also distinguish the Ras scores of the pediatric tumors. Adapted with permission from (51).

Figure 1.8 – Whole-exome sequencing identifies mutations in histone H3.3 and chromatin remodelling genes in pediatric glioblastomas



a

Sample ID	*H3F3A*	*ATRX/DAXX*	*TP53*	*IDH1*	*NF1*	*PDGFRA*
PGBM1	K27M	C1122fs	P152fs;R306X	WT	WT	WT
PGBM2	K27M	WT	R213X*	WT	WT	K385I
PGBM3	K27M	WT	N131del	WT	WT	WT
PGBM4	K27M	K1057fs*	G262fs*	WT	WT	WT
PGBM5	K27M	WT	WT	WT	Y2264fs	WT
PGBM6	K27M	M1800T*	WT	WT	F1247fs;V2230del	WT
PGBM8	K27M	WT	R273C	WT	WT	WT
PGBM9	K27M	WT	R273P*	WT	WT	WT
PGBM10	K27M	WT	WT	WT	T990fs	WT
PGBM11	G34R	S1394fs*	Y163C*	WT	WT	WT
PGBM12	G34R	E1727fs	R342X;R175H	WT	WT	Y849D*
PGBM13	G34R	R1739X*	T256fs*	WT	WT	WT
PGBM14	G34R	E1757X*	R273C;R248Q	WT	WT	WT
PGBM15	G34R	H2254R*	S51delins	WT	WT	WT
PGBM16	G34V	R2111X*	R342X*	WT	WT	K385M
PGBM17	WT	G1589V*	Y220C	R132H	WT	WT
PGBM18	WT	R1426X*	R273C;R196X	R132H	C622X;L1489fs	WT
PGBM19	WT	K1584fs†	R267Q;T230I	WT	R440X	WT
PGBM20	WT	N2443D*	R248W*	WT	WT	WT
PGBM21	WT	R238X (DAXX)	R267W;P152L	WT	R1947X	WT
PGBM22	WT	R1302fs;K1584fs‡	R337C;R175H	WT	R2616X;R461Xs‡	WT
PGBM23	WT	WT	I254S	R132H	WT	WT
PGBM24	WT	WT	R196X*	WT	WT	WT
PGBM25	WT	WT	R342X*	WT	G1526fs*	WT
PGBM26	WT	WT	R175H*	WT	Y2264fs*	WT
PGBM27	WT	WT	I251L*	WT	WT	WT
PGBM28	WT	WT	R273H	WT	T676fs	WT
PGBM29	WT	WT	V10G	R132H	WT	WT
PGBM30	WT	WT	G245S*	WT	WT	WT
PGBM31	WT	WT	WT	WT	WT	WT
PGBM32	WT	WT	WT	WT	T1627S	WT
PGBM33	WT	WT	WT	WT	Splicing	WT
PGBM34	WT	WT	WT	WT	WT	D842_I853delinsV
PGBM35	WT	WT	WT	WT	WT	WT
...	...	...	...	...	...	...
PGBM49	WT	WT	WT	WT	WT	WT

*Homozygous mutations.
†Sample PGBM19 additionally has a DAXX mutation C629Sfs, whereas PGBM21 has no ATRX mutation but has the DAXX mutation shown.
‡Sample PGBM22 has a third ATRX mutation, p.D2136N, and a third NF1 mutation, p.A887T.

(A) Most frequent somatic mutations in 48 paediatric glioblastoma tumors. Mutations identified in genes listed in this table were confirmed by Sanger sequencing, and were not present in dbSNP nor in the 1000 Genomes data set (October 2011), except for the *TP53* SNP at R273, which is associated with cancer.

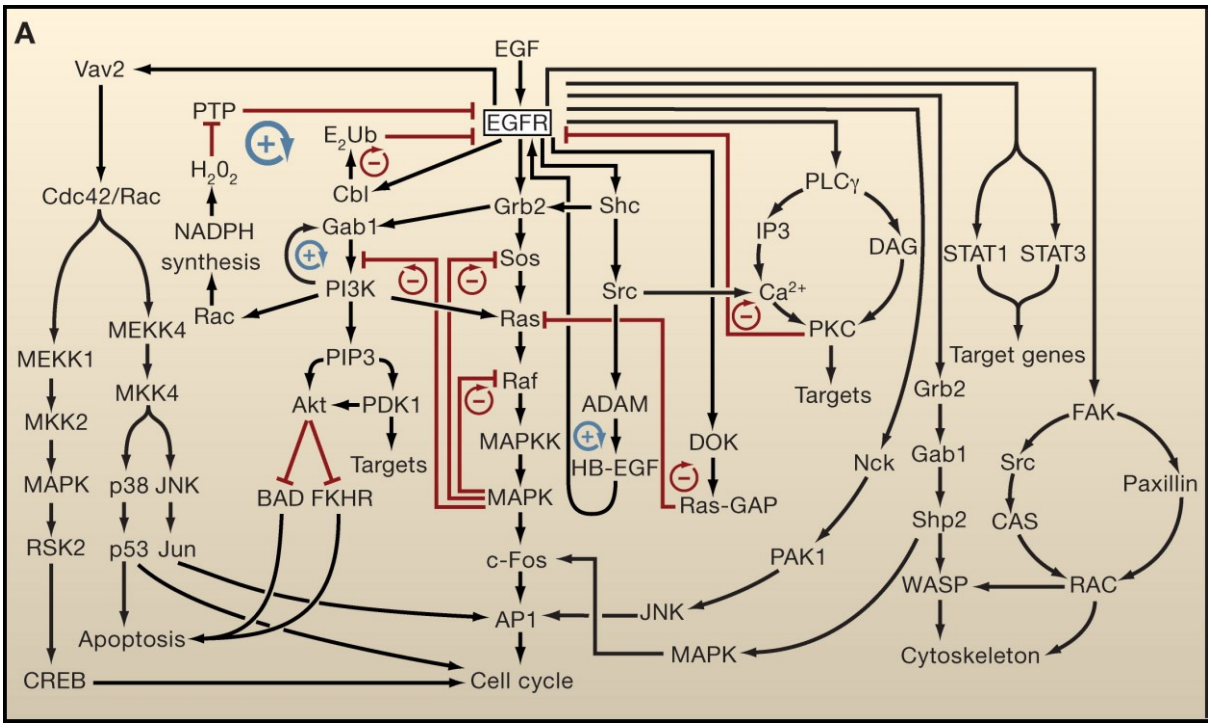
(B) Three recurrent non-synonymous single nucleotide variants (SNVs) were observed in *H3F3A*. The K27M, G34R and G34V mutations are shown in the context of the common post-translational modifications of the H3.3 N-terminal tail, which regulates the histone code. H3.3 has 136 amino acids, and is highly conserved across species from mammals to plants,

including the residues subject to mutation in paediatric GBM (see multiple alignment of amino acids 11 to 60).

(C) Schematic of the mutations observed in ATRX in the 48 WES samples.

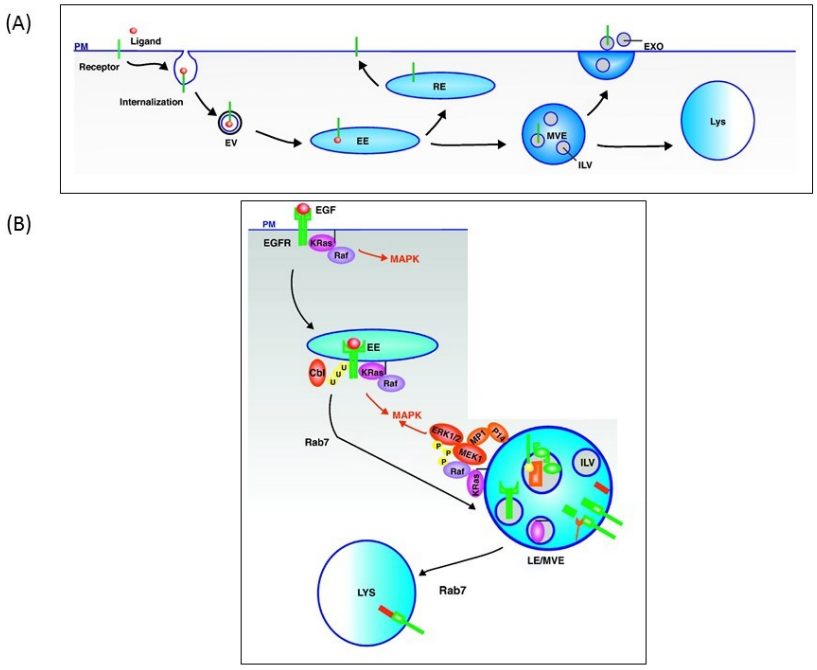
(D) Schematic of the overlap between mutations affecting *ATRX-DAXX*, *H3F3A* and *TP53*. Eight samples had all three mutations. Adapted with permission from (1).

Figure 1.9 – Intracellular signalling networks activated by EGFR



A subset of intracellular signalling components influenced by epidermal growth factor receptor (EGFR) activation are intertwined in a complex network. Through a combination of stimulatory (black arrows) or inhibitory (red lines) signals, several key positive feedback loops (blue circular arrows) and negative feedback loops (red circular arrows) emerge in the network and exert significant influence on its behavior. For example, inhibition of Ras by Ras-GAP or EGFR by protein kinase C (PKC) serves a negative feedback function. On the other hand, H_2O_2 inhibits protein tyrosine phosphatases (PTPs) and thus prolongs or increases activity of EGFR by a positive feedback mechanism. Adapted with permission from (65).

Figure 1.10 – Receptor-mediated endocytosis and signalling



(A) Schematic overview of the basic sorting pathways in the endocytic system. Subsequent to internalization, receptors shuttle via endocytic vesicles (EV) to early endosomes (EE). The receptor can be sorted back to the plasma membrane (PM) via recycling endosomes (RE) or it can be sequestered in intraluminal vesicles (ILV) of multivesicular endosomes (MVE). MVEs can fuse with the lysosome (LYS), where the ILVs are degraded, while specialized forms of MVEs fuse with the plasma membrane to release exosomes (EXO).

(B) Epidermal growth factor receptor (EGFR) initiates K-Ras/Raf-mediated MAPK-signalling at the plasma membrane (PM) and from early endosomes (EE). EGFR is sorted to late endosomes (LE) and multivesicular endosomes (MVE) after being modified with ubiquitin (U) by Cbl. This results in degradation of EGFR in lysosomes (LYS). Still, phosphorylated (P) Raf and K-Ras can facilitate ERK1/2-mediated signalling from LE/MVE when associating with the p14/MP1/MEK1-complex. Adapted with permission from (79).

Chapter 2: Sorting Nexin 3 overexpression disrupts EGFR and MET endosomal trafficking promoting cell proliferation and tumorigenicity in pediatric glioblastoma

Dong-Anh Khuong-Quang[1]*, Takrima Haque[1]*, Helene Delhommelle[2], Djihad Hadjadj[2], Tiffany Lee[3], Hani Halabi[4], Steffen Albrecht[5], Brian Meehan[6], Zhifeng Dong[7], Damien Faury[6], Peter Siegel[7], Stephan Pfister[8], Janusk Rak[6,9], and Nada Jabado[1,6,9]

1. Department of Human Genetics, McGill University, Montreal, Canada
2. Magistere de Genetique, Universite Paris-Diderot, France
3. Department of Anatomy, McGill University, Montreal, Canada
4. Department of Radiation Oncology, Montreal General Hospital, McGill University Health Centre, Canada.
5. Department of Pathology, Montreal Children's Hospital, McGill University Health Centre, Montreal, Canada
6. Montreal Children's Hospital Research Institute, McGill University Health Centre, Montreal, Canada
7. Rosalind and Morris Goodman Cancer Research Centre, McGill University, Montreal, Quebec, Canada
8. Division of Pediatric Neuro-oncology, German Cancer Research Center (DKFZ), Heidelberg, Germany.
9. Department of Pediatrics, McGill University, Montreal, Quebec, Canada

Manuscript in preparation, 2016

2.1 Abstract

Amplification/mutation of receptors-tyrosine-kinases (RTK) plays a major role in gliomagenesis. Using microarray data we generated, we identified overexpression of Sorting Nexin 3 (SNX3), a protein involved in the endosomal trafficking of RTK including EGFR as a potential key player of RTK dysregulation. We hypothesized that dysregulated expression of SNX3 delays RTK degradation and promotes a sustained intracellular activation through these receptors, mimicking RTK amplification seen in adult GBM (aGBM) events in a subset of pediatric GBM (pGBM). We stably overexpressed cMyc-tagged-SNX3 in pGBM (SF188 and SJG2) and aGBM (U87) cell lines. Parallel SNX3 knockdown experiments were performed in these cell lines. Effects of overexpression/silencing of SNX3 were investigated on EGFR and MET activation, cell signalling and cell proliferation *in vitro* and *in vivo* (xenograft model NOD/SCID mice). SNX3 overexpression delayed EGFR and MET degradation following RTK engagement, which consequently increased and sustained activation of Ras and JNK pathways and cell proliferation *in vitro*. Importantly, it promoted tumor formation in NOD-SCID mice. Our results indicate that SNX3-overexpression disrupts physiological trafficking of multiple membrane receptors including EGFR and MET leading to sustained activation of signalling pathways that promote cellular proliferation and tumorigenicity in pGBM.

2.2 Introduction

Brain tumors are the largest group of solid neoplasias in children, accounting for approximately 20% of new pediatric cancer cases in Canada (3), and are currently the leading cause of cancer-related mortality and morbidity in the pediatric years (4). Within this group, astrocytomas account for 45% of all pediatric brain tumors (8). According to the World Health Organization classification, they are divided into low grade (grade I/II) and high-grade (HGA) (grade III/IV) tumors. Grade IV tumors (glioblastoma, GBM) are rare in children representing 2.2% of all pediatric central nervous system (CNS) tumors (14). They are characterized by specific histopathologic criteria: necrosis (pseudopalissading necrosis), microvascular proliferation, high cellular proliferation, and nuclear atypia (5). Although recent high throughput genetic analyses of pediatric high-grade astrocytomas (pHGA) identified hallmark genetic and epigenetic alterations specific to the childhood tumors compared to their adult counterparts (1), current treatments in children are still driven from adult studies and similarly show little therapeutic success (15). This is despite the presence of numerous clinical trials associating conventional therapies (surgery, radiation and chemotherapy mainly temozolomide) to adjuvant therapies including anti-angiogenic and/or anti-receptor tyrosine kinase agents in both adults and children. Today, GBM is still the most treatment-refractory CNS tumor in pediatrics and adults with a 5-year overall survival rate ranging from 5 to 20% in children (4, 15, 42).

Receptor Tyrosine Kinases (RTKs) play a prominent role in the control of a variety of intracellular processes including many metabolic and physiological processes in the brain. In cancer, somatically acquired mutations in RTKs are widespread; several studies on pediatric and adult HGAs highlighted their importance in gliomagenesis (15, 74). The most studied RTK in HGA is EGFR which was considered as a potential therapeutic target because of its

overexpression in the majority of cases in both adult and pediatric tumors (45). Although this gene is frequently amplified or mutated in adults, studies in children have shown infrequent genetic alterations (54, 126). Liang *et al.* determined the protein expression profile of several RTKs in pediatric high-grade gliomas (pHGGs) including astrocytomas, and confirmed that wild type EGFR was overexpressed in 58% of cases as well as PDGFRα (45%) and PDGFRβ (32%) using a cohort of 42 patients (47). MET overexpression has not been investigated to date in pGBM; however in adults, overexpression of this RTK was detected in 29% of 62 patients (127) and was associated with poor prognosis (78). In a recent integrated analysis of pGBM, 19% of 58 tumors showed *PDGFRA* amplification, 1.7% *MET* amplification, and no tumor showed *EGFR* amplification (55). Even though *PDGFRA* amplification rate is high in this cohort and was more frequent in relapsed GBM, the authors reported that many tumors that lacked amplification of *PDGFRA* still showed PDGFRα overexpression (55), confirming that overexpression of RTK in pGBM is not only due to gene amplification.

Dysregulated activation of many RTKs through mechanisms including mutation, overexpression, structural rearrangements, disruption of autocrine/paracrine loops, and inactivation of regulatory constraints is implicated in multiple human neoplasias (67). More precisely, most chromosomal translocations result in RTK variants that are located in the cytosol and are activated independently of ligand binding. These variants fail to enter the endocytic pathway and thus escape downregulation through lysosomal degradation (68). In aGBM, somatically acquired mutations of EGFR confer to this RTK the ability to evade ubiquitin-mediated sorting to degradation (69). In addition to mutations that prevent RTKs from an efficient downregulation, there is a growing evidence to support a role for disruption of the endocytic pathway in oncogenesis. In fact, the major mechanism for downregulation of EGFR is the endocytic pathway (70). Receptor trafficking plays a role in the regulation of receptor signalling by controlling the magnitude or the specificity of the response (71), thus

the targeting to lysosome is important to prevent a sustained activation that can lead to cell transformation (67). Moreover, RTKs can still signal from endosomes after internalisation, and this signalling will end when these RTKs are dephosphorylated and degraded (68). Therefore, aberrant endocytosis that could delay the internalization, trafficking and/or degradation of RTKs may promote oncogenesis through sustained signalling.

Our group has published the first study of gene expression profiling in pGBM (51) and has confirmed that, even though they share common aberrations of major signal transduction and cell cycle pathways, pediatric tumors are molecularly distinct from their adult counterparts. Transcriptional profiling analyses distinguished two genetically and prognostically different subsets of pGBM based on their differing association with active Ras/Akt pathways (51, 52).

Our laboratory identified Sorting Nexin 3 (SNX3) to be specifically overexpressed in pGBM associated with Ras/Akt pathway activation (51). SNX3 belongs to a family of proteins characterized by a Phox homology (PX) domain and is involved in the regulation of endosomal trafficking of membrane receptors, including RTKs (99, 100, 108-112). SNX3 is a small protein of 162 amino acids and its PX domain binds directly and selectively to phosphatidylinositol-3-phosphates (PtIns(3)P) on early endosomes (108, 113). Moreover, the generation of PtIns(3)P is required for SNX3 membrane recruitment (108, 114). Overexpression of SNX3 leads to the expansion of tubular-vesicular structures that are characteristic of early, recycling and late endosomes, and delays EGFR degradation with internalized complexes being retained in the swollen structures rather than being targeted to the lysosome for degradation (108, 110). Up to now, no link has been shown between SNX3 and cancer.

In this study, we further investigate the role of SNX3 overexpression in the dysregulation of the trafficking of RTKs in pGBM, leading to an increased and sustained signalling of downstream pathways, promoting tumorigenesis both *in vitro* and *in vivo*.

2.3 Material and methods

Cells, transfections and antibodies

Human pediatric glioblastoma cell line SF188 and adult glioblastoma cell line U87 (kind gifts of Dr. Del Maestro, Montreal Neurologic Institute, Canada) were both used in our *in vitro* experiments. They were routinely maintained in Eagle's minimal essential medium and Dulbecco's modified Eagle's medium respectively (Wisent Inc, Canada) supplemented with 10% fetal bovine serum (Wisent). Cells were stably transfected with wild type Myc-tagged SNX3 plasmid (kind gift of Dr. W. Hong, Institute of Molecular and Cell Biology, Singapore) using Lipofectamine 2000 (Invitrogen, USA) followed by G418 selection according to manufacturer's instructions. Control cells were stably transfected with an empty vector using the same procedure. Transfection efficiency was assessed using immunofluorescence and western blot analysis.

Primary antibodies were from Cell Signaling Technology, USA (pMEK, pERK, pJNK, Myc-tag, β-actin) and from Santa Cruz Biotechnology, USA (SNX3, EGFR, EEA1, LAMP1). Other researchers provided the following antibodies: anti-SNX3 used for immunohistochemistry from Dr. W. Hong (Institute of Molecular and Cell Biology, Singapore), anti-MET from Dr. M. Park (Goodman Cancer Centre, Canada).

qRT-PCR

qRT-PCR was performed as previously described (51). Primers used for SNX3 were as follows: TGCGTCTTGTTTGCTTCTTG (forward primer) and TCCTCCAGGCTTGTAATACCC (reverse primer).

Tissue microarray

Tissue microarrays (4 µm) of 107 GBM tumor cores and 63 pilocytic astrocytomas were immunohistochemically stained for SNX3, EGFR and MET proteins. Unstained sections were subjected to antigen retrieval in 10 mM citrate buffer (pH 6.0) for 10 min at sub-boiling temperatures. Individual slides were incubated overnight at 4 °C with rabbit anti-SNX3 (kind gift of Dr. W. Hong, 1/200), rabbit anti-EGFR (Epitomics, USA, 1/500), or rabbit anti-MET (Assay Designs, USA, 1/50) antibodies. After incubation with the primary antibody, secondary biotin-conjugated donkey anti-rabbit antibodies (Jackson ImmunoResearch, USA) were applied for 30 min. After washing with PBS, slides were developed with diaminobenzidine (Dako, Denmark) as the chromogen. All slides were counterstained using Harris haematoxylin. Immunohistochemistry staining on TMA was scored by three individuals independently, including a pathologist (SA).

Indirect immunofluorescence

Indirect immunofluorescence was performed on cells fixed in 4% paraformaldehyde. Permeabilization was performed in 0.5% Triton-X100 at room temperature (RT) for 20 min before blocking in 5% BSA in PBS supplemented with 10% of normal goat serum for 1h at RT. Cells were incubated in block solution with primary antibodies overnight at 4°C then incubated with secondary antibodies for 1h at RT. Samples were mounted using ProLong ® Gold Antifade reagent (Invitrogen, USA). Confocal images were acquired using an Axiovert

200 M microscope coupled to the Zeiss Laser Scanning Microscope 510 Pascal system (Life Sciences Complex Advanced BioImaging Facility, McGill University, Canada). Experiments were performed three times.

Western blot analysis

Cultured cells were lysed in EBC cell lysis buffer (50 mM Tris, pH 8.0, 120 mM NaCl, 0.5% Igepal) supplemented with phosphatase inhibitors (50 mM NaF and 1 mM b-glycerolphosphate) and protease inhibitors (Roche, Switzerland). Protein lysates were cleared by centrifugation at >12,000g, and the supernatant was collected for analysis.

For western blot analysis, protein samples were separated by 10% SDS-PAGE and electrotransferred onto PDVF membranes (EMD Millipore, USA). After blocking, blotted membranes were incubated with primary antibodies in block solution overnight at 4°C. Horseradish peroxidase-conjugated secondary antibodies were incubated with blots for 1 h at RT. Antibodies were detected by chemiluminescence using ECL Plus (Amersham Biosciences). Experiments were performed three times.

Monolayer proliferation assay

Using the xCELLigence Real Time Cell Analyzer (RTCA) instrument (Roche Diagnosis, Germany) placed in a humidified incubator at 37°C and 5% CO2, cell proliferation was performed using modified 16-well plates (E-16 plate). 500 cells were seeded into the wells. Each clone was tested in triplicate within the same experiment. Cells were incubated during 7 days in growth medium and monitored by the RTCA. Electronic impedance was read every 30 min and plotted as cell index after change in impedance was calculated based on mathematical algorithms. Experiments were performed three times.

Soft agar assay

10,000 SF188 or U87 cells were suspended in 0.3% agarose base supplemented with complete culture medium. This suspension was layered over 0.5% agarose base in a six well plate. Cells were treated with different concentrations of erlotinib (Selleck Chemicals, USA), SU11274 (Selleck Chemicals, USA) or DMSO, which were added to the agarose base. Once a week, liquid culture media with the appropriate concentration of inhibitor was poured on top of the gel. After six weeks, colonies over 50μm were manually counted. Each condition was tested in triplicate within the same experiment. Experiments were performed three times.

Mouse experiments

Six-week old female SCID and NOD-SCID mice were purchased from Harlan Laboratories, (USA). SF188 or U87 cells transfected with Myc-tagged SNX3 vector or empty vector were injected subcutaneously (10^6 cells in 0.1 mL of BD Matrigel (BD Biosciences, USA)). Tumor growth was directly assessed biweekly by measurement. Mice were sacrified when tumors reached 18mm in diameter and then tumors were collected. Animal experiments were performed according to protocols approved by the University Animal Care Committee of the McGill University Health Center.

Statistical analysis

All experiments were performed at least three times unless otherwise specified. Statistical analyses were performed using Prism® (GraphPad, USA). Error bars represented in graphs denote the standard error of the mean (SEM). Numerical data were processed for significance using two-tailed Student t-test with the threshold p-value of 0.05 except for the gene expression microarray analysis (Welsch t-test with Benjamini and Hochsberg multiple testing correction).

2.4 Results

SNX3 is specifically overexpressed in pediatric glioblastoma

From the gene-expression analysis of pediatric glioblastomas we published (51), we sought to investigate potential oncogenes involved in its tumorigenicity. SNX3 was specifically overexpressed in the subset of tumors with active Ras/Akt ($p<0.001$) (Table S2.1A). This subset of tumors also showed an overexpression of EGFR by qRT-PCR (Table S2.1A). SNX3 overexpression was confirmed by RT-PCR (data not shown). In addition, overexpression of other members of the sorting nexin family was observed at moderate levels (Table S2.1B). Given the role of SNX3 in receptor trafficking and regulation (108), we considered it as a potential oncogene since it increased and sustained the activity of receptor tyrosine kinases. We first confirmed the overexpression of the protein by IHC in an independent cohort of 107 pediatric GBM samples, using a tissue-microarray. Compared to a cohort of 63 pediatric pilocytic astrocytomas, SNX3 was overexpressed in 69% of pGBMs and this was highly specific to the high-grade tumors ($p<0.001$) (Figure 2.1A). To move further in our analysis, we used the same tissue-microarray to assess MET and EGFR expression. We observed a strong positive correlation between MET and SNX3 expression (Pearson coefficient 0.601) (Figure 2.1B) and a moderate positive correlation between EGFR and SNX3 (Pearson coefficient 0.318) (data not shown). MET and EGFR were respectively overexpressed in 71% and 50% of cases.

SNX3 overexpression expands the endosomal compartment and delays EGF/EGFR degradation

To assess the relevance of our hypothesis, we first examined the effects of SNX3 overexpression in glioblastoma cell lines. After confirmation by qRT-PCR that SF188 and U87 cell lines did not overexpress SNX3, MET, or EGFR (data not shown), we stably

transfected them with Myc-tagged SNX3 plasmid (Figure S2.1A). The transfected cells did not show any morphologic change perceptible under the light microscope (data not shown).

As previously described (108), immunofluorescence experiments showed an expansion of the endosomal compartment where SNX3 is overexpressed, mostly in EEA1-positive early endosomes but also in late endosomes labelled with LAMP1 (Figure S2.1B). To further determine the effect of SNX3 overexpression on EGFR trafficking, EGFR was activated and internalized upon EGF binding. Kinetic experiments tracking the trafficking of labeled-EGF within the cell confirmed that the endocytosis of receptor tyrosine kinases was not affected by SNX3 overexpression since EGF colocalized with EEA1 at early time points in both control and SNX3 overexpressed cells. Upon further incubation, EGF degradation was delayed when SNX3 was overexpressed (Figure 2.2 A and B). Previous studies have shown that EGF half-dissociation value pH is 5.6 (128), thus EGF remains bound to EGFR after internalization until the degradation of both ligand and receptor in the lysosome (129), allowing us to indirectly track EGFR through EGF signal. These data confirm that, in our model, SNX3 overexpression delays EGFR trafficking and degradation by disrupting the endosomal pathway.

We performed the reverse experiment, looking at the effects of SNX3 silencing on the trafficking of the EGF/EGFR complex. Our data showed that SNX3 silencing has marginal effects on EGFR endosomal pathway (Figure S2.2), as previously described in HeLa cells (110, 130).

Overexpressing SNX3 increases and sustains receptor tyrosine kinase signalling

To determine the impact of the delayed degradation of EGFR on RTK signalling, we analyzed the activation of the Ras pathway by immunoblot, using pMEK and pERK as markers of its

activation. Upon stimulation by its ligand TGFα, chosen over EGF because of its more potent mitogenic action (131), EGFR was activated leading to the downstream phosphorylation of MEK and ERK. Moreover, in comparison to control cells, their levels were heightened and sustained up to 4 hours after stimulation (Figure 2.3A). In the light of the trafficking data, we can conclude that the increased and sustained activation of the Ras pathway results from the signalling of internalized EGFR (68).

SNX3 has been described to be involved in the regulation of the trafficking of several distinct membrane receptors. Although EGFR was the sole RTK described to date (108, 132), we sought to look at other RTKs. MET is an RTK involved in many cancers and has been described to cross-talk with EGFR (133), including in adult glioblastoma (134). We first looked at Ras activation upon MET stimulation by HGF and found the same profile than the one following EGFR activation (Figure 2.3B). We also looked at another pathway known to be activated by MET, the JNK pathway. JNK has been proposed to be a potential therapeutic target in glioblastoma through its role in the maintenance of stem-like glioblastoma cells (135), or its control in temozolomide resistance through MGMT expression (136, 137). We observed the same pattern of increased and sustained activation of pJNK (Figure 2.3C). To note, we did not look for Akt signalling in SF188 cells, because these cells present a constitutively active Akt.

PDGRA is the most frequent target of focal amplification in pGBM (55) and DIPG (56), and is also mutated in up to 15% of pHGG (138). Therefore, we studied this RTK in our model but we did not find any alteration in the Ras signalling pathway when SNX3 was overexpressed (Figure S2.3).

Overall, these data shed light on the role of SNX3 in the regulation of EGFR and MET, showcasing that its overexpression led to an increased and sustained activation of these receptors with downstream effects on Ras and JNK pathways.

SNX3 overexpression promotes cell proliferation and tumor formation in vitro and in vivo and affects cell sensitivity to tyrosine kinase inhibitors

The Ras pathway has been shown to be a crucial player in tumorigenesis by promoting cell proliferation. To further understand the potential role of SNX3 in pGBM formation, we assessed its role in cell proliferation and tumor formation.

Monolayer proliferation assays confirmed our hypothesis that SNX3 overexpression conferred to cells a proliferation advantage (Figure 2.4A). Moreover, this was showed by soft agar assays where these SNX3-overexpressing cells were able to grow and to form colonies in anchorage-independent conditions compared with EV transfectants (Figure 2.4B).

To complete our study, we showed that, *in vivo*, SF188 cells were tumorigenic when overexpressing SNX3 in immunocompromised mice. We used two different strains of mouse with two different types of immunodeficiency, Severe Combined ImmunoDeficiency (SCID) and Non-Obese Diabetic – Severe Combined ImmunoDeficiency (NOD-SCID), and in both models, we observed tumor formation (respectively in 3/5 and 5/6 mice). SF188 cells are known to be poorly tumorigenic (unpublished data), and we did not observe tumor formation with our empty vector control. These tumors were macroscopically vascularized and showed pathological features of glioblastoma with neovascularization, necrosis and nuclear atypia. They were positive for glial (GFAP) and astrocytic (S100) markers (Figure 2.5).

Lastly, based on previous reports of poor response to single-agent targeted therapy because of coactivation of RTKs (124), we tested the effect of EGFR and MET small molecule inhibitors

in vitro. We first confirmed that erlotinib and SU11274 inhibit EGFR and MET respectively with a downstream effect on the Ras pathway in our cell lines (Figure S2.4A). Using soft agar assays, we observed a critical dose-dependent response to SU11274, a MET inhibitor in SF188 cells, whether or not they overexpressed SNX3 (Figure 2.4B). Both EGFR and MET inhibition were effective to decrease colony formation in soft agar, and, as expected, dual inhibition of both MET and EGFR was superior to single agent inhibition. Although these inhibitors were efficient on SNX3 overexpressing cells, this overexpression did not strongly affect the response to these inhibitors, compared to the control (Figure 2.6).

2.5 Discussion

Our study shed light on a novel role of Sorting Nexin 3 in promoting pediatric glioblastoma genesis through dysregulation of RTK trafficking. In normal conditions, close regulation of receptor tyrosine kinase activity is required to prevent uncontrolled cell proliferation. Several mechanisms of dysregulation are known to be involved in tumor formation, including trafficking dysregulation. Our data are the first to suggest SNX3 involvement in cancer (Figure 2.7).

The sorting nexin family is an emergent group of players in the regulation of protein trafficking. They were first described in the late 90s and are unified by a Phox homology (PX) domain. To date, there are 33 mammalian sorting nexin described. Sorting Nexin 6 (139) has been involved in breast cancer by promoting transcriptional repression. Sorting Nexin 2 was proposed as a potential drug target in lung cancer by controlling the localization and expression of MET (140). It has also been described as a fusion partner with *ABL1* in acute lymphoblastic leukemia and cells harbouring this translocation to have a poor sensitivity to imatinib and dasatinib (141). Nishimura *et al.* have studied the role of SNX1 promoting MET

activation and disruption of EGFR trafficking in gefitinib resistant human lung cancer cell lines (142, 143).

Sorting Nexin 3 is a small protein which bears only a single recognizable domain, the SNX-PX domain, and it localizes to the endosome through the binding of this domain to PtIns(3)P in the endosome membrane (108). In our study, SNX3 promotes cell proliferation and tumor formation by delaying the degradation of EGFR and MET. This is most interesting in pediatric glioblastoma because several receptor tyrosine kinases including EGFR and MET are overexpressed at the protein level without gene amplification/alteration. Disruption of endosomal trafficking could be the reason as to why these proteins are overexpressed.

Because of the recurrent mutations in RTKs, mostly EGFR and PDGFR, and the *in vitro* and *in vivo* evidence of their role in gliomagenesis, several clinical trials of RTK inhibitors were conducted in both adult and pediatric GBMs with little effect (76). In 2007, Stommel *et al*. demonstrated that co-activation of multiple RTKs in glioblastoma decreases tumor cell response to targeted therapies, and described the concept of kinase switching (124). Furthermore, a phase II study with gefitinib showed that this treatment is associated with an efficient dephophorylation of EGFR in post-treatment tissue although not leading to any modulatory effect on downstream signalling pathways (77), highlighting the robustness of this signalling network. Here, we rationalize that pGBM cells are able to overcome RTK targeted treatment by promoting the upregulation of other RTKs through the dysregulation of their trafficking and degradation pathway. A cross talk has been shown between EGFR and MET in lung cancer, leading to EGFR inhibitor resistance (144) through MET amplification. SNX3 overexpression is also a potential factor of drug resistance, by enhancing other RTKs signalling.

Because of the recurrence of *PDGFRA* amplification or mutation (55, 138) and its likely oncogenic role in pHGG (145, 146), we studied the effect of SNX3 on PDGFRα signalling. Our data suggest that it is not affected by SNX3 overexpression. However, because dependent on the level of stimulation, PDGFR can be sorted into the clathrin-mediated pathway where SNX3 is involved (54) or the raft/caveolin-mediated endocytosis (147), further experiments are required before we can definitively rule out any role of SNX3 in its trafficking. Thus, further studies identifying how the cargo of SNX3-positive endosomes is sorted and regulated would be required to fully understand its mechanism.

Moreover, further experiments are required to understand what triggers SNX3 dysregulation. The promoter analysis from the ENCODE project show that some of its transcription factors' binding sites are involved in cell cycle progression and control (E2F1 and MYC) (148). However, further studies are required to assess the role of these transcription factors in the promotion and regulation of SNX3 transcription.

Sorting Nexin 3 is widely expressed in humans (99), including in the brain. Mizutani *et al.* reported that SNX3 is strongly expressed during embryogenesis in the mouse central nervous system and during the neonatal period, mostly in the cerebral cortex and cerebellum (149), therefore SNX3 is likely to have an important role in the genesis and development of these organs. Although the role of SNX3 in the brain development remains unclear, this supports the hypothesis that glioblastoma is a neurodevelopmental disease.

Recent publications have shed light on a novel role of SNX3. Aside from its role in the endosomal trafficking of membrane proteins in early endosomes, SNX3 is involved in a distinct retromer pathway from the classical one involving SNX1-SNX2 and SNX5-SNX6. The first cargo protein described was Wntless (Wls). Wls is a transporter of Wnt proteins, which play a central role in development and tissue homeostasis. SNX3 is crucial in the

retrograde transport of Wntless between endosomes and trans-Golgi network, in this distinct pathway, and is required for Wnt secretion (150, 151). However, to our knowledge, EGFR and other receptor tyrosine kinases have not been identified as retromer cargo proteins (152, 153).

In recent years, evidence has been increasing surrounding the critical role of MET in oncogenesis (154), including through its endocytic pathway (155). Although MET is not found to be altered at the gene level in the majority of pGBMs, and was not overexpressed in the cell lines we used, the fact that it was overexpressed in 70% of pGBM on our tissue microarray and the results of *in vitro* MET inhibition warrant further studies in order to validate MET as a potential target in pGBM.

Growing evidence emphasizes the role of RTK localization in the regulation of the mitogenic signalling, thus in tumorigenesis (156). This knowledge may identify novel therapeutic targets and also shed light on the regulation of response to cancer therapy (157). This paper shows that SNX3 is a crucial factor in the formation of a subset of pediatric glioblastoma with a worse prognosis, and may play a role in the resistance of these tumors to conventional therapies as well as targeted therapies against RTK.

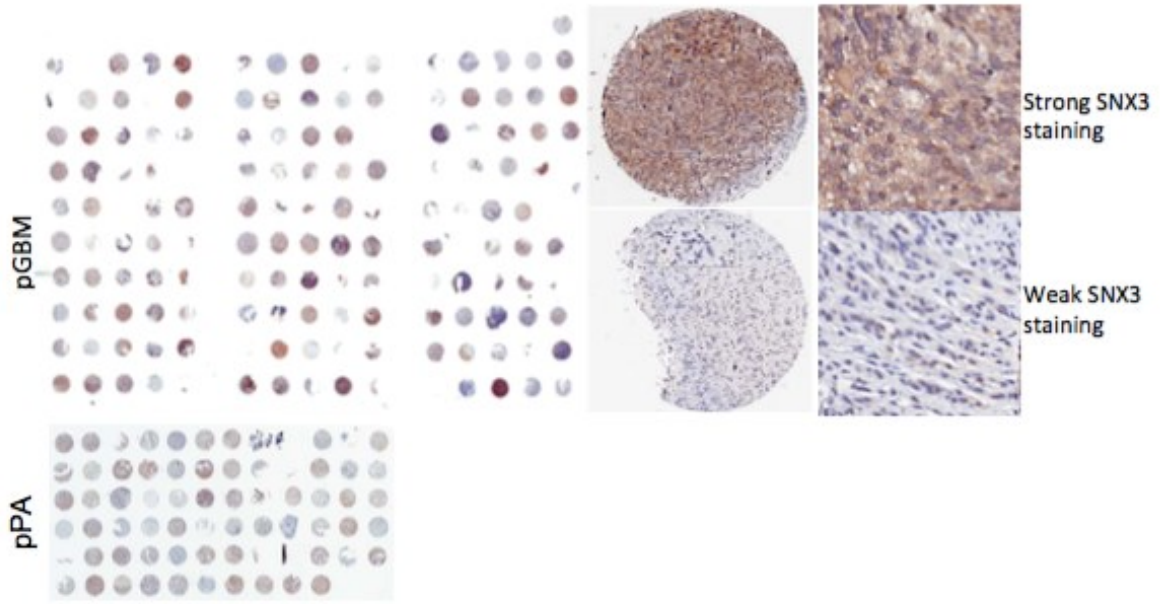
2.6 Acknowledgements

The authors would like to thank Dr. Morag Park (McGill University) and her lab for their help in conducting MET experiments. The authors also would like to acknowledge the services provided by the Animal Facility at the Montreal Children's Hospital and by the Goodman Cancer Research Centre Image Core Facility.

2.7 Figures

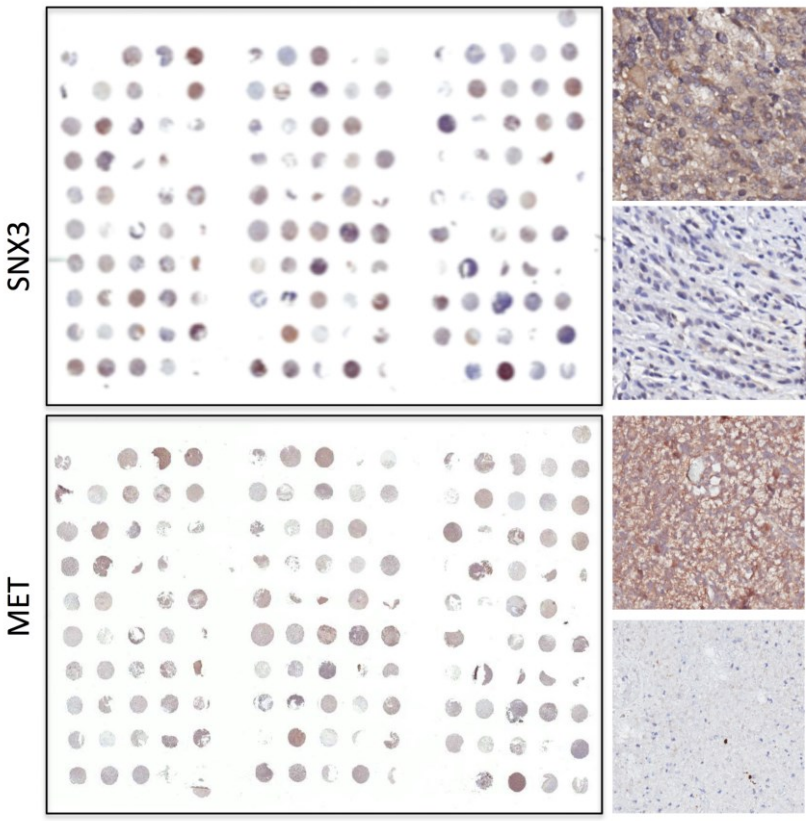
Figure 2.1 – SNX3 overexpression is specific to pGBM and correlates with MET expression

A



SNX3 staining	pGBM (n=130)	pPA (n=74)	P
weak	30%	49%	
strong	65%	16%	p<0.001
NA	5%	35%	

B



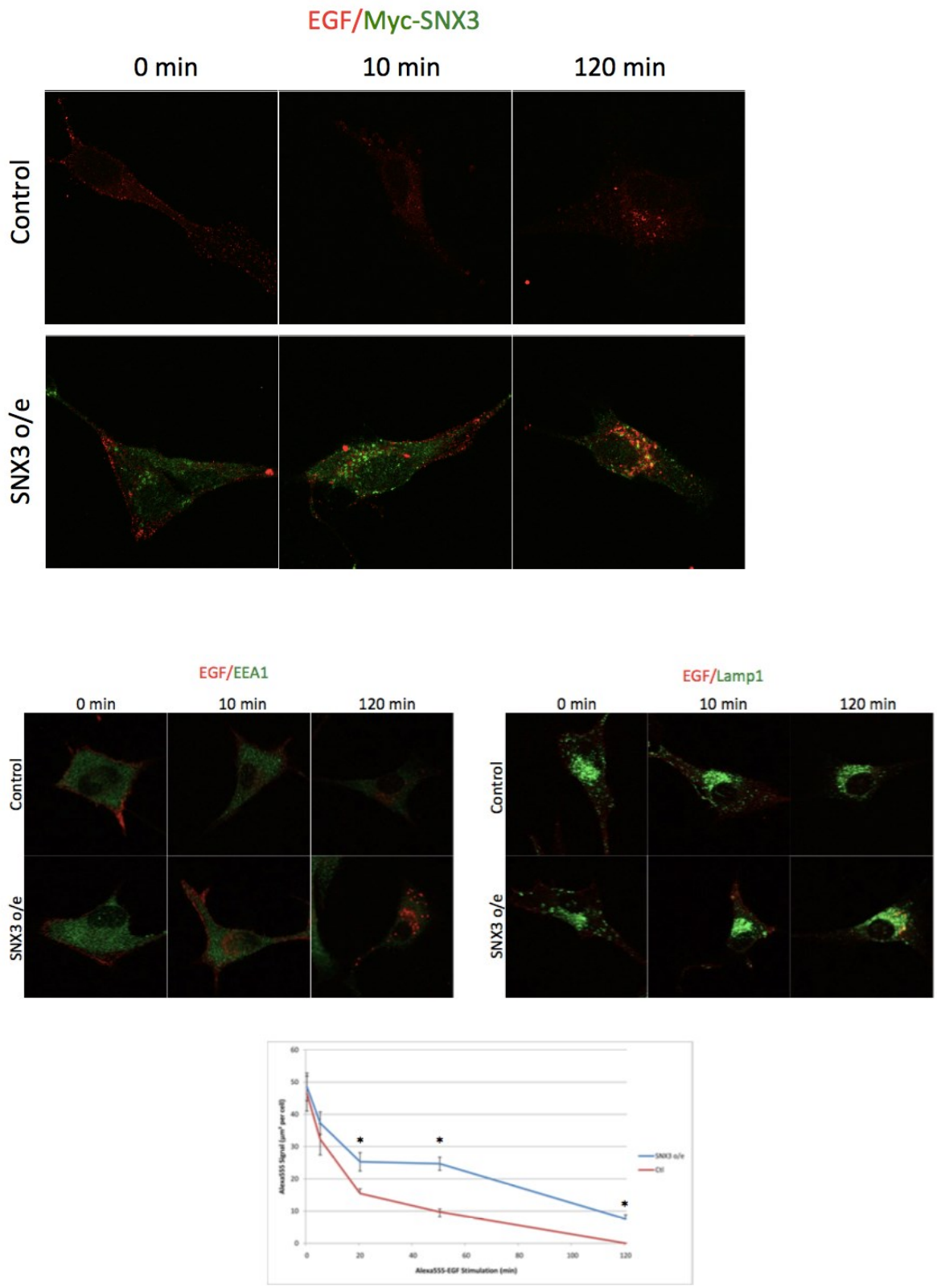
Scoring	SNX3 (n=116)	MET (n=118)
0	36	34
1	40	42
2	38	39
3	2	3
Pearson Correlation	r = 0.601	

(A) SNX3 immunohistochemistry was performed on TMA slides of pGBM and pediatric pilocytic astrocytomas (pPA), as mentioned in the material and methods section. Scoring was done independently by two different individuals using the staining on normal controls as

baseline (data not shown). Compared to pPA, SNX3 was significantly overexpressed in pGBM ($p<0.001$, two-tailed Student t-test).

