PhD degree in Molecular Medicine

European School of Molecular Medicine (SEMM),

University of Milan and University of Naples "Federico II"

Faculty of Medicine

Settore disciplinare: MED/04

FUNCTIONAL DISSECTION OF HISTONE H3 LYSINE 27 METHYLATION IN GLIOMAGENESIS

Elena Signaroldi

IFOM-IEO Campus, Milan

Matricola n. R08404

Supervisor: Dr. Giuseppe Testa

IFOM-IEO Campus, Milan

Added co-Supervisor: Dr. Gioacchino Natoli

IFOM-IEO Campus, Milan

Anno accademico 2011-2012

Index

Figures and tables index

List of abbreviations

4-OHT: 4-hydroxytamoxifen
ADCF: animal derived component free
AR: ankyrin repeat
ARF: alternate reading frame
BME: Basal Medium Eagle
Bmpr1b: bone morphogenetic protein receptor 1 b
bp: base pair
BSA: bovine serum albumin
$CaCl_2$: calcium chloride
CCD: charge-coupled device
CDK: cyclin-dependent kinase
cDNA: complementary DNA
ChIP-seq: chromatin immunoprecipitation followed by deep sequencing
CNS: central nervous system
CO_2: carbon dioxide
Cre: cyclization recombination
CSC: cancer stem cell
Cy3: cyanine 3
DAPI: 4',6-diamidino-2-phenylindole
DEG: differentially expressed gene
DMSO: dimethylsufoxide
DNA: deoxyribonucleic acid
DNMT: DNA methyl transferase
dNTP: deoxynucleoside 5'-triphosphate
DPBS: Dulbecco's phosphate buffered saline
EBD: estrogen binding domain
ECL: enhanced chemiluminescence
EDTA: ethylenediaminetetraacetic acid
EGF: epidermal growth factor
EGFR: epidermal growth factor receptor
EGFRvIII: epidermal growth factor receptor variant III
Cre-ERT2: Cre-estrogen receptor tamoxyfen-inducible
ESC: embryonic stem cell
Ezh2: enhancer of Zeste homolog 2
FACS: fluorescence-activated cell sorting
FBS: fetal bovine serum
FDR: false discovery rate
FGF-2: fibroblast growth factor
FITC: fluorescein isothiocyanate
GBM: glioblastoma multiforme
GFAP: glial fibrillary acidic protein
GO: gene ontology
GPC: glioma propagating cells
HBS: HEPES buffered saline
HBSS: Hank's balanced salt solution
HDAC: histone deacetylase
HEPES: N-2-hydroxyethyl-1-piperazineethanesulfonic acid
hgDMEM: high glucose Dulbecco's modified Eagle medium
HIV: human immunodeficiency virus
HSP90: heat shock protein 90
ID: identification data
IF: immunofluorescence

INK4a: cyclin-dependent kinase inhibitor
KCl: potassium chloride
KEGG: Kyoto encyclopedia of genes and genomes
KO: knockout
LB: Luria broth
LOF: loss of function
loxP: locus of crossover in P1
MADM: mosaic analysis with double markers
$Na_2HPO_4 \cdot 7H_2O$: sodium phosphate dibasic eptahydrate
NaCl: sodium chloride
NAFBS: North America fetal bovine serum
NG2: neuronal-glial antigen 2
NP: neural precursor
NPM: nucleophosmin
NSC: neural stem cell
OPC: oligodendroglial precursor cell
PcG: Polycomb group
PCR: polymerase chain reaction
PDGFB: platelet derived growth factor b
PDGFRA: platelet derived growth factor receptor a
PDL: poly-D-lysine
PFA: paraformaldehyde
PMSF: phenylmethanesulfonylfluoride
PRC: Polycomb repressive complex
RIPA: radioimmunoprecipitation assay
RNA: ribonucleic acid
rRNA: ribosomal RNA
RTK: receptor tyrosine kinase
SDS: sodium dodecyl sulphate
SDS-PAGE: SDS-polyacrylamide gel electrophoresis
SET: SU(VAR)3-9, E(z), trx
SVZ: subventricular zone
TAE: Tris base, acetic acid, EDTA
TAT: trans-activator of transcription
TBS-T: tris-buffered saline with Tween 20
TCGAN: The Cancer Genome Atlas Network
TE: Tris, EDTA
TPC: tumor propagating cell
tRNA: transfer RNA
WHO: World Health Organization