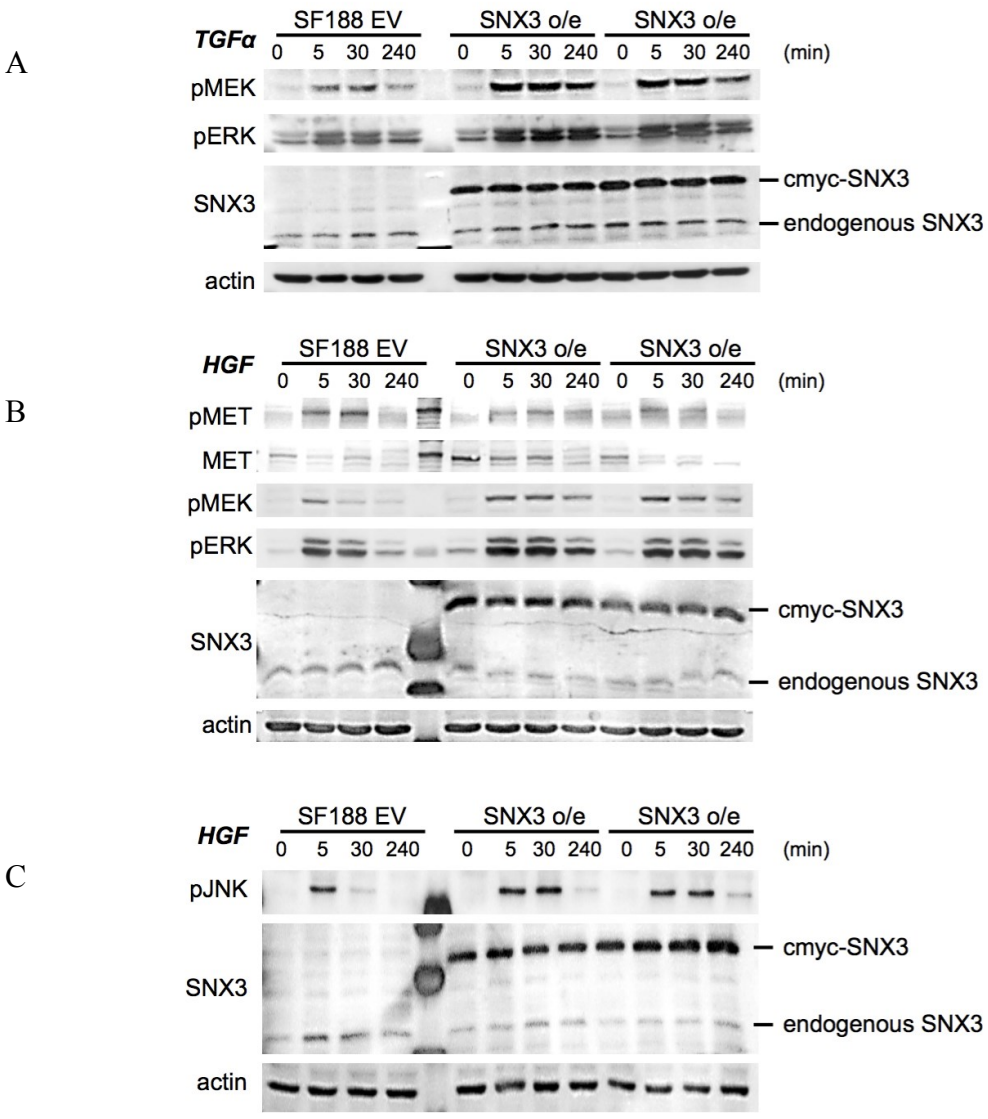
(B) MET immunohistochemistry was performed on the same TMA slides. Scoring was done blindly by two independent observers, and correlation was analyzed using Pearson's correlation.

Figure 2.2 – SNX3 overexpression delays EGF degradation



Kinetic immunofluorescence experiments were performed after incubation of U87 cells with Alexa555 EGF (50ng/mL of media) showing that EGF, which is coupled with EGFR before internalization, remains longer in early and late endosomes compared to EV. Images were recorded on Zeiss Laser Scanning Microscope Pascal on Axiovert 200 Confocal Microscope at the McGill Cancer Centre. Quantification of the Alexa555 signal was performed using Zeiss AIM software. Values are means of ten distinct cells ± SEM; *$p<0.001$ (two-tailed Student t-test).

Figure 2.3 – SNX3 overexpression increases and sustains signalling from downstream pathways after RTK activation



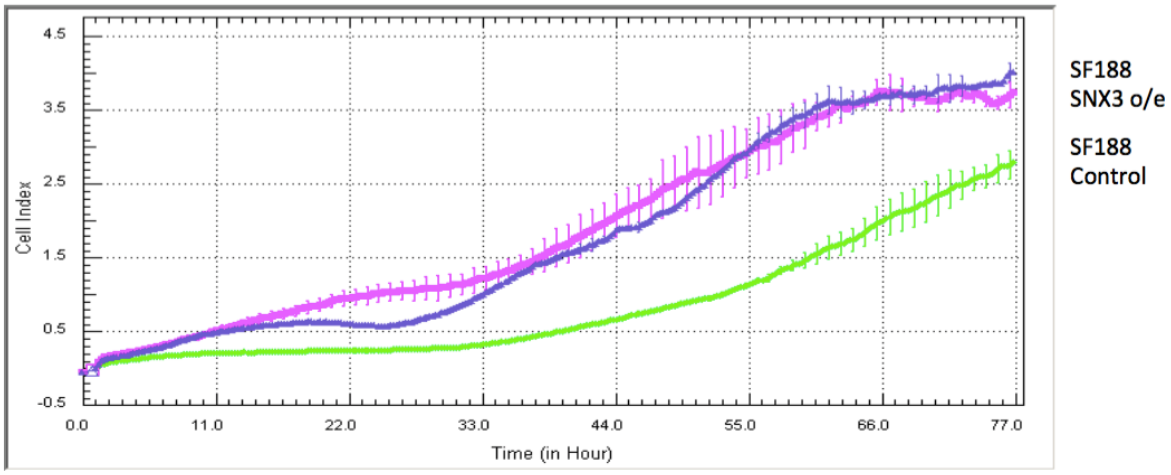
(A) SNX3 overexpression leads to an increased and sustained activation of the Ras pathway after TGFα stimulation in SF188 cells. After overnight starvation, cells were stimulated with TGFα (100ng/mL). The western blot analysis of whole cells lysates shows an increased phosphorylation of MEK and ERK (markers of Ras activation) at 5 min in SNX3 overexpressing cells that persisted at 4 hours. Fifth lane corresponds to the protein ladder. This result has been observed in three separate experiments.

(B) SNX3 overexpression leads to an increased and sustained activation of the Ras pathway after HGF stimulation in SF188 cells. After overnight starvation, cells were stimulated with HGF (50ng/mL). The western blot analysis of whole cells lysates shows an increased phosphorylation of MEK and ERK (markers of Ras activation) at 5 min in SNX3 overexpressing cells that persisted at 4 hours. Fifth lane corresponds to the protein ladder. This result has been observed in three separate experiments.

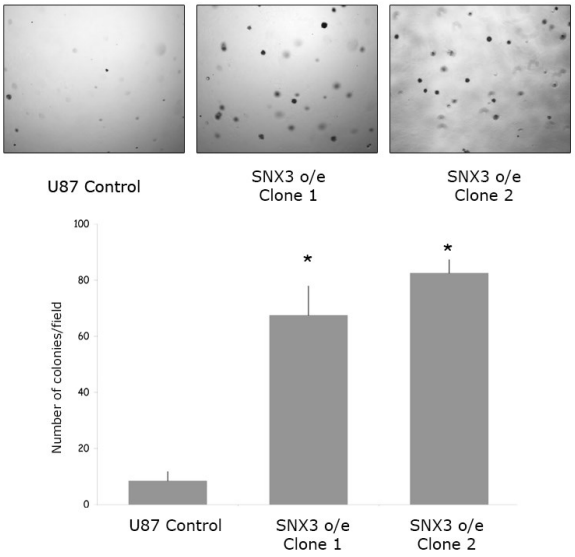
(C) SNX3 overexpression leads to a sustained activation of the JNK pathway after HGF stimulation in SF188 cells. The immunoblot of whole cell lysates shows that aside from the Ras pathway activation, SNX3 is involved in other pathways activation such as JNK, under the same experimental conditions. Fifth lane corresponds to the protein ladder. This result has been observed in three separate experiments.

Figure 2.4 – SNX3 overexpression gives cells a growth advantage

A



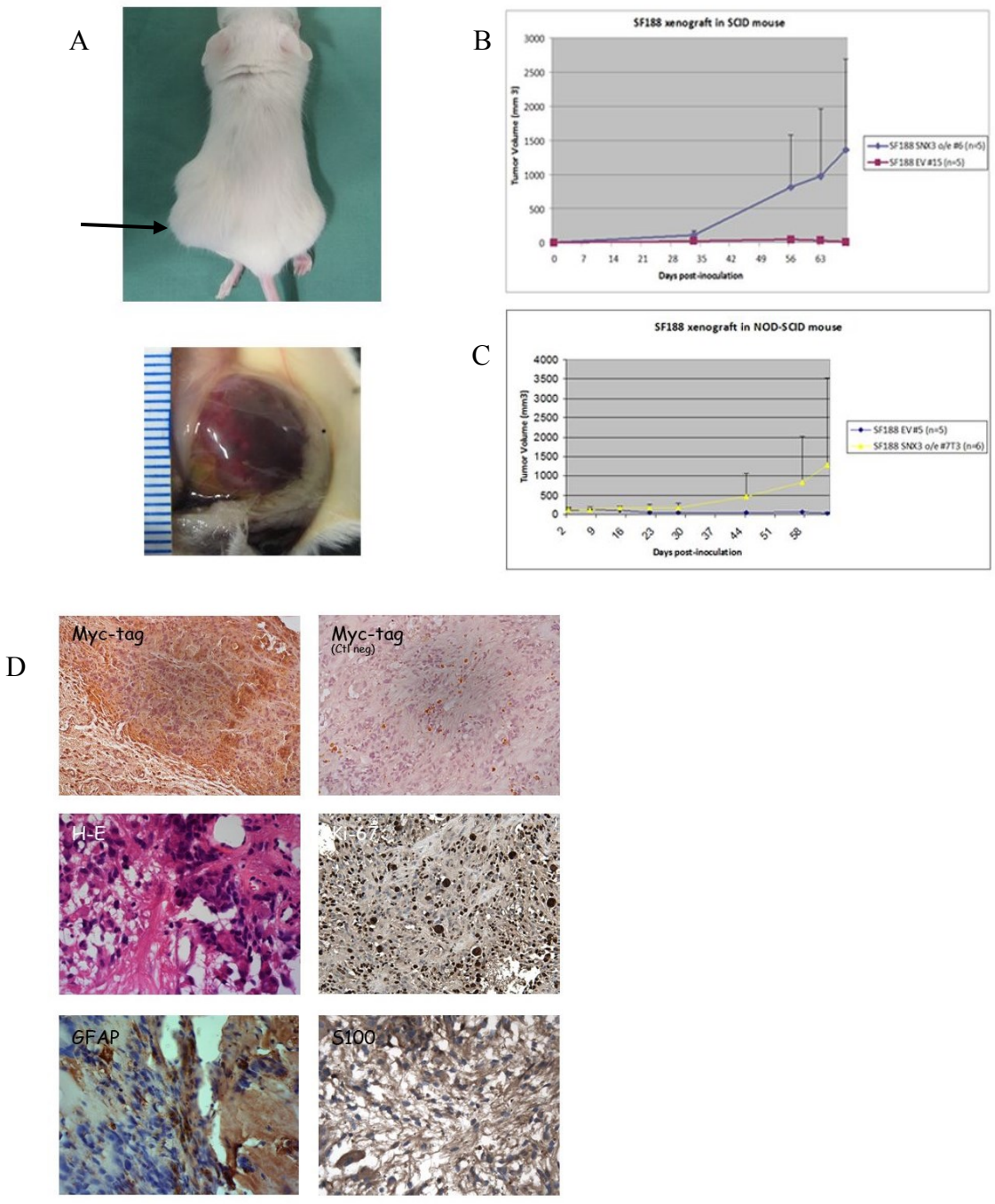
B



(A) SNX3 overexpression gives cells a growth advantage in monolayer proliferation. Assay was performed using xCELLigence RTCA system; impedance was measured every 30 minutes and plotted on the graph as cell index. Cell index ± SEM of two different clones of SNX3 overexpressing SF188 cells (pink and purple) are shown and compared to a control clone (green). This experiment was performed three times independently.

(B) SNX3 overexpression enhances anchorage-independent growth ability in U87 cells. U87 cells were suspended in 0.3% agar containing 10% of foetal bovine serum and incubated for 6 weeks. Colonies over 50µm were manually counted. Magnification x4. Values are means of three separate experiments ± SEM; *$p<0.001$ (two-tailed Student t-test).

Figure 2.5 – SNX3 overexpression induces tumor formation in immunocompromised mice



(A) Macroscopic observation of tumor formation. SF188 cells were inoculated into the flank of immunocompromised mice and tumor formation was observed after 4 to 5 weeks. As

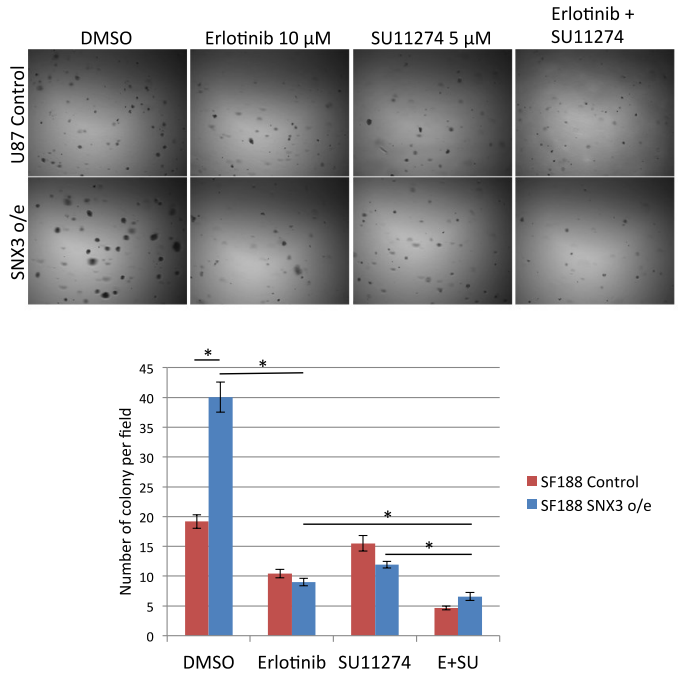
shown on the image, tumors were vascularized and presented hemorrhagic features.

(B) SF188 SNX3 overexpressing cells promote tumor formation in NOD-SCID. $10x10^6$ cells in matrigel were injected in the subcutaneous fat pad of SCID mouse (5 mice for each clone). Tumor measurement was done twice weekly. In clone #6, tumor formation was observed in 3 mice out of 5, whereas control cells did not promote any tumor formation.

(C) SF188 SNX3 overexpressing cells promote tumor formation in NOD-SCID. $10x10^6$cells in matrigel were injected in the subcutaneous fat pad of NOD-SCID mouse (5 mice for each clone). Tumor measurement was done twice weekly. In clone #7T3, tumor formation was observed in 3 mice out of 5, whereas control cells did not promote any tumor formation.

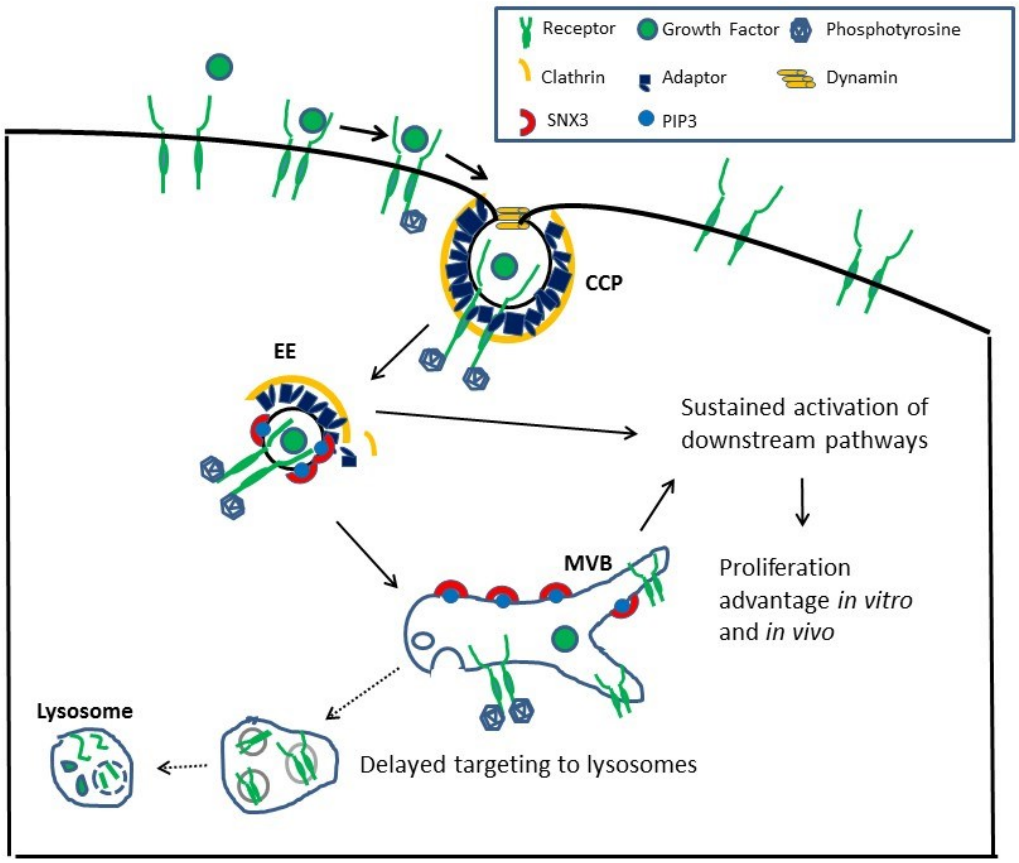
(D) Tumor immunohistochemical staining was performed to characterize xenograft tumor, using hematoxylin and eosin staining (H-E), Ki67 a cellular marker for proliferation, GFAP and S100 which are astrocytoma markers, and Myc the tag used for SNX3 transfection. Anatomopathologic analysis was consistent with glioblastoma.

Figure 2.6 – Synergetic efficiency of combined targeted therapy decreasing colony formation in soft agar



SF188 cells were suspended in 0.3% agarose containing either vehicle (DMSO 1%), erlotinib (EGFR inhibitor) 10µM, SU11274 (MET inhibitor) 5µM or addition of erlotinib 10µM and SU11274 5µM. Cells were incubated for 6 weeks then colonies over 50µm were manually counted. Magnification x4. Values are means of three separate experiments ± SEM; *p<0.001 (two-tailed Student t-test).

Figure 2.7 – Schematic representation of SNX3 role in pediatric glioblastoma cell



Upon ligand binding, EGFR or MET is phosphorylated and activated, inducing its internalisation in clathrin-coated pits (CCP). The overexpression of SNX3 on early endosome (EE) membrane delays RTK targeting to lysosomes, thus sustaining the activation of downstream signalling pathways by the continuous signalling of these RTKs in early and multivesicular bodies (MVB). The cell acquires then a proliferation advantage both *in vitro* and *in vivo*.

2.8 Supplementary material

Supplementary table S2.1 – Ras activation is significantly associated with SNX3 expression in pGBM

A

Sample	Ras	EGFR *p*-fold value	SNX3 *p*-fold value
66	+	1.306	2.288
1013	+	1.63	2.201
429	+	7.954	2.156
434	+	34.64	2.25
P143	+	0.73	3.721
P4011	+	1.039	2.448
P4211	+	2.269	3.171
63	-	0.555	0.888
151	-	0.999	1.436
184	-	2.164	0.537
272	-	0.953	1.25
3948	-	2.64	0.599
612	-	0.999	2.172
P4450	-	0.972	0.622

B

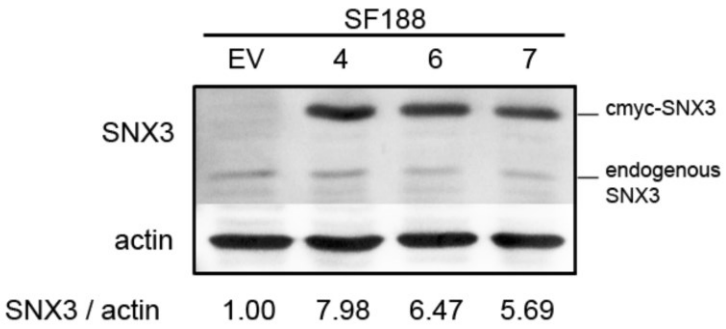
SNX	*p*-fold value (compared to normal)	*p*-fold value (comparing Ras + and Ras -)
SNX3	1.56	2.9
SNX 27	1.34	1.172
SNX 17	1.3	
SNX 2	1.26	
SNX 4	1.14	
SNX12	1.06	
SNX14	1.06	
SNX15	1.02	
SNX13	1.01	
SNX 11	0.99	
SNX1	0.96	
SNX6	0.9	
SNX8	0.91	
SNX 9	0.82	
SNX 10	0.69	

(A) *P*-fold value of EGFR and SNX3 in different pGBM samples analyzed by gene expression profiling (51). Samples are classified based on their Ras pathway activation. SNX3 is significantly overexpressed in the group where Ras is activated ($p < 0.0001$, Welsch t-test, Benjamini and Hochsberg multiple testing correction).

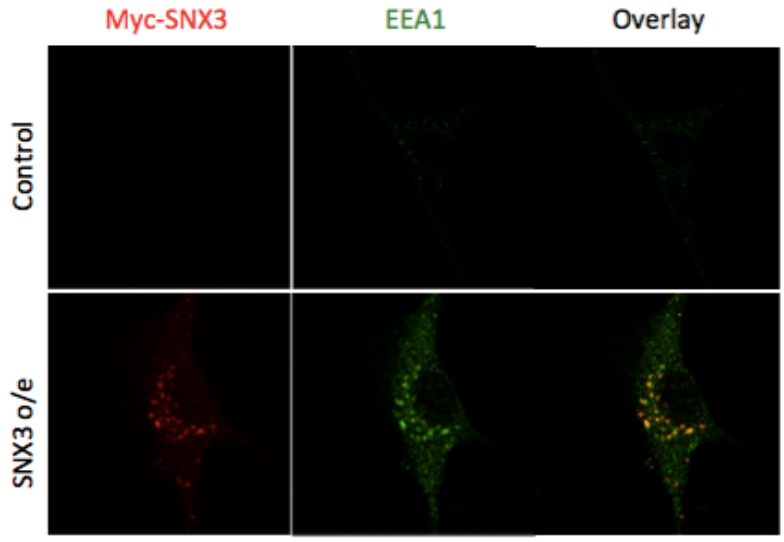
(B) Mean fold-change values of several SNXs from the cohort of 14 pGBM samples used in (A). On the left column, shown data when compared to normal brain. On the right column, *p*-fold values of SNX3 and SNX27 when comparing samples associated with Ras activation and samples without Ras activation ($p < 0.0001$, Welsch t-test, Benjamini and Hochsberg multiple testing correction)

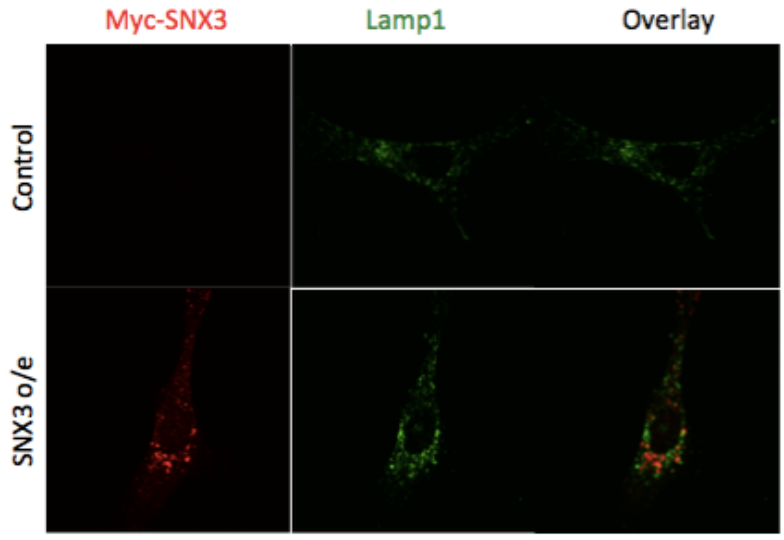
Supplementary figure S2.1 – SNX3 expression levels in SF188 transfectants

A



B

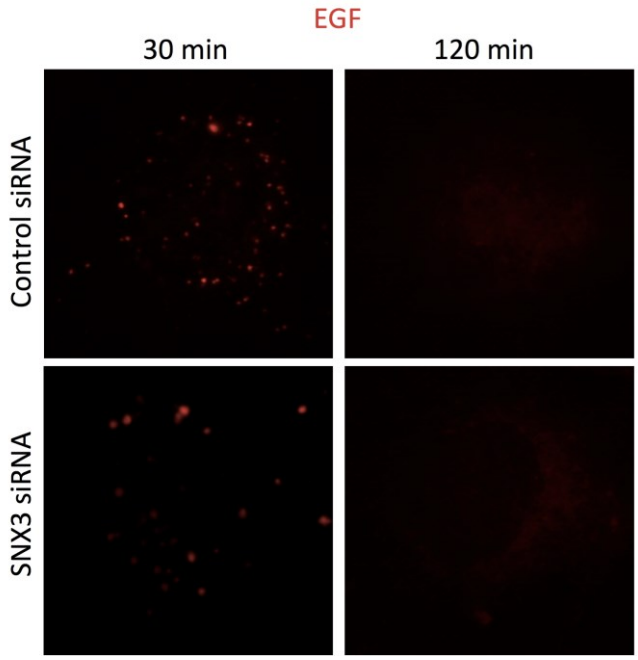


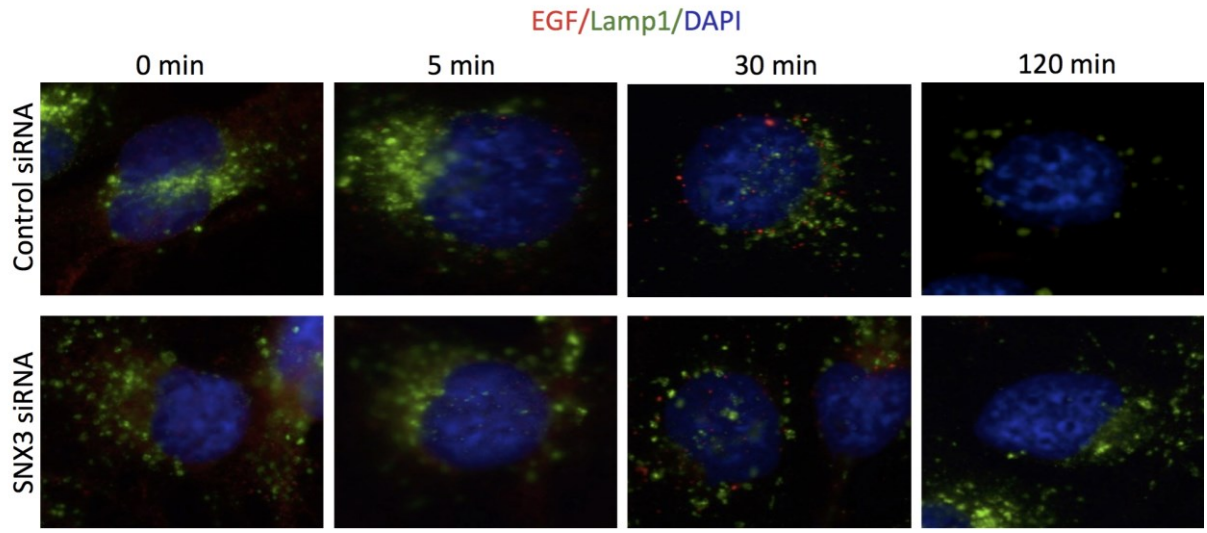


(A) SNX3 overexpression in SF188 cells. Cells were stably transfected by lipofection using a Myc-tag vector, and positive clones selected with G418 antibiotic. The clones used in this study were selected after western blot analysis. Western blot band quantification was performed using ImageQuantTM software. The ratio SNX3/actin in SF188 transfectants was normalized to the control (Empty Vector).

(B) SNX3 overexpression induces enlarged endosomes compared to control cells. SNX3 colocalizes with EEA1 and Lamp1 in the endosomal pathway. Using Myc, EEA1 and LAMP1 antibodies, indirect immunofluorescence was performed in U87 transfectants. Images were recorded on Zeiss Laser Scanning Microscope Pascal on Axiovert 200 Confocal Microscope at the McGill Cancer Centre.

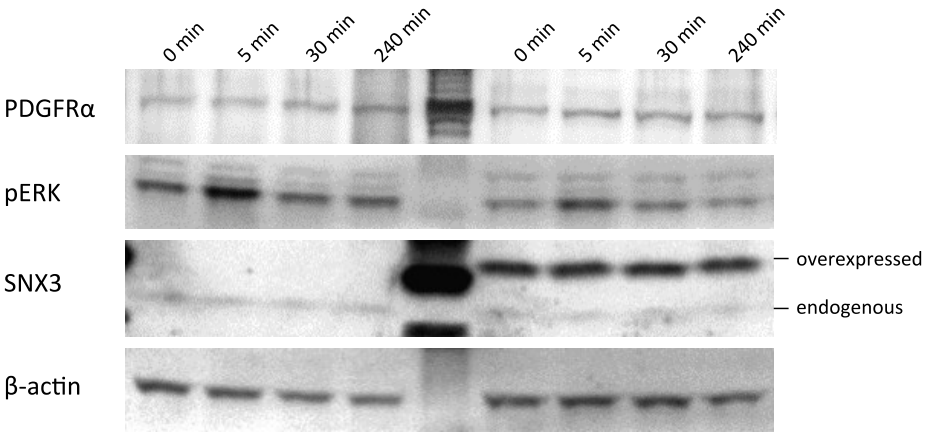
Supplementary figure S2.2 – SNX3 silencing has a marginal effect on EGFR endosomal pathway





Kinetic immunofluorescence experiments were performed after incubation of SF188 cells with Alexa555 EGF (50ng/mL of media) showing that silencing SNX3 did not affect EGF internalization and trafficking. Images were recorded on Zeiss Laser Scanning Microscope Pascal on Axiovert 200 Confocal Microscope at the McGill Cancer Centre.

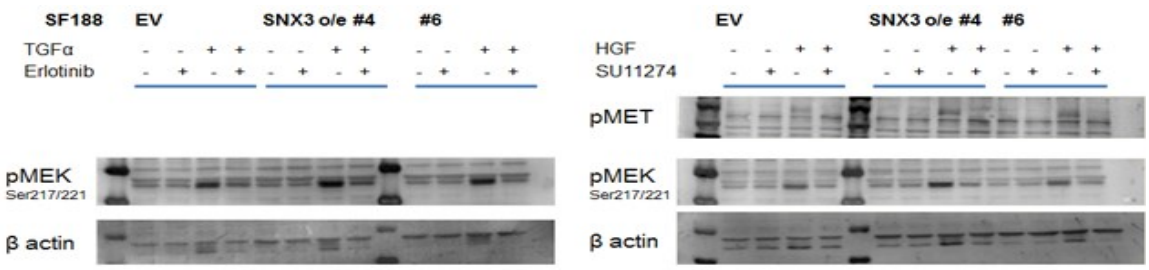
Supplementary figure S2.3 – Signalling SNX3 and PDGFRα



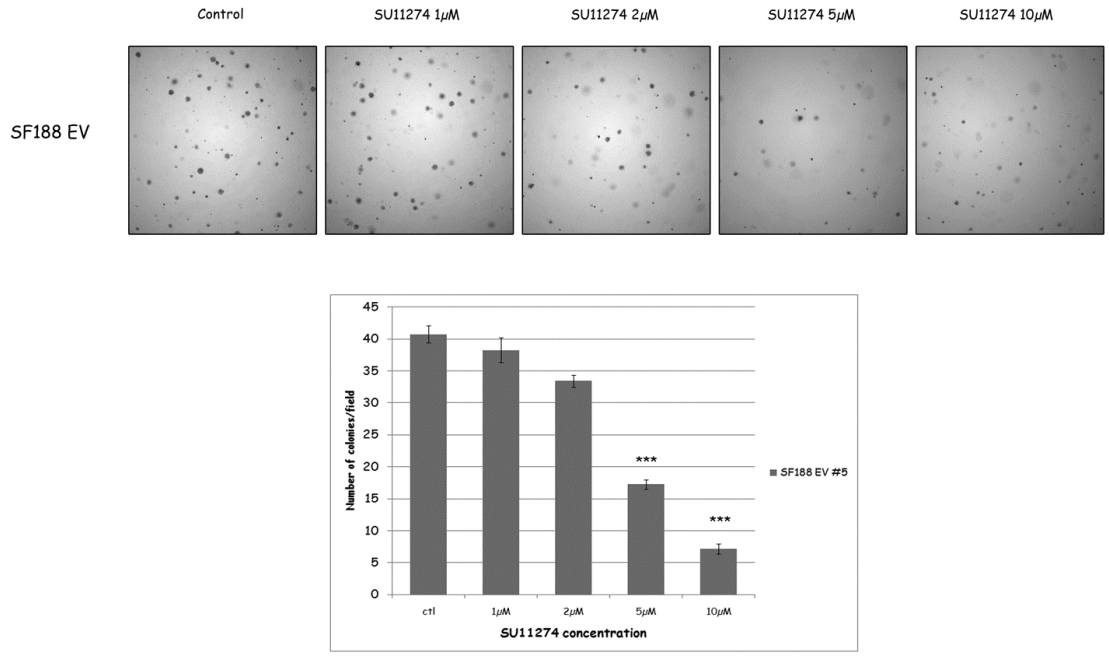
After overnight starvation, cells were stimulated with PDGFA (50ng/mL). The western blot analysis of whole cells lysates did not show any significant difference in the phosphorylation level of ERK (marker of Ras activation) both in early and late time points. This result has been observed in three separate experiments.

Supplementary figure S2.4 – Efficient targeting of tyrosine kinase inhibitors against their specific RTK

A



B



(A) SF188 cells were treated with erlotinib (10µM) before EGFR activation by TGFα, and with SU11274 (5µM) before HGF activation, western blot analysis of whole cell lysates was performed to assess the efficiency of these inhibitors to prevent the activation of their respective RTK. This result has been observed in three separate experiments.

(B) SF188 dose-dependent sensitivity to SU11274. Cells were suspended in 0.3% agarose containing either vehicle (DMSO 0.1%), or SU11274 at various concentration (1, 2, 5,

10μM). Cells were incubated for 6 weeks then colonies over 50μm were manually counted. Magnification x4. Values ± SEM; ***p<0.001.

2.8 Connecting text between Chapter 2 and Chapter 3

Understanding the molecular mechanisms leading to glioblastoma formation opens avenues for novel therapeutical approaches where conventional therapies have failed. Over the last few years, the role of receptor tyrosine kinases has been studied thoroughly mostly in adult tumors, giving clinicians good evidence of their potential as therapeutical targets. Unfortunately, tyrosine kinase inhibitors (TKI) have shown moderate efficacy in tumor control. In the particular case of pGBM, we have demonstrated that the increased and sustained activation of RTK is due to the dysregulation of the endosomal pathway. This report is the first one that gives a mechanistic approach of RTK overexpression and activation in pGBM, explaining the absence of genetic amplification in that context. Up to now, TKI have shown limited efficacy in pGBM, only tested in early phase clinical trials. We think that nowadays, identification of potential therapeutic targets should be considered at the individual scale, in order to optimize treatment and chance of response, and in this algorithm, assessment of SNX3 overexpression would be useful when TKI are used. However, more recent data on pGBM have shed light on a likely critical role of H3.3 mutations in pGBM formation. Our lab and others have described recurrent mutations in H3.3 and ATRX/DAXX complex (1, 60). In this context of identification of novel key players in tumorigenesis, we aim to characterize a specific subgroup of malignant gliomas, the DIPGs. These tumors have been described to overexpress PDGFRα and EGFR, not necessarily with associated gene amplification (57, 158); therefore they would be good candidates to further study SNX3 role in their formation. However, in the light of these recent findings we directed the second chapter of this PhD into genetic characterization of histone H3.3 and ATRX in DIPGs, with copy number alteration and clinical correlation.

Chapter 3: K27M mutation in histone H3.3 defines clinically and biologically distinct subgroups of pediatric diffuse intrinsic pontine gliomas

Dong-Anh Khuong-Quang[1]*, Pawel Buczkowicz[2,5,11]*, Patricia Rakopoulos[2,5], Xiao-Yang Liu[1], Adam M Fontebasso[3], Eric Bouffet[4], Ute Bartels[4], Steffen Albrecht[8], Jeremy Schwartzentruber[6], Louis Letourneau[6], Mathieu Bourgey[6], Guillaume Bourque[6], Alexandre Montpetit[6], Genevieve Bourret[6], Pierre Lepage[6], Adam Fleming[7], Peter Lichter[9], Marcel Kool[10], Andreas von Deimling[12], Dominik Sturm[10], Andrey Korshunov[12], Damien Faury[7], David T Jones[10], Jacek Majewski[1,6], Stefan M Pfister[10,13], Nada Jabado[1,3,7,#,] Cynthia Hawkins[2,5,11,#].

Departments of Human Genetics[1] and Experimental Medicine[3], McGill University, Montreal, Canada.

The Arthur and Sonia Labatt Brain Tumour Research Centre[2], Divisions of Haematology-Oncology[4] and Pathology[11], The Hospital for Sick Children, Toronto, Canada.

Department of Laboratory Medicine and Pathobiology[5], Faculty of Medicine, University of Toronto, Toronto, Canada.

Departments of Paediatrics[7] and Pathology[8], Montreal Children's Hospital, McGill University Health Center, Montreal, Quebec, H1P 2P3, Canada.

Genome Quebec Innovation Centre[6], McGill University, Montreal, Canada.

Divisions of Molecular Genetics[9] and Pediatric Neuro-oncology[10], Clinical Cooperation Unit Neuropathology[12], the German Cancer Research Center (DKFZ), Heidelberg, Germany.

Department of Hematology and Oncology[13], Heidelberg University Hospital, Heidelberg, Germany

*These authors contributed equally to this work.

# These authors are co-senior and co-corresponding authors.

Published in and adapted from (159)

Acta Neuropathologica. 2012 Sep;124(3):439-47. doi: 10.1007/s00401-012-0998-0.

3.1 Abstract

Pediatric glioblastomas (GBM) including diffuse-intrinsic-pontine-gliomas (DIPG) are devastating brain tumors with no effective therapy. Here, we investigated clinical and biological impacts of histone H3.3 mutations. Forty-two DIPGs were tested for H3.3 mutations. Wild type versus mutated (K27M-H3.3) subgroups were compared for *HIST1H3B, IDH, ATRX* and *TP53* mutations, copy number alterations and clinical outcome. K27M-H3.3 occurred in 71%, *TP53* mutations in 77% and *ATRX* mutations in 9% of DIPGs. *ATRX* mutations were more frequent in older children ($p<0.0001$). No G34V/R-H3.3, IDH1/2 or H3.1 mutations were identified. K27M-H3.3 DIPGs showed specific copy-number changes, including all gains/amplifications of *PDGFRA* and *MYC/PVT1* loci. K27M-H3.3 mutation defines clinically and biologically distinct subgroups and is prevalent in DIPG, which will impact future therapeutic trial design. K27M- and G34V- H3.3 have location-based incidence (brainstem/cortex) and potentially play distinct roles in pediatric GBM pathogenesis. K27M-H3.3 is universally associated with short survival in DIPG while patients wild type for H3.3 show improved survival. Based on prognostic and therapeutic implications, our findings argue for H3.3-mutation testing at diagnosis, which should be rapidly integrated into the clinical decision-making algorithm, particularly in atypical DIPG.

Key words: DIPG, H3.3, ATRX, TP53, survival, targeted therapy.

3.2 Introduction

High-grade astrocytomas (anaplastic astrocytoma and glioblastoma (GBM)) are the most biologically aggressive form of cancer and a leading cause of cancer-related mortality and/or morbidity in the pediatric years (4, 160, 161). They account for 20% of all brain tumors in children and occur mainly supratentorially in the cortex or thalamus, or in the brainstem where they are called diffuse intrinsic pontine gliomas (DIPGs). Cortical GBM can be amenable to complete surgical resection, however up to 85% will die within 2 years of diagnosis (161-163). DIPGs cannot be surgically removed, because of their location and the infiltrative nature of the disease. They have a median survival of less than one year, with fewer than 10% of children surviving for more than 2 years (164, 165). In addition, based on their infiltrative nature and location within the brainstem, DIPGs are often diagnosed clinically based on a combination of neurological signs, duration of symptoms and specific neuro-imaging findings. Currently, biopsy of these tumors is controversial because the findings do not alter therapy if the child presents with classic clinical and imaging features. However, biopsy may be helpful if biologic information gleaned from the tissue may guide therapy or provide additional prognostic information.

Despite aggressive therapeutic approaches and decades of clinical trials evaluating numerous chemotherapeutic and radiation therapy regimens, there has been no improvement in survival for children with GBM. The impediment to treatment is the invasive capacity of these high-grade astrocytomas within the brain and their inherent resistance to adjuvant therapies. In addition, permanent damage inflicted to a developing brain by current life-saving therapies severely impacts the quality of life of surviving children (166-168).

A number of recent comprehensive studies have reported differences at both the copy number and expression levels that distinguish pediatric DIPG from both their adult and pediatric supratentorial GBM counterparts, indicating that they may be separate biologic entities (56, 57, 169). These studies also identified frequent up-regulation of receptor tyrosine kinases (RTKs) in DIPGs, in particular PDGFR-alpha, MET and IGF1R. These RTKs are also over-expressed in supratentorial GBM, albeit at much lower levels. These findings spearheaded several ongoing clinical trials targeting these RTKs, however initial results show similarly poor response rates to those seen to previous, more standard therapies (170). This suggests that RTK-inhibition alone may not be sufficient to combat DIPG.

We (1) and others (60) recently identified mutations in histone H3.3 (gene name *H3F3A*) at either amino acid 27, resulting in replacement of lysine by methionine (K27M), or at amino acid 34, resulting in replacement of glycine by valine or arginine (G34V/R), as molecular drivers of a subgroup of pediatric and young adult GBM. We also showed in supratentorial pediatric GBM that H3.3 mutations significantly overlapped with mutations in *TP53* and in *ATRX* (alpha-thalassemia/mental-retardation syndrome-X-linked) (1), which encodes a subunit of a chromatin remodelling complex required for H3.3 incorporation at pericentric heterochromatin and telomeres (171, 172). K27M mutations in H3.3, or in the related H3.1, were additionally found in 60% and 18% of DIPGs, respectively (60). Here we investigate the frequency of these mutations in a large series of 42 DIPGs. We additionally assess whether *ATRX* mutations are prevalent in DIPG, and whether they overlap with histone H3.3 and/or *TP53* mutations similar to our findings in supratentorial GBM (1). Lastly, we investigate the clinical and biologic features of DIPG subgroups based on histone H3.3 mutation status.

3.3 Patients and Methods

Patients and samples

Patient biological material was collected from the Hospital for Sick Children in Toronto, Canada, The Montreal Children's Hospital/McGill University Health Center in Montreal, Canada, and from the German Cancer Research Center (DFKZ) in Heidelberg, Germany. The study was approved by the Institutional Review Boards of the respective hospitals. Patients were included if they had classic DIPG MRI findings and clinical presentation, including short duration of symptoms (classic DIPG), or had atypical MRI findings and/or clinical presentation (atypical DIPG) but had biopsies demonstrating high-grade astrocytoma. Cases were independently reviewed by senior pediatric neuropathologists (CH, SA, AVD) according to the WHO guidelines. Sixteen of the DIPG samples were pre-treatment biopsies, twenty-five were post-treatment autopsy specimens and one sample was collected at autopsy from an untreated patient (DIPG02). The mean age of diagnosis was 7.12 years (range, 0 to 17 years) with a median survival of 0.83 years (Figure 3.1A). Clinical characteristics of patients are summarized in Table 3.1. All patients were considered and treated as DIPGs in their respective centres. Forty patients had astrocytomas (38 high-grade and 2 grade II). The other two cases had no immunohistochemical evidence of glial differentiation and were labeled as primitive neuroectodermal tumors based on autopsy. Clinical characteristics of the 48 pediatric supratentorial GBMs were previously described (1).

Sanger sequencing

Coding exons of *H3F3A*, *HIST1H3B*, *ATRX*, *TP53* and *IDH1* and *2* were sequenced using Sanger fluorescent sequencing after amplification by polymerase chain reaction using standard methods, at The Hospital for Sick Children or McGill University/Genome Quebec Centre (primer sequences in Supplementary Table 3.1). The *TP53* gene was sequenced for the

entire coding sequence (exons 2-11) and the spanning intron-exon junctions with primers as previously described(173). Sequences were analyzed using Applied Biosystems' 3730xl DNA Analyzer technology.

Array hybridization and data analysis

Twenty samples were hybridized to the Genome-Wide Human SNP Array 6.0 and three to the Human Mapping 250 SNP Nsp Array from Affymetrix (Santa Clara, CA) (Table 3.1). The sample preparation, including DNA extraction, digestion, labelling and hybridization were performed as directed by the manufacturer. Data was analyzed using Partek Genomics Suite v6.4 (Partek Incorporated, St. Louis, MO) and Genotyping Console 4.1 (Affymetrix), GISTIC2.0 (Broad Institute, Cambridge, MA).

Immunohistochemistry

Formalin-fixed, paraffin-embedded (FFPE) sections were immunohistochemically stained for nuclear ATRX as previously described (1). 5-μm sections were cut from paraffin blocks and mounted on positively charged microscope slides. Following an overnight incubation at 60^{o}C, the slides were de-waxed in xylene and hydrated by washes in decreased concentration of ethanol in distilled water. Sections were heat treated in 10 mM citrate buffer for the purpose of antigen retrieval and blocked for endogenous biotin and peroxidase. The tissue sections were incubated at 4^{o}C overnight with rabbit anti-human ATRX (HPA 001906; Sigma-Aldrich, St. Louis, MO) at a 1:600 dilution. Immunodetection utilized 3,3'diaminobenzidine (DAB) and counterstaining was conducted with hematoxylin. The sections were scored for nuclear ATRX positivity by two independent observers blinded to the clinical data.

Statistical analysis of clinical and molecular data

Analyses were performed using GraphPad Prism 5 software (La Jolla, CA, USA). When appropriate, two group comparisons were analyzed with two-sided Fisher's exact test, and continuous scale data were analyzed with unpaired two-tailed Student t-test. Overall survival curves were analyzed using the Kaplan-Meier method and the log-rank test was used to make univariate assessments of Kaplan-Meier plots. P-value ≤ 0.05 was considered significant. Multivariate analysis was done using multivariate Cox proportional hazards models and significance testing ($\alpha = 0.05$) based on the Wald test. Analysis of significant focal amplifications and deletions was conducted using GISTIC 2.0 (Broad Institute, Cambridge, MA) with significance being assigned to regions with false discovery rate less than or equal to 5% ($q \leq 0.05$).

3.4 Results

Histone H3.3 mutations are frequent in DIPG

We sequenced *H3F3A* in 42 DIPG samples comprising either biopsy material prior to any treatment (n=16) or autopsy samples (n=26, one sample from untreated patient at autopsy; DIPG02). We identified the recurrent mutation in Histone H3.3 leading to K27M amino acid substitution in 30/42 (71%) DIPGs (Table 3.1). K27M-H3.3 was identified in pre-treatment biopsy samples as well as autopsy material indicating that it is present at diagnosis and not induced by therapy. No *HIST1H3B* (0/29, including 12 of the H3.3 wild-type patients tested for this mutation), *IDH1* (0/23) or *IDH2* (0/20) mutations were identified in our sample set. G34V/R-H3.3 previously identified in 13% of pediatric and young adult supratentorial GBM was absent in DIPG. The K27M-H3.3 mutation was more prevalent in DIPGs (71%) compared to supratentorial GBMs (14%) (Figure 3.2A, $p<0.0001$).

ATRX mutations are associated with older patient age

We previously showed that G34V/R-H3.3 GBM samples universally also carried *ATRX* and *TP53* mutations (13/13) while K27M-H3.3 GBM samples had significant, albeit lower, overlap with *ATRX* and *TP53* mutations (respectively 30% and 60% of the 10 samples investigated) (1). In the current study, using sequencing and/or immunohistochemical analysis, we identified 2/22 DIPG samples with *ATRX* mutation or loss of immunostaining (suggestive of an underlying mutation) (Table 3.1). Both cases with ATRX mutation also harboured the K27M-H3.3 mutation. In DIPGs *ATRX* mutations tended to be found in older children (mean age 11.82 (±1.18) versus 5.20 (±0.81) years, p=0.02 for *ATRX* mutant versus wild-type DIPGs; Figure 3.2B). The same age-associated distribution of *ATRX* mutation was identified across the entire GBM cohort including both DIPGs and supratentorial GBM (mean age 16.91 (±2.11) versus 8.00 (±0.69) years, $p<0.0001$ for *ATRX* mutant versus wild-type cases; Figure 3.2C).

TP53 mutations are frequent in both H3.3 mutant and wild-type DIPGs

As previously described, a significant number of DIPG samples carried mutations in *TP53,* (17/22, 77%). Fourteen of these samples carrying *TP53* mutations were also mutant for K27M-H3.3 (Table 3.1). However, even though there was overlap between K27M-H3.3 and *TP53* mutations in DIPGs, there was no significant difference in the frequency of *TP53* mutations between the K27M-H3.3 mutated and wild-type groups (78% and 75%, respectively).

DIPG subgroups based on H3F3A mutation status show differing copy number alterations

Analysis of DNA copy number alterations in K27M-H3.3 versus H3.3 wild-type DIPG samples showed areas of overlap but also major differences between both groups. Large

chromosomal copy number alterations common to both groups included loss of 10q, 11p, 13q and 14q as well as gains of 1q and 19q (Figure 3.3A). H3.3 wild-type tumors showed large scale gains of chromosomes 2p and 7p as well as losses of chromosome 9p and 12q. Samples carrying the K27M-H3.3 mutation exhibited common loss of chromosome 5q, 6q, 17p and 21q (Figure 3.3A). Gains in the mutated group included 19p.

Focal recurrent gains and deletion in both groups were further analyzed using GISTIC2.0. H3.3 wild-type patients had significant focal gains/amplifications of regions 2p25.1 (q = 0.028) and 2p24.3 (q = 0.028) including the *ASAP2* and *MYCN* genes, respectively (Figure 3.3B). No significant recurrent focal copy number losses were observed in this group (Figure S3.1). Analysis of frequent focal copy number alterations in the K27M-H3.3 group revealed amplification of 4q12 (q=0.00015) and 8q24.21 (q=0.033) corresponding to *PDGFRA* and *MYC/PVT1* locus gains/amplifications, respectively (Figure 3.3B). Interestingly, *PDGFRA* gains/amplifications were found in 40% (6/15; 1 pre-treatment, 5 post-treatment) of patients in the H3.3 mutant group, while no gains/amplifications of *PDGFRA* were detected in the H3.3 wild-type group. Significant areas of focal deletion in K27M-H3.3 samples included 4p16.3 (q=0.021), 11p15.4 (q=0.000028), 11q22.1 (q=0.022) and 15q24.1 (q=0.016) (Figure S3.1).

Histone H3.3 wild type status is associated with better overall survival

H3.3 mutational status and survival data was available for 39 DIPG patients, 27 of which (69%) carried the K27M-H3.3 mutation. The mean overall survival for patients with K27M-H3.3 mutated tumors was 0.73 years ((±0.48)) versus 4.59 years (±5.55) (p=0.0008) for patients with wild-type tumors. Kaplan-Meier survival analysis revealed significantly worse overall survival of DIPG patients carrying the K27M-H3.3 mutation (Log rank p=0.0027 versus wild-type patients) (Figure 3.1B). Similarly, when patients were stratified by

underlying histologic diagnosis (anaplastic astrocytoma versus GBM) Kaplan-Meier survival analysis also demonstrated significantly worse overall survival of DIPG patients carrying the K27M-H3.3 mutation (Log rank p=0.013 versus wild-type patients). All of the long-term survivors were included in the H3.3 wild-type group. The mean age of diagnosis of patients with K27M-H3.3 mutations was 8.13 (±3.75) years versus 4.57 (±4.07) years for the wild-type patients (p=0.010). The distribution of age of diagnosis based on K27M-H3.3 mutational status is shown in Figure 3.1C. Multivariate analysis (Cox Regression), including age, histologic diagnosis and H3.3 mutation status, demonstrated H3.3 mutation status to be the only significant predictor of overall survival with a hazard ratio of 4.3 (95% confidence intervals 1.3-14.5, p=0.019) (table 3.2).

3.5 Discussion

Our findings confirm that the K27M mutation in histone H3.3 is a frequent event in pediatric DIPG. We further show that the both the type of H3.3 mutation and their association with *ATRX* mutations are age and location dependent. The G34V/R-H3.3 mutation was not found in DIPGs whereas it represents 13% of H3.3 mutations in supratentorial GBM. A recent report similarly did not find mutations at the G34 residue of H3.3 in DIPG (60). Similarly, *ATRX* mutations were infrequent in DIPG but were present in 29% of supratentorial GBM. This may be an age-related phenomenon as the mean age of our DIPG cohort was 7.1*vs* 12.24 years for the supratentorial GBM patient cohort. In support of this, G34V/R-H3.3 was seen in older patients (mean age 19.66 years (±1.56)) and almost exclusively in hemispheric GBM(1). Similarly *ATRX* mutation-positive patients were significantly older than wild-type patients. We did not identify K27M-H3.1, which has been recently identified in 18% (9 samples) of DIPGs (60). This difference in frequency may be due to sampling bias, however, our findings

support H3.3 as the major histone to be targeted in pediatric GBM and K27 the major residue affected in DIPG.

H3.3 is the major histone to be loaded on chromatin during brain development. This histone variant is known to modulate specific chromatin changes and gene expression profiles and to be associated with active chromatin and translation. Histone lysine methylation has emerged as an important player in regulating gene expression and chromatin function (174). K27 is a critical residue in all seven histone 3 variants and the subject of post-translational histone modifications as it can be both methylated and acetylated (174-176). Acetylation may induce active transcription while mono, bi or tri-methylation of K27 is associated with a repressive mark on chromatin and gene expression. Abrogation of acetylation and/or potential mimicry of a methylated lysine through the methionine substitution are likely to interfere with chromatin function, inducing defects in chromatin remodelling and tumorigenesis. This is supported by our observation of specific copy number changes associated with mutant K27M-H3.3. Further studies aiming to model this mutation are required to precisely determine the effect of this mutation in chromatin remodelling in pediatric GBM.

ATRX mutations were only identified in 9% of DIPGs compared to 29% of supratentorial pediatric GBM. Notably, presence of the *ATRX* mutation significantly overlapped with *TP53* mutations in GBM ($p=0.01$) regardless of the location within the brain and with G34V/R mutants in supratentorial GBM ($p<0.0001$), and was age-dependent as it mainly occurred in older children ($p<0.0001$). The requirement for *ATRX* mutations in GBM may thus be due to tumor location and/or the age of the patient. This is potentially indicative of a different cell of origin or age-related plasticity of the tumor, similar to differences in genetic alterations seen based on age in infant MLL-positive leukemia (177). *TP53* mutations are associated with the vast majority of both K27M and G34V/R H3.3 mutations identified in pediatric and young

adult GBM. In both DIPG and supratentorial GBM, *TP53* alterations were commonly identified (77% and 54% respectively). Interestingly, in DIPGs the K27M-H3.3 and wild-type H3.3 subsets had similarly high *TP53* mutation and allelic loss rates. In K27M-H3.3-mutated tumors this may thus represent an important second hit however, our data is also indicative of an important role of *TP53* mutations in the pathogenesis of GBM independent of H3.3 mutational status.

One of most common copy number gains reported in multiple genomic studies of DIPG and pediatric GBM is that of *PDGFRA* (56, 57, 138). Here we report *PDGFRA* gain or amplification to be seen exclusively in patients carrying K27M-H3.3 mutations, where it is present in 40% of cases. We also identify gains and amplifications in a gene locus containing *MYC/PVT1*, here also exclusively in K27M-H3.3 mutants. PVT-1 is an oncogene and a Myc protein target known to be over-expressed in transformed cells (178). Amplification of *MYC/PVT1* has been shown to contribute to the pathogenesis of ovarian and breast cancer, and is part of the chromosome 8q24 prostate cancer risk locus (179). The finding of these copy number changes in a subset of K27M-H3.3 mutants suggests that *PDGFRA* and *MYC/PVT1* locus gains/amplifications are subsequent to K27M-H3.3 mutations. The addition of histone modifying agents to RTK inhibitors may thus be of therapeutic benefit in this group of patients.

A clinically significant finding of this study is the fact that patients who harbour the K27M-H3.3 mutation have worse overall survival when compared to patients that are wild type for H3.3. This association with survival was independent of patient age and histologic diagnosis. The only attributable histologic feature exclusive to the K27M mutant group is glial differentiation. However, not all of the samples mutated for K27M-H3.3 met criteria for GBM. The hypothesis that mutation status identifies distinct subtypes is further supported by

the differences in copy number profiles and age distribution between these patient groups. Importantly, long-term survivors were only identified in the group of patients that are wild type for this gene. Some of these patients had an atypical clinical presentation (longer duration of symptoms or atypical radiology) and thus were biopsied, demonstrating high-grade histologic features. Despite high-grade histology, this group of H3.3 wild-type patients did not follow the expected clinical course of what can be considered classic for children with DIPG, suggesting a potentially different clinical and molecular entity which should be added to the group of "atypical" DIPG when considering clinical trial design. These H3.3 wild-type tumors may be more heterogeneous in terms of histology and biological features than K27M-H3.3 tumors, and may perhaps, with larger numbers, be further divided into different sub-groups. One uncommon subgroup of H3.3 wild-type brainstem tumors, which is hinted at by our series is the PNETs. Interestingly, the two PNET patients in our cohort presented with "classical" DIPG features were treated as DIPG patients and both had a poor outcome despite the H3.3 wild-type status. This raises the hypothesis that H3.3 wild type status may suggest better outcome only in glial neoplasms.

In contrast, the H3.3 mutated group contained patients who, at autopsy, had tumors, which, if classified by WHO guidelines, would be considered diffuse astrocytoma, grade 2. Nevertheless, these patients had the short survival expected of classic DIPG. Conversely, some patients with high-grade histology such as those with features of GBM were in the wild-type group. Thus, mutational status of H3.3 may be more helpful than histologic appearance alone in identifying patients expected to have a poor clinical outcome at presentation.

Our findings support performing a stereotactic biopsy, particularly for atypical clinical cases of DIPG. The finding of the K27M-H3.3 mutation can be considered as diagnostic of classic DIPG with its expected poor outcome. These patients may benefit from agents targeted at

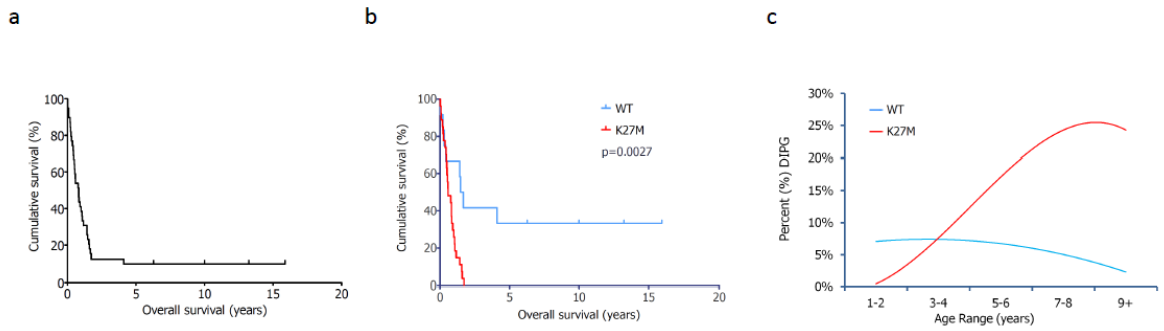
chromatin remodelling and/or histone post-translational modifications with an additional RTK inhibitor. Although not universally indicative of better clinical behaviour, wild-type patients should perhaps be considered as atypical DIPG, and if coupled with atypical radiology and/or clinical presentation may warrant a different therapeutic approach.

3.6 Acknowledgments

This work was supported by the Canadian Institutes of Health Research (CIHR, MOP 115004), the Cole Foundation, and was funded in part by a Genome Canada/CIHR grant (co-funding from Genome BC, Genome Quebec, CIHR-ICR (Institute for Cancer Research) and C17, through the Genome Canada/CIHR joint ATID Competition (project title: The Canadian Paediatric Cancer Genome Consortium: Translating next generation sequencing technologies into improved therapies for high-risk childhood cancer (NJ, CH). This work was partially funded by the ICGC project PedBrain Tumor (#108456) granted by the Bundesministerium für Bildung und Forschung (BMBF) and the Deutsche Krebshilfe (AK, PL, and SMP). D.A. Khuong-Quang is the recipient of a studentship from the Foundation of Stars. X. Liu and A. Fontebasso are the recipients of studentships from CIHR. N. Jabado is the recipient of a Chercheur Boursier Award from Fonds de Recherche en Santé du Québec.

3.7 Figures

Figure 3.1 – K27M-H3.3 is associated with worse overall survival and higher age of diagnosis in DIPG

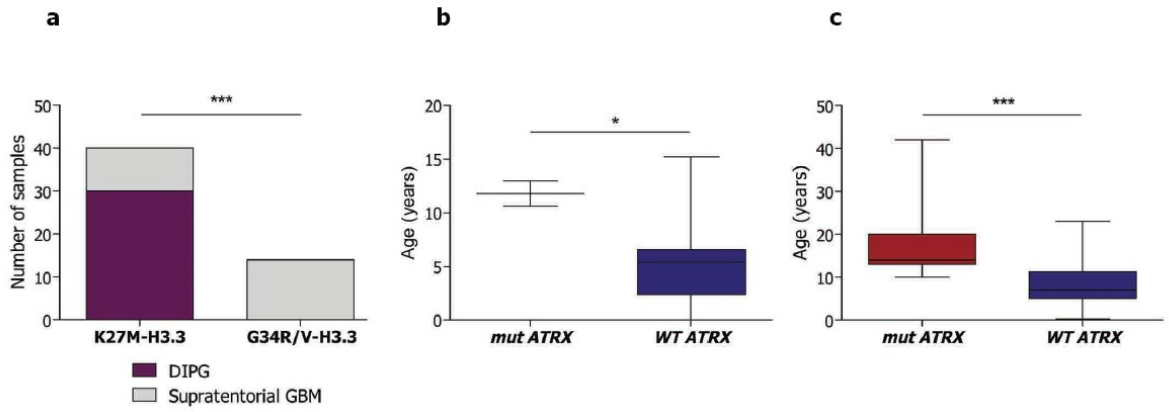


(A) Kaplan-Meier curve of overall survival for all DIPG patients (n=39).

(B) DIPG patients carrying K27M-H3.3 mutation have worse overall survival compared to patients wild type for this histone as determined by Kaplan-Meier analysis (Log-rank, p=0.0027). Notably, all long-term survivors were wild type for *H3F3A*.

(C) Age distribution of DIPG patients based on K27M-H3.3 mutational status. DIPG patients mutated for K27M-H3.3 have a higher age of diagnosis 8.13 years (±3.75) as compared to wild-type patients (4.57 years (±4.07), p=0.010).

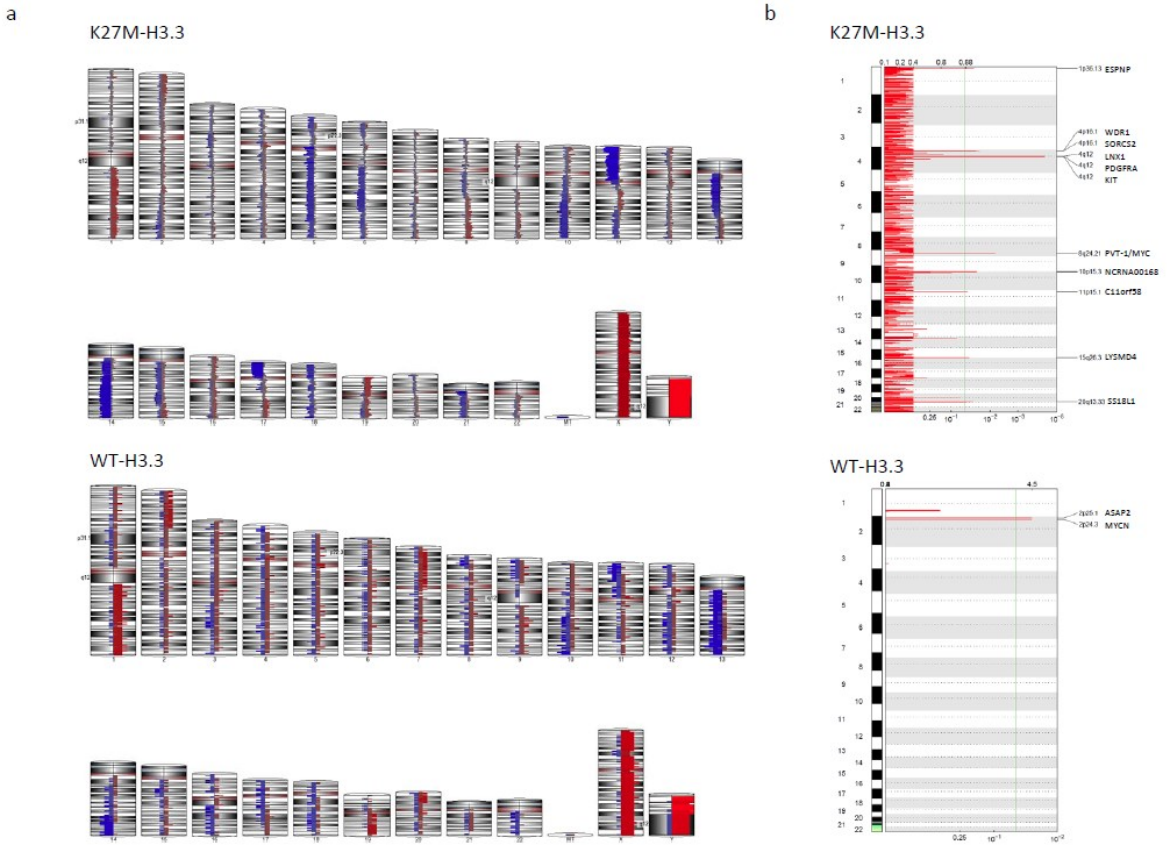
Figure 3.2 – K27M-H3.3 is prevalent in DIPG and is associated with ATRX mutations mainly in older children



(A) Distribution of DIPG and supratentorial GBM based on H3.3 mutations suggests prevalence of K27M-H3.3 in DIPG.

(B-C) *ATRX* mutations in DIPG (B) and all location pediatric GBM (C) are significantly more prevalent in tumors from older children (mean ages 11.82 years ±1.18 and 16.91±2.11, respectively) as compared to children with no *ATRX* mutation (mean ages of 5.20 years ±0.81 and 8.00±0.69, respectively) ($p=0.02$ and $p<0.0001$, respectively).

Figure 3.3 – Whole chromosome view of copy number alterations (CNA) in K27M-H3.3 mutants and wild-type DIPG samples



(A) Similarities in CNA between both groups included loss of 10q, 11p, 13q and 14q as well as gains of 1q and 19q. However, major differences in copy number were identified with samples wild type for K27M-H3.3 exhibiting gains of chromosome 2p and 7p as well as losses of chromosome 9p and 12q while samples mutants for K27M-H3.3 commonly exhibited loss of chromosome 5q, 6q, 17p and 21q.

(B) Focal recurrent amplifications determined by GISTIC 2.0 analysis ($q \leq 0.05$) show significant differences in focal gains between samples carrying K27M-H3.3 and samples wild type for H3.3. This included *PDGFRA* (4q12), *MYV/PVT1* locus (8q24.21) gains and amplifications, which were exclusively identified in K27M-H3.3 mutants and *ASAP2* (2p25.1)

and *MYCN* (2p24.3) gains and amplifications, which were exclusively identified in wild-type patients.

3.8 Tables

Table 3.1 – Patient characteristics and mutational status of samples for H3.1, H3.3, ATRX and TP53

Abbreviations: Age Dx – Age of Diagnosis in years; OS – overall survival in years; GBM – glioblastoma multiforme; AA – anaplastic astrocytoma; LGA – low grade astrocytoma; K27M – lysine to methionine at residue 27 of H3.3; WT – wild-type; (*) patient still alive at last follow-up; (#) *ATRX* mutations tested by immunohistochemistry.

Sample ID	Age Dx	Gender	OS (y)	Path Dx	Treatment	H3.3	H3.1	*ATRX*	*TP53*	*IDH1/2*	SNP array
DIPG04	10.6	M	0.85	GBM	post	K27M	WT	MUT	MUT	WT	SNP6.0
DIPG05	8.9	F	1.00	GBM	post	K27M	WT	-	MUT	-	SNP6.0
DIPG06	7.2	M	0.50	GBM	post	K27M	WT	WT	MUT	WT	SNP6.0
DIPG07	5.0	M	0.55	GBM	post	K27M	WT	WT	MUT	WT	SNP6.0
DIPG08	5.8	F	0.44	GBM	post	K27M	WT	WT	MUT	WT	SNP6.0
DIPG10	7.6	M	0.23	GBM	pre	K27M	-	-	-	-	250k
DIPG13	6.6	F	0.83	GBM	post	K27M	WT	WT	MUT	WT	SNP6.0
DIPG14	5.2	M	1.57	AA	post	K27M	-	-	-	-	-
DIPG15	11.3	M	0.26	GBM	post	K27M	-	WT[b]	-	-	-
DIPG17	6.4	F	0.46	GBM	post	K27M	WT	WT	MUT	WT	SNP6.0
DIPG18	8.0	F	0.03	GBM	post	K27M	-	-	-	-	-
DIPG20	6.4	M	0.92	GBM	pre	K27M	-	WT[b]	-	-	-
DIPG21	2.4	F	1.43	AA	post	K27M	WT	WT	WT	WT	SNP6.0
DIPG22	3.9	F	0.59	LGA	post	K27M	WT	WT	WT	WT	SNP6.0
DIPG23	6.6	M	0.78	GBM	post	K27M	WT	WT	MUT	WT	SNP6.0
DIPG24	7.8	M	0.37	AA	post	K27M	WT	-	MUT	WT	SNP6.0
DIPG26	3.9	M	0.09	GBM	post	K27M	WT	-	WT	WT	SNP6.0
DIPG27	4.4	F	0.19	LGA	post	K27M	WT	-	-	-	-
DIPG29	8.1	M	1.73	GBM	post	K27M	WT	-	MUT	WT	SNP6.0
DIPG30	7.6	F	1.06	GBM	post	K27M	WT	-	MUT	WT	SNP6.0
DIPG31	14.2	F	1.60	GBM	post	K27M	WT	-	-	-	SNP6.0
DIPG32	14.0	F	0.52	GBM	post	K27M	-	-	-	-	-
DIPG34	8.0	M	0.83	GBM	pre	K27M	-	-	-	-	-
DIPG35	11.0	F	1.17	GBM	pre	K27M	WT	-	-	-	-
DIPG36	6.0	F	-	GBM	pre	K27M	-	-	WT	WT[c]	-

DIPG37	13.0	F	1.08	GBM	pre	K27M	-	MUT	MUT	WT	-
DIPG39	17.0	M	0.58	AA	pre	K27M	-	-	-	WT	-
DIPG40	4.0	M	0.08	GBM	pre	K27M	-	WT	-	WT	-
DIPG41	16.3	F	-	GBM	pre	K27M	-	-	MUT	WT[c]	-
DIPG42	7.0	M	-	GBM	pre	K27M	-	-	MUT	WT[c]	-
DIPG01	5.0	F	0.42	PNET	post	WT	WT	WT	MUT	WT	SNP6.0
DIPG02	0.0	F	0.03	PNET	pre	WT	WT	WT	-	-	250k
DIPG03	4.6	M	1.48	GBM	post	WT	WT	-	-	-	SNP6.0
DIPG09	6.5	M	4.11	GBM	post	WT	WT	-	MUT	WT	SNP6.0
DIPG11	1.7	M	6.28[a]	GBM	pre	WT	WT	WT[b]	-	-	250k
DIPG12	0.3	F	1.68	AA	pre	WT	WT	WT[b]	-	-	-
DIPG16	5.8	M	0.31	GBM	pre	WT	WT	WT	MUT	WT	SNP6.0
DIPG19	3.1	F	0.21	AA	post	WT	WT	-	-	-	
DIPG25	7.1	F	1.43	AA	post	WT	WT	WT	WT	WT	SNP6.0
DIPG28	15.2	F	15.88[a]	AA	pre	WT	WT	WT[b]	-	-	-
DIPG33	2.5	M	13.23[a]	AA	pre	WT	WT	WT[b]	-	-	-
DIPG38	3.0	M	10.00[a]	GBM	pre	WT	WT	WT	-	-	-

Table 3.2 – Multivariate Cox regression analysis

Variable	HR	P value	Lower 95% CI	Upper 95% CI
K27-H3.3	4.277	0.019	1.264	14.472
Histology	1.93	0.246	0.636	5.862
Age Dx	0.96	0.608	0.823	1.121

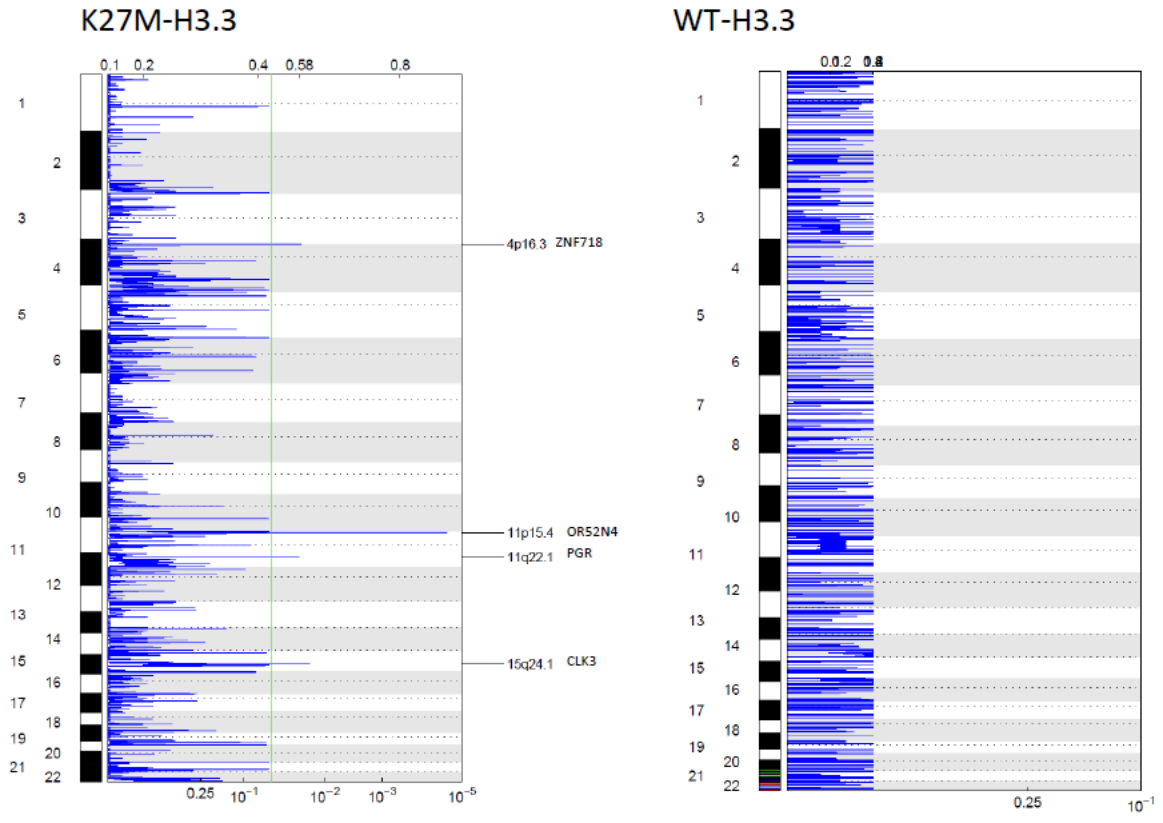
3.9 Supplementary material

Supplementary table S3.1 – Primer sequences used for validation of mutations in H3F3A, ATRX, HIST1H3B, IDH1 and IDH2

Sup Table 1 - Primer design for *H3F3A*, *HIST1H3B*, *ATRX*, *IDH1* and *IDH2*.

Gene	Exon No.	Forward Primer	Reverse Primer
H3F3A	2	TTGATTTTTCAATGCTGGTAGG	TTTAAGCAGTAGTTAAGTGTTCAAATG
H3F3A	2	AAGCAACTGGCTACAAAAGC	AAGAGAGACTTTGTCCCATTTT
H3F3A	3	GAAGAATTGTTGGGTGGCTTA	TCCATAAACCAAACCAGAGGA
H3F3A	4	GCTGCTTTTGAGCTTTGTCC	GTCTCACGTCTCAGATATAAAGTGAC
HIST1H3B	-	TTGGTGGTCTGACTCTATAAAAGAA	CGGTAACGGTGAGGCTTTT
ATRX	1	ATGGTCCTGTGAATGCCATC	GGCATTTAAGGGGACCAAAC
ATRX	2	ACCTTGGGAAATCCCGAATA	CAATGACTATCCATCCCTCCA
ATRX	3	GTTGGCAAATGGAAGGATTC	AGTAGGGGGTGGAGGGTACA
ATRX	4	GGGTGAAAAGGGTGTTTTGTT	GCATAGGGAACCCTCAACAA
ATRX	5	GGTTTTAGTTTCTAGTACAGTTGACCA	GGGGAATGTGTTCCTAAAACC
ATRX	6	GCAAAATTGCTGATGAGTTTTT	CATTTTATTATCCTTGAAAAATTCTGA
ATRX	7	AAGAAATGAATTCTCTGAACTCTTGA	CCAACTTTGTTTCCCTCTCTG
ATRX	8	TGATGAGCAAGGTGGAAAATC	TGACACTGTTTTGCAACCTGA
ATRX	9	AAATCCTGCTGGGATTTTTG	GGGTAGTTTTGTTTCTTTTGTTGC
ATRX	10	CCCCATGGGTAGGTCTTTTT	TTGCTTGTATTGGCCTAGCA
ATRX	11_12	TCAGTCCTTCCTCAGCTCGT	TCCATGATAAAGGCAACATTCA
ATRX	13	GCTTCTCTACACTGCCAAAAGTG	TTCTGCTTCCAATAGATGCTTT
ATRX	14	TTGTGGGTTTAGAAAGGGTAAA	TGCAAAACTGAAAAAGAACAACA
ATRX	15	TGAGCATTTCATTGGGGAAT	TGAAAGAGCGGGAAAGAAAA
ATRX	16	TTAACCAAATACGGGAGCAGA	TTTCACAGCAGACTAAGATGAACC
ATRX	17	GGCAAGAGGGATTAAAAGATGA	TGGCGACATTAAGGGTGATT
ATRX	18	TTGGAAATTCTGGCCGTTTA	TTCCCACTGAAATATGCATCAC
ATRX	19	TCTTCAGCCCCTACGACTGT	GAAGGAAAGTCCCCCTGTTC
ATRX	20	CAATTGGATTTGTGGTGTGG	CCACCCCACTCACCAATTTA
ATRX	21	CCACCTTTTCCTGCTGTGTT	CATTAATAGAAATAAATTAAGG
ATRX	22	TGGAACAGAGAGGTAACAGCA	AACAAACCTCCCCTCAGGAT
ATRX	23	TGCTCTGTTTTAATGTCGAGTCA	TGAAGGCATGGTCATTCAGA
ATRX	24	CAGCTTCCCAAAGTGCTAGG	CGAGGCATTTTAAAGGCTGA
ATRX	25	TCTATTGGCACATTTATTTCT	TTGGCCTCCCAAAGTCCTGAGATT
ATRX	26	TTTGGAGTCCAGAGTTTAGACC	AACTTGCAGGAAGACTGTGAGCGA
ATRX	27	CATCTGATGCTGAGGAAAGTTCTG	GAGATCCCTGATACTGAATACTAGC
ATRX	27	TGAATCTTCATCTGATGGCACTGA	TTTCTGTTCATCGCTGCTTCCCTC
ATRX	27	AGGAATGGATAATCAAGGGCACA	TCCTTTCCCTGTTGACTTCTCAGC
ATRX	27	GGATAAGCGTAATTCTTCTGACAGTGC	AGCACTTGCTTGCTGCTTCTTAGG
ATRX	27	CCGGTGGTGAACATAAGAAATCTG	AACTGTGACTCATCCTGCTCACCT
ATRX	27	CTTGTTCAGTTCCACTGCTGCCAT	CCTGTTCTGGCTCTGTAACCTACT
ATRX	27	CCAATGCAAGATGAGCCTTC	GAGTAAGCAGATGACCTAAATTACCAC
ATRX	28	CACACCAGTGTCCTGGAGATTT	AGGAAACACTGAATGTTAGCTCATCT
ATRX	29	TGCCAAGGTTGTCATGTGCTTAG	GAAGTCTTCCAAGGGCAGATACCA
ATRX	30	CCAGCAATGTTGGCTTTATCTGAACTG	AAGCACATCCGATTTTCCAA
ATRX	31	TTCCTTGTTGAGACCCACTGCTCA	GCCATGTTTGGTCGTTTGTACATAGT
ATRX	32	GCTAATTGTAGGGATGCCGTTTCG	CTCAGAATAGTGGTTGACATGAGTTCAG
ATRX	33	AGTGTGAGAATGGGTTTGTGGAGT	TGGGTATCAGTAGCCTTCGACACA
ATRX	34	GGGCTTCTATAAAGCTTGCTAATCTGTC	ACACCCACAACTGTAACATTTCCC
ATRX	35	TGTGCTTTGGAGGAGGTAGCCAAT	TAAGCAACACACAGGCCTAACCCA
IDH1	4	ACCAAATGGCACCATACGA	TTCATACCTTGCTTAATGGGTGT
IDH2	4	GCTGCAGTGGGACCACTATT	TGTGGCCTTGTACTGCAGAG

Supplementary figure S3.1 – Focal significant recurrent deletions in H3.3-K27M mutant and wild-type patients as determined by GISTIC 2.0 analysis (q ≤ 0.05)



3.10 Connecting text between Chapter 3 and Chapter 4

Thus far in this thesis, we have identified prevalent H3.3-K27M mutations in DIPG, which are associated with *TP53* mutations, specific copy-number alterations and poorer outcome, identifying a subgroup of tumors defined by this specific H3.3 mutation. Glioblastoma is a heterogenous tumor, and previous data from TCGA (25) have classified adult tumors into Proneural, Neural, Classical and Mesenchymal subgroups based on their gene expression profiling. Further to these results, the same TCGA consortium used genome-wide DNA methylation arrays to identify a glioma-CpG island methylator phenotype belonging to the proneural subgroup and tightly associated with *IDH1* mutation (31). Data from chapter 3 has identified an H3.3-K27M subgroup with specific genetic and clinical features. Based on these results and inspired by the methodological approach used by Noushmehr and TCGA, we aim to characterize adult and pediatric GBMs using an integrative analysis combining DNA methylation, gene-expression, copy-number alterations and genetic mutations, as well as clinical presentation and outcome, and more specifically we aim to investigate the impact of H3.3 mutations on the epigenome.