Abstract

Gliomas represent 80% of the central nervous system tumors. Historically, they have been classified by the World Health Organization according to their supposed cell of origin and to their malignancy grade. Among gliomas, astrocytomas are the most common and are classified histologically with a four grade system where grade IV corresponds to the most malignant form, also known as glioblastoma multiforme (GBM). Furthermore, in recent years a molecular characterization has been defined for GBM. Different subtypes have been identified, each of them over expressing a specific group of genes.

For decades, it has been assumed that cancer was caused only by genetic alterations. Now the view is changed and cancer is considered both a genetic and epigenetic disease. It was demonstrated indeed that epigenetic silencing of genes through DNA or histone methylation at their promoters can be an alternative way to achieve their loss of function; in addition, DNA demethylation of constitutive heterochromatin can promote genome instability. In this context, we refer to epigenetics as the sum of heritable changes in phenotype and/or gene expression without altering the primary DNA sequence.

Methylation of lysine 27 on histone H3 is a post-translational modification that is mediated by the histone methyltransferase complex known as Polycomb Repressive Complex 2 (PRC2) through its active subunit, Enhancer of Zeste Homolog 2 (EZH2). This facultative heterochromatin mark promotes the recruitment of Polycomb Group (PcG) proteins to achieve gene silencing.

PcG proteins have been shown to play a major role in embryonic development and adult somatic cell differentiation. Initial studies on embryonic stem cells (ESCs) showed that Polycomb Complexes are required to maintain stem cell identity. However further investigation showed that this process is much more elaborated and the current model proposes that PcG proteins function dynamically during development and differentiation to lock off the expression of alternative cell fate regulators in any particular lineage.

Moreover EZH2 and other Polycomb members have been found to be dysregulated in a variety of cancer types. EZH2 is overexpressed in tumors like prostate, breast and bladder, and BMI1, a member of Polycomb Repressive Complex 1 (PRC1), is over expressed in GBM, causing aberrant expression of neural stem cell (NSC) markers and preventing apoptosis. It was also demonstrated that genes that are directly regulated by Polycomb in ESCs are up to 12-fold more likely to have cancer specific DNA hypermethylation at their promoters than other genes.

In my project I have pursued two main questions. First, I wanted to assess whether EZH2 is required for glioma initiation. Moreover, since tumor-initiating cells have been found also in GBM, I wanted to check if EZH2 is required to maintain this stem cell pool.

In order to address those questions, I chose a well established animal model of GBM which relies on the loss of *Ink4a/Arf* together with the over expression of EGFRvIII. A conditional knockout allele for *Ezh2* was introduced in order to remove this protein at different stages of the disease. With this work I could demonstrate that while EZH2 is required for the establishment of the tumor, it can be dispensable for its maintenance. The Polycomb axis acts early on during tumor formation, causing the relocation of H3K27me3 in an instructive manner and this process seems to be fundamental in order to achieve the full transformation of the cells. I was able to show that EZH2 is dispensable for glioma maintenance, pointing to a unique window of Polycomb sensitivity that characterizes the primary phase of gliomagenesis prior to the establishment of the glioma propagating cell (GPC) compartment that is able to reconstitute tumors.

Introduction

1 Gliomas

Glial tumors or gliomas, as the name implies, are tumors thought to originate from the glial compartment of the brain. Gliomas represent the most common central nervous system (CNS) malignancies in adults, accounting for about 80% of the CNS cancers. They include astrocytomas, oligodendrogliomas, oligoastrocytomas and ependymomas and the nomenclature originally reflected the morphology of the cells found in the tumor mass. This thesis is focused on astrocytomas and oligodendrogliomas and I do not discuss ependymomas further.