Chapter 4: Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma

Dominik Sturm[1, *], Hendrik Witt[1, 7, *], Volker Hovestadt[2, *], Dong-Anh Khuong-Quang[11, *], David T.W. Jones[1], Carolin Konermann[3], Elke Pfaff[1], Martje Tönjes[2], Martin Sill[4], Sebastian Bender[1], Marcel Kool[1], Marc Zapatka[2], Natalia Becker[4], Manuela Zucknick[4], Thomas Hielscher[4], Xiao-Yang Liu[11], Adam M. Fontebasso[12], Marina Ryzhova[13], Steffen Albrecht[14], Karine Jacob[11], Marietta Wolter[15], Martin Ebinger[16], Martin U. Schuhmann[17], Timothy van Meter[18], Michael C. Frühwald[19], Holger Hauch[20], Arnulf Pekrun[21], Bernhard Radlwimmer[2], Tim Niehues[22], Gregor von Komorowski[23], Matthias Dürken[23], Andreas E. Kulozik[7], Jenny Madden[24], Andrew Donson[24], Nicholas K. Foreman[24], Rachid Drissi[25], Maryam Fouladi[25], Wolfram Scheurlen[26], Andreas von Deimling[5, 9], Camelia Monoranu[27], Wolfgang Roggendorf[27], Christel Herold-Mende[8], Andreas Unterberg[8], Christof M. Kramm[28], Jörg Felsberg[15], Christian Hartmann[29], Benedikt Wiestler[10], Wolfgang Wick[10], Till Milde[6, 7], Olaf Witt[6, 7], Anders M. Lindroth[3], Jeremy Schwartzentruber[30], Damien Faury[11], Adam Fleming[11], Magdalena Zakrzewska[31], Pawel P. Liberski[31], Krzysztof Zakrzewski[32], Peter Hauser[33], Miklos Garami[33], Almos Klekner[34], Laszlo Bognar[34], Sorana Morrissy[35], Florence Cavalli[35], Michael D. Taylor[35], Peter van Sluis[36], Jan Koster[36], Rogier Versteeg[36], Richard Volckmann[36], Tom Mikkelsen[37], Kenneth Aldape[38], Guido Reifenberger[15], V. Peter Collins[39], Jacek Majewski[40], Andrey Korshunov[5], Peter Lichter[2], Christoph Plass[3, #], Nada Jabado[11, #], Stefan M. Pfister[1, 7, #]

[1] Division of Pediatric Neurooncology, German Cancer Research Center (DKFZ) Heidelberg, 69120 Heidelberg, Germany

[2] Division of Molecular Genetics, German Cancer Research Center (DKFZ) Heidelberg, 69120 Heidelberg, Germany

[3] Division of Epigenetics and Cancer Risk Factors, German Cancer Research Center (DKFZ) Heidelberg, 69120 Heidelberg, Germany

[4] Division of Biostatistics, German Cancer Research Center (DKFZ) Heidelberg, 69120 Heidelberg, Germany

[5] Clinical Cooperation Unit Neuropathology, German Cancer Research Center (DKFZ) Heidelberg, 69120 Heidelberg, Germany

[6] Clinical Cooperation Unit Pediatric Oncology, German Cancer Research Center (DKFZ) Heidelberg, 69120 Heidelberg, Germany

[7] Department of Pediatric Oncology, Hematology, and Immunology, Heidelberg University Hospital, 69120 Heidelberg, Germany

[8] Department of Neurosurgery, Heidelberg University Hospital, 69120 Heidelberg, Germany

[9] Department of Neuropathology, Heidelberg University Hospital, 69120 Heidelberg, Germany

[10] Department of Neurooncology, Heidelberg University Hospital, 69120 Heidelberg, Germany

[11] Departments of Pediatrics and Human Genetics, McGill University, McGill University Health Center Research Institute, Montreal, QC H3Z 2Z3, Canada

[12] Division of Experimental Medicine, Montreal Children's Hospital, McGill University Health Center Research Institute, Montreal, QC H3Z 2Z3, Canada

[13] NN Burdenko Neurosurgical Institute, Moscow, 125047, Russia

[14] Department of Pathology, Montreal Children's Hospital, McGill University Health Center Research Institute, Montreal, QC H3H 1P3, Canada

[15] Department of Neuropathology, Heinrich-Heine-University, 40225 Düsseldorf, Germany

[16] Department of Hematology and Oncology, Children's University Hospital Tübingen, 72076 Tübingen, Germany

[17] Department of Neurosurgery, University Hospital Tübingen, 76076 Tübingen, Germany

[18] Virginia Commonwealth University, Richmond, VA 23298, USA

[19] Pediatric Hospital, Klinikum Augsburg, 86156 Augsburg, Germany

[20] Pediatric Hospital, Klinikum Heilbronn, 74078 Heilbronn, Germany

[21] Prof.-Hess-Kinderklinik, Klinikum Bremen-Mitte, 28177 Bremen, Germany

[22] Children's Hospital, Helios Clinics, 47805 Krefeld, Germany

[23] Children's University Hospital, 68135 Mannheim, Germany

[24] Department of Pediatrics, University of Colorado Denver, Aurora, CO 80045, USA

[25] Division of Oncology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229, USA

[26] Cnopf'sche Kinderklinik, Nürnberg Children's Hospital, 90419 Nürnberg, Germany

[27] Department of Neuropathology, Institute of Pathology, University Würzburg, 97080 Würzburg, Germany

[28] University Children's Hospital, Martin Luther University Halle-Wittenberg, 06120 Halle, Germany

[29] Institute of Pathology, Department of Neuropathology, Hannover Medical School, 30175 Hannover, Germany

[30] McGill University and Genome Quebec Innovation Centre, Montreal, QC H3A 1A4, Canada

[31] Department of Molecular Pathology and Neuropathology, Medical University of Lodz 92-216 Poland

[32] Department of Neurosurgery, Polish Mother's Memorial Hospital Research Institute, Lodz 93-338 Poland

[33] 2nd Department of Paediatrics, Semmelweis University, Budapest H-1094 Hungary

[34] Department of Neurosurgery, Medical and Health Science Center, University of Debrecen, H-4032 Debrecen, Hungary

[35] Program in Developmental and Stem Cell Biology, Division of Neurosurgery, Arthur and Sonia Labatt Brain Tumour Research Centre, Hospital for Sick Children, University of Toronto, Toronto, ON M4N 1X8, Canada

[36] Department of Oncogenomics, AMC, University of Amsterdam, Amsterdam 1105 AZ, The Netherlands

[37] Departments of Neurology and Neurosurgery, Henry Ford Hospital, Detroit, MI 48202, USA

[38] Department of Neuro-Oncology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

[39] Division of Molecular Histopathology, Department of Pathology, University of Cambridge, Cambridge, CB2 0QQ, United Kingdom

[40] Department of Human Genetics, McGill University, Montreal, QC H3Z 2Z3, Canada

*These authors contributed equally to this work.

#These authors contributed are co-corresponding and co-senior authors.

Published in and adapted from (7)

Cancer Cell. 2012 Oct 16;22(4):425-37. doi: 10.1016/j.ccr.2012.08.024.

4.1 Abstract

Glioblastoma (GBM) is a brain tumor that carries a dismal prognosis and displays considerable heterogeneity. We have recently identified recurrent H3F3A mutations affecting two critical amino acids (K27 and G34) of histone H3.3 in one-third of pediatric GBM. Here, we show that each H3F3A mutation defines an epigenetic subgroup of GBM with a distinct global methylation pattern, and that they are mutually exclusive with IDH1 mutations, which characterize a third mutation-defined subgroup. Three further epigenetic subgroups were enriched for hallmark genetic events of adult GBM and/or established transcriptomic signatures. We also demonstrate that the two H3F3A mutations give rise to GBMs in separate anatomic compartments, with differential regulation of transcription factors OLIG1, OLIG2, and FOXG1, possibly reflecting different cellular origins.

4.2 Introduction

Glioblastoma (GBM; World Health Organization [WHO] grade IV), the most common primary brain tumor, carries a universally dismal prognosis in children and adults (5). With evidence emerging recently of age-specific molecular and genetic differences, it is now becoming apparent that pediatric GBM is largely biologically distinct from adult GBM. Based on similarities in recurrent genomic aberrations (23, 53-56, 180), it was long thought that pediatric GBMs more closely resembled adult "secondary" GBMs, which arise from a preceding lower-grade lesion. However, stepwise transformation from less-malignant gliomas into GBMs rarely occurs in children (168). Furthermore, IDH1 or IDH2 mutations, which are found in up to 98% of adult secondary GBMs, are very rare in childhood GBMs (<10%) (27, 55, 180-185).

We recently identified two recurrent somatic mutations in the *H3F3A* gene, affecting highly

conserved residues of its encoded protein, the replication-independent histone 3 variant H3.3, in one-third of pediatric GBMs (1). Mutations in a protein complex comprised of H3.3 and ATRX/DAXX were detected in 45% of cases, and were shown to be associated with TP53 mutations and alternative lengthening of telomeres (ALT). The H3.3 mutations result in amino acid substitutions at K27 or G34—at or near residues targeted by key post-translational modifications that regulate H3.3's activity in governing gene expression (186), and were shown to be linked to distinct transcriptional profiles (1). Methylation of K27 and K36 is also disrupted by elevated levels of the onco-metabolite 2-hydroxyglutarate (2-HG) resulting from gain-of-function mutations in IDH1 (39, 187), which was previously shown to be associated with a distinct Glioma-CpG-Island Methylator Phenotype (G-CIMP) (31).

In the present study, we further investigate the heterogeneity of glioblastoma across the entire age spectrum, and elucidate the impact of *H3F3A* mutations on the GBM epigenome.

4.3 Methods

Patients and tumor samples

Primary tumor samples for methylation (n = 136; Table S4.1), mutation (n = 460; Table S4.2), and gene expression (n = 69) analysis and all clinical data were collected at the DKFZ (Heidelberg, Germany) and at McGill University (Montreal, Canada). Paraffin-embedded samples (n = 143; Table S4.4) for TMA analysis were collected from the Burdenko Neurosurgical Institute (Moscow, Russia) and from the Department of Neuropathology, University of Würzburg (Germany). Patient clinical details can be found in Table S4.1 for the methylation analysis data set and in Table S4.4 for the TMA cohort. All of the tumors were banked at the time of primary diagnosis between 1994 and 2011 in accordance with research ethics board approval from the respective institutes. Informed consent was obtained from all

patients included in this study. An overview of all samples included in different data collections is given in figure S4.1A. All of the samples were independently reviewed by senior pediatric neuropathologists (S.A. and A.K.) according to the WHO guidelines. Detailed information about samples provided by TCGA can be found elsewhere (http://cancergenome.nih.gov).

DNA methylation profiling

For genome-wide assessment of DNA methylation GBM samples (n = 136) and controls (n = 10; four fetal and two adult samples of non-neoplastic cerebellum; two samples of Whole-Genome Amplified DNA (unmethylated control; two samples of M.SssI-treated DNA [100% methylated control]) were arrayed using the Illumina HumanMethylation450 BeadChip according to the manufacturer's instructions at the DKFZ. Methylation data of additional adult glioblastoma samples (n = 74) were obtained from the TCGA website (https://tcga-data.nci.nih.gov; available data from TCGA batches 79 and 111).

The following filtering criteria were applied: Removal of probes targeting the X and Y chromosomes (n = 11,551), removal of probes containing a single-nucleotide polymorphism (dbSNP132 Common) within five base pairs of and including the targeted CpG-site (n = 24,536), and probes not mapping uniquely to the human reference genome (hg19) allowing for one mismatch (n = 9,993). In total, 438,370 probes were kept for analysis.

For a subset of differentially methylated genes from the 450k array, MassARRAY technology (Sequenom, San Diego, CA, USA) was used to validate our results, allowing us to compare DNA methylation levels at 29 individual CpG-sites investigated by both techniques. DNA methylation measurements of those 29 CpG dinucleotides were highly correlated (median Pearson's correlation: r = 0.96; range: 0.71–1.00).

Gene expression profiling

Glioblastoma samples for which RNA of sufficient quantity and quality was available (n = 69) were analyzed on the Affymetrix GeneChip Human Genome U133 Plus 2.0 Array at the Microarray Department of the University of Amsterdam, the Netherlands. Sample library preparation, hybridization, and quality control were performed according to the manufacturer's protocols. Expression data were normalized using the MAS5.0 algorithm of the GCOS program (Affymetrix Inc). Gene expression data of additional adult glioblastoma samples (n = 74) were obtained from the TCGA website (https://tcga-data.nci.nih.gov; available data from TCGA batches 79 and 111). Predictive analysis of microarrays was used to assign TCGA methylation and gene expression subgroups to each of the samples in the present study.

Detection of CNAs

Copy-number aberrations were detected from the 450k Infinium methylation array in a custom approach using the sum of both methylated and unmethylated signals (Figure S4.1D). For the detection of EGFR and PDGFRA high-level amplifications, homozygous CDKN2A deletions, and CNAs affecting chromosomes 7 and 10 (as depicted in Figure 4.1), automatic scoring was verified by manual curation of the respective loci for each individual profile, and compared with results obtained from SNP profiling and fluorescence in situ hybridization (FISH) analysis where available. The three methodologies showed very high concordance.

Statistical analysis and measurement of differential DNA methylation and gene expression

For unsupervised consensus clustering we used the 8,000 most variable methylated probes (by standard deviation) across the data set (R package: clusterCons) (188, 189). The consensus matrix was calculated using the k-means algorithm (10 random starting sets, maximum of

1,000 iterations) on a fraction of probes (0.8) in 1,000 iterations. The significance analysis of microarrays (SAM) method was used to identify genes that are differentially methylated or differentially expressed between subgroups. Correction for multiple testing was performed using the Benjamini-Hochberg method. Genes were considered significantly differentially methylated/expressed between two subgroups when displaying an adjusted p value < 0.01 and a methylation difference of 0.2 or a 2-fold change in expression.

Statistical analysis of clinical and molecular data

Kaplan-Meier analysis was performed to estimate the survival time of different GBM subgroups and a log rank test was used to test for differences of more than one survival curve. Comparisons of binary and categorical patient characteristics between subgroups were performed by the use of a two-sided Fisher's exact test. An unpaired t test was used to test for differences between the mean values for continuous variables in GBM subgroups.

Immunohistochemistry and FISH

Hematoxylin and eosin stained sections from all 143 paraffin blocks were prepared to define representative tumor regions for inclusion in the TMA. Antibodies against the following antigens were applied: OLIG2 (Millipore, Billerica, MA, USA; AB9610; dilution 1:250), FOXG1 (Abcam, Cambridge, UK; ab18259; dilution 1:50), ATRX (Sigma, HPA001906; dilution 1:750), and mutated IDH1 (R132H; (190); dianova, DIA H09). Multicolor interphase FISH analysis for PDGFRA, EGFR, and CDKN2A was performed as described (191). Telomere-specific FISH was done using a standard formalin-fixed paraffin-embedded FISH protocol (59) using a FITC peptide nucleic acid telomere probe from Dako (Glostrup, Denmark).

Genomic sequencing

Targeted sequencing of *H3F3A* (first coding exon), *IDH1* (exon 4), and *TP53* (all exons) was performed by QIAGEN (Hilden, Germany) in both forward and reverse directions using purified PCR products. PCR procedures were as previously described (192). Primer sequences are available upon request.

4.4 Results

Integrated molecular classification of glioblastoma

We used an integrative approach based on epigenetic, copy-number, expression, and genetic analyses to investigate the heterogeneity of glioblastoma across all age groups. An overview of all GBM samples subjected to various analyses is given in Figure S4.1A available online.

We investigated a cohort of GBMs from children (n = 59) and adult patients (n = 77) for genome-wide DNA methylation patterns using the Illumina 450k methylation array, and complemented our data with unpublished profiles of 74 adult GBM samples generated by The Cancer Genome Atlas (TCGA) Consortium (23) (Table S4.1). Consensus clustering using the 8,000 most variant probes across the data set robustly identified six distinct DNA methylation clusters (Figures 4.1 and S4.1B). Based on correlations with mutational status, DNA copy-number aberrations, and gene expression signatures, as outlined below, we have labeled these subgroups “IDH,” “K27,” “G34,” “RTK I (PDGFRA),” “Mesenchymal,” and “RTK II (Classic).”

A striking finding of this integrated analysis is that H3F3A K27 and G34 mutations were exclusively distributed to the K27 (18/18) and G34 (18/18) clusters, respectively ($p < 0.001$; Fisher’s exact test) (Figure 4.1). The IDH group contained 88% of IDH1-mutated tumors

(23/26) ($p < 0.001$) and displayed concerted, global hypermethylation (Figures 4.1, 4.2A, and 4.2B), thereby expanding the previously described link between IDH1 mutation and G-CIMP$^+$ tumors to a pediatric setting (31). In contrast, tumors in the G34 cluster specifically showed widespread hypomethylation across the whole genome, and especially in nonpromoter regions, when compared with all other subgroups (Figures 4.2A and 4.2B). This suggests the existence of a more global version of a CpG hypomethylator phenotype (CHOP), as proposed for a small number of genes in gastric cancer (193). More detailed mapping of differentially methylated regions revealed that the hypomethylation observed in *H3F3A* G34-mutated tumors was particularly prominent at chromosome ends (Figures 4.2C and 4.2D), potentially linking subtelomeric demethylation to alternative lengthening of telomeres, which is most frequently observed in this subgroup (1).

Of note, all mutations in *H3F3A* and *IDH1* were mutually exclusive ($p < 0.001$) (Figure 4.1). To further test this observation, we extended the targeted sequencing analysis of *H3F3A* and *IDH1* to include 460 GBM samples from patients covering a broad age range (Figure S4.1C; Table S4.2). Even in this expanded series, no co-occurring mutations in *H3F3A* and *IDH1* were detected ($p < 0.001$), and the age distribution confirmed reported associations of certain mutations with GBM in children (H3F3A K27), adolescent patients (H3F3A G34), and young adult patients (IDH1) (1, 2, 27) (Figure S4.1C; Table S4.2). As we have shown, *TP53* mutations largely overlap with *H3F3A* mutations in pediatric GBM (1), similar to the association of *TP53* and *IDH1* mutations in adults (27). This observation also holds true in our larger cohort, with a high enrichment of *TP53* mutations in the G34 (18/18), IDH (22/24), and K27 (13/18) clusters ($p < 0.001$) (Figure 4.1).

Since pediatric GBMs have been shown to display a distinct spectrum of focal copy-number aberrations (CNAs) compared with their adult counterparts (53-55), we integrated DNA

methylation clusters with copy-number data derived from the methylation arrays (Figures 4.1 and S4.1D). Interestingly, PDGFRA amplification was significantly more common in the RTK I "PDGFRA" cluster than any other subgroup (11/33; $p < 0.001$), hence our proposed name for this group. The RTK II "Classic" cluster demonstrated a very high frequency of whole chromosome 7 gain (50/56; $p < 0.001$) and whole chromosome 10 loss (56/56; $p < 0.001$), as well as frequent homozygous deletion of CDKN2A (35/56; $p < 0.001$) and amplification of EGFR (39/56; $p < 0.001$) (Figures 4.1 and S4.1D)—all hallmark CNAs of adult GBM (5), as reflected by the complete absence of pediatric patients in this cluster. Overall, tumors from the IDH, K27, and G34 clusters were mostly devoid of the detected CNAs associated with the other GBM subgroups (amplifications of PDGFRA and EGFR, deletion of CDKN2A, chromosome 7 gain, and chromosome 10 loss) (Figure 4.1; Table S4.1), in keeping with a previously reported finding in G-CIMP$^+$ tumors (31).

To additionally place the methylation subgroups proposed here into the context of previous GBM classification systems, we used the gene expression signature described by the TCGA to classify 122 of the above tumors with available transcriptome data into one of four gene expression subtypes: Proneural, Neural, Mesenchymal, and Classical (25) (Figure 4.1; Table S4.1). This further confirmed the prototypic nature of tumors in the RTK II "Classic" cluster, which was clearly enriched for "Classical" expression profiles ($p < 0.001$). The RTK I "PDGFRA" cluster was highly enriched for "Proneural" expression ($p = 0.01$), further substantiating the previously reported association of PDGFRA amplification with this expression subtype (25). As expected, all tumors in the IDH cluster displayed "Proneural" expression patterns. Interestingly, the K27 cluster also showed a clear enrichment of tumors with a "Proneural" signature ($p < 0.01$), indicating that this expression subtype can be divided into subgroups harbouring distinct genomic aberrations based on methylation profiling and

targeted gene sequencing. “Mesenchymal” gene expression was mostly restricted to one methylation subgroup ($p < 0.001$) that showed a much lower incidence of typical GBM-related CNAs, generally fewer copy-number changes, and no characteristic point mutations. We therefore termed this methylation cluster, which displayed the largest similarity with normal brain methylation patterns, “Mesenchymal.” Copy-number aberrations in these samples were, however, observed at a similar amplitude as in other cases, indicating an absence of excess stromal contamination.

Our finding of six GBM methylation clusters is different from a TCGA study using Illumina 27k arrays, which identified three methylation clusters in an adult GBM cohort (31). Applying their signature to our data set, however, showed that two clusters (G-CIMP$^+$ and Cluster #3) mapped almost exactly to two of our subgroups (IDH and RTK II “Classic”, respectively, $p < 0.001$) (Figure 4.1). By adding pediatric cases to the study cohort, we demonstrate that TCGA methylation Cluster #2 can be further divided into four biologically distinct subgroups, defined by a clear enrichment for mutations (K27, G34), CNAs (PDGFRA), and/or gene expression signatures (Mesenchymal). The same DNA methylation clusters were apparent when restricting our analyses to the pediatric population, with the exception of the RTK II “Classic” cluster, which is not represented in the pediatric population (Figure S4.1E). Notably, by analyzing tumors from patients spanning a broad age spectrum, we further observed a clear age-dependent increase in overall DNA methylation levels (Figure S4.2A), even after adjusting our analysis to exclude tumors with age-related mutations in *IDH1* or *H3F3A* (Figure S4.2B).

GBM subgroups show correlations with clinicopathological variables

The DNA methylation clusters described here were closely associated with specific age groups, pointing toward the biological diversity of epigenetic GBM subgroups (Figure 4.1).

While the K27 cluster predominantly consisted of childhood patients (median age 10.5 years, range 5–23 years), patients in the G34 cluster were found mostly around the threshold between the adolescent and adult populations (median age 18 years, range 9–42 years), as previously suggested (1). The RTK I "PDGFRA" cluster also harboured a proportion of pediatric patients (median age 36 years, range 8–74 years), in line with reports of PDGFRA CNAs being more prevalent in childhood high-grade gliomas (53-55). The Mesenchymal cluster displayed a widespread age distribution (median age 47, range 2–85 years). The IDH and RTK II "Classic" clusters were mostly comprised of younger adult (median age 40 years, range 13–71 years) and older adult (median age 58, range 36–81 years) patients, respectively, reflecting the established differences in patient age between IDH1-mutated/G-CIMP^{+} and IDH1 WT adult GBM (27, 31).

The epigenetic GBM subgroups identified here also showed mutation-specific patterns of tumor location in the central nervous system (Figure 4.3A). While K27-mutated tumors were predominantly seen in midline locations, e.g., thalamus, pons, and the spinal cord (21/25 cases with available data), tumors from all other subgroups almost exclusively arose in the cerebral hemispheres (86/92, $p < 0.001$). To further investigate this association of mutation type and location, we investigated the transcriptomic profiles of H3F3A-mutated samples ($n = 13$). Gene signatures characteristic for K27 and G34 mutant GBMs were applied to a published series of 1,340 transcriptomic profiles representing multiple regions of the developing and adult human brain (194) (Figure S4.3). The G34 mutant signature appeared to be most strongly expressed in early embryonic regions and early- to mid-fetal stages of neocortex and striatum development. In contrast, the K27 signature most closely matched with mid- to late-fetal stages of striatum and thalamus development. Thus, G34 and K27 mutant GBMs seem to show expression patterns of early developmental stages correlating with their subsequent

tumor location, possibly indicating different cellular origins and/or time of tumor initiation for these two subgroups.

Correlating our proposed methylation clusters with patient survival indicated differences between mutation-defined subgroups, but this was somewhat restricted by the low number of patients with available survival data in each subgroup (Figure 4.3B). We therefore enlarged our survival analysis to include all tumors with known H3F3A and IDH1 mutation status (Figure 4.3C). As expected, patients with IDH1 mutant tumors had a significantly longer overall survival (OS) than patients with H3F3A and IDH1 WT tumors ($p < 0.001$) (27, 28, 31). Notably, G34 mutant GBM patients also showed a trend toward a better OS than WT tumor patients, with marginal statistical significance ($p = 0.05$). In contrast, patients with K27 mutations tended toward an even shorter OS than patients with WT tumors, although this did not reach statistical significance ($p = 0.12$). Comparing the two H3F3A mutations, patients harbouring G34-mutated tumors clearly had a longer OS than patients with tumors carrying the K27 mutation ($p < 0.01$). While this association may be partly linked to G34-mutated tumors being more accessible to surgery than the midline K27-mutated tumors, the better prognosis of G34 versus K27 was independent of location for those cases where both mutation type and tumor site information were available ($p = 0.02$; HR = 0.20, 95% CI = 0.05–0.77; Cox proportional hazards model).

Integrating methylome and transcriptome data identifies marker genes of GBM subgroups

A combined analysis of DNA methylation and gene expression data was used to identify subgroup-specific differentially regulated genes (Figures 4.4A–4.4C and S4.4A–4.4C; Table S4.3). This analysis revealed Oligodendrocyte Lineage Genes 1 and 2 (OLIG1 and OLIG2) and the neural development gene FOXG1 as top candidates for further analysis in H3F3A-mutated GBMs (Figure 4.4A–4.4C). DNA hypermethylation across the OLIG1 and OLIG2

loci occurred exclusively in G34-mutated tumors, which concurrently displayed significantly lower OLIG1 and OLIG2 gene expression (Figure 4.4D–4.4F). Interestingly, this pattern closely mimics that of embryonic stem cells, where epigenetic inactivation of OLIG1 and OLIG2 has been proposed as a mechanism to prevent neural lineage commitment (174). Expression of FOXG1 was significantly lower in K27-mutated tumors than in all other subgroups, accompanied by higher levels of promoter methylation (Figures S4.4A, S4.4D, and S4.4E). This comparative analysis also further supported our suggestion of a CHOP-like phenotype in G34 tumors, as most of the differentially methylated genes were found to be hypomethylated (1653/1946, 85%, Figure S4.4B) in this subgroup, in contrast to the hypermethylator G-CIMP pattern observed in the IDH subgroup (Figure S4.4C). Hypermethylation and concurrent downregulation of TP73 antisense RNA 1 (TP73-AS1) was identified as a unique characteristic of this IDH/G-CIMP^{+} cluster (Figures S4.4D and S4.4F). Interestingly, inactivation of this gene by promoter methylation has been reported as a common mechanism in a high proportion of oligodendrogliomas, 80% of which are also known to harbor IDH1 mutations (195).

Immunohistochemical analysis correctly subclassifies mutation-defined GBM subgroups

In an attempt to subgroup GBM samples based on differential protein expression—a method which is likely to be more suitable for possible clinical application—we used commercially available antibodies against OLIG2, FOXG1, and mutated IDH1 (R132H) to stain a tissue-microarray (TMA) with cores from 143 pediatric GBMs, and classified tumors according to their protein expression patterns (Figures 4.5A and 4.5B; Table S4.4). The resulting fractions of tumors with predicted mutations in IDH1 (IDH1^{R132H}, n = 6) and H3F3A (OLIG2^{+}/FOXG1^{-} for K27, n = 37, and OLIG2^{-}/FOXG1^{+} for G34, n = 21) were consistent with the frequency of each mutation in the pediatric population as detected by targeted gene

sequencing (Figure S4.1C). Our approach correctly classified GBMs with known H3F3A and IDH1 mutation status, and revealed a frequent association between OLIG2^{-}/FOXG1^{+} tumors (assumed to be G34-mutated), loss of ATRX protein expression, and an ALT phenotype (Figures 4.5B–D), as previously reported for H3F3A G34-mutated tumors (1). The putative H3F3A mutant groups also did not overlap with tumors harbouring IDH1 (R132H) mutations, and only one case with EGFR amplification and homozygous CDKN2A deletion was detected therein (Figure S4.5A). The correlation with clinicopathological variables, such as tumor location and patient survival, also reflected our findings from the array-based analysis (Figures S4.5B and S4.5C). Of note, rare tumors represented on the TMA occurring in the basal ganglia and the spinal cord were almost always found in the OLIG2^{+}/FOXG1^{-} subgroup (and therefore predicted to harbor the H3F3A K27 mutation), further strengthening our hypothesis of the H3F3A K27 mutation as a unifying characteristic of midline GBM.

4.5 Discussion

We have identified six biological subgroups of GBM based on global DNA methylation patterns, which correlate with specific molecular-genetic alterations and key clinical parameters. Our findings suggest that at least 30%–40% of pediatric/young adult GBMs are likely characterized by disrupted epigenetic regulatory mechanisms, associated with recurrent and mutually exclusive mutations in either H3F3A or IDH1, and aberrant DNA methylation patterns. Placing these subgroups into the context of previous molecular GBM classification schemes described by the TCGA (25, 31) revealed a clear correlation with DNA methylation clusters and a corresponding enrichment for previously established expression signatures in different epigenetic subgroups. We have also demonstrated that our proposed classification can refine that described by the TCGA for adult GBM, to give a stratification system that is applicable across all ages, and defines additional biologically meaningful subgroups. A

simplified graphical summary of the key molecular and biological characteristics of the GBM subgroups as identified by our integrated classification strategy is given in Figure 6.

We and others have recently described a high frequency of H3F3A K27 mutations in thalamic GBMs and in diffuse intrinsic pontine gliomas (DIPGs), suggesting that the latter likely represent an anatomically-defined subset of K27 mutant GBM (1, 2, 60). We now extend this observation to a larger subgroup of GBM, characterized by the K27M mutation, which almost exclusively occurs in midline locations, including rare tumors in the basal ganglia and the spinal cord. This is in line with a recent study by Puget *et al*. (58) in which gene expression patterns of brainstem gliomas were found to resemble midline/thalamic tumors, indicating a closely related origin. The K27 subgroup also displays markedly lower expression of the ventral telencephalic marker FOXG1 than other subgroups. Conversely, non-K27 tumors were restricted to hemispheric locations, further underlining the biological divergence of epigenetic GBM subgroups. While recurrent focal amplification of PDGFRA has been suggested as a key oncogenic event in pediatric DIPGs in some studies (56-58), midline-associated tumors in the K27 or OLIG2^{+}/FOXG1^{-} subgroups (including ten brainstem gliomas with known PDGFRA copy-number status) lacked this common feature in our series. PDGFRA amplification was, however, enriched in a subgroup of supratentorial hemispheric GBMs. In part, this discrepancy may be explained by the use of autopsy (and therefore post radio/chemotherapy) material in previous study cohorts of DIPGs, which might have been confounded by the higher incidence of PDGFRA amplifications observed in radiation-induced gliomas (55). Nevertheless, amplifications of PDGFRA have also been detected in small numbers of pretreatment samples (55, 56, 58), and post-treatment samples were not found to show increased widespread genomic instability (57). This particularly clinically challenging subset of tumors clearly warrants further investigation, underlining the importance of routine

stereotactic biopsy of DIPGs at the time of primary diagnosis.

OLIG2 has previously been reported as a universal marker for diffuse gliomas (196), and OLIG2-positive progenitor-like cells of the subventricular zone have been suggested as potential glioma-initiating cells (197). There is also evidence that OLIG2-mediated modification of p53 function is required for complete inactivation of the latter in malignant gliomas, which typically show indirect loss of p53 activity through MDM2 amplification or $p14^{ARF}$ deletion (198). Here, we describe a distinct subgroup of GBM, harbouring the H3F3A G34 mutation, in which OLIG1 and OLIG2 protein expression is absent. Given the ~100% mutation frequency of TP53 in this subgroup, this may indicate a different pathogenesis of G34-mutated GBM, in which direct inactivation of p53 is required rather than via an OLIG2-dependent mechanism.

The previously reported association of H3F3A mutations, particular the G34 mutation, with loss of ATRX and ALT (1) is further expanded upon here. Interestingly, the global CHOP that we observed in G34 mutants was particularly pronounced in subtelomeric regions, suggesting a possible mechanistic link with ALT in these tumors (199). Whether this is a more general phenotype that can be observed in clinically and etiologically distinct subgroups of other human cancers, remains to be investigated.

The close link between H3F3A mutation type, tumor location, and differential expression of key neuronal lineage markers leads us to speculate that there may be differences in the cell of origin and/or the time of tumor development between these GBM subgroups. Although supported by the differential expression of mutant-specific gene signatures at different stages of human brain development, this remains to be formally shown. Also requiring further validation in larger, prospective cohorts is the association of the G34 mutation with better

overall survival compared with H3F3A and IDH1 WT tumors, and that of K27 mutation with potentially poorer prognosis, as observed in our series and a recent cohort of pediatric DIPGs (2).

Given the location of the H3F3A mutations at or near critical regulatory histone residues, and their distinct methylation profiles, we consider it likely that the H3.3 mutations are directly involved in producing widespread aberrant DNA methylation and deregulation of gene expression. This has recently been shown for IDH1 mutations, which alone are sufficient to induce the global epigenetic reprogramming of the G-CIMP phenotype in normal astrocytes (38). Overproduction of the oncometabolite 2-hydroxyglutarate in IDH1-mutated cells inhibits the TET family of 5-methlycytosine hydroxylases and H3K27-specific demethylases. This is thought to lead to decreased 5-hydroxylmethycytosine and increased H3K27 methylation (39), resulting in aberrant DNA and histone methylation, and a block to differentiation (35, 117, 200).

Seminal studies have shown that DNA methylation patterns are tightly linked to histone 3 lysine K27 trimethylation (H3K27me3) patterns (119, 120), and in high-CpG-density promoters, loss of H3K4me3 and retention of H3K4me2 or H3K27me3 is correlated with an increase in DNA methylation (174). Therefore, mutations affecting H3K27 methylation are likely to affect DNA methylation. In addition, mutations in ATRX have been shown to give rise to changes in the patterns of DNA methylation of several highly repeated sequences, which further supports the link between chromatin remodelling machinery and DNA methylation (201). Based on our current knowledge, the incorporation of H3.3 variants into the genome, and the subsequent effects on gene regulation, involve the H3.3 chaperone complex (including ATRX and DAXX) in a replication independent manner (202, 203) but other factors also likely play a role. The exact mechanism by which the H3F3A mutations

might be inducing epigenetic reprogramming requires further elucidation.

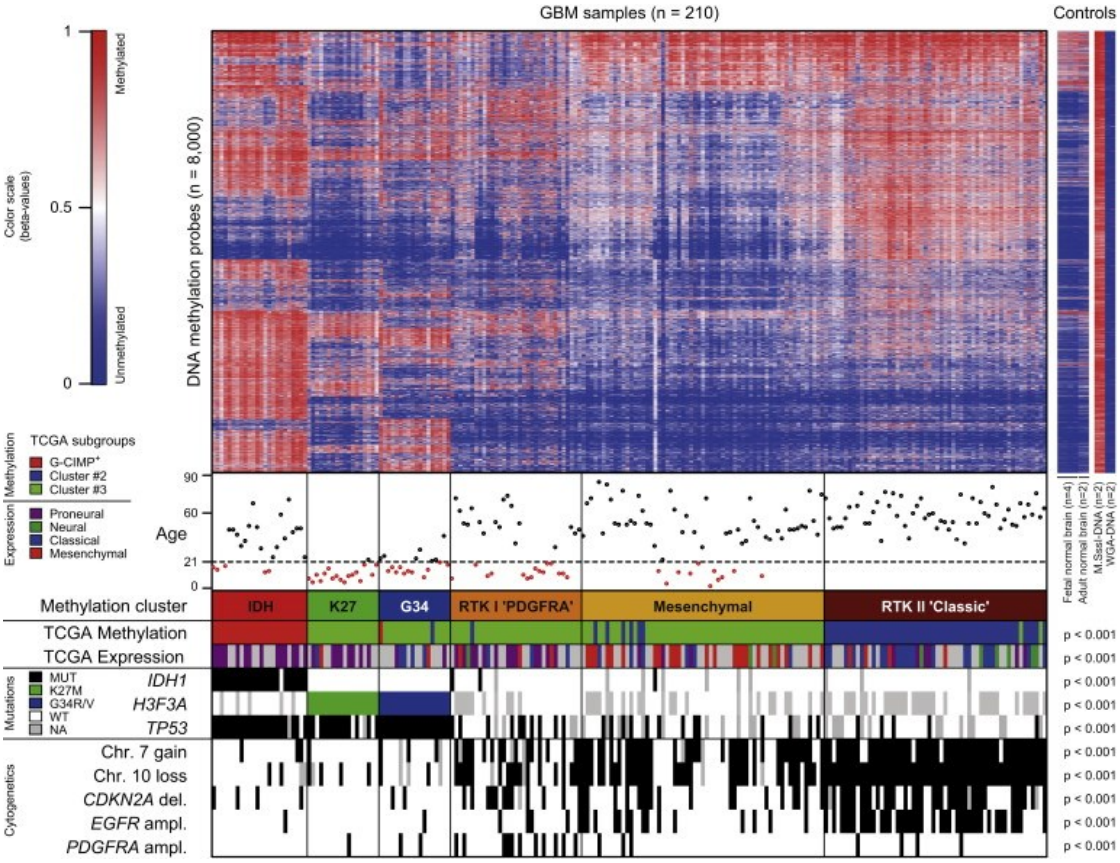
In conclusion, this study describes a number of findings that enhance our understanding of the heterogeneity of GBM, as well as shedding light on potential cellular origins and oncogenic pathways leading to gliomagenesis. We have identified potential prognostic biomarkers, which may be further exploited for molecular diagnostic purposes, and also provided a focus for future work at a basic and translational/targeted therapeutic level, particularly in a pediatric and young adult setting.

4.6 Acknowledgments

We thank Andrea Wittmann and Laura Sieber from the Division of Pediatric Neurooncology at the DKFZ for excellent technical support and Matthias Schick and Roger Fischer from the DKFZ Genomics and Proteomics Core Facility for performing the microarray analyses to a very high standard. We also acknowledge the outstanding technical assistance of Hannelore Schraut (University of Würzburg) and Leonore Senf (Nürnberg Children's Hospital). The project was supported by grants from the German Cancer Aid (109252 and 108456); the BMBF (to A.K., B.R., O.W., P.L., S.M.P., G.R., and J.F.; ICGC PedBrain, NGFN^{Plus} #01GS0883); Koningin Wilhelmina Fonds (UvA-2010-4713) and KIKA (to M.K.); the Cole Foundation; the Canadian Institute of Health Research; the Institute of Cancer Research, Génome Canada and Génome Quebec (to N.J.); and the Alfred-Müller Award for Neuro-Oncology to S.M.P.

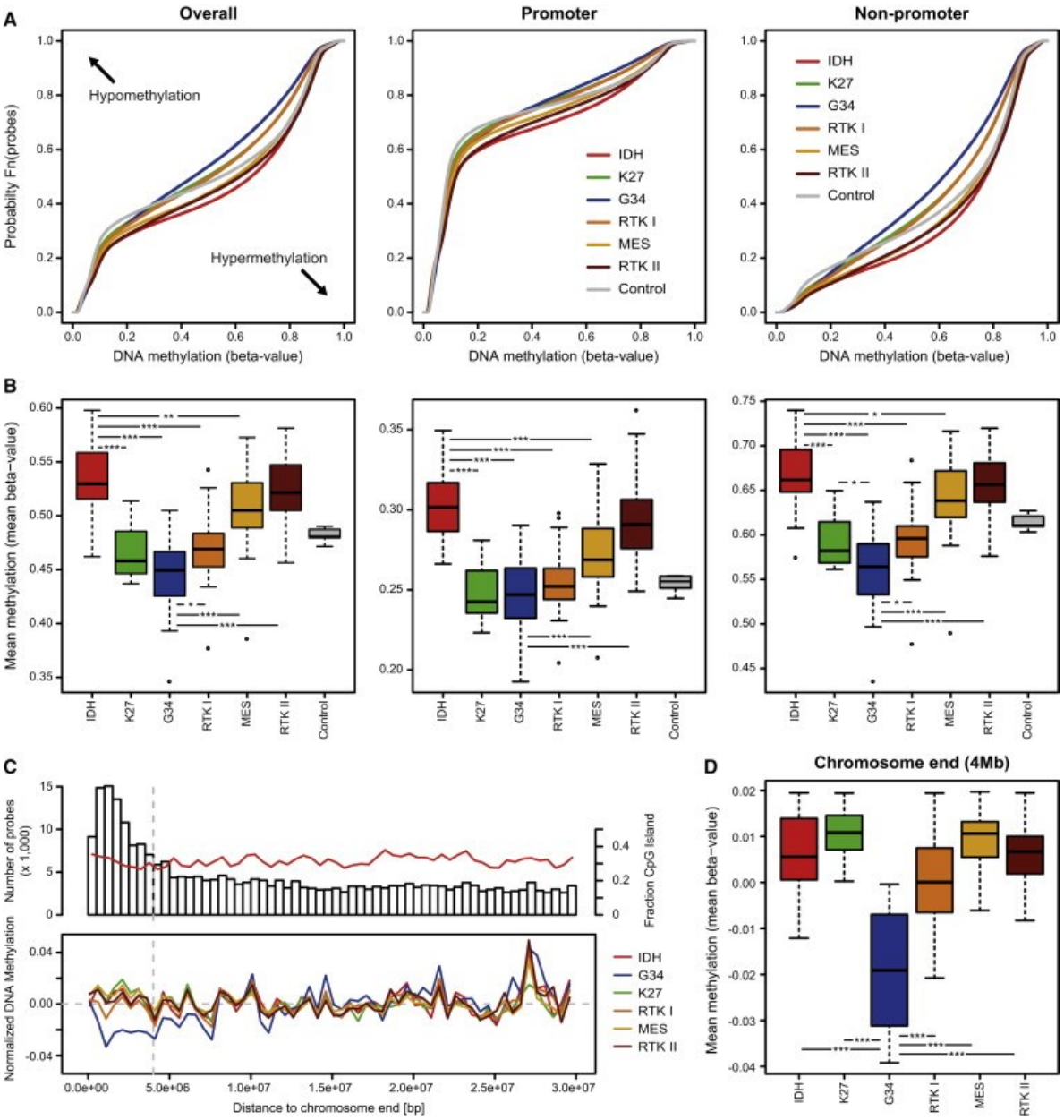
4.7 Figures

Figure 4.1 – Methylation profiling reveals the existence of six epigenetic GBM subgroups



Heatmap of methylation levels in six GBM subgroups identified by unsupervised k-means consensus clustering, and control samples as indicated. Each row represents a probe; each column represents a sample. The level of DNA methylation (beta-value) is represented with a color scale as depicted. For each sample (n = 210), patient age, subgroup association, mutational status, and cytogenetic aberrations are indicated.

Figure 4.2 – Global DNA methylation patterns in GBM subgroups



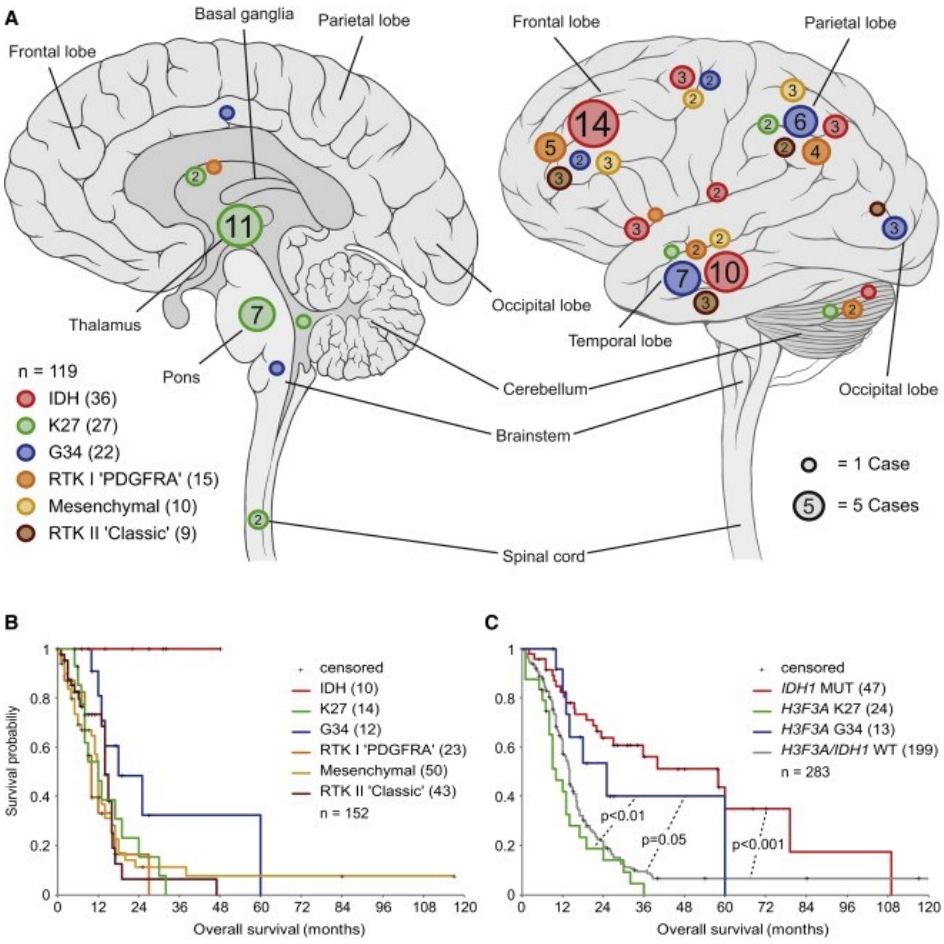
(A) Distinct patterns of global DNA methylation in GBM subgroups as identified by consensus clustering. The empirical cumulative distribution function for DNA methylation levels (beta-values) is plotted individually for each subgroup.

(B) Overall DNA methylation levels (mean beta-values) of individual GBM methylation subgroups. Significant differences ($^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$) to IDH and G34 subgroups are indicated.

(C) Upper panel: Probe density in respect of distance to chromosome end. The fraction of probes located within CpG-Islands (red line) remains stable. Lower panel: Mean methylation value per subgroup within windows of 500kb, normalized to control samples. Individual samples are normalized by the mean overall methylation value.

(D) Mean methylation value within 4 Mb to the chromosome end normalized to the mean overall methylation value and to control samples. Significant differences (***$p < 0.001$) between subgroups compared to G34 tumors are indicated. MES, Mesenchymal.

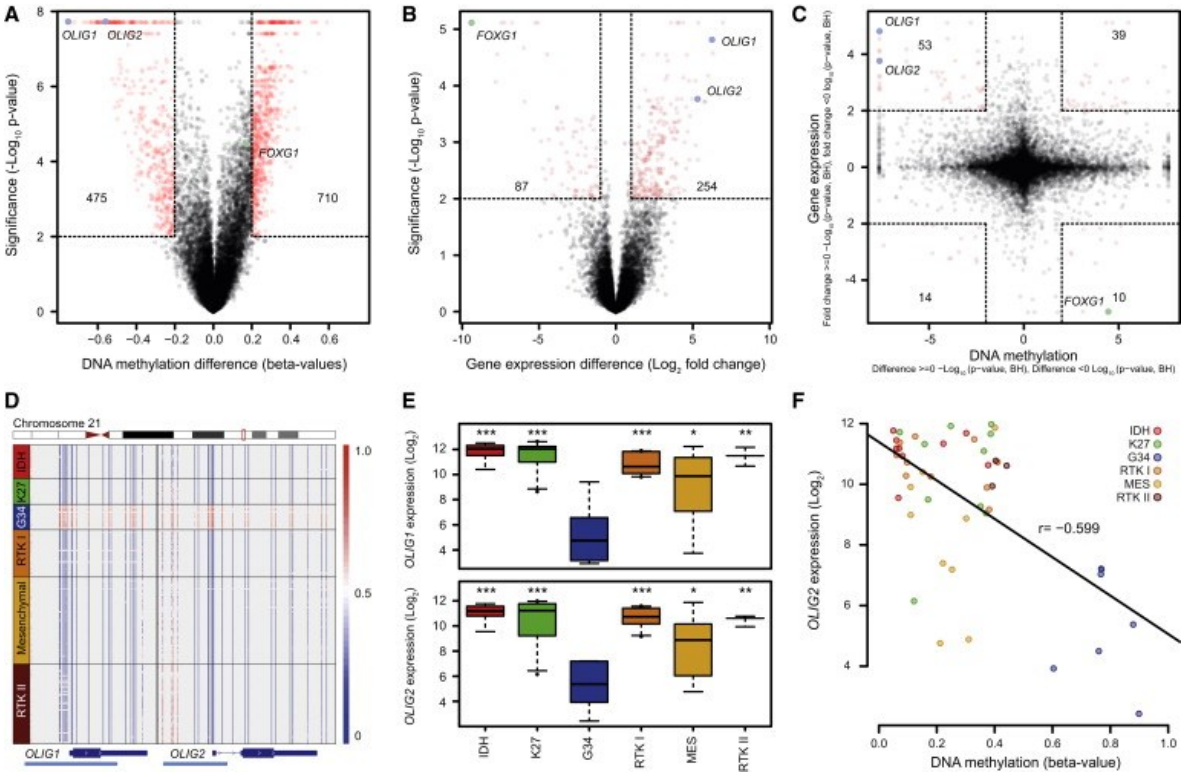
Figure 4.3 – Epigenetic subgroups of GBM correlate with distinct clinical characteristics



(A) Location of 119 GBMs in the human central nervous system grouped by methylation clusters. The number of cases in each group is indicated within the circles. Circles without numbers represent single cases. Different colors indicate methylation cluster affiliation. Tumors occurring in midline locations are depicted in the sagittal view (left panel), tumors occurring in the cerebral and cerebellar hemispheres are depicted in the exterior view (right panel).

(B and C) Kaplan-Meier survival curves for GBM subgroups defined by methylation profiling (B), and mutation analysis (C). The p values were computed by log rank tests between subgroups.

Figure 4.4 – Identification of marker genes affected by differential methylation and expression in GBM subgroups



(A and B) Volcano plots illustrating differences in DNA methylation (A) and gene expression (B) between tumors from the K27 and G34 subgroups. Difference in beta-values (A) and Log_2 fold change in gene expression values (B) are plotted on the x-axis, adjusted p values calculated using the SAM method are plotted on the y-axis.

(C) Starburst plot integrating DNA methylation (x-axis) and gene expression (y-axis) data.

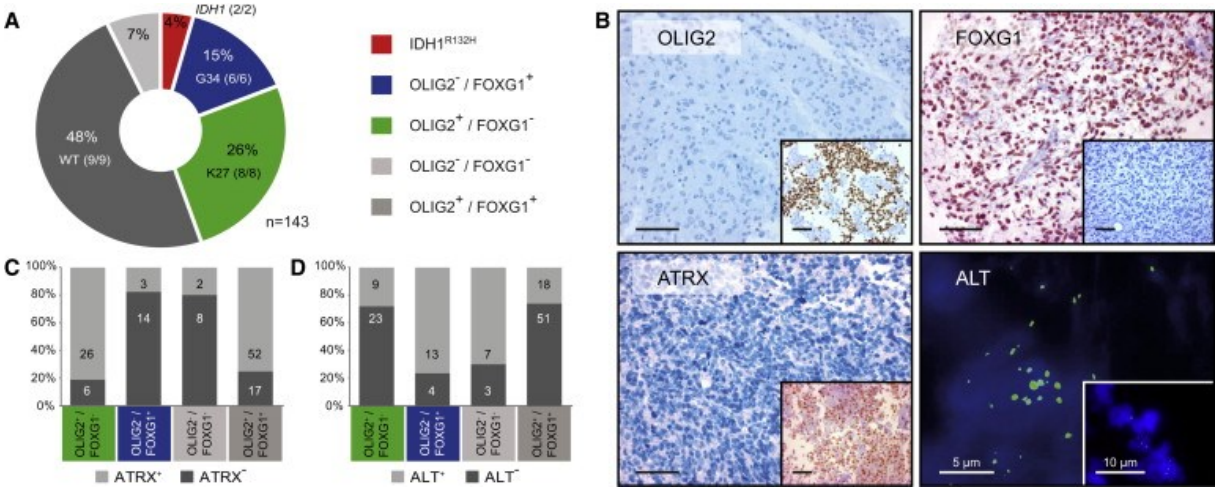
(D) Methylation levels at the OLIG1 and OLIG2 loci across all 210 GBM samples investigated. Each row represents one sample; each vertical bar represents one CpG-site. Light blue bars indicate promoter regions. Methylation levels are represented by a color scale as indicated.

(E) Mean gene expression levels of OLIG1 (upper panel) and OLIG2 (lower panel) across GBM subgroups (n = 48). Significant differences ($^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$) between

subgroups compared to G34 tumors are indicated.

(F) Inverse correlation of promoter methylation (x-axis) and gene expression (y-axis) of OLIG2 across GBM methylation subgroups (n = 48; Pearson's correlation coefficient, r). MES, Mesenchymal.

Figure 4.5 – Identification of H3F3A-mutated GBMs by differential protein expression patterns

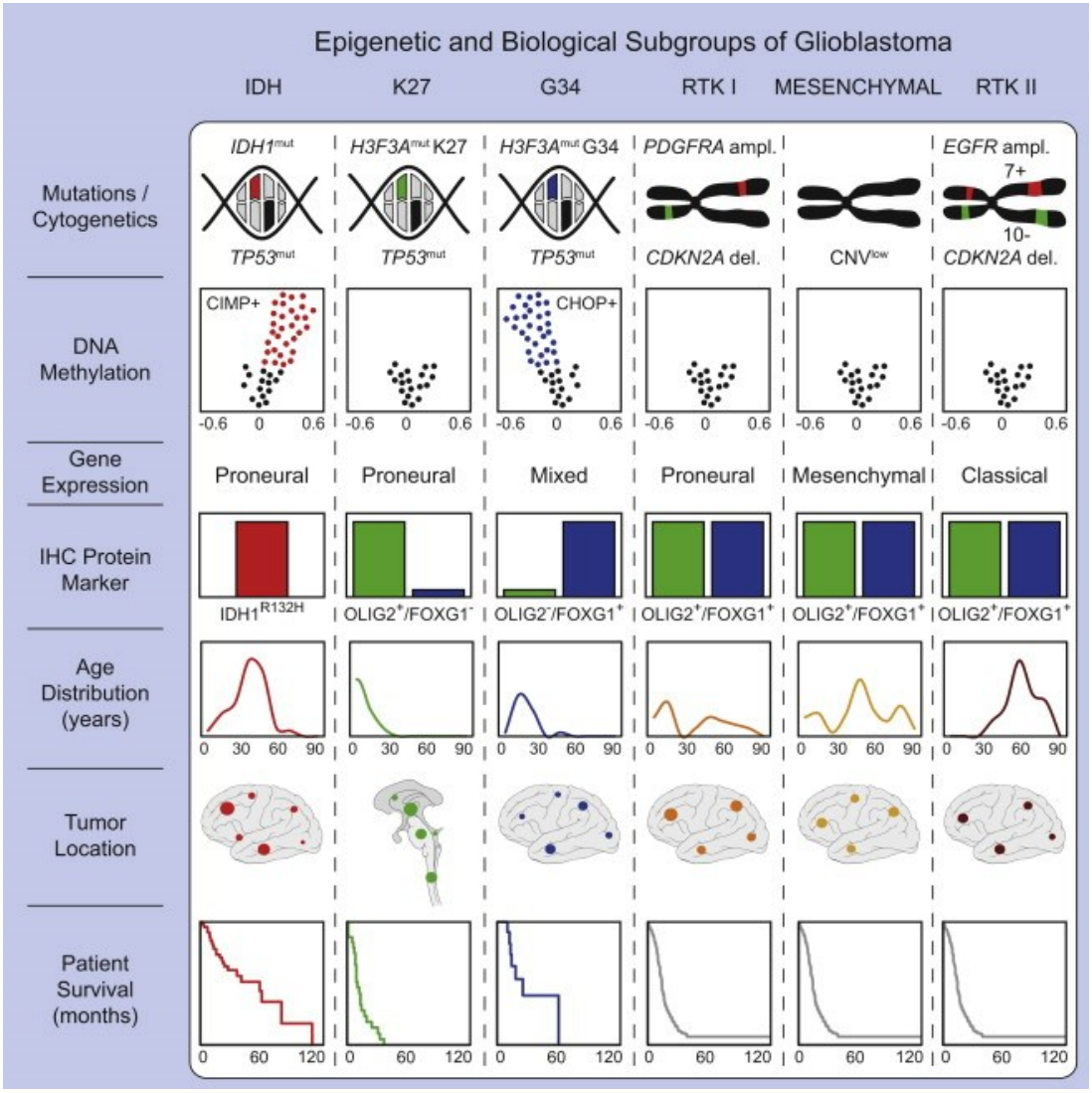


(A) Classification of 143 pediatric GBMs according to protein expression of OLIG2, FOXG1, and mutated IDH1 ($IDH1^{R132H}$). Numbers in brackets indicate samples with known H3F3A and IDH1 mutation status as predicted by immunohistochemistry and verified by targeted gene sequencing, respectively.

(B) Typical pattern of $OLIG2^{-}/FOXG1^{+}$ cells with concomitant loss of ATRX protein expression and ALT as observed in G34-mutated GBMs. Insets show contrasting staining results for comparison. Scale bars represent 100 µm unless indicated differently.

(C and D) Correlation of GBMs as classified in (A) with ATRX loss (C), and ALT (D).

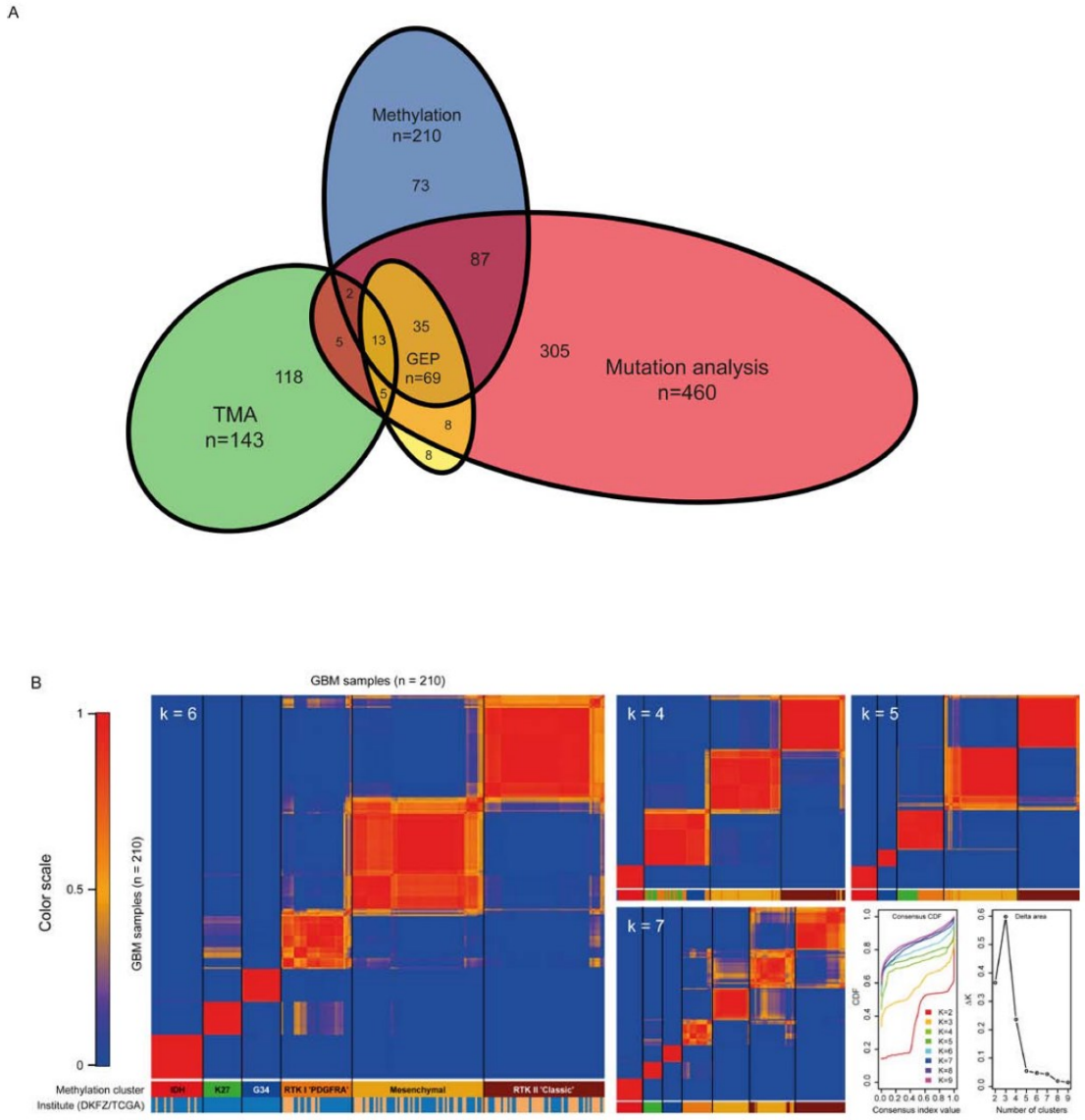
Figure 4.6 – Graphical summary of key molecular and biological characteristics of GBM subgroups

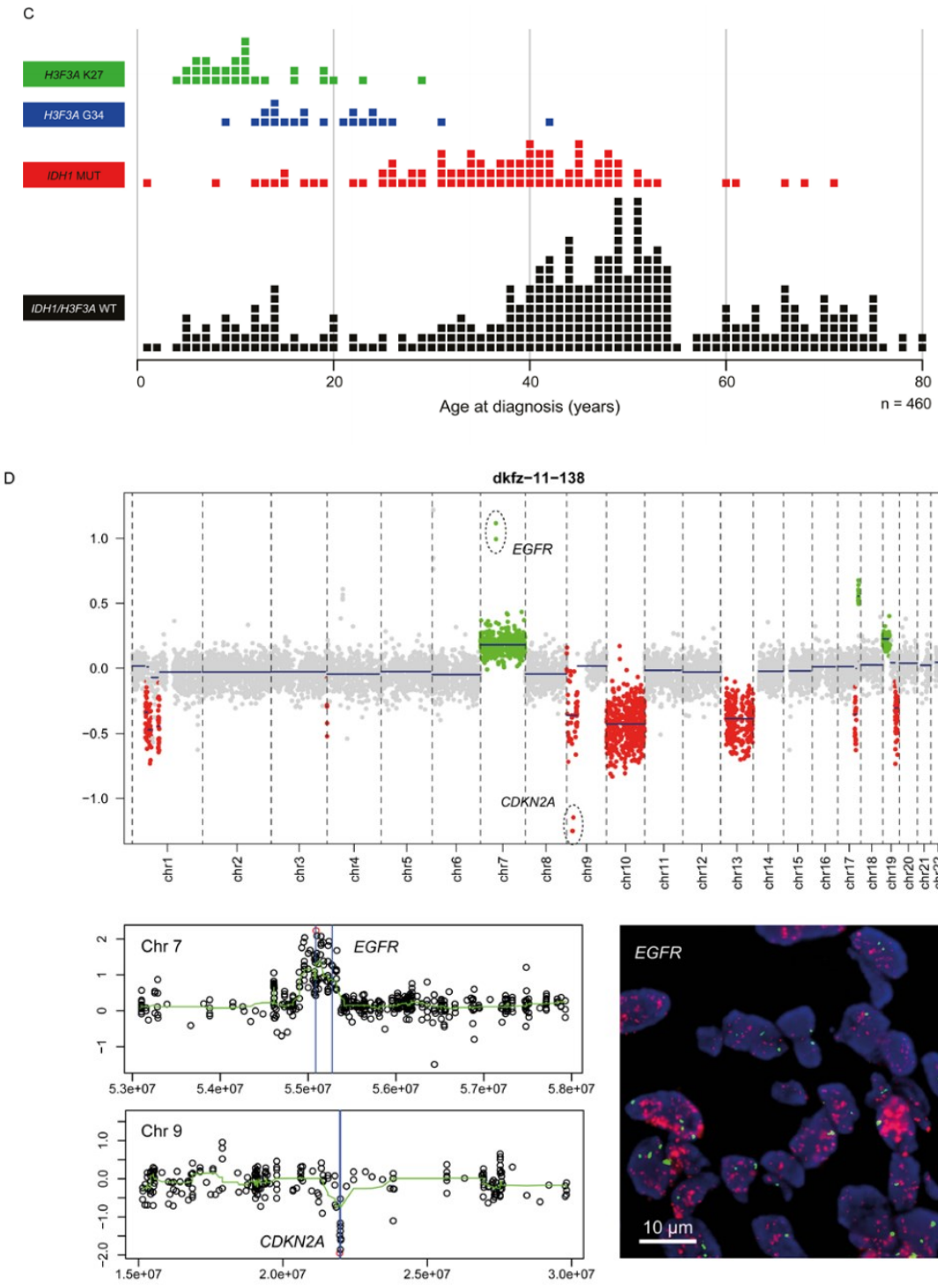


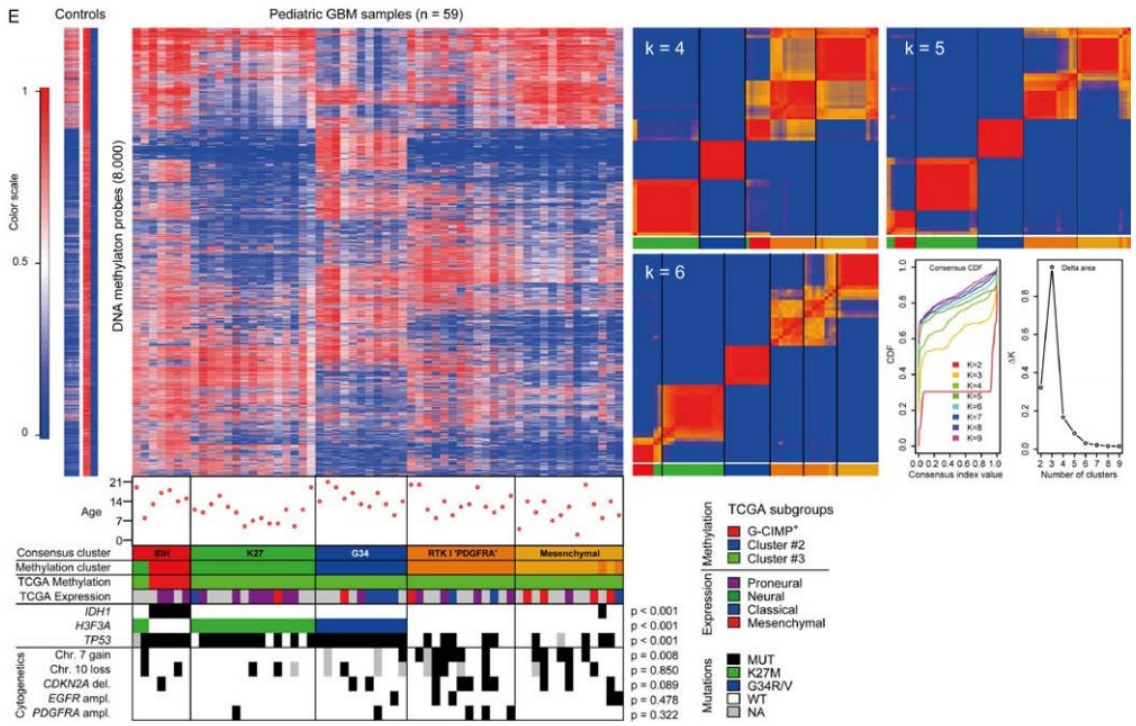
Simplified schematic representation of key genetic and epigenetic findings in six GBM subgroups as identified by methylation profiling and correlations with clinical patient data.

4.8 Supplementary material

Supplementary figure S4.1, related to Figure 4.1 – Integrated molecular classification of glioblastoma




C
H3F3A K27
H3F3A G34
IDH1 MUT
IDH1/H3F3A WT
0
20
40
60
80
Age at diagnosis (years)
n = 460
D
dkfz-11-138
EGFR
CDKN2A
1.0
0.5
0.0
−0.5
−1.0
0.72
0.12
-0.24
-0.96
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
chr15
chr16
chr17
chr18
chr19
chr20
chr21
chr22
Chr 7
EGFR
5.3e+07
5.4e+07
5.5e+07
5.6e+07
5.7e+07
5.8e+07
Chr 9
CDKN2A
1.5e+07
2.0e+07
2.5e+07
3.0e+07
EGFR
10 µm



(A) Overview of all GBM samples included in this study. Relations between data collections of GBMs investigated by different techniques. Numbers indicate the number of samples (non) overlapping between different data collections.

(B) Consensus matrix (*k-means* clustering; K=6) of six methylation clusters identified by unsupervised consensus clustering of 210 GBM samples using the 8,000 most variant methylation probes across the sample set. GBM samples are represented in corresponding order on the x and y axes. Consensus index values are represented by a color scale as indicated. The institute where each sample was profiled (DKFZ/TCGA) is color-coded below. The sample order is retained in the methylation heatmap shown in Figure 4.1. Consensus matrices for different numbers of clusters (K=4, K=5, and K=7) are annotated with the cluster affiliation for each sample (as identified for K=6), as indicated by color code. Empirical cumulative distribution (ECDF) functions and consecutive differences of areas under the ECDF curves from *k-means* clustering for different numbers of methylation clusters are indicated.

(C) Age-distribution of *H3F3A* and *IDH1* mutations in the mutation analysis cohort. Each box represents one GBM sample, and GBM samples are displayed according to the patient age at primary diagnosis. For each sample the mutational status of *H3F3A* and *IDH1* is indicated by different colored boxes as shown.

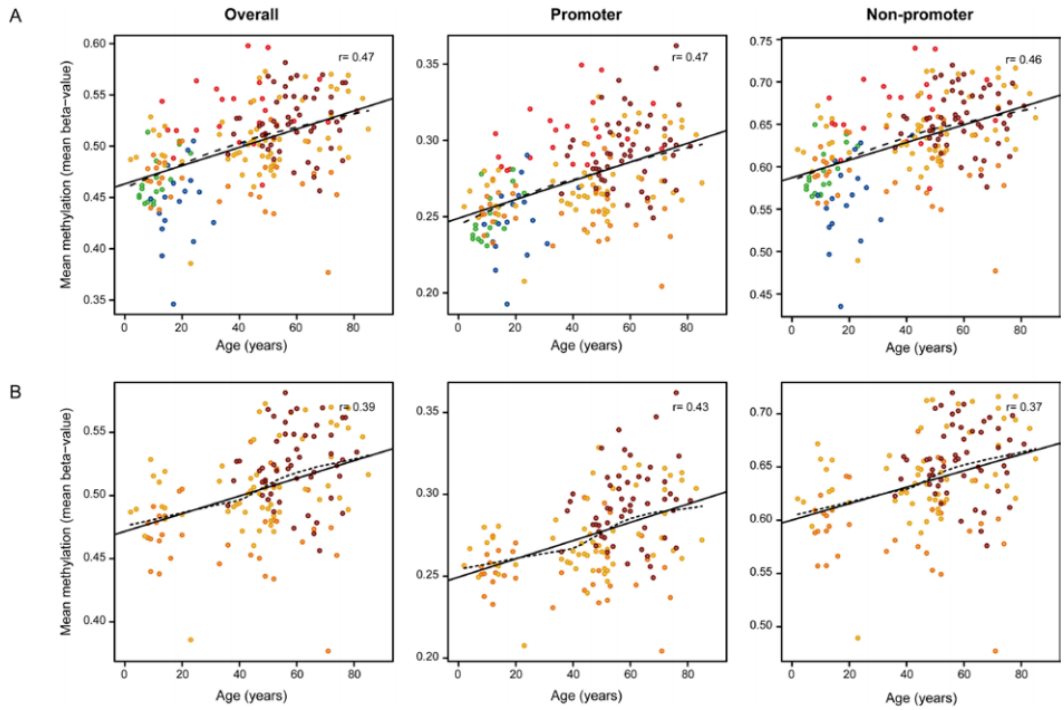
(D) Exemplary copy-number profile of a GBM sample harbouring *EGFR* amplification and *CDKN2A* deletion investigated on the Illumina 450k methylation array (upper panel). Scoring of the *CDKN2A* deletion and *EGFR* amplification (lower left panel), as validated by *EGFR* Fluorescence *In Situ* Hybridization (FISH) (lower right panel).

(E) Heatmap of methylation levels in five pediatric GBM subgroups identified by unsupervised consensus clustering of only pediatric samples (age<22; n=59), four fetal normal brain controls, as well as M.SssI-treated and WGA-DNA (order of controls as in Figure 4.1). Subgroup annotations, mutational status, and cytogenetic aberrations are indicated as in Figure 4.1. Consensus matrices from *k-means* clustering for different numbers of clusters (K=4, K=5, and K=6), using the 8,000 most variant methylation probes across the pediatric sample set. Cluster affiliation from the complete sample set is indicated for each sample. Empirical cumulative distribution (ECDF) functions and consecutive differences of areas under the ECDF curves from *k-means* clustering for different numbers of methylation clusters are indicated.

Supplementary table S4.2, related to Figure 4.1 – Summarized results of the targeted sequencing analysis of H3F3A and IDH1 in 460 GBM samples from patients across all ages

Age at diagnosis (years)	0-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	Median age	Range
***H3F3A* K27**	16	12	2	0	0	0	0	0	10	4-29
***H3F3A* G34**	1	11	8	1	1	0	0	0	18	9-42
***IDH1* MUT**	2	8	12	31	27	5	3	1	38	1-71
***IDH1/H3F3A* WT**	20	28	13	40	96	58	38	26	47	1-80
***IDH1/H3F3A* MUT**	0	0	0	0	0	0	0	0	NA	NA

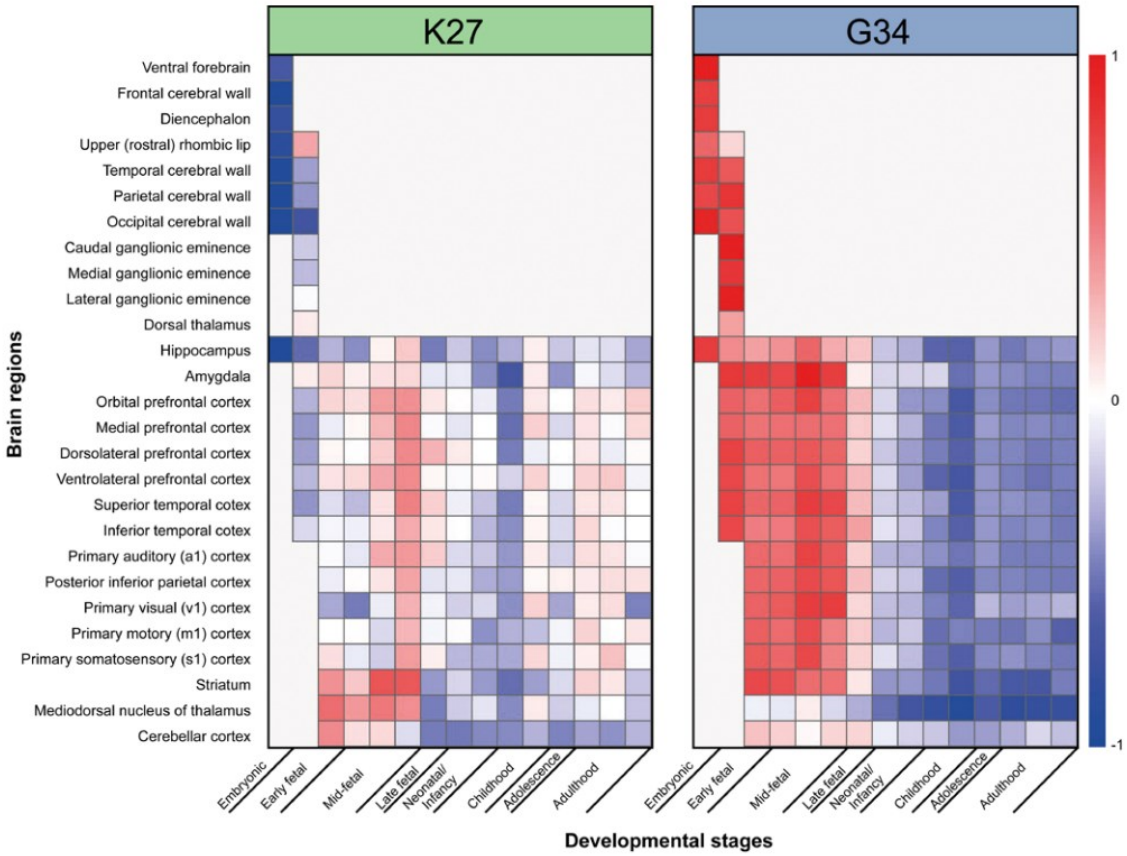
Supplementary figure S4.2, related to Figure 4.2 – Age-dependent global DNA methylation patterns in GBM



(A) DNA methylation levels (mean beta-values) of GBM samples according to age. The Pearson's correlation coefficient (*r*), as well as the linear and lowess regression curves are indicated.

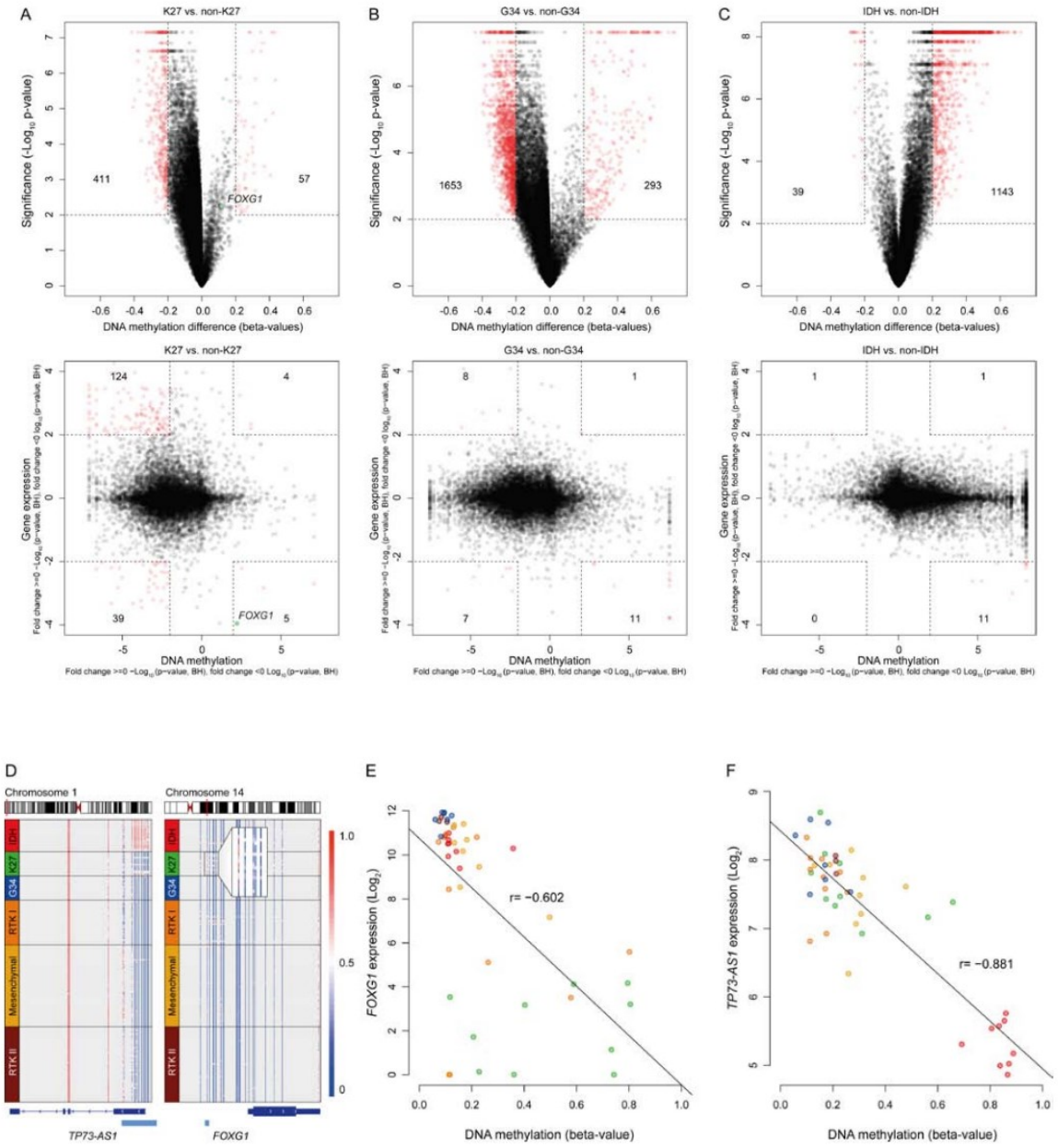
(B) The same analysis as in (A), except that GBM samples carrying age-related mutations in *H3F3A* or *IDH1* were excluded from the analysis to account for their influence on global DNA methylation.

Supplementary figure S4.3, related to Figure 4.3 – Correlation of H3F3A mutations with transcriptomic profiles



Heatmap showing the accordance of K27 and G34 gene signatures with the spatio-temporal transcriptome of the human brain (194). The values indicated by the heatmap are a measure of the relative expression of the combined K27- or G34-specific gene signatures per brain region and developmental stage, compared with the average across all stages/regions. Normal brain transcriptomes were not available for each brain region and developmental stage, as indicated by light grey areas.

Supplementary figure S4.4, related to Figure 4.4 – Differential DNA methylation and gene expression in GBM subgroups

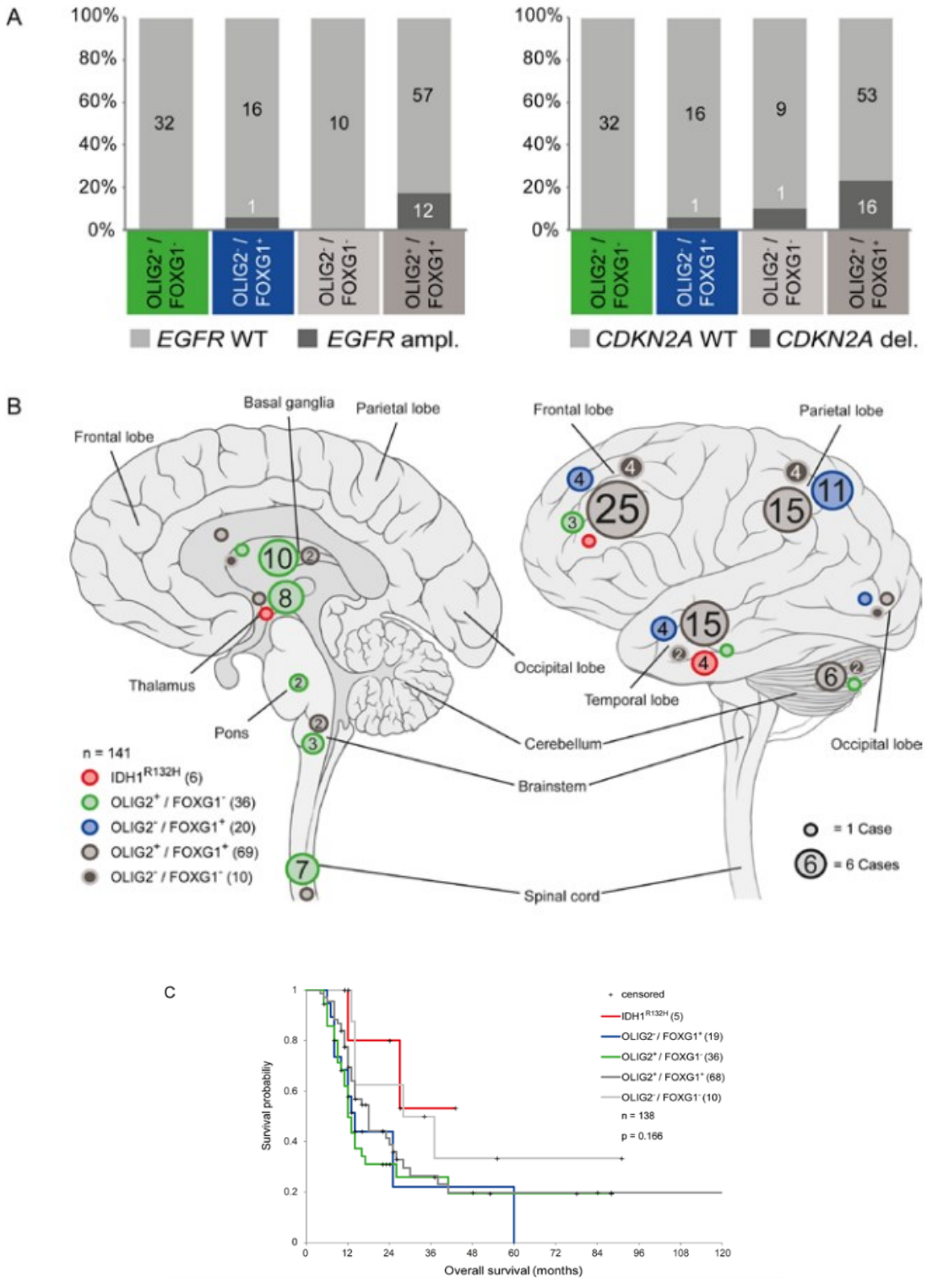


(A-C) Volcano plots (upper panel) and starburst plots (lower panel) of combined methylation probes and genes comparing K27 (A), G34 (B), and IDH subgroups (C) with all other samples. Upper panel: Differences in DNA methylation are plotted on the x-axis, and adjusted p-values are plotted on the y-axis. Lower panel: Integrated DNA methylation (x-axis) and gene expression data (y-axis).

(D) Methylation levels at the *TP73-AS1* and *FOXG1* loci across all 210 GBM samples investigated. Each row represents one sample; each vertical bar represents one CpG-site. Light blue bars indicate regions used for correlations in (E) and (F). Only one differentially methylated CpG-site was used in (E) (see also inserted magnification). Methylation levels are represented by a color scale as indicated.

(E and F) Inverse correlation of DNA methylation (x-axis) and gene expression (y-axis) of *FOXG1* (E) and *TP73-AS1* (F) across GBM methylation subgroups (n=48; Pearson‘s correlation coefficient, *r*).

Supplementary figure S4.5, related to Figure 4.5 – Clinical and molecular characteristics of GBM subgroups identified by differential protein expression



(A) Correlation of 143 GBMs classified by protein expression patterns of OLIG2 and FOXG1 (see also Figure 4.5) with *EGFR* amplification (left panel) and homozygous *CDKN2A* deletion (right panel).

(B) Location of 141 GBMs in the human central nervous system grouped by protein expression patterns of OLIG2, FOXG1, and mutated IDH1 (R132H). The number of cases in each group is indicated within the circles. Circles without numbers represent single cases. Different colors indicate IHC subgroup affiliation. Tumors occurring in midline locations are depicted in the sagittal view (left panel), tumors occurring in the cerebral and cerebellar hemispheres are depicted in the exterior view (right panel).

(C) Kaplan-Meier survival curves for GBM subgroups defined by immunohistochemistry, and p-value computed by log-rank test over all subgroups.

Supplementary table S4.1 – Molecular and clinical patient characteristics of 210 GBM samples included in the methylation profiling study, related to Figure 4.1

Table S4.1 includes subgroup annotations, mutational status, cytogenetic aberrations (as depicted in Figure 4.1), and clinical patient characteristics for each GBM sample included in the methylation profiling study.