1.1 Histopathological classification

The first comprehensive classification of tumors of the nervous system dates back to 1979, when the World Health Organization (WHO) defined the guidelines for the histological typing ["Histological typing of tumors of the central nervous system" - Geneva, 1979]. Several advances have been made in past few decades, including immunohistochemical typing and genetic profiling but also epidemiology, clinical symptoms associated to the disease, imaging data, prognosis and predictive factors [Louis et al. 2007]. The latest edition was issued by the WHO in 2007.

Gliomas are part of a broader group of tumors arising from neuroepithelial tissue. Historically, they have been classified according to the supposed cell of origin and they are divided into astrocytomas, oligodendrogliomas and oligoastrocytomas (table 1). The classification is accompanied by a grading scheme, which defines four grades of malignancy (Grade I-IV) with grade I being the least and grade IV the most malignant form of the tumor. The grading helps in predicting the behavior of the disease, the prognosis and in choosing the most appropriate therapy (table 1).

Astrocytic tumors
Pilocytic astrocytoma
Pilomyxoid astrocytoma
Subependymal giant cell astrocytoma
Pleomorphic xanthoastrocytoma
Diffuse astrocytoma
Fibrillary astrocytoma
Gemistocytic astrocytoma
Protoplasmic astrocytoma
Anaplastic astrocytoma
Glioblastoma
Giant cell glioblastoma
Gliosarcoma
Gliomatosis cerebri
Oligodendrocytic tumors
Oligodendroglioma
Anaplastic oligodendroglioma
Oligoastrocytic tumors
Oligoastrocytoma
Anaplastic oligoastrocytoma

	I	II	III	IV
Astrocytic tumors				
Subependymal giant cell astrocytoma	x			
Pilocytic astrocytoma	x			
Pilomyxoid astrocytoma		x		
Diffuse astrocytoma		x		
Pleomorphic xanthoastrocytoma		x		
Anaplastic astrocytoma			x	
Glioblastoma				x
Giant cell glioblastoma				x
Gliosarcoma				x
Oligodendrocytic tumors				
Oligodendroglioma		x		
Anaplastic oligodendroglioma			x	
Oligoastrocytinc tumors				
Oligoastrocytoma		x		
Anaplastic oligoastrocytoma			x	

Table 1. WHO classification and grading of gliomas [adapted from "WHO classification of tumours of the central nervous system" - Lyon, 2007]

Particular attention has been dedicated at the grading of astrocytic tumors, which can span the entire grade spectrum. This special attention is related to the biology that underlies this class of tumors as they are the most frequent form of glioma. Current models propose two ontogenic ways through which the most malignant stage, also known as glioblastoma multiforme (GBM), is achieved. In one case, astrocytomas arise as low grade tumors (grade II) and by stepwise accumulation of genetic alterations progress to anaplastic astrocytoma (grade III) and eventually to glioblastoma (grade IV). These GBMs are classed as secondary glioblastomas and are formed within 5-10 years from the initial astrocytoma diagnosis. By contrast, primary GBMs arise without any previous clinical history, thus it is likely that they arise *de novo*. These type of tumors usually are found in older patients compared to secondary GBM and they remain clinically undetectable until the massive growth rate causes impairment of brain activity with overt neurological signs, often due to compression of adjacent anatomical structures. Primary and secondary GBMs show almost identical histopatological features and are both characterized by a very high cell proliferation with widespread invasion. Tumor cells infiltrate the parenchyma and migrate along

white matter tracts and beneath subdural sheets. The invasion of the perivascular space and the sustained angiogenesis can cause hemorrhages. Patients diagnosed with either primary or secondary GBM usually die within 12-16 months from the time of diagnosis. This dismal prognosis is mostly due to the lack of therapies that can eliminate residual tumor infiltration after surgical resection and the difficulties of dealing with a tumor within such an important organ. Recurrence of the tumor associated with high resistance to conventional chemotherapy and radiotherapy strategies have so far prevented effective disease control.