Sample ID	Methylation Subgroup (dkfz)	Methylation Subgroup (Noushmehr)	Gene Expression Subgroup (TCGA)	*H3F3A* Mutation Status	*IDH 1* Mutation Status	*TP53* Mutation Status	Chr. 7 Gain	Chr. 10 Loss	*CDKN2A* Deletion	*EGFR* Amplification	*PDGFRA* Amplification	Tumor Location	Age at Diagnosis (years)	Gender	Death	OS (months)	Progression	PFS (months)
dkfz_11-001	G34	Cluster #3	NA	G34R	WT	MUT	0	0	0	0	0	temporal lobe	14	M	1	10	1	7
dkfz_11-002	RTK I 'PDGFRA'	Cluster #3	NA	WT	WT	WT	1	NA	1	0	0	cerebellar	12	F	1	8	1	6
dkfz_11-003	IDH	G-CIMP+	NA	WT	MUT	MUT	0	0	0	0	0	fronto-temporal	13	M	0	14	0	14
dkfz_11-004	Mesenchymal	Cluster #3	Proneural	WT	WT	WT	1	1	1	0	0	frontal lobe	12	F	0	25	1	12
dkfz_11-005	K27	Cluster #3	Classical	K27M	WT	WT	0	0	0	0	0	thalamic	11	M	1	5	1	4
dkfz_11-006	K27	Cluster #3	Proneural	K27M	WT	MUT	0	0	0	0	0	thalamic	10	M	1	8	1	5
dkfz_11-007	Mesenchymal	Cluster #3	NA	WT	WT	MUT	1	0	0	0	0	NA	20	M	NA	NA	1	12
dkfz_11-008	K27	Cluster #3	NA	K27M	WT	WT	0	0	0	0	0	thalamic	11	M	1	32	1	25
dkfz_11-009	IDH	G-CIMP+	NA	WT	MUT	MUT	0	1	0	0	0	temporal lobe	14	M	0	27	0	27
dkfz_11-010	Mesenchymal	Cluster #3	NA	WT	WT	WT	0	0	0	0	0	NA	2	M	0	117	0	117

dkfz_11-011	RTK I 'PDGFRA'	Cluster #3	NA	WT	WT	WT	0	0	0	0	0	frontal lobe	12	M	1	2	0	2
dkfz_11-012	RTK I 'PDGFRA'	Cluster #3	NA	WT	WT	MUT	1	1	1	0	1	temporal lobe	12	M	1	17	1	15
dkfz_11-013	K27	Cluster #3	Proneural	K27M	WT	MUT	0	0	0	0	1	ventricular	10	F	1	24	1	14
dkfz_11-014	RTK I 'PDGFRA'	Cluster #3	Classical	WT	WT	MUT	0	0	1	0	0	frontal lobe	10	F	1	10	1	2
dkfz_11-015	RTK I 'PDGFRA'	Cluster #3	NA	WT	WT	WT	0	0	1	0	1	frontal lobe	13	F	1	16	1	11
dkfz_11-016	RTK I 'PDGFRA'	Cluster #3	NA	WT	WT	WT	1	1	1	1	0	parietal lobe	8	M	1	9	1	5
dkfz_11-017	G34	Cluster #3	Classical	G34R	WT	MUT	0	0	0	1	0	temporal lobe	9	M	0	9	1	6
dkfz_11-018	G34	Cluster #3	NA	G34R	WT	MUT	1	NA	0	0	0	temporal lobe	14	M	0	27	0	27
dkfz_11-019	K27	Cluster #3	NA	K27M	WT	MUT	0	0	0	0	0	DIPG	13	F	1	13	1	5
dkfz_11-020	G34	Cluster #3	Classical	G34R	WT	MUT	0	0	0	0	0	frontal lobe, praecentral	13	M	1	18	0	11
dkfz_11-021	IDH	G-CIMP+	Proneural	WT	MUT	MUT	0	0	0	0	0	NA	15	M	NA	NA	NA	NA
dkfz_11-022	RTK II 'Classic'	Cluster #2	NA	WT	WT	MUT	1	1	0	0	1	NA	66	F	NA	NA	NA	NA
dkfz_11-023	Mesenchymal	Cluster #3	Mesenchymal	WT	WT	MUT	0	0	1	1	0	temporal lobe	14	M	1	14	1	5
dkfz_11-024	K27	Cluster #3	Mesenchymal	K27M	WT	MUT	0	1	0	0	0	thalamic	6	F	1	12	1	7
dkfz_11-025	RTK I 'PDGFRA'	Cluster #3	NA	WT	MUT	MUT	0	0	0	0	0	parietal lobe	8	M	NA	NA	NA	NA
dkfz_11-026	K27	Cluster #3	Proneural	K27M	WT	WT	0	0	0	0	0	thalamic	11	F	1	19	1	11
dkfz_11-027	Mesenchymal	Cluster #3	Mesenchymal	WT	WT	WT	0	0	0	0	0	parietal lobe	14	M	0	84	1	12
dkfz_11-028	G34	Cluster #3	NA	G34R	WT	MUT	0	0	0	0	0	temporal lobe	24	M	NA	NA	NA	NA
dkfz_11-029	G34	Cluster #3	Proneural	G34R	WT	MUT	0	0	0	0	0	parietal lobe	13	M	1	13	1	9
dkfz_11-030	G34	Cluster #3	NA	G34R	WT	MUT	0	0	0	0	0	gyrus cinguli	26	M	NA	NA	NA	NA
dkfz_11-031	Mesenchymal	Cluster #3	Mesenchymal	WT	WT	WT	0	1	0	0	0	frontal lobe	13	F	1	38	1	18
dkfz_11-032	G34	Cluster #3	NA	G34R	WT	MUT	NA	NA	0	0	0	brain stem	21	F	NA	NA	NA	NA
dkfz_11-033	Mesenchymal	Cluster #3	NA	WT	WT	MUT	1	NA	0	0	0	frontal lobe, praecentral	10	F	NA	NA	NA	NA
dkfz_11-034	G34	G-CIMP+	NA	G34V	WT	MUT	1	1	0	0	0	occipital lobe	24	F	NA	NA	NA	NA
dkfz_11-035	Mesenchymal	Cluster #3	Mesenchymal	WT	WT	WT	1	1	1	0	0	parietal lobe	14	F	1	14	1	12
dkfz_11-036	K27	Cluster #3	NA	K27M	WT	MUT	0	0	0	0	0	NA	5	M	0	5.7	NA	NA
dkfz_11-037	K27	Cluster #3	NA	K27M	WT	MUT	1	1	0	0	0	NA	8	F	NA	NA	NA	NA
dkfz_11-038	RTK II 'Classic'	Cluster #2	Classical	WT	WT	MUT	1	1	0	1	0	NA	38	M	NA	NA	1	22

dkfz_11-039	IDH	G-CIMP+	Proneural	WT	MUT	MUT	0	0	0	1	0	NA	40	F	NA	NA	0	31
dkfz_11-040	IDH	G-CIMP+	Proneural	WT	MUT	MUT	0	0	0	0	0	NA	32	F	NA	NA	0	36
dkfz_11-041	Mesenchymal	Cluster #3	Mesenchymal	WT	WT	WT	1	1	0	0	0	NA	46	F	NA	NA	1	9
dkfz_11-042	Mesenchymal	Cluster #3	Mesenchymal	WT	MUT	WT	0	0	0	0	0	NA	34	M	NA	NA	0	33
dkfz_11-043	RTK I 'PDGFRA'	Cluster #3	Mesenchymal	WT	WT	WT	0	0	0	0	0	fronto-temporal	20	F	NA	NA	0	8
dkfz_11-044	RTK I 'PDGFRA'	Cluster #3	Proneural	WT	WT	WT	1	1	0	0	0	NA	36	M	NA	NA	0	10
dkfz_11-045	RTK I 'PDGFRA'	Cluster #3	Proneural	WT	WT	WT	0	0	0	0	0	parietal lobe	19	M	NA	NA	1	8
dkfz_11-046	IDH	G-CIMP+	NA	WT	MUT	WT	1	NA	1	0	0	NA	48	F	NA	NA	0	38
dkfz_11-047	Mesenchymal	Cluster #3	Mesenchymal	WT	WT	WT	1	1	0	1	0	NA	36	M	NA	NA	0	7
dkfz_11-048	K27	Cluster #3	Proneural	K27M	WT	MUT	0	1	0	0	0	thalamic	23	F	NA	NA	0	12
dkfz_11-049	IDH	G-CIMP+	Proneural	WT	MUT	MUT	0	0	0	0	0	NA	33	M	NA	NA	0	38
dkfz_11-050	RTK I 'PDGFRA'	Cluster #3	Proneural	WT	WT	WT	1	0	0	0	0	cerebellar	20	F	NA	NA	0	6
dkfz_11-051	Mesenchymal	Cluster #3	Mesenchymal	WT	WT	WT	1	1	0	0	0	NA	38	M	NA	NA	0	7
dkfz_11-052	IDH	G-CIMP+	Proneural	WT	MUT	MUT	1	0	0	0	0	NA	48	M	NA	NA	1	48
dkfz_11-053	G34	Cluster #3	Proneural	G34R	WT	MUT	0	0	0	0	0	temporal lobe	42	F	1	14	1	12
dkfz_11-054	RTK I 'PDGFRA'	Cluster #3	Proneural	WT	WT	MUT	0	1	0	0	1	NA	33	M	NA	NA	0	11
dkfz_11-055	RTK II 'Classic'	Cluster #2	Classical	WT	WT	WT	1	1	0	1	0	NA	40	M	NA	NA	0	10
dkfz_11-056	IDH	G-CIMP+	Proneural	WT	MUT	MUT	0	NA	0	0	0	NA	25	F	NA	NA	1	17
dkfz_11-057	IDH	G-CIMP+	Proneural	WT	MUT	MUT	0	0	0	0	0	NA	38	M	NA	NA	1	16
dkfz_11-058	IDH	G-CIMP+	Proneural	WT	MUT	MUT	0	0	0	0	0	frontal lobe	18	M	NA	NA	1	17
dkfz_11-059	RTK II 'Classic'	Cluster #2	Classical	WT	WT	WT	1	1	1	1	0	NA	36	M	NA	NA	0	8
dkfz_11-060	IDH	G-CIMP+	Proneural	WT	MUT	MUT	0	0	1	0	0	frontal lobe	17	M	NA	NA	NA	NA
dkfz_11-061	RTK I 'PDGFRA'	Cluster #3	Proneural	WT	WT	MUT	1	1	0	0	0	parietal lobe	14	M	1	5	1	5
dkfz_11-062	RTK I 'PDGFRA'	Cluster #3	Classical	WT	WT	MUT	1	0	0	1	0	temporal lobe	9	F	1	10	1	9
dkfz_11-063	RTK I 'PDGFRA'	Cluster #3	Classical	WT	WT	WT	0	0	0	1	0	frontal lobe	9	M	1	12	1	10
dkfz_11-064	K27	Cluster #3	Proneural	K27M	WT	MUT	0	NA	0	0	0	parietal lobe	5	M	1	6	1	4
dkfz_11-065	K27	Cluster #3	NA	K27M	WT	MUT	0	0	0	0	0	temporal lobe	19	M	1	8	NA	NA
dkfz_11-066	G34	Cluster #3	NA	G34R	WT	MUT	1	0	0	0	0	occipital lobe	23	F	0	10	0	10

dkfz_11-067	G34	Cluster #3	NA	G34R	WT	MUT	0	0	0	0	1	temporal lobe	31	F	1	25	NA	NA
dkfz_11-068	K27	Cluster #3	Proneural	K27M	WT	MUT	0	0	0	0	0	cerebellopontine angle	8	F	1	7	1	5
dkfz_11-069	Mesenchymal	Cluster #3	NA	WT	WT	WT	0	0	0	0	0	frontal lobe, praecentral	7	F	1	12	1	8
dkfz_11-070	RTK I 'PDGFRA'	Cluster #3	Proneural	WT	WT	WT	0	0	0	0	1	ventricular	16	F	1	27	1	18
dkfz_11-071	Mesenchymal	Cluster #3	NA	WT	WT	WT	0	0	0	0	0	NA	4	M	1	11	1	10
dkfz_11-072	G34	Cluster #3	Mesenchymal	G34R	WT	MUT	0	1	0	0	0	parietal lobe	15	M	0	17	0	17
dkfz_11-073	K27	Cluster #3	NA	K27M	WT	NA	0	0	0	0	0	NA	19	F	NA	NA	NA	NA
dkfz_11-074	IDH	G-CIMP+	NA	WT	MUT	MUT	0	0	0	0	0	NA	25	F	NA	NA	NA	NA
dkfz_11-075	RTK I 'PDGFRA'	Cluster #3	NA	WT	WT	MUT	NA	NA	0	0	0	NA	11	M	NA	NA	NA	NA
dkfz_11-091	K27	Cluster #3	NA	K27M	WT	MUT	0	0	0	0	0	thalamic	12	F	1	30	1	26
dkfz_11-092	Mesenchymal	Cluster #3	NA	WT	WT	WT	0	0	0	0	0	NA	44	F	1	16	1	14
dkfz_11-093	IDH	G-CIMP+	NA	WT	MUT	MUT	0	0	1	0	0	NA	49	M	0	48	0	48
dkfz_11-094	RTK I 'PDGFRA'	Cluster #3	NA	WT	MUT	WT	0	0	0	0	0	NA	53	M	0	17	0	17
dkfz_11-095	Mesenchymal	Cluster #3	NA	WT	WT	MUT	1	1	1	1	0	NA	52	F	1	12	1	8
dkfz_11-096	RTK II 'Classic'	Cluster #2	NA	WT	WT	NA	0	1	1	1	0	NA	54	F	NA	NA	NA	NA
dkfz_11-097	Mesenchymal	Cluster #3	NA	WT	WT	WT	0	1	1	1	0	NA	52	F	0	4	0	4
dkfz_11-098	Mesenchymal	Cluster #3	NA	WT	WT	NA	1	1	1	0	0	NA	47	M	NA	NA	NA	NA
dkfz_11-099	RTK I 'PDGFRA'	Cluster #3	NA	WT	WT	WT	1	1	0	1	0	NA	49	F	0	3	0	3
dkfz_11-100	Mesenchymal	Cluster #3	NA	WT	WT	WT	1	1	1	0	0	NA	45	F	1	17	1	14
dkfz_11-101	Mesenchymal	Cluster #3	NA	WT	WT	MUT	NA	1	0	0	0	NA	45	F	1	12	1	4
dkfz_11-102	RTK II 'Classic'	Cluster #2	NA	WT	WT	NA	1	1	1	0	0	NA	48	F	1	14	1	12
dkfz_11-103	RTK I 'PDGFRA'	Cluster #3	NA	WT	WT	NA	0	1	0	0	0	NA	49	F	NA	NA	NA	NA
dkfz_11-104	Mesenchymal	Cluster #3	NA	WT	WT	WT	1	1	1	0	0	NA	40	F	1	10	1	8
dkfz_11-105	RTK II 'Classic'	Cluster #2	NA	WT	WT	NA	0	1	1	1	0	NA	48	M	1	14	1	13
dkfz_11-106	Mesenchymal	Cluster #3	NA	WT	WT	WT	0	0	1	0	0	NA	41	M	0	12	0	12
dkfz_11-107	Mesenchymal	Cluster #3	NA	WT	WT	NA	1	1	1	1	0	NA	42	M	1	12	1	8
dkfz_11-108	Mesenchymal	Cluster #3	NA	WT	WT	NA	1	1	1	0	0	NA	48	M	1	13	1	8
dkfz_11-109	RTK II 'Classic'	Cluster #2	NA	WT	WT	NA	0	1	0	1	0	NA	48	F	NA	NA	NA	NA

dkfz_11-110	Mesenchymal	Cluster #3	NA	WT	WT	MUT	0	0	0	0	0	NA	43	M	0	7	0	7
dkfz_11-111	Mesenchymal	Cluster #3	NA	WT	WT	MUT	1	1	0	0	0	NA	44	F	1	12	1	8
dkfz_11-112	RTK II 'Classic'	Cluster #2	NA	WT	WT	NA	1	1	1	1	0	NA	49	M	NA	NA	NA	NA
dkfz_11-113	RTK II 'Classic'	Cluster #2	NA	WT	WT	WT	1	1	1	1	0	NA	39	M	1	16	1	8
dkfz_11-114	RTK I 'PDGFRA'	Cluster #3	NA	WT	WT	NA	1	1	0	0	1	NA	44	M	NA	NA	NA	NA
dkfz_11-115	Mesenchymal	Cluster #3	NA	WT	WT	WT	NA	NA	0	0	0	NA	33	M	0	12	0	12
dkfz_11-116	IDH	G-CIMP+	NA	WT	MUT	MUT	0	0	0	0	0	NA	47	F	0	27	0	27
dkfz_11-117	RTK I 'PDGFRA'	Cluster #3	NA	WT	WT	WT	1	1	1	1	0	NA	47	M	0	9	1	8
dkfz_11-118	RTK II 'Classic'	Cluster #2	NA	WT	WT	NA	1	1	0	0	0	NA	49	M	0	5	0	5
dkfz_11-119	Mesenchymal	Cluster #3	NA	WT	WT	NA	1	1	0	1	0	NA	54	M	NA	NA	NA	NA
dkfz_11-120	Mesenchymal	Cluster #3	NA	WT	WT	WT	1	0	0	0	0	NA	50	F	1	12	1	8
dkfz_11-121	IDH	G-CIMP+	NA	WT	MUT	MUT	1	0	0	0	0	NA	34	M	0	32	0	32
dkfz_11-122	Mesenchymal	Cluster #3	NA	WT	WT	WT	1	1	1	1	0	NA	49	F	1	18	1	14
dkfz_11-123	Mesenchymal	Cluster #3	NA	WT	WT	NA	1	1	1	1	0	NA	51	M	1	5	1	3
dkfz_11-124	IDH	G-CIMP+	NA	WT	MUT	MUT	0	0	0	0	0	NA	47	M	0	22	0	22
dkfz_11-125	Mesenchymal	Cluster #3	NA	WT	WT	WT	1	1	0	0	0	NA	52	F	1	7	1	3
dkfz_11-126	RTK II 'Classic'	Cluster #2	NA	WT	WT	WT	1	1	1	1	0	NA	51	M	0	10	0	10
dkfz_11-127	Mesenchymal	Cluster #3	NA	WT	WT	NA	1	1	0	1	0	NA	51	F	1	6	1	3
dkfz_11-128	RTK II 'Classic'	Cluster #2	NA	WT	WT	WT	1	1	1	1	0	NA	47	F	1	15	1	12
dkfz_11-129	RTK II 'Classic'	Cluster #2	NA	WT	WT	WT	1	1	1	0	0	parietal lobe	53	M	1	6.6	1	6.2
dkfz_11-130	Mesenchymal	Cluster #3	NA	WT	WT	WT	0	1	0	0	0	parietal lobe	54	M	1	19.9	1	16.8
dkfz_11-131	Mesenchymal	Cluster #3	NA	WT	WT	WT	1	1	1	0	0	frontal lobe	51	M	1	11.6	1	10.8
dkfz_11-132	RTK II 'Classic'	Cluster #2	NA	WT	WT	NA	1	1	1	1	0	parietal lobe	52	M	1	6.3	0	NA
dkfz_11-133	RTK II 'Classic'	Cluster #2	NA	WT	WT	WT	1	1	1	1	0	temporal lobe	52	W	1	16.1	1	8.3
dkfz_11-134	RTK II 'Classic'	Cluster #2	NA	WT	WT	NA	1	1	0	1	0	occipital lobe	55	W	NA	NA	NA	NA
dkfz_11-135	RTK II 'Classic'	Cluster #2	NA	WT	WT	WT	1	1	1	0	0	frontal lobe	69	M	NA	NA	NA	NA
dkfz_11-136	IDH	G-CIMP+	NA	WT	MUT	NA	0	0	0	0	0	temporal lobe	71	W	NA	NA	NA	NA
dkfz_11-137	IDH	G-CIMP+	NA	WT	MUT	MUT	0	0	0	0	0	parietal lobe	68	M	NA	NA	NA	NA

dkfz_11-138	RTK II 'Classic'	Cluster #2	NA	WT	WT	NA	1	1	1	1	1	frontal lobe	75	W	NA	NA	NA	NA
dkfz_11-139	RTK I 'PDGFRA'	Cluster #3	NA	WT	WT	NA	1	0	1	0	1	frontal lobe	71	W	NA	NA	NA	NA
dkfz_11-140	RTK II 'Classic'	Cluster #3	NA	WT	WT	WT	1	1	1	1	0	temporal lobe	68	M	NA	NA	NA	NA
dkfz_11-141	RTK II 'Classic'	Cluster #2	NA	WT	WT	NA	1	1	1	1	0	temporal lobe	72	W	NA	NA	NA	NA
dkfz_11-142	RTK II 'Classic'	Cluster #3	NA	WT	WT	NA	1	1	0	0	0	frontal lobe	57	M	1	15	1	6
dkfz_11-143	G34	Cluster #2	NA	G34R	WT	MUT	0	0	0	0	0	parietal lobe	22	F	0	19	0	NA
dkfz_11-144	G34	Cluster #3	NA	G34R	WT	MUT	0	0	0	0	0	temporal lobe	19	F	1	60	1	46
dkfz_11-145	K27	Cluster #3	Proneural	K27M	WT	MUT	0	1	0	0	0	thalamic	7	F	1	9	1	5
dkfz_11-146	G34	Cluster #3	Classical	G34R	WT	MUT	NA	NA	0	0	1	frontal lobe	17	F	NA	NA	NA	NA
dkfz_11-147	G34	Cluster #3	Classical	G34R	WT	MUT	0	0	1	0	0	NA	12	F	1	12	1	1
dkfz_11-148	K27	Cluster #3	Proneural	K27M	WT	WT	0	0	0	0	0	DIPG	6	F	1	17	1	9
dkfz_11-149	G34	Cluster #3	NA	G34R	WT	MUT	0	0	1	0	0	parietal lobe	17	F	NA	NA	1	NA
dkfz_11-150	K27	Cluster #3	NA	K27M	WT	MUT	0	0	0	0	0	NA	16	M	NA	NA	NA	NA
dkfz_11-151	Mesenchymal	Cluster #3	NA	WT	WT	NA	1	0	0	0	0	temporal lobe	9	M	NA	NA	NA	NA
TCGA-06-0650	Mesenchymal	Cluster #3	Mesenchymal	NA	WT	WT	0	0	0	1	0	NA	39	F	1	23	1	11
TCGA-06-1804	RTK II 'Classic'	Cluster #2	Classical	NA	WT	WT	1	1	0	0	0	NA	81	F	1	13	0	NA
TCGA-06-5408	Mesenchymal	Cluster #2	Classical	NA	WT	MUT	1	1	0	1	1	NA	54	F	1	11	1	5
TCGA-06-5410	Mesenchymal	Cluster #3	Mesenchymal	WT	NA	NA	0	0	0	0	0	NA	72	F	1	3	0	NA
TCGA-06-5411	RTK I 'PDGFRA'	Cluster #3	Neural	NA	WT	WT	1	1	0	0	0	NA	51	M	1	8	1	6
TCGA-06-5412	Mesenchymal	Cluster #3	Mesenchymal	NA	WT	WT	0	0	0	0	0	NA	78	F	1	4	1	2
TCGA-06-5413	RTK II 'Classic'	Cluster #2	Neural	NA	WT	WT	1	1	0	1	0	NA	67	M	0	8	1	6
TCGA-06-5414	RTK II 'Classic'	Cluster #2	Classical	NA	WT	WT	1	1	0	1	0	NA	61	M	0	8	1	5
TCGA-06-5415	RTK II 'Classic'	Cluster #2	Classical	NA	WT	WT	1	1	1	1	0	NA	60	M	0	8	0	NA
TCGA-06-5416	Mesenchymal	Cluster #3	Proneural	NA	NA	NA	0	0	0	0	0	NA	23	F	0	6	0	NA
TCGA-06-5417	IDH	G-CIMP+	Proneural	NA	MUT	MUT	0	0	0	0	0	NA	45	F	0	5	0	NA
TCGA-06-5418	Mesenchymal	Cluster #3	Mesenchymal	NA	WT	WT	1	1	1	0	0	NA	75	F	1	2	0	NA
TCGA-06-5856	RTK II 'Classic'	Cluster #2	Classical	NA	WT	WT	1	1	0	1	0	NA	58	M	1	3	0	NA
TCGA-06-5858	Mesenchymal	Cluster #3	Classical	WT	WT	MUT	1	0	0	0	0	NA	45	F	0	6	1	3

TCGA-06-5859	RTK II 'Classic'	Cluster #2	Neural	WT	WT	WT	1	1	0	0	0	NA	63	M	0	4	0	NA
TCGA-06-6389	IDH	G-CIMP+	Proneural	NA	MUT	MUT	0	0	0	0	0	NA	50	F	0	7	0	NA
TCGA-06-6390	Mesenchymal	Cluster #3	Classical	NA	WT	WT	1	1	0	0	0	NA	58	M	1	5	0	NA
TCGA-06-6391	RTK I 'PDGFRA'	Cluster #3	Proneural	NA	WT	WT	1	0	0	0	0	NA	44	F	1	1	0	NA
TCGA-12-5295	RTK II 'Classic'	Cluster #2	Classical	NA	WT	WT	0	1	1	1	0	NA	60	F	1	14	1	13
TCGA-12-5299	RTK II 'Classic'	Cluster #2	Classical	NA	WT	WT	1	1	1	1	0	NA	56	F	1	3	0	NA
TCGA-12-5301	RTK II 'Classic'	Cluster #2	Neural	NA	WT	WT	1	1	1	0	0	NA	59	M	1	2	0	NA
TCGA-14-0781	Mesenchymal	Cluster #3	Mesenchymal	NA	WT	WT	0	0	0	0	0	NA	49	M	1	0	0	NA
TCGA-15-1444	IDH	G-CIMP+	Proneural	NA	MUT	MUT	0	0	0	0	0	NA	NA	NA	NA	NA	0	NA
TCGA-19-5947	Mesenchymal	Cluster #3	Mesenchymal	NA	NA	NA	1	1	0	0	0	NA	47	F	1	6	1	1
TCGA-19-5950	RTK II 'Classic'	Cluster #2	Classical	NA	NA	NA	1	1	NA	1	0	NA	52	F	0	11	0	NA
TCGA-19-5951	RTK II 'Classic'	Cluster #2	Classical	NA	NA	NA	0	1	1	1	0	NA	76	M	1	8	0	NA
TCGA-19-5952	Mesenchymal	Cluster #2	Classical	NA	NA	NA	1	1	1	1	0	NA	62	M	1	18	0	NA
TCGA-19-5954	RTK II 'Classic'	Cluster #2	Classical	NA	NA	NA	1	1	1	1	0	NA	72	F	0	7	0	NA
TCGA-19-5955	Mesenchymal	Cluster #2	Mesenchymal	NA	NA	NA	0	0	0	0	0	NA	83	M	1	1	0	NA
TCGA-19-5956	RTK I 'PDGFRA'	Cluster #3	Proneural	NA	NA	NA	0	0	1	0	0	NA	53	F	0	6	0	NA
TCGA-19-5958	RTK II 'Classic'	Cluster #2	Classical	NA	WT	WT	1	1	NA	0	0	NA	56	M	0	5	0	NA
TCGA-19-5959	RTK II 'Classic'	Cluster #2	Classical	NA	WT	WT	1	1	1	1	0	NA	77	F	0	5	0	NA
TCGA-19-5960	RTK II 'Classic'	Cluster #2	Proneural	NA	WT	WT	1	1	0	0	0	NA	56	M	0	5	0	NA
TCGA-26-1442	IDH	G-CIMP+	Proneural	NA	MUT	MUT	0	0	1	0	0	NA	43	M	0	31	0	NA
TCGA-26-5132	RTK II 'Classic'	Cluster #2	Classical	NA	WT	WT	1	1	0	1	1	NA	74	M	0	9	0	NA
TCGA-26-5133	IDH	G-CIMP+	Proneural	NA	WT	MUT	0	1	0	0	0	NA	59	M	0	14	1	12
TCGA-26-5134	RTK I 'PDGFRA'	Cluster #3	Proneural	NA	WT	WT	0	1	1	0	1	NA	74	M	0	5	0	NA
TCGA-26-5135	RTK I 'PDGFRA'	Cluster #3	Proneural	NA	WT	WT	1	1	0	0	1	NA	72	F	1	8	0	NA
TCGA-26-5136	Mesenchymal	Cluster #3	Mesenchymal	NA	WT	MUT	1	1	0	0	0	NA	78	F	0	1	0	NA
TCGA-26-5139	RTK II 'Classic'	Cluster #2	Mesenchymal	NA	WT	WT	1	1	1	1	0	NA	65	F	0	1	0	NA
TCGA-28-2501	Mesenchymal	Cluster #2	Mesenchymal	NA	WT	WT	0	0	0	1	0	NA	NA	NA	NA	NA	0	NA
TCGA-28-2510	Mesenchymal	Cluster #3	Neural	NA	WT	WT	0	0	0	0	0	NA	NA	NA	NA	NA	0	NA

TCGA-28-5204	RTK II 'Classic'	Cluster #2	Neural	NA	WT	WT	1	1	1	1	0	NA	72	M	1	14	0	NA
TCGA-28-5207	Mesenchymal	Cluster #3	Mesenchymal	NA	WT	MUT	1	1	0	0	0	NA	71	M	1	11	0	NA
TCGA-28-5208	RTK II 'Classic'	Cluster #2	Mesenchymal	NA	WT	WT	0	1	1	1	0	NA	52	M	0	15	1	4
TCGA-28-5209	RTK II 'Classic'	Cluster #2	Mesenchymal	NA	WT	WT	1	1	1	1	0	NA	66	F	0	2	0	NA
TCGA-28-5213	Mesenchymal	Cluster #3	Mesenchymal	NA	WT	WT	1	1	0	0	0	NA	72	M	0	9	0	NA
TCGA-28-5214	RTK II 'Classic'	Cluster #2	Mesenchymal	NA	WT	WT	1	1	0	1	0	NA	53	M	0	14	1	8
TCGA-28-5215	RTK I 'PDGFRA'	Cluster #3	Mesenchymal	NA	WT	MUT	NA	1	0	0	0	NA	62	F	1	10	1	5
TCGA-28-5216	RTK I 'PDGFRA'	Cluster #3	Mesenchymal	NA	WT	MUT	0	1	NA	0	1	NA	52	M	0	13	0	NA
TCGA-28-5218	Mesenchymal	Cluster #3	Mesenchymal	NA	WT	WT	1	0	0	0	0	NA	63	M	1	5	0	NA
TCGA-28-5219	Mesenchymal	Cluster #3	Classical	NA	WT	MUT	1	1	0	0	0	NA	47	F	0	8	1	8
TCGA-28-5220	RTK II 'Classic'	Cluster #2	Classical	NA	WT	WT	1	1	0	0	0	NA	67	M	0	10	1	8
TCGA-28-6450	RTK II 'Classic'	Cluster #2	Classical	NA	WT	WT	1	1	1	0	0	NA	60	M	1	5	0	NA
TCGA-32-1979	RTK II 'Classic'	Cluster #2	Neural	NA	WT	WT	1	1	1	1	0	NA	69	F	1	19	1	8
TCGA-32-1980	Mesenchymal	Cluster #3	Neural	NA	NA	NA	0	0	0	0	0	NA	72	M	1	1	0	NA
TCGA-32-5222	RTK II 'Classic'	Cluster #2	Proneural	NA	WT	WT	1	1	1	1	1	NA	66	M	0	5	1	3
TCGA-41-5651	RTK II 'Classic'	Cluster #2	Proneural	NA	WT	MUT	1	1	0	0	0	NA	59	F	0	11	1	6
TCGA-76-4925	RTK II 'Classic'	Cluster #2	Proneural	NA	WT	MUT	1	1	1	0	0	NA	76	M	1	4	1	2
TCGA-76-4926	Mesenchymal	Cluster #3	Mesenchymal	NA	WT	WT	1	1	1	1	0	NA	68	M	1	4	1	1
TCGA-76-4927	RTK II 'Classic'	Cluster #2	Classical	NA	WT	WT	1	1	0	1	0	NA	58	M	1	17	0	NA
TCGA-76-4928	Mesenchymal	Cluster #3	Classical	NA	WT	WT	1	1	0	0	1	NA	85	F	1	3	0	NA
TCGA-76-4929	RTK II 'Classic'	Cluster #2	Neural	NA	WT	MUT	1	1	NA	0	0	NA	76	F	1	3	0	NA
TCGA-76-4931	RTK II 'Classic'	Cluster #2	Classical	NA	WT	WT	1	1	1	1	0	NA	NA	NA	NA	NA	0	NA
TCGA-76-4932	RTK II 'Classic'	Cluster #2	Proneural	NA	WT	WT	1	1	1	1	0	NA	50	F	1	47	0	NA
TCGA-76-4934	RTK I 'PDGFRA'	Cluster #3	Proneural	NA	WT	MUT	NA	1	1	0	1	NA	66	F	1	2	0	NA
TCGA-76-4935	RTK I 'PDGFRA'	Cluster #3	Proneural	NA	WT	WT	1	1	1	0	1	NA	52	F	0	10	1	10
TCGA-76-6191	RTK II 'Classic'	Cluster #2	Proneural	NA	WT	WT	1	1	1	0	0	NA	57	M	1	16	1	10
TCGA-76-6192	Mesenchymal	Cluster #2	Proneural	NA	WT	WT	1	1	1	0	0	NA	74	M	1	3	1	2
TCGA-76-6193	Mesenchymal	Cluster #2	Mesenchymal	NA	WT	MUT	0	0	0	0	1	NA	78	M	1	2	1	2

TCGA-76-6282	Mesenchymal	Cluster #3	Mesenchymal	NA	WT	WT	1	1	0	1	0	NA	63	F	1	17	1	15
TCGA-76-6285	RTK I 'PDGFRA'	Cluster #2	Proneural	NA	WT	WT	1	1	1	0	0	NA	64	F	1	8	0	NA
TCGA-81-5910	RTK II 'Classic'	Cluster #2	Classical	NA	WT	WT	1	1	1	1	0	NA	64	M	1	1	0	NA
TCGA-87-5896	RTK II 'Classic'	Cluster #2	Classical	NA	WT	WT	1	1	1	1	0	NA	50	F	0	0	0	NA

Supplementary table S4.3 – Statistical analysis and measurement of differential DNA methylation and gene expression, related to Figure 4.4

Table S4.3 includes a complete list of genes analyzed for DNA methylation and gene expression. Combined methylation values per promoter and gene expression levels were averaged for available samples per subgroup. Difference in methylation and Log_2 fold change in gene expression, as well as associated adjusted p values calculated using the Significance Analysis of Microarrays (SAM) method are included for indicated subgroup comparisons.

Because of space issue, this list is not included in this thesis and is available online (http://www.cell.com/cancer-cell/fulltext/S1535-6108(12)00364-9).

Supplementary table S4.4 – Molecular and clinical patient characteristics of 143 GBM samples included in the immunohistochemistry study, related to Figure 4.5

Table S4.4 includes IHC subgroup annotation, mutational status, and clinical patient characteristics (as depicted in Figures 4.5 and S4.5) for each GBM sample included in the immunohistochemistry study.

TMA Sample ID	Immunohistochemistry Subgroup	*H3F3A* Mutation Status	*IDH 1* Mutation Status	*TP53* Mutation Status	Tumor Location	Age at Diagnosis (years)	Gender	Death	OS (months)	Progression	PFS (months)
1	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	12	M	1	23	1	18
2	OLIG2+/FOXG1-	NA	NA	NA	thalamic	13	F	0	88	0	88
3	OLIG2-/FOXG1+	NA	NA	NA	parietal lobe	15	M	1	6	1	4
4	OLIG2-/FOXG1-	NA	NA	NA	frontal lobe	16	M	1	28	1	22
5	OLIG2+/FOXG1-	NA	NA	NA	brain stem	5	M	0	23	1	15
7	OLIG2-/FOXG1+	NA	NA	NA	occipital lobe	8	F	0	22	0	22
8	OLIG2+/FOXG1+	NA	NA	NA	basal ganglia	4	F	1	12	1	11
9	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	9	F	1	14	1	8
10	OLIG2+/FOXG1-	NA	NA	NA	basal ganglia	11	F	1	16	1	14
11	OLIG2+/FOXG1+	NA	NA	NA	temporal lobe	13	M	1	25	1	18
12	IDH1 mut (R132)	NA	NA	NA	temporal lobe	14	F	0	24	1	8
14	OLIG2-/FOXG1-	NA	NA	NA	temporal lobe	10	F	0	12	0	12
15	OLIG2-/FOXG1+	NA	NA	NA	frontal lobe	7	M	1	8	1	6
16	OLIG2-/FOXG1-	NA	NA	NA	cerebellar	11	F	1	14	1	11
17	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	17	M	1	13	1	9
18	OLIG2+/FOXG1-	NA	NA	NA	basal ganglia	16	F	1	6	1	4
19	OLIG2+/FOXG1+	NA	NA	NA	parietal lobe	11	M	1	18	1	12
20	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	6	M	1	12	1	3
21	OLIG2-/FOXG1+	NA	NA	NA	parietal lobe	1	M	1	8	1	6
22	OLIG2-/FOXG1-	NA	NA	NA	lateral ventricle	13	F	0	91	0	91
23	OLIG2+/FOXG1+	NA	NA	NA	cerebellar	16	M	0	22	0	22
24	OLIG2+/FOXG1+	NA	NA	NA	parietal lobe	16	F	1	13	1	6
27	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	16	M	0	14	0	14
28	OLIG2+/FOXG1-	NA	NA	NA	basal ganglia	4	F	1	12	1	8
29	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	12	M	1	16	1	14
30	OLIG2-/FOXG1-	NA	NA	NA	frontal lobe	13	F	1	13	1	11
31	OLIG2+/FOXG1+	NA	NA	NA	temporal lobe	10	M	1	8	1	8
32	OLIG2+/FOXG1+	NA	NA	NA	parietal lobe	13	F	1	11	1	5
33	OLIG2-/FOXG1+	NA	NA	NA	parietal lobe	7	M	0	16	0	16
34	OLIG2+/FOXG1+	NA	NA	NA	basal ganglia	7	M	1	18	1	16
35	OLIG2-/FOXG1+	NA	NA	NA	parietal lobe	3	M	1	25	1	10
37	IDH1 mut (R132)	NA	NA	NA	temporal lobe	17	M	0	27	1	15

40	OLIG2+/FOXG1+	NA	NA	NA	parietal lobe	5	F	1	12	1	6
43	OLIG2-/FOXG1-	NA	NA	NA	occipital lobe	11	F	1	14	1	11
44	OLIG2+/FOXG1-	NA	NA	NA	basal ganglia	1	M	1	13	1	10
47	OLIG2-/FOXG1+	NA	NA	NA	parietal lobe	15	M	0	13	1	11
48	IDH1 mut (R132)	NA	NA	NA	temporal lobe	17	M	0	43	0	43
49	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	12	M	1	11	1	8
50	OLIG2-/FOXG1+	NA	NA	NA	frontal lobe	14	M	0	22	0	22
51	OLIG2+/FOXG1+	NA	NA	NA	temporal lobe	15	M	1	10	1	7
52	OLIG2+/FOXG1+	WT	NA	WT	parietal lobe	14	F	1	14	1	12
53	OLIG2+/FOXG1+	NA	NA	NA	corpus callosum	4	F	1	9	1	4
54	OLIG2+/FOXG1+	WT	NA	WT	parietal lobe	15	M	0	84	1	12
55	OLIG2+/FOXG1-	NA	NA	NA	spinal cord	6	M	1	11	1	6
56	OLIG2-/FOXG1-	NA	NA	NA	temporal lobe	16	M	0	34	1	22
57	OLIG2+/FOXG1+	WT	NA	MUT	temporal lobe	14	M	1	14	1	5
59	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	14	F	0	14	0	14
60	OLIG2+/FOXG1+	NA	NA	NA	temporal lobe	5	M	0	10	0	10
61	OLIG2+/FOXG1-	K27M	NA	WT	thalamic	6	F	1	12	1	7
62	OLIG2+/FOXG1-	NA	NA	NA	thalamic	6	F	0	10	0	10
63	OLIG2+/FOXG1+	NA	NA	NA	spinal cord	8	F	8	8	0	8
64	OLIG2+/FOXG1+	NA	NA	NA	temporal lobe	6	M	1	11	1	7
65	OLIG2+/FOXG1+	NA	NA	NA	cerebellar	6	F	1	8	1	4
66	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	13	M	1	41	1	10
67	OLIG2-/FOXG1+	G34R	NA	MUT	parietal lobe	13	M	1	13	1	9
68	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	17	F	0	48	1	4
69	OLIG2-/FOXG1+	NA	NA	NA	parietal lobe	12	M	1	7	1	4
70	OLIG2+/FOXG1+	NA	NA	NA	temporal lobe	16	M	0	12	0	12
71	OLIG2+/FOXG1+	WT	NA	WT	frontal lobe	13	F	1	38	1	18
72	OLIG2+/FOXG1-	NA	NA	NA	spinal cord	9	M	1	12	1	9
73	OLIG2+/FOXG1+	NA	NA	NA	cerebellar	15	F	1	18	1	14
74	OLIG2-/FOXG1-	NA	NA	NA	frontal lobe	10	F	0	55	0	55
75	OLIG2+/FOXG1+	NA	NA	NA	cerebellar	13	F	1	14	1	10
76	OLIG2+/FOXG1-	NA	NA	NA	brain stem	5	M	1	41	1	2
80	OLIG2+/FOXG1+	NA	NA	NA	parietal lobe	2	M	1	5	1	3
81	OLIG2+/FOXG1+	NA	NA	NA	temporal lobe	16	M	0	14	0	14
82	OLIG2+/FOXG1-	NA	NA	NA	basal ganglia	10	M	0	53	1	4
84	OLIG2+/FOXG1+	NA	NA	NA	parietal lobe	6	M	1	12	1	6
86	OLIG2-/FOXG1-	NA	NA	NA	cerebellar	6	F	0	11	0	11
87	OLIG2-/FOXG1-	NA	NA	NA	frontal lobe	16	M	1	37	1	20
88	OLIG2-/FOXG1+	NA	NA	NA	parietal lobe	10	F	1	14	1	10
89	OLIG2-/FOXG1+	NA	NA	NA	parietal lobe	14	F	1	8	1	4
90	OLIG2+/FOXG1-	NA	NA	NA	spinal cord	7	M	1	6	1	3
91	OLIG2+/FOXG1-	NA	NA	NA	basal ganglia	15	M	1	26	1	13
92	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	4	M	0	48	0	48
93	OLIG2+/FOXG1-	NA	NA	NA	thalamus	5	M	0	78	1	74
94	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	5	M	0	88	1	73
96	IDH1 mut (R132)	NA	NA	NA	thalamus	5	M	1	12	1	2
98	OLIG2+/FOXG1+	NA	NA	NA	parietal lobe	17	F	0	14	0	14
99	OLIG2-/FOXG1+	NA	NA	NA	temporal lobe	10	M	0	12	0	12
100	OLIG2+/FOXG1-	NA	NA	NA	brain stem	4	F	1	11	1	8
101	OLIG2+/FOXG1-	NA	NA	NA	pons	8	F	1	10	1	6
102	OLIG2-/FOXG1+	NA	NA	NA	frontal lobe	8	F	0	14	0	14
103	OLIG2+/FOXG1+	NA	NA	NA	brain stem	16	M	0	10	0	10
104	OLIG2+/FOXG1+	NA	NA	NA	temporal lobe	5	M	1	18	1	7
106	OLIG2+/FOXG1-	NA	NA	NA	frontal lobe	4	M	1	13	1	11
107	OLIG2+/FOXG1+	NA	NA	NA	temporal lobe	5	M	1	6	1	3
108	OLIG2-/FOXG1+	NA	NA	NA	parietal lobe	12	F	1	12	1	8
109	OLIG2+/FOXG1-	NA	NA	NA	basal ganglia	13	M	0	37	0	37
110	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	16	F	0	14	0	14

111	OLIG2+/FOXG1-	NA	NA	NA	lateral ventricle	7	M	1	8	1	2
112	OLIG2+/FOXG1+	NA	NA	NA	cerebellar	3	F	1	11	1	8
113	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	13	F	1	26	1	11
114	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	14	M	0	10	0	10
115	OLIG2+/FOXG1+	NA	NA	NA	parietal lobe	11	M	0	12	0	12
116	OLIG2+/FOXG1+	NA	NA	NA	parietal lobe	14	M	1	12	1	8
117	OLIG2+/FOXG1-	NA	NA	NA	basal ganglia	11	M	1	5	1	3
118	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	15	F	1	13	1	10
119	OLIG2+/FOXG1-	NA	NA	NA	basal ganglia	1	M	0	8	0	8
120	OLIG2+/FOXG1-	NA	NA	NA	frontal lobe	13	M	1	6	1	4
121	OLIG2+/FOXG1+	NA	NA	NA	temporal lobe	9	M	0	158	1	156
122	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	8	M	1	8	1	5
123	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	6	M	0	14	0	14
124	OLIG2+/FOXG1+	NA	NA	NA	brain stem	5	F	0	26	0	26
125	OLIG2-/FOXG1+	NA	NA	NA	temporal lobe	6	F	0	12	1	10
126	OLIG2+/FOXG1-	NA	NA	NA	spinal cord	15	M	1	14	1	12
127	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	12	F	1	10	1	8
128	OLIG2-/FOXG1+	G34R	NA	MUT	temporal lobe	14	M	1	10	1	7
129	OLIG2+/FOXG1-	NA	NA	NA	spinal cord	6	M	0	24	0	24
130	OLIG2+/FOXG1-	NA	NA	NA	thalamic	7	M	0	22	1	7
131	OLIG2+/FOXG1+	WT	NA	WT	frontal lobe	12	F	0	25	1	12
132	OLIG2+/FOXG1+	NA	NA	NA	cerebellar	1	M	0	12	1	3
134	OLIG2+/FOXG1-	NA	NA	NA	spinal cord	13	F	0	22	0	22
135	OLIG2+/FOXG1+	NA	NA	NA	thalamic	1	M	0	11	0	11
136	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	5	F	0	17	0	17
137	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	7	F	1	8	1	5
138	OLIG2+/FOXG1-	K27M	NA	WT	temporal lobe	11	M	1	5	1	4
139	OLIG2+/FOXG1-	K27M	NA	WT	thalamic	10	M	1	8	1	5
140	OLIG2+/FOXG1-	NA	NA	NA	basal ganglia	8	M	0	5	1	3
142	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	7	M	0	16	0	16
143	OLIG2+/FOXG1+	NA	NA	NA	temporal lobe	6	F	1	4	1	2
144	OLIG2+/FOXG1-	NA	NA	NA	frontal lobe	8	F	1	12	1	10
145	OLIG2+/FOXG1+	NA	NA	NA	occipital lobe	15	F	1	28	1	9
146	OLIG2+/FOXG1+	NA	NA	NA	parietal lobe	17	F	0	10	0	10
147	OLIG2+/FOXG1+	NA	NA	NA	parietal lobe	12	M	0	14	0	14
148	OLIG2+/FOXG1+	NA	NA	NA	temporal lobe	13	M	0	12	1	3
150	OLIG2+/FOXG1-	NA	NA	NA	thalamic	10	F	1	14	1	12
151	OLIG2+/FOXG1+	NA	NA	NA	frontal lobe	17	M	0	16	0	16
152	OLIG2-/FOXG1+	G34R	WT	NA	parietal lobe	14	M	0	NA	1	1
153	OLIG2-/FOXG1+	G34R	WT	NA	frontal lobe	17	F	NA	NA	NA	NA
154	OLIG2+/FOXG1+	WT	WT	NA	parietal lobe	17	M	1	30	1	21
155	OLIG2-/FOXG1+	G34R	WT	NA	NA	12	F	1	12	1	1
156	IDH1 mut (R132)	WT	R132H	NA	frontal lobe	15	F	1	27	1	2
157	OLIG2+/FOXG1-	K27M	WT	NA	spinal cord	20	M	NA	NA	NA	NA
158	OLIG2+/FOXG1-	K27M	WT	NA	cerebellar	7	M	1	9	1	8
159	OLIG2+/FOXG1+	WT	WT	NA	temporal lobe	5	M	NA	NA	NA	NA
160	OLIG2+/FOXG1+	WT	WT	NA	temporal lobe	5	M	1	24	1	14
162	OLIG2+/FOXG1-	K27M	WT	NA	pons	6	F	1	17	1	9
163	OLIG2+/FOXG1-	K27M	WT	NA	NA	9	F	1	9	1	8
165	OLIG2-/FOXG1+	G34R	WT	NA	temporal	19	F	1	60	1	46
166	IDH1 mut (R132)	WT	R132H	NA	temporal lobe	19	F	NA	NA	NA	NA
167	OLIG2+/FOXG1+	WT	WT	NA	parietal lobe	16	F	0	149	1	NA
168	OLIG2+/FOXG1-	K27M	WT	NA	thalamic	7	F	1	9	1	5

4.9 Accession Numbers

The complete CpG methylation values have been deposited in NCBI's Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo) and are accessible through GEO Series accession number GSE36278. The complete gene expression values are accessible through GEO Series accession numbers GSE36245 and, as part of a previously reported series, GSE34824 (1).