1.2 The problem of the cell of origin

The historical classification of gliomas was based on the features of the cells found in the tumor, which were considered to be reminiscent of their cell of origin. In this context, the term "cell of origin" will be used to indicate the normal cell in which transforming events occur that will finally result in tumor formation. Some features of GBMs have led to the proposal of a more complex situation in which cells other than astrocytes or oligodendrocytes could act as initiators. First, the difference between primary and secondary GBMs (i.e., *de novo* versus progressive disease) opened the possibility that two different cells could be targets of distinct genetic or epigenetic alterations, resulting in two different manifestations of the disease. Second, the intra- and inter-tumor heterogeneity of GBM, in which mixed cytological subtypes can be found, together with a non uniform pattern of genetic lesions and expression profiles, reinforced the debate about the real cell of origin. It also remains unclear whether the cell of origin can be ascribed retrospectively based solely on the appearance of the tumor, as the mutations and changes during tumor progression may alter the cellular phenotype.

For astrocytomas, astrocytes were thought to originate the tumor since they were believed to be the only replicative cell population within the brain, before the discovery of neural stem cells (NSCs). Moreover, gliomas are highly positive for glial fibrillary acidic protein (GFAP), a well established marker of astrocytic compartment. GBMs are characterized by poorly differentiated

cells, thus the hypothesis of astrocytes acting as cell of origin implies a process of dedifferentiation through which fully differentiated cells regain immature progenitor features. One study in particular demonstrated that mouse transformed cortical astrocytes were able to revert to a more immature state [Bachoo et al. 2002]. However, one difficulty has been the problem of defining astrocyte identity [Kimelberg 2004, Garcìa-Marqués et al. 2012]. Several types of astrocytes have been identified in the brain and it is possible that only a subset of them is able to go back to former stages of differentiation upon transformation. Another problem is represented by the GFAP marker; it has been shown that also adult NSCs express this type of intermediate filament, making the identification of mature astrocytes less precise [Doetsch et al. 1999]. In addition, *in vitro* cultures of rodent cortical astrocytes often contain progenitor cells [Laywell et al. 2000, Sergent-Tanguy et al. 2006] raising the hypothesis that this population (and not astrocytes) is the one from which experimental tumors originate. The hypothesis of adult astrocytes acting as cell of origin for brain tumors is therefore plausible but still lacks proper evidence.

Another possibility has centered around the adult NSCs of the subventricular zones (SVZ) or other types of progenitor cells found in the brain, like oligodendroglial precursor cells (OPCs). With two different models, Bachoo and Uhrbom showed that not only astrocytes but also transformed NSCs were able to originate tumors with full penetrance [Bachoo et al. 2002, Uhrbom et al. 2002], and their findings were supported by other works using different model systems [Zhu et al. 2005, Kwon et al. 2008, Zheng et al. 2008]. Emerging evidences in the literature are pointing out a different paradigm and designate OPCs as the ones serving as cell of origin [Lei et al. 2011, Liu et al. 2011, Sugiarto et al. 2011]. Indeed, in an elegant experiment, using the mosaic analysis with double markers (MADM), Liu et al. demonstrated that the cell of origin belonged to the OPC compartment [Liu et al. 2011]. Nevertheless, some problems presented for astrocytes exist also for those models. Cell expressing the proteoglycan NG2 (neuron-glial antigen 2), a marker for OPCs, were shown to generate also gray matter astrocytes [Zhu et al. 2008]; moreover, platelet derived growth factor receptor alpha (PDGFRA) is expressed from both OPCs and NSCs [Jackson et

al. 2006], highlighting once more the difficulty of achieving unequivocal, prospective identification of glioma-relevant cell populations.

Finally, a recent paper by Friedmann-Morvinski et al. demonstrated that gliomas can originate from NSCs, astrocytes but also neurons [Friedmann-Morvinski et al. 2012]. In this study the authors restore the idea of dedifferentiation of terminally differentiated cells, like cortical neurons. Indeed, they found that tumor cells expressed high levels of neural stem or progenitor cells. Moreover, they discovered that tumors originated through dedifferentiation of differentiated cells (either astrocytes or neurons) matched to the mesenchymal molecular subtype (see section 1.4). One potential explanation to resolve these differing outcomes is that GBM, as is now realized comprises several distinct subtypes. Therefore, one possible explanation could be that there are different cell of origins for distinct subtypes of the disease.

So far, all the possibilities presented by several studies were not able to provide a definitive answer about the real cell of origin. All hypotheses tested showed some still unresolved caveats, especially concerning the identification of proper markers to discriminate among different cell population.

1.3 Tumor- propagating cells in glioma

As is the case for other types of tumors, gliomas have been reported to harbor a subpopulation of cells that can be defined as “glioma stem cells”. Nevertheless, the term “cancer stem cells” (CSCs) raises heated debate due to the various uncertainties implicit in this description. Since the molecular mechanisms driving stem cell identity and self-renewal are still being elucidated, a clearer molecular definition remains out of reach. In 2006 the American Association for Cancer Research stated that CSCs are a subpopulation of cells found in a tumor with self-renewal capacity and able to give rise to the heterogeneous cell composition found in the tumor [Clarke et al. 2006]. This pool of cells is able to reconstitute and propagate the tumor and is considered to provide the most plausible explanation for recurrence after surgical resection. In my thesis I refer

to this population as “tumor-propagating cells” (TPCs) as suggested by several reviews [Visvader 2011, Chen et al. 2012]. Several studies demonstrated the presence of TPCs in gliomas [Ignatova et al. 2002, Singh et al. 2003, Hemmati et al. 2003, Kondo et al. 2004, Singh et al. 2004, Yuan et al. 2004, Galli et al. 2004] and now they are considered to be a faithful model system used to study this type of tumor either in culture or in xenotransplantation experiments. Nonetheless, many controversies affect this field of research.

One issue is the way in which the tumor is maintained; three are the models proposed currently [reviewed in Chen et al. 2012]: (i) the classical hierarchical cancer stem cell model, in which very few cells inside the tumor are endowed with the continuous self-renewal capacity and are able to generate other different cell types responsible for the intrinsic heterogeneity of this disease; (ii) the stochastic model, in which all the tumor cells are considered to be equivalent in their ability to sustain the tumor growth; the heterogeneity in this case is justified by different clones growing together; (iii) the evolutionary model, in which the hierarchy of cancer stem cells is constantly evolving due to new genomic alterations that confer growth advantage to the cells in which they occurr. In this view, distinct subclones are generated through a Darwinian evolutionary process within the tumor.

Another issue regards the definition of robust and reliable markers that univocally identify the TPC population. One of the first cell surface markers identified was CD133/prominin 1 [Singh et al. 2003, Singh et al. 2004]. However, it is now clear that $CD133^-$ cells can also display the tumor propagating behaviour described for $CD133^+$ cells. Other markers have been proposed such as A2B5 [Ogden et al. 2008, Tchoghandjian et al. 2010], CD15/LeX [Son et al. 2009] and integrin alpha-6 [Lathia et al. 2010], but none of these alone is able to unequivocally identify the TPCs. In addition, the patient to patient variability and the variety of GBM subtypes further complicates the situation. Similarly to the haematopoietic field, it seems likely that a large panel of multiple markers is the only robust means to purify TPCs.

In summary, there exist many differences among gliomas: the histological classification, the malignancy grade, the genetic signature lesion, the type of disease progression and the molecular

classification (discussed in the following section). Given this diversity, it could be difficult to achieve a universal paradigm, supposing its existence, both for the cell of origin and for the cancer stem cell theory.

1.4 Molecular classification of gliomas

The problem of heterogeneity of gliomas has now been extensively explored also at the molecular level.