4.10 Supplementary experimental procedures

Tumor samples

Snap-frozen primary tumor samples for methylation (n=136; Table S4.1), mutation (n=460; Table S4.2), and gene expression (n=69) analysis and all clinical data were collected at the DKFZ (Heidelberg, Germany) and at McGill University (Montreal, Canada). Paraffin-embedded samples (n=143; Table S4.4) for TMA analysis were collected from a single study at the Burdenko Neurosurgical Institute (Moscow, Russia) and from the Department of Neuropathology, University of Würzburg (Germany). Patient clinical details can be found in Table S4.1 for the methylation analysis data set and in Table S4.4 for the tissue-microarray (TMA) cohort. All tumors were banked at the time of primary diagnosis between 1994 and 2011 in accordance with research ethics board approval from the respective institutes. Informed consent was obtained from all patients. An overview of all samples included in different data collections is given in Figure S4.1A. At least 80% tumor cell content was estimated in all TMA samples by staining each sample with hematoxylin and eosin. Diagnoses of all cases were confirmed by histopathologic assessment by at least two neuropathologists, including a central pathology review that utilized the 2007 WHO classification for CNS tumors. Scoring of tumor location was performed blind to molecular data by reviewing neuro-surgical as well as neuro-pathological reports from the respective

institutions, and only when sufficient and unambiguous information was available. Detailed information about samples provided by The Cancer Genome Atlas (TCGA) can be found elsewhere (http://cancergenome.nih.gov).

Nucleic acid isolation

Extraction of high molecular weight DNA from frozen tumor samples was carried out as previously described (191). Total RNA from biopsy samples was isolated after milling of the frozen sample in a Micro-Dismembrator S (B. Braun, Melsungen, Germany), using TRIzol reagent (Invitrogen, Carlsbad, USA), and subsequently passed over an RNeasy Mini Spin Column (Qiagen, Hilden, Germany) for the removal of small fragments. DNA quality was assessed on a 1% agarose gel, and integrity control and sample quantitation of total extracted RNA was performed using the Agilent 2100 Bioanalyzer (Agilent Technologies, Wilmington, DE, USA).

Methylation array processing

For genome-wide assessment of DNA methylation we used the HumanMethylation450 BeadChip (Illumina, San Diego, USA) (204).Methylation analysis of glioblastoma samples (n=136) and controls (n=10; four fetal and two adult samples of non-neoplastic cerebellum; two samples of Whole-Genome Amplified (WGA-)DNA (unmethylated control); two samples of M.SssI-treated DNA (100% methylated control)) was performed according to the manufacturer's instructions at the DKFZ Genomics and Proteomics Core Facility (Heidelberg, Germany). The complete CpG methylation values have been deposited in NCBI's Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo) and are accessible through GEO Series accession number GSE36278. Methylation data of additional adult glioblastoma samples (n=74) were obtained from the TCGA website (https://tcga-data.nci.nih.gov; available data from TCGA batches 79 and 111).

The following filtering criteria were applied: Removal of probes targeting the X and Y chromosomes (n=11,551), removal of probes containing a single-nucleotide polymorphism (dbSNP132 Common) within five base pairs of and including the targeted CpG-site (n=24,536), and probes not mapping uniquely to the human reference genome (hg19) allowing for one mismatch (n=9,993). In total 438,370 probes were kept for analysis.

Detection of Copy-Number Aberrations (CNAs)

Copy number aberrations were detected from the 450k Infinium methylation array in a custom approach using the sum of both methylated and unmethylated signals. Probes found to be highly variant in the six normal cerebellum samples were excluded from the analysis according to the following criteria: Removal of probes not within the 0.05 and 0.85 quantile of median summed values or over the 0.8 quantile of the median absolute deviation. Log-ratios of samples to the median value of control samples were calculated, and sample noisiness was determined as the median absolute deviation of adjacent probes. Probes were then combined by joining 20 adjacent probes, and resulting genomic windows less than 100kb in size were iteratively merged with adjacent windows of smaller size. Windows of more than 5Mb were excluded from analysis, resulting in a total of 8,654 windows throughout the genome. For each window, the median probe value was calculated and shifted to minimize the median absolute deviation from all windows to zero for every sample. Segmentation was performed by applying the circular binary algorithm using the following settings: min.width=10, nperm=32000, alpha=0.001, undo.splits="sdundo", undo.SD=2.2 (205). The median value of windows contained in each segment was calculated, and classified as homozygous or hemizygous deletion, neutral, gain or high-level amplification by the following arbitrary thresholds: -0.96, -0.24, 0.12 and 0.72. For the detection of *EGFR* and *PDGFRA* high-level amplifications, homozygous *CDKN2A* deletions, and CNAs affecting chromosomes 7 and 10 (as depicted in Figure 4.1), automatic scoring was verified by manual

curation of the respective loci for each individual profile, and compared with results obtained from SNP profiling and FISH analysis where available (Figure S4.1D).

Gene expression array processing

Glioblastoma samples for which RNA of sufficient quantity and quality was available (n=69) were analyzed on the Affymetrix GeneChip® Human Genome U133 Plus 2.0 Array at the Microarray Department of the University of Amsterdam, the Netherlands. Sample library preparation, hybridization, and quality control were performed according to the manufacturer's protocols. Expression data were normalized using the MAS5.0 algorithm of the GCOS program (Affymetrix Inc). Target intensity was set to 100 ($\alpha 1$= 0.04 and $\alpha 2$ = 0.06). Detection p-values were assigned to each probe set using the MAS5.0 algorithm (trimmed mean 96 = 100). Quality of the arrays was ensured by inspection of the *ACTB* and *GAPDH* 5′-3′ ratios as well as the percentage of present calls generated by the MAS5.0 algorithm. For subsequent analysis only one probe was used to represent each gene (R2, http://r2.amc.nl, HugoOnce=yes). Probes not showing a present call in any samples were discarded. The complete gene expression values have been deposited in NCBI's Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo) and are accessible through GEO Series accession numbers GSE36245 and GSE34824. Gene expression data of additional adult glioblastoma samples (n=74) were obtained from the TCGA website (https://tcga-data.nci.nih.gov; available data from TCGA batches 79 and 111).

Statistical analysis and measurement of differential DNA methylation and gene expression

Missing values were imputed using a k-nearest neighbor algorithm (206). For unsupervised consensus clustering we used the 8,000 most variable methylated probes (by standard deviation) across the dataset (R package: clusterCons) (188, 189). The consensus matrix was calculated using the *k-means* algorithm (10 random starting sets, maximum of 1000 iterations)

on a fraction of probes (0.8) in 1000 iterations. Strongest statistical support for six subgroups is illustrated by using empirical cumulative distribution function (ECDF) plots, differences between areas under the cumulative distribution (CDF) plots (K=2 to K=9) generated from the consensus matrix, and the corresponding consensus matrix plots for visual assessment (Figure S1B). For illustrative purposes the consensus matrix was reordered within individual subgroups using hierarchical clustering. Sample order was retained in the corresponding DNA methylation heatmap (x-axes). Additionally, hierarchical clustering was performed on DNA methylation probes (y-axes). Consensus clustering of pediatric cases only (age<22) (Figure S4.1E) was performed using the same parameters.

Ensembl gene annotations (GRCh37.p5) and Affymetrix probe localizations were obtained from MartView (http://www.biomart.org). Probes within−1500/+500bp of the transcription start site of all protein-coding transcripts were considered to be located within the promoter of the corresponding gene (11.1±7.0) and were collapsed by taking the mean of the 50% most variant probes (by SD). The Significance Analysis of Microarrays (SAM) method was used to identify genes, which are differentially methylated or differentially expressed between subgroups (including 6 samples without methylation data but with known mutations in *H3F3A* or *IDH1*). Correction for multiple testing was performed using the Benjamini-Hochberg method. Genes were considered significantly differentially methylated/expressed between two subgroups when displaying an adjusted p-value <0.01 and a methylation difference of 0.2 or a 2-fold change in expression.

Classification according to described DNA methylation and gene expression signatures

Classification of samples studied on the Illumina 450k methylation array into three DNA methylation clusters described by Noushmehr and colleagues (31) was performed by using the Predictive Analysis of Microarrays (PAM) method (207).Since many of the described 1,503

most variable DNA methylation probes are also present on the 450k array, we used the overlapping set of probes (n=1,444) to train the classifier with data from the original publication (91 samples) and predict sample assignments accordingly.

A similar approach was taken to classify samples profiled for gene expression according to the 840 gene signature described by Verhaak and colleagues (25).Samples from the original publication with data from the Affymetrix U133A platform available on the TCGA website were used as a training set (101 samples, excluding samples with TSS code 02). Gene expression data for all 74 TCGA samples analyzed in this study was also obtained from the TCGA website. Matching of gene symbols was performed to integrate TCGA data with samples profiled in this study (analyzed on the Affymetrix U133 plus 2.0 platform, 785 gene overlap). Samples were then quantile normalized and expression values were normalized per gene by *z*-score transformation across all patient samples, as described in the original publication. Subsequently, classifier training and subgroup prediction into the four described subtypes using the PAM method was performed.

Comparison of mutation-derived gene signatures and brain regions/developmental stages

Genes differentially expressed between *H3F3A* mutated versus all other GBM samples were compared to the spatio-temporal transcriptome of the ("normal") human brain (GSE25219) (194).Mutation-specific signatures for up- or downregulated genes were generated by selecting genes differentially expressed between K27-mutated ($p<0.05$) or G34-mutated ($p<0.01$) GBM samples compared to all other pediatric GBM samples wild type for *H3F3A*, generating four gene signatures (G34-up, G34-down, K27-up and K27-down). Normal brain expression data for the respective genes was extracted from R2, where one reporter was used to represent a gene symbol (HugoOnce=yes). In this way, a match on gene symbol could be made between the two different platforms used (normal brain analyzed on the Affymetrix

Exon 1.0 platform; tumor samples analyzed on the U133 Plus 2.0 platform). Gene expression values were normalized per gene by *z*-score transformation and averaged per sample, effectively generating 4 average *z*-score values per normal brain sample per gene signature (G34-up, G34-down, K27-up and K27-down). To visualize accordance, we combined the up- and down-regulated gene signatures by calculating ($score^{up}$ - $score^{down}$)/2, leaving a single score per sample for both the K27 and the G34 comparisons. All the samples within the normal brain dataset were subsequently categorized into unique groups of brain region and developmental stage, of which the average scores were calculated and represented in a heatmap. Thus, the heatmap represents the level of expression of the respective mutation-specific gene signatures per brain region/development stage, with a higher score (corresponding to a red colour) representing higher-than-average expression of the signature compared with other region/stage combinations. Samples with non-unique locations (m1c/sc1 tissue combi) were discarded.

Statistical analysis of clinical and molecular data

Kaplan-Meier analysis was performed to estimate the survival time of different GBM subgroups and a log-rank test was used to test for differences of more than one survival curve. Comparisons of binary and categorical patient characteristics between subgroups were performed by the use of a two-sided Fisher's exact test. An unpaired t-test was used to test for differences between the mean values for continuous variables in GBM subgroups. All statistical computations were performed with the statistical software environment R, version 2.14.0 (208).

Immunohistochemistry (IHC) and Fluorescence In Situ Hybridization (FISH)

Hematoxylin and eosin (HE) stained sections from all 143 paraffin blocks were prepared to define representative tumor regions. In addition, 10 samples of non-neoplastic brain tissues

were included as control samples, as described previously (209). All tissue specimens were arrayed into a recipient block as previously described (210). Antibodies against the following antigens were applied: OLIG2 (Millipore, Billerica, MA, USA; AB9610; dilution 1:250), FOXG1 (Abcam, Cambridge, UK; ab18259; dilution 1:50), ATRX (Sigma, HPA001906; dilution 1:750), and mutated IDH1 (R132H; (190); dianova, DIA H09)). TMA staining was performed, evaluated, and scored as published (211). Multicolor interphase FISH analysis for *PDGFRA*, *EGFR,* and *CDKN2A* was performed as described (191). Telomere-specific FISH was done using a standard formalin-fixed paraffin-embedded FISH protocol (59), using a FITC peptide nucleic acid telomere probe from Dako (Glostrup, Denmark).

Genomic sequencing

Targeted gene sequencing of *H3F3A* (first coding exon), *IDH1* (exon 4), and *TP53* (all exons) was performed by Qiagen (Hilden, Germany) in both forward and reverse directions using purified PCR products. Sequence alignment was carried out using the Staden package Pregap4 v1.5 and Gap v4.10 software. PCR procedures were as previously described (192). Primer sequences are available upon request.

Chapter 5: General discussion

Although high-grade astrocytomas including DIPGs are rare, they remain the main cause of mortality within the group of patients affected with a brain tumor. Current therapeutic approaches have been proven insufficient to improve the patient's prognosis over the last decades, urging the scientific community to combine their efforts into understanding the molecular basis of this disease. Because these tumors are very rare, and in the case of DIPGs, rarely biopsied and always fatal, tissue material is scarce and international collaboration is required to collect enough samples to have meaningful results. During my PhD, our lab has participated into this international effort, and in collaboration with groups within Canada (Dr. Jacek Majewski's laboratory from Genome Quebec, Dr. Cynthia Hawkins's laboratory from Toronto SickKids) and overseas (Dr. Stephan Pfister's group from DKFZ, Heidelberg, Germany), our work has provided new insights into pGBM oncogenesis. RTK inhibitors (TKI) have been investigated for several years now in the treatment of aGBM and pGBM without any significant success (76, 212). It was the first targeted therapy assessed in this disease, based on frequent overexpression of EGFR, and its failure highlighted the difficulty to treat a tumor with multiple genetic alterations. Our work has given evidence as to why current therapies against RTKs are inefficient, outlining the importance of redundant pathways and RTK activation. More importantly, the last two chapters of my dissertation have shed light on what we think are critical genetic and epigenetic alterations in pGBM and DIPG, and we hope that this work will be the basis for the development of potential novel therapeutic targets that will be relevant to tumors with a dismal prognosis.

5.1 Underestimated importance of trafficking regulation of signalling proteins

RTKs are frequently overexpressed in both adult and pediatric GBM, and TCGA data report that 88% of adult tumors have an altered RTK/Ras/PI3K signalling pathway, including *EGFR*

mutation or amplification in 45% (24). Therefore, it made sense that EGFR inhibitors were the first targeted therapy investigated in aGBM, unfortunately without any significant treatment benefit. One hypothesis for this lack of response is the activation of redundant pathways upon inhibition of one target, and one way to answer that would be to simultaneously target multiple signalling pathways. In pGBM, EGFR is overexpressed in up to 80% of tumors (45), although most often without gene alteration (54, 126). In this context, EGFR inhibitors have also been tested, with no further success (212). In the first objective of this dissertation, we addressed the two questions as to why EGFR and other RTKs are frequently overexpressed at the protein level with rare gene alteration, and why single targeted agent is not enough to induce a sustainable tumor response. We have shown that the dysregulated expression of RTKs was related to a disrupted endocytic pathway through SNX3 overexpression, a protein involved in early endosome sorting. Moreover, our *in vitro* and *in vivo* results confirmed the role of RTK activation and signalling in cell proliferation and tumor formation in pGBM. To our knowledge, this is the first study on the role of SNX3 in cancer. Sorting nexins are an emerging family of proteins which play a crucial role in trafficking regulation, in the endosomal pathway as well as in the retromer-mediated retrograde trafficking (152). Over the last decade, cumulative evidence has shown the role of endocytosis to regulate RTK signalling, more particularly in cancer (66, 213). Overall, our data shed light on an underestimated mechanism involved in pediatric glioblastoma oncogenesis: the disruption of key RTK trafficking, increasing their signalling and delaying their degradation. Moreover, we showed that this mechanism is not specific to one RTK, therefore confirming that single targeted therapy is unlikely to be able to block pathways that are redundant, with RTK activation being able to switch from one to another during this process. Because of its frequent overexpression in pGBM and because this expression can be easily assessed by immunohistochemistry, we think SNX3 is a useful biomarker to predict

tumor response to TKI. Unfortunately, there is no drug that targets EGFR endocytosis and is approved for human use (213). Furthermore, because endocytosis is a universal mechanism critical for the normal cells, one should be cautious before using it in patients, especially children who have not yet completed their growth and maturation. Presently, the U.S. National Institutes of Health's website, https://clinicaltrials.gov, which is a registry of clinical studies conducted around the world, lists several recruiting trials combining treatment with at least one receptor tyrosine kinase inhibitor in pediatric high-grade astrocytomas, emphasizing the importance of this target in novel treatment strategies.

5.2 Pediatric high-grade astrocytomas: an epigenetic disease

Previous studies have shown that GBM in both adults and pediatrics are clinically and biologically very heterogenous, although they appear similar under the microscope. Therefore, in order to improve therapeutic approaches, it is critical to distinguish the various subgroups and characterize them as best as possible to identify possible biomarkers and therapeutic targets.

In 2012, we and others have described recurrent histone H3.3 mutations in pGBM. It was the first time that somatic histone mutations were described in human diseases, and this finding has changed our understanding of pediatric high-grade astrocytomas. Previous studies on adult secondary GBM described these tumors to harbor frequent mutation of *IDH1/2*, conferring to the tumor DNA a specific methylation profile called the Glioma-CpG Island Methylator Phenotype (G-CIMP). This was the first clue on how critical epigenetic dysregulation was likely to be involved in tumor formation. *H3F3A* mutations in pGBM extend this hypothesis to their pediatric counterparts, confirmed by our study reported in Chapter 5. Using an integrative approach, we described 6 subgroups of pediatric and adult

GBMs based on their DNA methylation profile as follows: “IDH” regrouping tumors harbouring *IDH1/2* mutation, “K27” regrouping all the tumors with *H3F3A* K27M mutation, “G34” regrouping tumors with *H3F3A* G34R/V mutations, “RTK I (PDGFRA)”, and “RTK II (classic)” regrouping adult primary GBM with *EGFR* amplification, and “mesenchymal”. These subgroups correlate with clinical patterns and specific genetic alterations. In 2013, TCGA reported their integrative analysis of aGBM (24), including DNA methylation data; they confirmed that aGBMs did not present mutations in histone 3 genes and showed that they clustered into 6 subtypes. When compared to our classification, “IDH” was assigned to the “C-GIMP” subtype, “mesenchymal” to M1-M2 subtypes, “RTK II (classic)” to M3-M4 subtypes, and “RTK I (PDGFRA)” to M6 subtype, confirming the robustness of our classification (24).

Our data showed that at least 30%–40% of pediatric and young adult GBMs are characterized by disrupted epigenetic regulatory mechanisms, associated with recurrent and mutually exclusive mutations in either *H3F3A* or *IDH1*, and with aberrant DNA methylation patterns. Interestingly, when defining tumors by the type of *H3F3A* mutation, we observe that they arise in distinct neuroanatomical compartments (Figure 5.1). On one hand, K27M-H3.3 mutations occur in tumors arising from the midline and hindbrain. This group of tumors includes DIPG, thalamic, cerebellar and spinal GBMs. More recent studies have better characterized this group, which also presents K27M mutation in genes encoding the canonical histone H3.1 (*HIST1H3B* and *HIST1H3C*) and recurrent *ACVR1* mutations (214-217). Since our study (2), it has been confirmed by several independent groups that the K27M mutation is an independent prognostic factor of poor outcome (2, 218, 219). On the other hand, H3.3-G34R/V tumors all arise in the cortical brain. These tumors also occurred in older patients (adolescent/young adult), and are associated with *TP53* and *ATRX* mutations.

From the data we have presented in this dissertation, further investigations are required to better understand the pathogenicity of GBM and propose new therapeutic avenues. We think that this effort should be directed into three different directions: (1) pathogenicity of mutated histone H3, (2) characterization of pGBM that are not mutated for histone H3, and (3) translational research to implement these findings as quickly as possible in a clinical setting. The latter will be discussed in this chapter.

Given the location of the *H3F3A* mutations at or near critical regulatory histone residues, and their distinct methylation profiles, H3.3 mutations are likely to be directly involved in producing widespread aberrant DNA methylation and deregulation of gene expression. We previously know from anterior studies that DNA methylation profiles are tightly linked to histone 3 lysine K27 trimethylation (H3K27me3) patterns (119, 120). Moreover, in high-CpG-density promoters, loss of H3K4me3 and retention of H3K4me2 or H3K27me3 is correlated with an increase in DNA methylation (174). Thus, mutations affecting H3K27 methylation are likely to affect DNA methylation.

The exact mechanism by which the *H3F3A* or *HIST3H1B* mutations might be inducing epigenetic reprogramming requires further elucidation. However, recent studies have reported that H3-K27M mutations in pGBM reduce global levels of H3K27me3 both in the tumor (220-223) and *in vitro* models (221) by inhibiting PRC2 methyltransferase activity through the interaction with its EZH2 subunit (221, 223). Because the H3-K27M mutation in primary tumors and cell lines is heterozygous, this indicated a dominant-negative effect on all H3 variants. H3K27me3 is a known transcriptional repressive mark (224), and analysis of chromatin immunoprecipitation followed by next-generation sequencing (ChIP-Seq) and whole-genome bisulfite sequencing showed that alterations in H3K27me3 occupancy in these tumors and global DNA hypomethylation are associated with a differential gene expression

(223). However, further work is needed to fully understand the mechanistic association between H3K27me3 and DNA hypomethylation and how these changes are targeted toward specific genetic loci.

The subset of tumors unified by H3.3-G34R/V mutations is characterized by a cortical location and an age older than for patients with an H3-K27M mutant tumor, alongside concurrent mutations in *TP53* and *ATRX* (1, 116). ChIP-Seq of H3K36me3 marks in a mutant cell line showed that its binding across the genome was redistributed with specific upregulation of MYCN (225). Our group has also reported an enrichment of mutations in the *SETD2* gene coding for a H3K36 trimethyltransferase in this subgroup (226), associated with global decrease in H3K36me3 levels.

Effort should also be put on the tumors that did not harbor mutations in histone H3 or IDH1. Based on our results, they are split between the group "RTK I 'PDGFRA'" and "mesenchymal", and are much more heterogenous both based on the DNA methylation profile and their genetic features (7). To add to the complexity and heterogeneity of pGBMs, Korshunov and colleagues have recently described within that group a subset of tumors with a methylation profile similar to those of low grade gliomas or pleiomorphic xanthoastrocytomas and associated with a better prognosis (218). As such, extensive and intensive research efforts must be focused to elucidate the mechanisms by which pGBMs exist and are labelled.

The close link between histone H3 mutation type, tumor location, and differential expression of key neuronal lineage markers leads us to speculate that there may be differences in the cell of origin and/or the time of tumor development between these GBM subgroups. Although supported by the differential expression of mutant-specific gene signatures at different stages of human brain development, this remains to be formally shown. To support that hypothesis,

Funato *et al.* created a model of DIPG by using human embryonic stem cells, and remarkably K27M-H3.3 was mitogenic only in neural progenitors derived from these cells, and not in undifferentiated stem cells or astrocytes also derived from these cells (227). The next steps would be to focus on the interface between genomics and epigenomics with ChIP-Seq and RNA sequencing techniques, and to use animal models to approach the complexity of the *in vivo* environmental biology of these tumors for functional investigation. Compound library screening would also be very useful to identify efficient therapies for patients with these tumors. The first zebrafish model of K27M-H3.3 mutation has already been published, confirming the *in vitro* data of loss of H3K27 methylation and derepression of polycomb target genes (228). There is no doubt that the development of targeted genome engineering technologies using site specific programmable nucleases such as zinc finger nucleases (ZFNs), transcription activator like effector nucleases (TALENs), and cluster regulatory interspaced short palindromic repeat (CRISPR) based RNA guided DNA endonucleases will be a useful approach to generate accurate animal models of pGBMs (229, 230).

5.3 DIPG in the global picture of pediatric high-grade gliomas

These recent years have tremendously brought forward our knowledge on DIPGs with the initial identification of a copy number gain of *PDGFRA* as the main recurrent genetic alteration found in these tumors as reported in several studies (56, 57). The results presented in this dissertation confirm that *PDGFRA* gain or amplification was present in 30% of our cohort, and exclusively in the patients carrying K27M-H3.3 mutations. Moreover it is interesting to note that 5 out of the 6 tumors with *PDGFRA* amplification were already treated with radiation. This suggests that *PDGFRA* gains/amplifications are subsequent to K27M-H3.3 mutations thus providing insight into the potential evolution of mutations in these tumors. Interestingly, our data from the comprehensive analysis of both adult and pediatric

GBM did not corroborate this. Ten brainstem samples with known *PDGFRA* status and K27M-H3.3 mutation were included in this study, and *PDGFRA* amplification was not a common feature of the subgroup with K27M-H3.3 but was enriched in the subgroup RTK I which encomprised supratentorial hemispheric GBMs. This discrepancy may be explained by the high incidence of *PDGFRA* amplification in radiation-induced gliomas (55), although tumors that were analyzed after treatment were not found to show increased widespread instability (57). This only confirms that DIPGs warrant further investigation to clarify and identify key oncogenic drivers, and encourages clinicians to perform stereotactic biopsy at the time of diagnosis. The safety of upfront stereotactic biopsy of diffuse pontine lesion has been demonstrated by Rouleau *et al.* on 24 patients without any procedure-related death or long-term morbidity (231). This was confirmed with the preliminary results from the BATS (Biology and Treatment Strategies) DIPG consortium phase II clinical trial which proposes upfront biopsy and treatment stratification for newly diagnosed DIPG and in which our group has a lead role (232).

Further integrative and comprehensive studies using next generation sequencing have been published by our group and others (214-217). These four publications were published in the same issue of *Nature Genetics* and reported recurrent activating mutations in *ACVR1* a growth factor receptor gene involved in BMP signalling, thus identifying novel therapeutic targets. Our group confirmed that thalamic GBMs and DIPGs belong to the same group of tumors unified by frequent H3.3-K27M mutation and by similar gene expression (58) and that they likely have a close related origin (214). Although about 80% of DIPG have a histone 3 mutation at the aminoacid 27 and a robust similar methylation profile, Cynthia Hawkins's group identified two other subgroups: "MYCN" with MYCN amplification and "silent" with silent genome based on sequencing and copy number analysis, with a low mutation rate (217).

Further functional analysis will be needed to fully understand the pathogenicity of these mutations, using *in vitro* and *in vivo* models, as mentioned earlier in this discussion. However, we can hypothesize that the combination of histone modifying agents to RTK inhibitors or BMP inhibitors may be of therapeutic benefit in this group of patients.

5.4 Translating into patient bedside

Until now, conventional approaches to treat glioblastoma have proven to be insufficient to cure patients. With the increasing knowledge of tumor biology and the development of biochemical and pharmaceutical methods allowing efficient rational drug designs, targeted therapy has become a vast avenue for novel approaches to treat cancer. Based on this, individualized targeted therapy will likely be the next step to treat patients with pGBM. Our study presented in Chapter 4 has shown that DNA methylation profiling with the 450kHumanMethylation (450HM) is a robust, reproducible and cost-effective tool to characterize tumors within the high-grade glioma group. Beyond the research sphere, we think that it should be implemented and used in a clinical setting. Several groups are already working on this (233), including our collaborators at the Heidelberg University Hospital (234). DNA methylation profiling is not only accurate to characterize GBM, but also other types of brain tumors, like medulloblastoma and CNS-PNET. 450HM has also been shown to be a good tool to reclassify tumors that could be difficult to diagnose just based on the pathological features (234).

One other aim of our research on pGBM is to develop biomarkers that will identify some tumor characteristics useful to guide the clinician. Although it is too early for therapeutic changes in pGBM, better characterization of a tumor is critical for the patient to receive adequate treatment. For aGBM, the most cited example is the methylation of *MGMT* promoter

as a predictive and prognostic factor of tumor response to temozolomide. Although used in adults, this is another example showcasing how pediatric tumors behave differently than their adult counterpart. MGMT is most often not expressed in pGBM but is an independent factor with a worse prognosis when overexpressed (235). However, data are conflicting regarding the role of MGMT in response to temozolomide. Methylation rates vary greatly from 50% (236) to less than 15% in most series (235, 237, 238). Although there are no clear data about the role of temozolomide in pHGG, it is still one of the most used drugs in a setting of a tumor with dismal prognosis. Nowadays, next generation sequencing is an increasingly reliable and affordable technique to use in clinical practice, and *H3F3A* Sanger sequencing is already performed in accredited laboratories (CLIA and others). That being said, immunohistochemistry remains the gold standard in many hospitals for diagnosis. Therefore, our group and others have tested and validated reliable antibodies to detect mutated histone H3 on fixed tissues (219, 239). Meanwhile, in the fourth chapter of this dissertation, we proposed to use OLIG2 and FOXG1 as biomarkers to distinguish K27M and G34R/V mutants. The immunohistochemistry detection of these two proteins is already routinely performed in several centers that may not have access to genomic facilities.

Based on this new knowledge about the critical role of epigenetics dysregulation in pGBM, novel targeted therapies should be tested soon in early clinical trials. Several epigenetic modulator drugs are already available for cancer treatment in humans, like BET-bromodomain inhibitors (240). Evidence of the oncogenic role of PRC2 inhibition mediated by mutant K27M-H3 proposes it as an excellent therapeutic target in midline pGBM. However, it is important to note that activating mutations of *EZH2* have been described in B-cell lymphomas (241), and therefore one must be cautious with that pathway, as it seems that the role of PRC2 is highly context dependent. Another possible avenue to target K27M-H3

mutant tumors is to increase H3K27 methylation by inhibiting demethylases. In that direction, preclinical data have shown promising results with the inhibition of histone demethylase JMJD3 (242). Overall, the main point to consider is the careful selection of patients for these targeted therapies based on the tumor biology. Previous studies both in adult and pediatric GBM have been disappointing because of this lack of selection (76). One other lesson that can be learnt from the past is the importance of multiple targets. In most cases, single-agent testing has led to tumor resistance because of the signalling pathway redundancy and the ability of cancer cells to overcome these obstacles. Therefore, inhibiting multiple targets with synergistic combination of drugs will most likely provide more effective therapies with more durable effects (243).

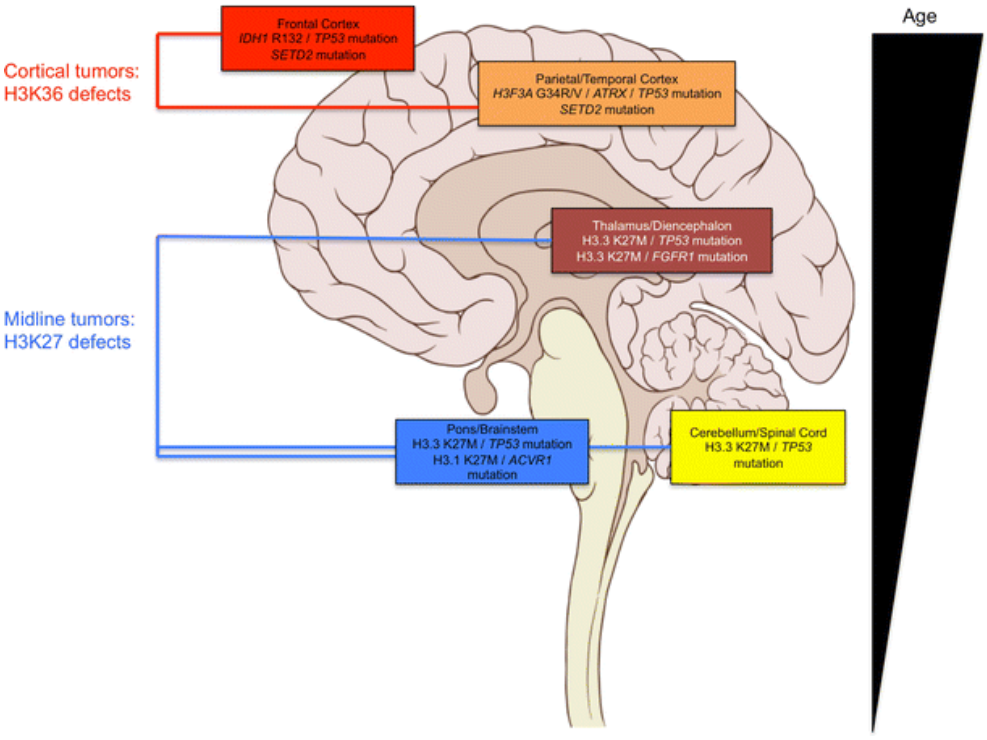
In the field of neuro-oncology, the *cas princeps* of integrative patient approach is in medulloblastoma patients. For them, GEP and DNA methylation profiling has identified four groups of tumor correlated with clinical presentation and molecular alterations. In the latest Saint-Jude protocol SJMB-12, the tumor classification has already been integrated in the treatment stratification, and patients in the SHH subgroup are treated with vismodegib, a hedgehog pathway inhibitor, during maintenance therapy (http://clinicaltrials.gov/show/NCT01878617). We hope that our current classification of pediatric and adult GBM classification will encourage the development of targeted therapies specific to each subgroup similarly. Targeted therapies directed against mutant IDH are currently in preclinical testing with various responses, but will unlikely to be sufficient by themselves (244-246). Interestingly, understanding the mechanism of action of the IDH1 enzyme has shed light into novel therapeutic avenues such as metformin which targets metabolic vulnerabilities induced by mutant IDH and therefore reduces cell growth, survival and self-renewal in mutant IDH1 cells (247).

5.5 Conclusion

In conclusion, these studies describe a number of findings that enhance our understanding of the heterogeneity of GBM and DIPG, as well as shed light on potential cellular origins and oncogenic pathways leading to gliomagenesis. We have identified potential prognostic biomarkers, which may be further exploited for molecular diagnostic purposes, and also provide a focus for future work at a basic and translational/targeted therapeutic level, particularly in a pediatric and young adult setting.

5.6 Figure

Figure 5.1 – Molecular subgroups of pediatric high-grade gliomas show neuroanatomical preferences



Schematic representation of a sagittal view of the human brain depicting neuroanatomical areas with observed alterations discussed herein. Age of patients harbouring these alterations is represented at the right. Adapted with permission from (125).

Bibliography

1. Schwartzentruber J, Korshunov A, Liu XY, Jones DT, Pfaff E, Jacob K, et al. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature. 2012;482(7384):226-31.

2. Khuong-Quang DA, Buczkowicz P, Rakopoulos P, Liu XY, Fontebasso AM, Bouffet E, et al. K27M mutation in histone H3.3 defines clinically and biologically distinct subgroups of pediatric diffuse intrinsic pontine gliomas. Acta Neuropathol. 2012;124(3):439-47.

3. Canada NCIo. Canadian Cancer Statistics 2008. Toronto, Canada2008.

4. Pollack IF. Pediatric brain tumors. Semin Surg Oncol. 1999;16(2):73-90.

5. Louis DN, Ohgaki H, Wiestler OD, Cavenee WK, Burger PC, Jouvet A, et al. The 2007 WHO classification of tumours of the central nervous system. Acta Neuropathol. 2007;114(2):97-109.

6. Kieran MW, Walker D, Frappaz D, Prados M. Brain tumors: from childhood through adolescence into adulthood. J Clin Oncol. 2010;28(32):4783-9.

7. Sturm D, Witt H, Hovestadt V, Khuong-Quang DA, Jones DT, Konermann C, et al. Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. Cancer Cell. 2012;22(4):425-37.

8. Maher EA, Furnari FB, Bachoo RM, Rowitch DH, Louis DN, Cavenee WK, et al. Malignant glioma: genetics and biology of a grave matter. Genes Dev. 2001;15(11):1311-33.

9. Ohgaki H, Kleihues P. Population-based studies on incidence, survival rates, and genetic alterations in astrocytic and oligodendroglial gliomas. Journal of neuropathology and experimental neurology. 2005;64(6):479-89.

10. Jones C, Perryman L, Hargrave D. Paediatric and adult malignant glioma: close relatives or distant cousins? Nature Reviews Clinical Oncology. 2012;9(7):400-13.

11. Ostrom QT, Gittleman H, Liao P, Rouse C, Chen Y, Dowling J, et al. CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2007-2011. Neuro Oncol. 2014;16 Suppl 4:iv1-63.

12. Ostrom QT, Gittleman H, Farah P, Ondracek A, Chen Y, Wolinsky Y, et al. CBTRUS statistical report: Primary brain and central nervous system tumors diagnosed in the United States in 2006-2010. Neuro Oncol. 2013;15 Suppl 2:ii1-56.

13. 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. 2012;14 Suppl 5:v1-49.

14. Kaderali Z, Lamberti-Pasculli M, Rutka JT. The changing epidemiology of paediatric brain tumors: a review from the Hospital for Sick Children. Childs Nerv Syst. 2009;25(7):787-93.

15. Rood BR, MacDonald TJ. Pediatric high-grade glioma: molecular genetic clues for innovative therapeutic approaches. J Neurooncol. 2005;75(3):267-72.

16. 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.

17. 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.

18. Friedman HS, Prados MD, Wen PY, Mikkelsen T, Schiff D, Abrey LE, et al. Bevacizumab alone and in combination with irinotecan in recurrent glioblastoma. J Clin Oncol. 2009;27(28):4733-40.

19. Kebudi R, Cakir FB. Management of diffuse pontine gliomas in children: recent developments. Paediatr Drugs. 2013;15(5):351-62.

20. Buczkowicz P, Bartels U, Bouffet E, Becher O, Hawkins C. Histopathological spectrum of paediatric diffuse intrinsic pontine glioma: diagnostic and therapeutic implications. Acta Neuropathol. 2014;128(4):573-81.

21. Schroeder KM, Hoeman CM, Becher OJ. Children are not just little adults: recent advances in understanding of diffuse intrinsic pontine glioma biology. Pediatr Res. 2014;75(1-2):205-9.

22. Walker DA, Liu J, Kieran M, Jabado N, Picton S, Packer R, et al. A multi-disciplinary consensus statement concerning surgical approaches to low-grade, high-grade astrocytomas and diffuse intrinsic pontine gliomas in childhood (CPN Paris 2011) using the Delphi method. Neuro Oncol. 2013;15(4):462-8.

23. Cancer Genome Atlas Research N. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature. 2008;455(7216):1061-8.

24. Brennan CW, Verhaak RG, McKenna A, Campos B, Noushmehr H, Salama SR, et al. The somatic genomic landscape of glioblastoma. Cell. 2013;155(2):462-77.

25. Verhaak RG, 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.

26. Phillips HS, Kharbanda S, Chen R, Forrest WF, Soriano RH, Wu TD, 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(3):157-73.

27. 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.

28. 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.

29. Watanabe K, Tachibana O, Sata K, Yonekawa Y, Kleihues P, Ohgaki H. Overexpression of the EGF receptor and p53 mutations are mutually exclusive in the evolution of primary and secondary glioblastomas. Brain Pathol. 1996;6(3):217-23; discussion 23-4.

30. Kleihues P, Ohgaki H. Primary and secondary glioblastomas: from concept to clinical diagnosis. Neuro Oncol. 1999;1(1):44-51.

31. 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.

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. Majewski J, Schwartzentruber J, Lalonde E, Montpetit A, Jabado N. What can exome sequencing do for you? J Med Genet. 2011;48(9):580-9.

34. Mardis ER, Ding L, Dooling DJ, Larson DE, McLellan MD, Chen K, et al. Recurring mutations found by sequencing an acute myeloid leukemia genome. N Engl J Med. 2009;361(11):1058-66.

35. 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.

36. Zhao S, Lin Y, Xu W, Jiang W, Zha Z, Wang P, et al. Glioma-derived mutations in IDH1 dominantly inhibit IDH1 catalytic activity and induce HIF-1alpha. Science. 2009;324(5924):261-5.

37. Aghili M, Zahedi F, Rafiee E. Hydroxyglutaric aciduria and malignant brain tumor: a case report and literature review. J Neurooncol. 2009;91(2):233-6.

38. Turcan S, Rohle D, Goenka A, Walsh LA, Fang F, Yilmaz E, et al. IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. Nature. 2012;483(7390):479-83.

39. Xu W, Yang H, Liu Y, Yang Y, Wang P, Kim SH, et al. Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. Cancer Cell. 2011;19(1):17-30.

40. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science. 2009;324(5929):930-5.

41. Delhommeau F, Dupont S, Della Valle V, James C, Trannoy S, Masse A, et al. Mutation in TET2 in myeloid cancers. N Engl J Med. 2009;360(22):2289-301.

42. Broniscer A, Gajjar A. Supratentorial high-grade astrocytoma and diffuse brainstem glioma: two challenges for the pediatric oncologist. Oncologist. 2004;9(2):197-206.

43. Pollack IF, Finkelstein SD, Burnham J, Holmes EJ, Hamilton RL, Yates AJ, et al. Age and TP53 mutation frequency in childhood malignant gliomas: results in a multi-institutional cohort. Cancer Res. 2001;61(20):7404-7.

44. Pollack IF, Finkelstein SD, Woods J, Burnham J, Holmes EJ, Hamilton RL, et al. Expression of p53 and prognosis in children with malignant gliomas. N Engl J Med. 2002;346(6):420-7.

45. Bredel M, Pollack IF, Hamilton RL, James CD. Epidermal growth factor receptor expression and gene amplification in high-grade non-brainstem gliomas of childhood. Clin Cancer Res. 1999;5(7):1786-92.

46. Heimberger AB, Hlatky R, Suki D, Yang D, Weinberg J, Gilbert M, et al. Prognostic effect of epidermal growth factor receptor and EGFRvIII in glioblastoma multiforme patients. Clin Cancer Res. 2005;11(4):1462-6.

47. Liang ML, Ma J, Ho M, Solomon L, Bouffet E, Rutka JT, et al. Tyrosine kinase expression in pediatric high-grade astrocytoma. J Neurooncol. 2008;87(3):247-53.

48. Bigner SH, Humphrey PA, Wong AJ, Vogelstein B, Mark J, Friedman HS, et al. Characterization of the epidermal growth factor receptor in human glioma cell lines and xenografts. Cancer Res. 1990;50(24):8017-22.

49. Pedersen MW, Tkach V, Pedersen N, Berezin V, Poulsen HS. Expression of a naturally occurring constitutively active variant of the epidermal growth factor receptor in mouse fibroblasts increases motility. International journal of cancer Journal international du cancer. 2004;108(5):643-53.

50. Moscatello DK, Holgado-Madruga M, Emlet DR, Montgomery RB, Wong AJ. Constitutive activation of phosphatidylinositol 3-kinase by a naturally occurring mutant epidermal growth factor receptor. J Biol Chem. 1998;273(1):200-6.

51. Faury D, Nantel A, Dunn SE, Guiot MC, Haque T, Hauser P, et al. Molecular profiling identifies prognostic subgroups of pediatric glioblastoma and shows increased YB-1 expression in tumors. J Clin Oncol. 2007;25(10):1196-208.

52. Haque T, Faury D, Albrecht S, Lopez-Aguilar E, Hauser P, Garami M, et al. Gene expression profiling from formalin-fixed paraffin-embedded tumors of pediatric glioblastoma. Clin Cancer Res. 2007;13(21):6284-92.