Several studies have tried to identify the genetic alterations involved in gliomagenesis and to subclassify tumors based on transcriptional state. Additionally, James et al. have demonstrated that multiple genetic lesions were occurring non-randomly in different glioma samples, providing evidence that the genotype, rather than the phenotype, could have been a better tool to understand the underlying molecular mechanisms of this cancer [James et al. 1988]. With the development of new technologies, the genomic landscape has been explored in a more comprehensive manner. Two complementary genome-wide studies were performed by The Cancer Genome Atlas Research Network (TCGAN) and the Vogelstein Laboratory and were reported in landmark papers in 2008. They were able to identify the most significant somatic mutations occurring in a large group of glioma samples, such as *Tp53*, *Pten*, *Nf1*, *Egfr*, *Rb* and *Pi3k*. This analysis revealed the presence of core signaling pathways commonly affected in gliomas: the p53 pathway, the RB pathway and the RTK pathway. Several other studies provided further information about mutations and copy number variations, thus helping to discover different patterns of alterations and their incidence. Interestingly, the Vogelstein Laboratory, due to extensive exome sequencing, was the first one to identify the mutation of the *Idh1* gene that are now diagnostic of secondary GBM. [The Cancer Genome Atlas Research Network 2008, Parsons et al. 2008].

Also the transcriptome of gliomas has been widely examined, revealing that tumors that are classified in the same histopathological group can belong to different molecular subtypes. Gene

expression profiles can be used to identify multiple subtypes of gliomas [Godard et al. 2003, Shai et al. 2003, Liang et al. 2005, Nigro et al. 2005], to predict survival [Nutt et al. 2003, Freije et al. 2004] and to unravel the differences between primary and secondary tumors [Maher et al. 2006, Tso et al. 2006]. In the last years, three main studies have tried to accurately define molecular subtypes based on the expression profiling and to correlate the subtypes with genetic alterations, core pathways, prognosis and survival [Phillips et al. 2006, Li et al. 2009, Verhaak et al. 2010]. More recently, another study identified six biological subgroups based on the integration of global DNA methylation data with genetic alterations, expression profiling and clinical parameters [Sturm et al. 2012]. Unfortunately there is a total lack of consensus on the number of subgroups and on the features characterizing each of them. Several aspects biased the interpretation of expression data, [reviewed in Marko et al. 2011]; differences in the sample cohort chosen for the analysis, sample preparation, analysis technique and bioinformatics and statistical analysis make the molecular classification lacking in a consensus. Moreover, it is possible a lot of the difficulties arise from the intra-tumoral heterogeneity, with multiple distinct subtypes present within a single tumor.

Notwithstanding this blurred picture arising from molecular studies, the importance of their findings is impacting the way of making clinical diagnosis and choosing therapies. Histology can be flanked by molecular data in order to better classify the tumor, improving the diagnosis and the treatment strategy applied to each patient cohort.

2 Epigenetics

2.1 General concepts about epigenetics

The evidence that some phenomena in cells or organisms were not simply relying on the DNA sequence came already in the 40's. The first one arose from studies carried out by several geneticists who had noted that the same DNA fragment could generate different phenotypes if placed in different positions in the genome (e.g., "position effect variegation" discovered by H.J.

Muller [Muller 1938] or "jumping genes" discovered by B. McClintock [McClintock 1950]). The second clue came from developmental studies, where it was evident that a multicellular organism is built of several different types of cells, all coming from the same unicellular zygote, which need to phenotypically diverge in order to accomplish different functions in the complete organism. More interestingly, once established, those differences in the phenotype could be transmitted to subsequent cell generations. A pioneer in this field of research was indeed a famous embryologist and geneticist, C.H. Waddington, who for the first time defined the concept of epigentics [Waddington 1942]. Work from J. Gurdon further demonstrated that during development all genes were retained in somatic cells, exemplified by their ability to generate the entire organism following nuclear transfer. The genes must therefore be turned on and off with exquisite control [Gurdon 1962]. After almost 80 years of research, the definition of epigenetics is still under debate; the most widely accepted refers to it as the sum of the changes in phenotype, including gene expression, that can be mitotically and/or meiotically inherited without changes in the underlying DNA sequence. There are three criteria used to define a molecular signal as epigenetic: (i) evidence of transmission (self-sustaining transmission to the progeny); (ii) mechanism for propagation (how the mark is maintained after DNA replication and cell division); (iii) effect on gene expression. According to this definition, two are the *bona fide* epigenetic marks: DNA methylation [reviewed in Bird 2002] and trimethylation of lysine 27 on histone H3 (H3K27me3) [Hansen et al. 2008, Margueron et al. 2009]. Despite these stringent criteria used to define *bona fide* epigenetic marks, other mechanisms are known to affect gene expression, particularly different histone posttranslational modifications, nucelosome repositioning and remodelling and small noncoding RNAs.