53. Bax DA, Mackay A, Little SE, Carvalho D, Viana-Pereira M, Tamber N, et al. A distinct spectrum of copy number aberrations in pediatric high-grade gliomas. Clin Cancer Res. 2010;16(13):3368-77.

54. Qu HQ, Jacob K, Fatet S, Ge B, Barnett D, Delattre O, et al. Genome-wide profiling using single-nucleotide polymorphism arrays identifies novel chromosomal imbalances in pediatric glioblastomas. Neuro Oncol. 2010;12(2):153-63.

55. Paugh BS, Qu C, Jones C, Liu Z, Adamowicz-Brice M, Zhang J, et al. Integrated molecular genetic profiling of pediatric high-grade gliomas reveals key differences with the adult disease. J Clin Oncol. 2010;28(18):3061-8.

56. Zarghooni M, Bartels U, Lee E, Buczkowicz P, Morrison A, Huang A, et al. Whole-genome profiling of pediatric diffuse intrinsic pontine gliomas highlights platelet-derived growth factor receptor alpha and poly (ADP-ribose) polymerase as potential therapeutic targets. J Clin Oncol. 2010;28(8):1337-44.

57. Paugh BS, Broniscer A, Qu C, Miller CP, Zhang J, Tatevossian RG, et al. Genome-wide analyses identify recurrent amplifications of receptor tyrosine kinases and cell-cycle regulatory genes in diffuse intrinsic pontine glioma. J Clin Oncol. 2011;29(30):3999-4006.

58. Puget S, Philippe C, Bax DA, Job B, Varlet P, Junier MP, et al. Mesenchymal transition and PDGFRA amplification/mutation are key distinct oncogenic events in pediatric diffuse intrinsic pontine gliomas. PloS one. 2012;7(2):e30313.

59. Heaphy CM, de Wilde RF, Jiao Y, Klein AP, Edil BH, Shi C, et al. Altered telomeres in tumors with ATRX and DAXX mutations. Science. 2011;333(6041):425.

60. Wu G, Broniscer A, McEachron TA, Lu C, Paugh BS, Becksfort J, et al. Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. Nature genetics. 2012;44(3):251-3.

61. Behjati S, Tarpey PS, Presneau N, Scheipl S, Pillay N, Van Loo P, et al. Distinct H3F3A and H3F3B driver mutations define chondroblastoma and giant cell tumor of bone. Nature genetics. 2013;45(12):1479-82.

62. Kallappagoudar S, Yadav RK, Lowe BR, Partridge JF. Histone H3 mutations-a special role for H3.3 in tumorigenesis? Chromosoma. 2015;124(2):177-89.

63. Filipescu D, Szenker E, Almouzni G. Developmental roles of histone H3 variants and their chaperones. Trends Genet. 2013;29(11):630-40.

64. Ullrich A, Schlessinger J. Signal transduction by receptors with tyrosine kinase activity. Cell. 1990;61(2):203-12.

65. Lemmon MA, Schlessinger J. Cell signaling by receptor tyrosine kinases. Cell. 2010;141(7):1117-34.

66. Parachoniak CA, Park M. Dynamics of receptor trafficking in tumorigenicity. Trends in cell biology. 2012;22(5):231-40.

67. Marmor MD, Yarden Y. Role of protein ubiquitylation in regulating endocytosis of receptor tyrosine kinases. Oncogene. 2004;23(11):2057-70.

68. Abella JV, Park M. Breakdown of endocytosis in the oncogenic activation of receptor tyrosine kinases. Am J Physiol Endocrinol Metab. 2009;296(5):E973-84.

69. Mosesson Y, Mills GB, Yarden Y. Derailed endocytosis: an emerging feature of cancer. Nat Rev Cancer. 2008;8(11):835-50.

70. Wells A, Welsh JB, Lazar CS, Wiley HS, Gill GN, Rosenfeld MG. Ligand-induced transformation by a noninternalizing epidermal growth factor receptor. Science. 1990;247(4945):962-4.

71. Wiley HS, Burke PM. Regulation of receptor tyrosine kinase signaling by endocytic trafficking. Traffic. 2001;2(1):12-8.

72. Vieira AV, Lamaze C, Schmid SL. Control of EGF receptor signaling by clathrin-mediated endocytosis. Science. 1996;274(5295):2086-9.

73. Cohen S. Isolation of a mouse submaxillary gland protein accelerating incisor eruption and eyelid opening in the new-born animal. J Biol Chem. 1962;237:1555-62.

74. Furnari FB, Fenton T, Bachoo RM, Mukasa A, Stommel JM, Stegh A, et al. Malignant astrocytic glioma: genetics, biology, and paths to treatment. Genes Dev. 2007;21(21):2683-710.

75. Humphrey PA, Wong AJ, Vogelstein B, Zalutsky MR, Fuller GN, Archer GE, et al. Anti-synthetic peptide antibody reacting at the fusion junction of deletion-mutant epidermal growth factor receptors in human glioblastoma. Proc Natl Acad Sci U S A. 1990;87(11):4207-11.

76. Lau D, Magill ST, Aghi MK. Molecularly targeted therapies for recurrent glioblastoma: current and future targets. Neurosurgical focus. 2014;37(6):E15.

77. Hegi ME, Diserens AC, Bady P, Kamoshima Y, Kouwenhoven MC, Delorenzi M, et al. Pathway analysis of glioblastoma tissue after preoperative treatment with the EGFR tyrosine kinase inhibitor gefitinib--a phase II trial. Molecular cancer therapeutics. 2011;10(6):1102-12.

78. Kong DS, Song SY, Kim DH, Joo KM, Yoo JS, Koh JS, et al. Prognostic significance of c-Met expression in glioblastomas. Cancer. 2009;115(1):140-8.

79. Platta HW, Stenmark H. Endocytosis and signaling. Curr Opin Cell Biol. 2011;23(4):393-403.

80. Le Roy C, Wrana JL. Clathrin- and non-clathrin-mediated endocytic regulation of cell signalling. Nat Rev Mol Cell Biol. 2005;6(2):112-26.

81. Bonifacino JS, Lippincott-Schwartz J. Coat proteins: shaping membrane transport. Nat Rev Mol Cell Biol. 2003;4(5):409-14.

82. Boucrot E, Ferreira AP, Almeida-Souza L, Debard S, Vallis Y, Howard G, et al. Endophilin marks and controls a clathrin-independent endocytic pathway. Nature. 2015;517(7535):460-5.

83. Parton RG, Simons K. The multiple faces of caveolae. Nat Rev Mol Cell Biol. 2007;8(3):185-94.

84. Sigismund S, Argenzio E, Tosoni D, Cavallaro E, Polo S, Di Fiore PP. Clathrin-mediated internalization is essential for sustained EGFR signaling but dispensable for degradation. Dev Cell. 2008;15(2):209-19.

85. Schmidt-Glenewinkel H, Vacheva I, Hoeller D, Dikic I, Eils R. An ultrasensitive sorting mechanism for EGF receptor endocytosis. BMC Syst Biol. 2008;2:32.

86. Ibanez CF. Message in a bottle: long-range retrograde signaling in the nervous system. Trends in cell biology. 2007;17(11):519-28.

87. Howe CL, Valletta JS, Rusnak AS, Mobley WC. NGF signaling from clathrin-coated vesicles: evidence that signaling endosomes serve as a platform for the Ras-MAPK pathway. Neuron. 2001;32(5):801-14.

88. Pennock S, Wang Z. Stimulation of cell proliferation by endosomal epidermal growth factor receptor as revealed through two distinct phases of signaling. Mol Cell Biol. 2003;23(16):5803-15.

89. Oksvold MP, Skarpen E, Wierod L, Paulsen RE, Huitfeldt HS. Re-localization of activated EGF receptor and its signal transducers to multivesicular compartments downstream of early endosomes in response to EGF. Eur J Cell Biol. 2001;80(4):285-94.

90. Grandal MV, Zandi R, Pedersen MW, Willumsen BM, van Deurs B, Poulsen HS. EGFRvIII escapes down-regulation due to impaired internalization and sorting to lysosomes. Carcinogenesis. 2007;28(7):1408-17.

91. Parachoniak CA, Luo Y, Abella JV, Keen JH, Park M. GGA3 functions as a switch to promote Met receptor recycling, essential for sustained ERK and cell migration. Dev Cell. 2011;20(6):751-63.

92. Carlton J, Bujny M, Rutherford A, Cullen P. Sorting nexins--unifying trends and new perspectives. Traffic. 2005;6(2):75-82.

93. Cullen PJ. Endosomal sorting and signalling: an emerging role for sorting nexins. Nat Rev Mol Cell Biol. 2008;9(7):574-82.

94. Worby CA, Dixon JE. Sorting out the cellular functions of sorting nexins. Nat Rev Mol Cell Biol. 2002;3(12):919-31.

95. Nguyen LN, Holdren MS, Nguyen AP, Furuya MH, Bianchini M, Levy E, et al. Sorting nexin 1 down-regulation promotes colon tumorigenesis. Clin Cancer Res. 2006;12(23):6952-9.

96. Fuchs U, Rehkamp G, Haas OA, Slany R, Konig M, Bojesen S, et al. The human formin-binding protein 17 (FBP17) interacts with sorting nexin, SNX2, and is an MLL-fusion partner in acute myelogeneous leukemia. Proc Natl Acad Sci U S A. 2001;98(15):8756-61.

97. Huang Z, Huang S, Wang Q, Liang L, Ni S, Wang L, et al. MicroRNA-95 Promotes Cell Proliferation and Targets Sorting Nexin 1 in Human Colorectal Carcinoma. Cancer Res. 2011;71(7):2582-9.

98. Kurten RC, Cadena DL, Gill GN. Enhanced degradation of EGF receptors by a sorting nexin, SNX1. Science. 1996;272(5264):1008-10.

99. Haft CR, de la Luz Sierra M, Barr VA, Haft DH, Taylor SI. Identification of a family of sorting nexin molecules and characterization of their association with receptors. Mol Cell Biol. 1998;18(12):7278-87.

100. Liu H, Liu ZQ, Chen CX, Magill S, Jiang Y, Liu YJ. Inhibitory regulation of EGF receptor degradation by sorting nexin 5. Biochem Biophys Res Commun. 2006;342(2):537-46.

101. Zheng B, Ma YC, Ostrom RS, Lavoie C, Gill GN, Insel PA, et al. RGS-PX1, a GAP for GalphaS and sorting nexin in vesicular trafficking. Science. 2001;294(5548):1939-42.

102. Zheng B, Tang T, Tang N, Kudlicka K, Ohtsubo K, Ma P, et al. Essential role of RGS-PX1/sorting nexin 13 in mouse development and regulation of endocytosis dynamics. Proc Natl Acad Sci U S A. 2006;103(45):16776-81.

103. Ernst T, Score J, Deininger M, Hidalgo-Curtis C, Lackie P, Ershler WB, et al. Identification of FOXP1 and SNX2 as novel ABL1 fusion partners in acute lymphoblastic leukaemia. Br J Haematol. 2011;153(1):43-6.

104. Sasaki T, Sasaki J, Sakai T, Takasuga S, Suzuki A. The physiology of phosphoinositides. Biol Pharm Bull. 2007;30(9):1599-604.

105. Panaretou C, Domin J, Cockcroft S, Waterfield MD. Characterization of p150, an adaptor protein for the human phosphatidylinositol (PtdIns) 3-kinase. Substrate presentation by phosphatidylinositol transfer protein to the p150.Ptdins 3-kinase complex. J Biol Chem. 1997;272(4):2477-85.

106. Foster FM, Traer CJ, Abraham SM, Fry MJ. The phosphoinositide (PI) 3-kinase family. J Cell Sci. 2003;116(Pt 15):3037-40.

107. Vanhaesebroeck B, Guillermet-Guibert J, Graupera M, Bilanges B. The emerging mechanisms of isoform-specific PI3K signalling. Nat Rev Mol Cell Biol. 2010;11(5):329-41.

108. Xu Y, Hortsman H, Seet L, Wong SH, Hong W. SNX3 regulates endosomal function through its PX-domain-mediated interaction with PtdIns(3)P. Nat Cell Biol. 2001;3(7):658-66.

109. Seet LF, Hong W. The Phox (PX) domain proteins and membrane traffic. Biochim Biophys Acta. 2006;1761(8):878-96.

110. Pons V, Luyet PP, Morel E, Abrami L, van der Goot FG, Parton RG, et al. Hrs and SNX3 functions in sorting and membrane invagination within multivesicular bodies. PLoS Biol. 2008;6(9):e214.

111. Gullapalli A, Garrett TA, Paing MM, Griffin CT, Yang Y, Trejo J. A role for sorting nexin 2 in epidermal growth factor receptor down-regulation: evidence for distinct functions of sorting nexin 1 and 2 in protein trafficking. Mol Biol Cell. 2004;15(5):2143-55.

112. Cavet ME, Pang J, Yin G, Berk BC. An epidermal growth factor (EGF) -dependent interaction between GIT1 and sorting nexin 6 promotes degradation of the EGF receptor. Faseb J. 2008;22(10):3607-16.

113. Mizutani R, Yamauchi J, Kusakawa S, Nakamura K, Sanbe A, Torii T, et al. Sorting nexin 3, a protein upregulated by lithium, contains a novel phosphatidylinositol-binding

sequence and mediates neurite outgrowth in N1E-115 cells. Cell Signal. 2009;21(11):1586-94.

114. Braun V, Wong A, Landekic M, Hong WJ, Grinstein S, Brumell JH. Sorting nexin 3 (SNX3) is a component of a tubular endosomal network induced by Salmonella and involved in maturation of the Salmonella-containing vacuole. Cell Microbiol. 2010.

115. Joshi AD, Loilome W, Siu IM, Tyler B, Gallia GL, Riggins GJ. Evaluation of tyrosine kinase inhibitor combinations for glioblastoma therapy. PloS one. 2012;7(10):e44372.

116. Sturm D, Bender S, Jones DT, Lichter P, Grill J, Becher O, et al. Paediatric and adult glioblastoma: multiform (epi)genomic culprits emerge. Nat Rev Cancer. 2014;14(2):92-107.

117. Lu C, Ward PS, Kapoor GS, Rohle D, Turcan S, Abdel-Wahab O, et al. IDH mutation impairs histone demethylation and results in a block to cell differentiation. Nature. 2012;483(7390):474-8.

118. McCabe MT, Brandes JC, Vertino PM. Cancer DNA methylation: molecular mechanisms and clinical implications. Clin Cancer Res. 2009;15(12):3927-37.

119. Statham AL, Robinson MD, Song JZ, Coolen MW, Stirzaker C, Clark SJ. Bisulfite sequencing of chromatin immunoprecipitated DNA (BisChIP-seq) directly informs methylation status of histone-modified DNA. Genome research. 2012;22(6):1120-7.

120. Brinkman AB, Gu H, Bartels SJ, Zhang Y, Matarese F, Simmer F, et al. Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. Genome research. 2012;22(6):1128-38.

121. Thorarinsdottir HK, Santi M, McCarter R, Rushing EJ, Cornelison R, Jales A, et al. Protein expression of platelet-derived growth factor receptor correlates with malignant histology and PTEN with survival in childhood gliomas. Clin Cancer Res. 2008;14(11):3386-94.

122. Jakacki RI, Hamilton M, Gilbertson RJ, Blaney SM, Tersak J, Krailo MD, et al. Pediatric phase I and pharmacokinetic study of erlotinib followed by the combination of erlotinib and temozolomide: a Children's Oncology Group Phase I Consortium Study. J Clin Oncol. 2008;26(30):4921-7.

123. Freeman BB, 3rd, Daw NC, Geyer JR, Furman WL, Stewart CF. Evaluation of gefitinib for treatment of refractory solid tumors and central nervous system malignancies in pediatric patients. Cancer Invest. 2006;24(3):310-7.

124. Stommel JM, Kimmelman AC, Ying H, Nabioullin R, Ponugoti AH, Wiedemeyer R, et al. Coactivation of receptor tyrosine kinases affects the response of tumor cells to targeted therapies. Science. 2007;318(5848):287-90.

125. Fontebasso AM, Gayden T, Nikbakht H, Neirinck M, Papillon-Cavanagh S, Majewski J, et al. Epigenetic dysregulation: a novel pathway of oncogenesis in pediatric brain tumors. Acta Neuropathol. 2014;128(5):615-27.

126. Bax DA, Gaspar N, Little SE, Marshall L, Perryman L, Regairaz M, et al. EGFRvIII deletion mutations in pediatric high-grade glioma and response to targeted therapy in pediatric glioma cell lines. Clin Cancer Res. 2009;15(18):5753-61.

127. Moriyama T, Kataoka H, Kawano H, Yokogami K, Nakano S, Goya T, et al. Comparative analysis of expression of hepatocyte growth factor and its receptor, c-met, in gliomas, meningiomas and schwannomas in humans. Cancer letters. 1998;124(2):149-55.

128. Ebner R, Derynck R. Epidermal growth factor and transforming growth factor-alpha: differential intracellular routing and processing of ligand-receptor complexes. Cell regulation. 1991;2(8):599-612.

129. Sorkin A, Waters CM. Endocytosis of growth factor receptors. BioEssays : news and reviews in molecular, cellular and developmental biology. 1993;15(6):375-82.

130. Danson C, Brown E, Hemmings OJ, McGough IJ, Yarwood S, Heesom KJ, et al. SNX15 links clathrin endocytosis to the PtdIns3P early endosome independently of the APPL1 endosome. J Cell Sci. 2013;126(Pt 21):4885-99.

131. Schreiber AB, Winkler ME, Derynck R. Transforming growth factor-alpha: a more potent angiogenic mediator than epidermal growth factor. Science. 1986;232(4755):1250-3.

132. Chiow KH, Tan Y, Chua RY, Huang D, Ng ML, Torta F, et al. SNX3-dependent regulation of epidermal growth factor receptor (EGFR) trafficking and degradation by aspirin in epidermoid carcinoma (A-431) cells. Cellular and molecular life sciences : CMLS. 2012;69(9):1505-21.

133. Jo M, Stolz DB, Esplen JE, Dorko K, Michalopoulos GK, Strom SC. Cross-talk between epidermal growth factor receptor and c-Met signal pathways in transformed cells. J Biol Chem. 2000;275(12):8806-11.

134. Velpula KK, Dasari VR, Asuthkar S, Gorantla B, Tsung AJ. EGFR and c-Met Cross Talk in Glioblastoma and Its Regulation by Human Cord Blood Stem Cells. Translational oncology. 2012;5(5):379-92.

135. Matsuda K, Sato A, Okada M, Shibuya K, Seino S, Suzuki K, et al. Targeting JNK for therapeutic depletion of stem-like glioblastoma cells. Scientific reports. 2012;2:516.

136. Vo VA, Lee JW, Lee HJ, Chun W, Lim SY, Kim SS. Inhibition of JNK potentiates temozolomide-induced cytotoxicity in U87MG glioblastoma cells via suppression of Akt phosphorylation. Anticancer research. 2014;34(10):5509-15.

137. Okada M, Sato A, Shibuya K, Watanabe E, Seino S, Suzuki S, et al. JNK contributes to temozolomide resistance of stem-like glioblastoma cells via regulation of MGMT expression. International journal of oncology. 2014;44(2):591-9.

138. Paugh BS, Zhu X, Qu C, Endersby R, Diaz AK, Zhang J, et al. Novel oncogenic PDGFRA mutations in pediatric high-grade gliomas. Cancer Res. 2013;73(20):6219-29.

139. Rivera J, Megias D, Bravo J. Sorting nexin 6 interacts with breast cancer metastasis suppressor-1 and promotes transcriptional repression. J Cell Biochem.111(6):1464-72.

140. Ogi S, Fujita H, Kashihara M, Yamamoto C, Sonoda K, Okamoto I, et al. Sorting nexin 2-mediated membrane trafficking of c-Met contributes to sensitivity of molecular-targeted drugs. Cancer science. 2013;104(5):573-83.

141. Tomita O, Iijima K, Ishibashi T, Osumi T, Kobayashi K, Okita H, et al. Sensitivity of SNX2-ABL1 toward tyrosine kinase inhibitors distinct from that of BCR-ABL1. Leukemia research. 2014;38(3):361-70.

142. Nishimura Y, Takiguchi S, Ito S, Itoh K. EGFstimulated AKT activation is mediated by EGFR recycling via an early endocytic pathway in a gefitinibresistant human lung cancer cell line. International journal of oncology. 2015.

143. Nishimura Y, Takiguchi S, Ito S, Itoh K. Evidence that depletion of the sorting nexin 1 by siRNA promotes HGF-induced MET endocytosis and MET phosphorylation in a gefitinib-resistant human lung cancer cell line. International journal of oncology. 2014;44(2):412-26.

144. Ohashi K, Maruvka YE, Michor F, Pao W. Epidermal growth factor receptor tyrosine kinase inhibitor-resistant disease. J Clin Oncol. 2013;31(8):1070-80.

145. Uhrbom L, Hesselager G, Nister M, Westermark B. Induction of brain tumors in mice using a recombinant platelet-derived growth factor B-chain retrovirus. Cancer Res. 1998;58(23):5275-9.

146. Dai C, Celestino JC, Okada Y, Louis DN, Fuller GN, Holland EC. PDGF autocrine stimulation dedifferentiates cultured astrocytes and induces oligodendrogliomas and oligoastrocytomas from neural progenitors and astrocytes in vivo. Genes Dev. 2001;15(15):1913-25.

147. De Donatis A, Comito G, Buricchi F, Vinci MC, Parenti A, Caselli A, et al. Proliferation Versus Migration in Platelet-derived Growth Factor Signaling: THE KEY ROLE OF ENDOCYTOSIS. Journal of Biological Chemistry. 2008;283(29):19948-56.

148. Consortium EP. The ENCODE (ENCyclopedia Of DNA Elements) Project. Science. 2004;306(5696):636-40.

149. Mizutani R, Nakamura K, Yokoyama S, Sanbe A, Kusakawa S, Miyamoto Y, et al. Developmental expression of sorting nexin 3 in the mouse central nervous system. Gene Expr Patterns. 2011;11(1-2):33-40.

150. Harterink M, Port F, Lorenowicz MJ, McGough IJ, Silhankova M, Betist MC, et al. A SNX3-dependent retromer pathway mediates retrograde transport of the Wnt sorting receptor Wntless and is required for Wnt secretion. Nat Cell Biol. 2011;13(8):914-23.

151. Zhang P, Wu Y, Belenkaya TY, Lin X. SNX3 controls Wingless/Wnt secretion through regulating retromer-dependent recycling of Wntless. Cell research. 2011;21(12):1677-90.

152. Cullen PJ, Korswagen HC. Sorting nexins provide diversity for retromer-dependent trafficking events. Nat Cell Biol. 2012;14(1):29-37.

153. Gallon M, Cullen PJ. Retromer and sorting nexins in endosomal sorting. Biochemical Society transactions. 2015;43(1):33-47.

154. Awad AJ, Burns TC, Zhang Y, Abounader R. Targeting MET for glioma therapy. Neurosurgical focus. 2014;37(6):E10.

155. Joffre C, Barrow R, Menard L, Calleja V, Hart IR, Kermorgant S. A direct role for Met endocytosis in tumorigenesis. Nat Cell Biol. 2011;13(7):827-37.

156. Mellman I, Yarden Y. Endocytosis and cancer. Cold Spring Harbor perspectives in biology. 2013;5(12):a016949.

157. Tomas A, Futter CE, Eden ER. EGF receptor trafficking: consequences for signaling and cancer. Trends in cell biology. 2013.

158. Gilbertson RJ, Hill DA, Hernan R, Kocak M, Geyer R, Olson J, et al. ERBB1 is amplified and overexpressed in high-grade diffusely infiltrative pediatric brain stem glioma. Clin Cancer Res. 2003;9(10 Pt 1):3620-4.

159. Khuong-Quang DA, Gerges N, Jabado N. [Mutations in histone H3.3 and chromatin remodeling genes drive pediatric and young adult glioblastomas]. Medecine sciences : M/S. 2012;28(10):809-12.

160. Saran F. Recent advances in paediatric neuro-oncology. Curr Opin Neurol. 2002;15(6):671-7.

161. Merchant TE, Pollack IF, Loeffler JS. Brain tumors across the age spectrum: biology, therapy, and late effects. Semin Radiat Oncol.20(1):58-66.

162. Bouffet E, Tabori U, Huang A, Bartels U. Possibilities of new therapeutic strategies in brain tumors. Cancer treatment reviews.36(4):335-41.

163. Rutka JT, Kuo JS, Carter M, Ray A, Ueda S, Mainprize TG. Advances in the treatment of pediatric brain tumors. Expert Rev Neurother. 2004;4(5):879-93.

164. Freeman CR, Perilongo G. Chemotherapy for brain stem gliomas. Childs Nerv Syst. 1999;15(10):545-53.

165. Maria BL, Rehder K, Eskin TA, Hamed LM, Fennell EB, Quisling RG, et al. Brainstem glioma: I. Pathology, clinical features, and therapy. J Child Neurol. 1993;8(2):112-28.

166. Gajjar A, Sanford RA, Heideman R, Jenkins JJ, Walter A, Li Y, et al. Low-grade astrocytoma: a decade of experience at St. Jude Children's Research Hospital. J Clin Oncol. 1997;15(8):2792-9.

167. Fisher PG, Tihan T, Goldthwaite PT, Wharam MD, Carson BS, Weingart JD, et al. Outcome analysis of childhood low-grade astrocytomas. Pediatric blood & cancer. 2008;51(2):245-50.

168. Broniscer A, Baker SJ, West AN, Fraser MM, Proko E, Kocak M, et al. Clinical and molecular characteristics of malignant transformation of low-grade glioma in children. J Clin Oncol. 2007;25(6):682-9.

169. Bartels U, Hawkins C, Vezina G, Kun L, Souweidane M, Bouffet E. Proceedings of the diffuse intrinsic pontine glioma (DIPG) Toronto Think Tank: advancing basic and translational research and cooperation in DIPG. J Neurooncol.105(1):119-25.

170. Jansen MH, van Vuurden DG, Vandertop WP, Kaspers GJ. Diffuse intrinsic pontine gliomas: a systematic update on clinical trials and biology. Cancer treatment reviews.38(1):27-35.

171. Dhayalan A, Tamas R, Bock I, Tattermusch A, Dimitrova E, Kudithipudi S, et al. The ATRX-ADD domain binds to H3 tail peptides and reads the combined methylation state of K4 and K9. Hum Mol Genet.20(11):2195-203.

172. Lewis PW, Elsaesser SJ, Noh KM, Stadler SC, Allis CD. Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. Proc Natl Acad Sci U S A.107(32):14075-80.

173. Shlien A, Tabori U, Marshall CR, Pienkowska M, Feuk L, Novokmet A, et al. Excessive genomic DNA copy number variation in the Li-Fraumeni cancer predisposition syndrome. Proc Natl Acad Sci U S A. 2008;105(32):11264-9.

174. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature. 2008;454(7205):766-70.

175. Reynolds N, Salmon-Divon M, Dvinge H, Hynes-Allen A, Balasooriya G, Leaford D, et al. NuRD-mediated deacetylation of H3K27 facilitates recruitment of Polycomb Repressive Complex 2 to direct gene repression. The EMBO journal.31(3):593-605.

176. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell. 2006;125(2):315-26.

177. Bueno C, Montes R, Catalina P, Rodriguez R, Menendez P. Insights into the cellular origin and etiology of the infant pro-B acute lymphoblastic leukemia with MLL-AF4 rearrangement. Leukemia.25(3):400-10.

178. Carramusa L, Contino F, Ferro A, Minafra L, Perconti G, Giallongo A, et al. The PVT-1 oncogene is a Myc protein target that is overexpressed in transformed cells. J Cell Physiol. 2007;213(2):511-8.

179. Meyer KB, Maia AT, O'Reilly M, Ghoussaini M, Prathalingam R, Porter-Gill P, et al. A functional variant at a prostate cancer predisposition locus at 8q24 is associated with PVT1 expression. PLoS Genet.7(7):e1002165.

180. Schiffman JD, Hodgson JG, VandenBerg SR, Flaherty P, Polley MY, Yu M, et al. Oncogenic BRAF mutation with CDKN2A inactivation is characteristic of a subset of pediatric malignant astrocytomas. Cancer Res. 2010;70(2):512-9.

181. Antonelli M, Buttarelli FR, Arcella A, Nobusawa S, Donofrio V, Oghaki H, et al. Prognostic significance of histological grading, p53 status, YKL-40 expression, and IDH1 mutations in pediatric high-grade gliomas. J Neurooncol. 2010;99(2):209-15.

182. 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.

183. De Carli E, Wang X, Puget S. IDH1 and IDH2 mutations in gliomas. N Engl J Med. 2009;360(21):2248; author reply 9.

184. Pollack IF, Hamilton RL, Sobol RW, Nikiforova MN, Lyons-Weiler MA, LaFramboise WA, et al. IDH1 mutations are common in malignant gliomas arising in adolescents: a report from the Children's Oncology Group. Childs Nerv Syst. 2011;27(1):87-94.

185. Setty P, Hammes J, Rothamel T, Vladimirova V, Kramm CM, Pietsch T, et al. A pyrosequencing-based assay for the rapid detection of IDH1 mutations in clinical samples. The Journal of molecular diagnostics : JMD. 2010;12(6):750-6.

186. Hyland EM, Molina H, Poorey K, Jie C, Xie Z, Dai J, et al. An evolutionarily 'young' lysine residue in histone H3 attenuates transcriptional output in Saccharomyces cerevisiae. Genes Dev. 2011;25(12):1306-19.

187. Chowdhury R, Yeoh KK, Tian YM, Hillringhaus L, Bagg EA, Rose NR, et al. The oncometabolite 2-hydroxyglutarate inhibits histone lysine demethylases. EMBO reports. 2011;12(5):463-9.

188. Monti S. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. Machine Learning. 2003;52(1/2):91-118.

189. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics. 2010;26(12):1572-3.

190. Capper D, Weissert S, Balss J, Habel A, Meyer J, Jager D, et al. Characterization of R132H mutation-specific IDH1 antibody binding in brain tumors. Brain Pathol. 2010;20(1):245-54.

191. Pfister S, Remke M, Benner A, Mendrzyk F, Toedt G, Felsberg J, et al. Outcome prediction in pediatric medulloblastoma based on DNA copy-number aberrations of chromosomes 6q and 17q and the MYC and MYCN loci. J Clin Oncol. 2009;27(10):1627-36.

192. Pfaff E, Remke M, Sturm D, Benner A, Witt H, Milde T, et al. TP53 mutation is frequently associated with CTNNB1 mutation or MYCN amplification and is compatible with long-term survival in medulloblastoma. J Clin Oncol. 2010;28(35):5188-96.

193. Kaneda A, Kaminishi M, Yanagihara K, Sugimura T, Ushijima T. Identification of silencing of nine genes in human gastric cancers. Cancer Res. 2002;62(22):6645-50.

194. Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M, et al. Spatio-temporal transcriptome of the human brain. Nature. 2011;478(7370):483-9.

195. Pang JC, Li KK, Lau KM, Ng YL, Wong J, Chung NY, et al. KIAA0495/PDAM is frequently downregulated in oligodendroglial tumors and its knockdown by siRNA induces cisplatin resistance in glioma cells. Brain Pathol. 2010;20(6):1021-32.

196. Ligon KL, Alberta JA, Kho AT, Weiss J, Kwaan MR, Nutt CL, et al. The oligodendroglial lineage marker OLIG2 is universally expressed in diffuse gliomas. Journal of neuropathology and experimental neurology. 2004;63(5):499-509.

197. Wang Y, Yang J, Zheng H, Tomasek GJ, Zhang P, McKeever PE, et al. Expression of mutant p53 proteins implicates a lineage relationship between neural stem cells and malignant astrocytic glioma in a murine model. Cancer Cell. 2009;15(6):514-26.

198. Mehta S, Huillard E, Kesari S, Maire CL, Golebiowski D, Harrington EP, et al. The central nervous system-restricted transcription factor Olig2 opposes p53 responses to genotoxic damage in neural progenitors and malignant glioma. Cancer Cell. 2011;19(3):359-71.

199. Gonzalo S, Jaco I, Fraga MF, Chen T, Li E, Esteller M, et al. DNA methyltransferases control telomere length and telomere recombination in mammalian cells. Nat Cell Biol. 2006;8(4):416-24.

200. Christensen BC, Smith AA, Zheng S, Koestler DC, Houseman EA, Marsit CJ, et al. DNA methylation, isocitrate dehydrogenase mutation, and survival in glioma. J Natl Cancer Inst. 2011;103(2):143-53.

201. Gibbons RJ, McDowell TL, Raman S, O'Rourke DM, Garrick D, Ayyub H, et al. Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. Nature genetics. 2000;24(4):368-71.

202. Drane P, Ouararhni K, Depaux A, Shuaib M, Hamiche A. The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. Genes Dev. 2010;24(12):1253-65.

203. Goldberg AD, Banaszynski LA, Noh KM, Lewis PW, Elsaesser SJ, Stadler S, et al. Distinct factors control histone variant H3.3 localization at specific genomic regions. Cell. 2010;140(5):678-91.

204. Sandoval J, Heyn H, Moran S, Serra-Musach J, Pujana MA, Bibikova M, et al. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. Epigenetics. 2011;6(6):692-702.

205. Olshen AB, Venkatraman ES, Lucito R, Wigler M. Circular binary segmentation for the analysis of array-based DNA copy number data. Biostatistics. 2004;5(4):557-72.

206. Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R, et al. Missing value estimation methods for DNA microarrays. Bioinformatics. 2001;17(6):520-5.

207. Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci U S A. 2002;99(10):6567-72.

208. Team RDC. R: A Language and Environment for Statistical Computing. Vienna, Austria: the R Foundation for Statistical Computing: R Foundation for Statistical Computing; 2011.

209. Witt H, Mack SC, Ryzhova M, Bender S, Sill M, Isserlin R, et al. Delineation of two clinically and molecularly distinct subgroups of posterior fossa ependymoma. Cancer Cell. 2011;20(2):143-57.

210. Pfister S, Remke M, Toedt G, Werft W, Benner A, Mendrzyk F, et al. Supratentorial primitive neuroectodermal tumors of the central nervous system frequently harbor deletions of the CDKN2A locus and other genomic aberrations distinct from medulloblastomas. Genes Chromosomes Cancer. 2007;46(9):839-51.

211. Neben K, Korshunov A, Benner A, Wrobel G, Hahn M, Kokocinski F, et al. Microarray-based screening for molecular markers in medulloblastoma revealed STK15 as independent predictor for survival. Cancer Res. 2004;64(9):3103-11.

212. Qaddoumi I, Kocak M, Pai Panandiker AS, Armstrong GT, Wetmore C, Crawford JR, et al. Phase II Trial of Erlotinib during and after Radiotherapy in Children with Newly Diagnosed High-Grade Gliomas. Frontiers in oncology. 2014;4:67.

213. Jo U, Park KH, Whang YM, Sung JS, Won NH, Park JK, et al. EGFR endocytosis is a novel therapeutic target in lung cancer with wild-type EGFR. Oncotarget. 2014;5(5):1265-78.

214. Fontebasso AM, Papillon-Cavanagh S, Schwartzentruber J, Nikbakht H, Gerges N, Fiset P-O, et al. Recurrent somatic mutations in ACVR1 in pediatric midline high-grade astrocytoma. Nature genetics. 2014;46(5):462-6.

215. Wu G, Diaz AK, Paugh BS, Rankin SL, Ju B, Li Y, et al. The genomic landscape of diffuse intrinsic pontine glioma and pediatric non-brainstem high-grade glioma. Nature genetics. 2014;46(5):444-50.

216. Taylor KR, Mackay A, Truffaux N, Butterfield YS, Morozova O, Philippe C, et al. Recurrent activating ACVR1 mutations in diffuse intrinsic pontine glioma. Nature genetics. 2014;46(5):457-61.

217. Buczkowicz P, Hoeman C, Rakopoulos P, Pajovic S, Letourneau L, Dzamba M, et al. Genomic analysis of diffuse intrinsic pontine gliomas identifies three molecular subgroups and recurrent activating ACVR1 mutations. Nature genetics. 2014;46(5):451-6.

218. Korshunov A, Ryzhova M, Hovestadt V, Bender S, Sturm D, Capper D, et al. Integrated analysis of pediatric glioblastoma reveals a subset of biologically favorable tumors with associated molecular prognostic markers. Acta Neuropathol. 2015;129(5):669-78.

219. Venneti S, Santi M, Felicella MM, Yarilin D, Phillips JJ, Sullivan LM, et al. A sensitive and specific histopathologic prognostic marker for H3F3A K27M mutant pediatric glioblastomas. Acta Neuropathol. 2014;128(5):743-53.

220. Chan KM, Fang D, Gan H, Hashizume R, Yu C, Schroeder M, et al. The histone H3.3K27M mutation in pediatric glioma reprograms H3K27 methylation and gene expression. Genes Dev. 2013;27(9):985-90.

221. Lewis PW, Muller MM, Koletsky MS, Cordero F, Lin S, Banaszynski LA, et al. Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. Science. 2013;340(6134):857-61.

222. Venneti S, Garimella MT, Sullivan LM, Martinez D, Huse JT, Heguy A, et al. Evaluation of histone 3 lysine 27 trimethylation (H3K27me3) and enhancer of Zest 2 (EZH2) in pediatric glial and glioneuronal tumors shows decreased H3K27me3 in H3F3A K27M mutant glioblastomas. Brain Pathol. 2013;23(5):558-64.

223. Bender S, Tang Y, Lindroth AM, Hovestadt V, Jones DT, Kool M, et al. Reduced H3K27me3 and DNA hypomethylation are major drivers of gene expression in K27M mutant pediatric high-grade gliomas. Cancer Cell. 2013;24(5):660-72.

224. Margueron R, Reinberg D. The Polycomb complex PRC2 and its mark in life. Nature. 2011;469(7330):343-9.

225. Bjerke L, Mackay A, Nandhabalan M, Burford A, Jury A, Popov S, et al. Histone H3.3. mutations drive pediatric glioblastoma through upregulation of MYCN. Cancer Discov. 2013;3(5):512-9.

226. Fontebasso AM, Schwartzentruber J, Khuong-Quang DA, Liu XY, Sturm D, Korshunov A, et al. Mutations in SETD2 and genes affecting histone H3K36 methylation target hemispheric high-grade gliomas. Acta Neuropathol. 2013;125(5):659-69.

227. Funato K, Major T, Lewis PW, Allis CD, Tabar V. Use of human embryonic stem cells to model pediatric gliomas with H3.3K27M histone mutation. Science. 2014;346(6216):1529-33.

228. Herz HM, Morgan M, Gao X, Jackson J, Rickels R, Swanson SK, et al. Histone H3 lysine-to-methionine mutants as a paradigm to study chromatin signaling. Science. 2014;345(6200):1065-70.

229. Ul Ain Q, Chung JY, Kim YH. Current and future delivery systems for engineered nucleases: ZFN, TALEN and RGEN. J Control Release. 2015;205:120-7.

230. Wijshake T, Baker DJ, van de Sluis B. Endonucleases: new tools to edit the mouse genome. Biochim Biophys Acta. 2014;1842(10):1942-50.

231. Roujeau T, Machado G, Garnett MR, Miquel C, Puget S, Geoerger B, et al. Stereotactic biopsy of diffuse pontine lesions in children. Journal of neurosurgery. 2007;107(1 Suppl):1-4.

232. Kieran M FA, Papillon-Cavanagh S, Schwartzenstruber J, Nikbakht H, Gerges N, Fiset PO, Bechet D, Faury D, De Jay N, Ramkissoon L, Corcoran A, Jones D, Sturm D, Johann P, Tomita T, Goldman S, Nagib M, Bendel A, Goumnerova L, Bowers DC, Leonard JR, Rubin JB, Alden T, DiPatri A, Browd S, Leary S, Jallo G, Cohen K, Prados MD, Banerjee A, Carret AS, Ellezam B, Crevier L, Klekner A, Bognar L, Hauser P, Garami M, Myseros J, Dong Z, Siegel PM, Gump W, Ayyanar K, Ragheb J, Khatib Z, Krieger M, Kiehna E, Robison N, Harter D, Gardner S, Handler M, Foreman N, Brahma B, MacDonald T, Malkin H, Chi S, Manley P, Bandopadhayay P, Greenspan L, Ligon A, Albrecht S, Pfister SM, Ligon KL, Majewski J, Gupta N, Jabado N. The BATS DIPG study: a national clinical trial of

upfront biopsy and treatment streatification for newly diagnosed diffuse intrinsic pontine glioma. Neuro Oncol. 2014;16(Suppl 1):i40-i59.

233. Danielsson A, Nemes S, Tisell M, Lannering B, Nordborg C, Sabel M, et al. MethPed: a DNA methylation classifier tool for the identification of pediatric brain tumor subtypes. Clin Epigenetics. 2015;7(1):62.

234. Jones D, Capper D, Sill M, Hovestadt V, Schweizer L, Lichter P, Zagzag D, Karajannis MA, Aldape KD, Korshunov A, von Deimling A, Pfister S. DNA methylation profiling identifies new tumor subgroups and aids diagnostic accuracy in a clinical neuropathology setting. Neuro-Oncology. 2014;16(suppl 1):i137-i45.

235. Pollack IF, Hamilton RL, Sobol RW, Burnham J, Yates AJ, Holmes EJ, et al. O6-methylguanine-DNA methyltransferase expression strongly correlates with outcome in childhood malignant gliomas: results from the CCG-945 Cohort. J Clin Oncol. 2006;24(21):3431-7.

236. Srivastava A, Jain A, Jha P, Suri V, Sharma MC, Mallick S, et al. MGMT gene promoter methylation in pediatric glioblastomas. Childs Nerv Syst. 2010;26(11):1613-8.

237. Buttarelli FR, Massimino M, Antonelli M, Lauriola L, Nozza P, Donofrio V, et al. Evaluation status and prognostic significance of O6-methylguanine-DNA methyltransferase (MGMT) promoter methylation in pediatric high-grade gliomas. Childs Nerv Syst. 2010;26(8):1051-6.

238. Nguyen A LM, Gaub MP, Pencreach E, Chenard MP, Guenot D, Entz-Werle N. MGMT (O6methylguanine DNA-methyltransferase) is not the key factor of response to alkylating agents in pediatric brain tumors. Neuro-Oncology. 2014;16(suppl 1):i137-i45.

239. Bechet D, Gielen GG, Korshunov A, Pfister SM, Rousso C, Faury D, et al. Specific detection of methionine 27 mutation in histone 3 variants (H3K27M) in fixed tissue from high-grade astrocytomas. Acta Neuropathol. 2014;128(5):733-41.

240. Chaidos A, Caputo V, Karadimitris A. Inhibition of bromodomain and extra-terminal proteins (BET) as a potential therapeutic approach in haematological malignancies: emerging preclinical and clinical evidence. Ther Adv Hematol. 2015;6(3):128-41.

241. McCabe MT, Ott HM, Ganji G, Korenchuk S, Thompson C, Van Aller GS, et al. EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. Nature. 2012;492(7427):108-12.

242. Hashizume R, Andor N, Ihara Y, Lerner R, Gan H, Chen X, et al. Pharmacologic inhibition of histone demethylation as a therapy for pediatric brainstem glioma. Nat Med. 2014;20(12):1394-6.

243. Vanan MI, Eisenstat DD. Management of high-grade gliomas in the pediatric patient: Past, present, and future. Neurooncol Pract. 2014;1(4):145-57.

244. Suijker J, Oosting J, Koornneef A, Struys EA, Salomons GS, Schaap FG, et al. Inhibition of mutant IDH1 decreases D-2-HG levels without affecting tumorigenic properties of chondrosarcoma cell lines. Oncotarget. 2015;6(14):12505-19.

245. Rohle D, Popovici-Muller J, Palaskas N, Turcan S, Grommes C, Campos C, et al. An inhibitor of mutant IDH1 delays growth and promotes differentiation of glioma cells. Science. 2013;340(6132):626-30.

246. Wang F, Travins J, DeLaBarre B, Penard-Lacronique V, Schalm S, Hansen E, et al. Targeted inhibition of mutant IDH2 in leukemia cells induces cellular differentiation. Science. 2013;340(6132):622-6.

247. Cuyas E, Fernandez-Arroyo S, Corominas-Faja B, Rodriguez-Gallego E, Bosch-Barrera J, Martin-Castillo B, et al. Oncometabolic mutation IDH1 R132H confers a metformin-hypersensitive phenotype. Oncotarget. 2015;6(14):12279-96.