2.2 Histones and histone-modifications

The basic unit of chromatin, also known as nucleosome, consist of 147 DNA base pairs wrapped around an octamer of histone proteins. This octamer is built with two molecules of each of the

core histone proteins, namely H2A, H2B, H3 and H4. Histone proteins are highly basic proteins, composed of a globular domain and a tail in their amino-terminal portion. The vast majority of histone posttranslational modification occurs at their tails, which stretch out from the nucleosome; the most common and well-studied covalent modifications include methylation, acetylation, phosphorylation and ubiquitylation. The information content of the modifications led to the formulation of the "histone code" hypothesis that adds another control level to the information potential of the DNA sequence. Modifications are able to influence different processes, among which chromatin organization, protein binding and transcriptional activity. It must be highlighted that all these mechanisms can be interconnected and several can be needed at once to achieve the final output.

Among the covalent modifications, methylation is considered to be the more complex. Indeed it can occur either at arginine or lysine residues and there are at least 24 known position on histone tails at which methylation can be established. Moreover, lysine can exist in multiple methylation states, namely unmethylated, mono- (me1), di- (me2) and tri- (me3) methylated and arginine can be unmethylated, mono- or di-methylated. A further level of complexity is represented by the fact that, differently from other modifications, the output of the methylation can be either an active or a silent state of gene expression. Paritcularly, methylation of lysine (K) has been shown to play a major role in switching transcription on and off during development and differentiation. Generally, methylated residues linked to transcriptionally active regions are H3K4, H3K36 and H3K79, while those associated to transcriptionally repressed regions are H3K9, H3K27 and H4K20. Despite this categorization, recent studies showed that marks with opposing effect can be found together in certain chromatin regions, as in the case of "bivalent domains" described by Bernstein et al. [Bernstein et al. 2006]. H3K4me3 and H3K27me3 lay together in these domains, keeping the genes expressed at very low levels and poised for definitive activation or repression, introducing an extra-layer of diversification for the control of the biological output. This finding was confirmed by the work by Azuara et al. [Azuara et al. 2006]. Thus, a complicated picture has emerged with no straightforward "histone code" being identified.

2.3 Polycomb mediated gene repression

As discussed in the previous section, the observation that adult animals were constituted by several hundred distinct cell types arising from one cell, the zygote, was a major unresolved problem in biology. It became clear that cell identity and function was maintained through a specific gene expression profile; but how is it possible to tightly control gene expression in a tissue specific manner? Moreover, cells are able to rapidly respond to external stimuli; how can the gene expression be modified precisely, robustly and in a time specific manner? And, how is it possible to pass the established pattern of active and repressed genes through cell division? The first major steps to addressing these questions came from classical genetic studies in the fruit fly, *Drosophila melanogaster*, and where then confirmed also in mammals.

In their lab, Pam and Ed Lewis identified a variety of fly mutants in which posterior segment of the abdomen showed an "anteriorized" phenotype [Lewis 1978]. It is now known that the anterior-posterior axis development of the body plan is controlled by expression of homeotic genes, also known as *Hox* genes. *Hox* genes are a family of homeodomain transcription factors that are expressed in a specific manner in each body segment. Lewis identified mutant fly males that had extra sex combs also on the second and third pair of legs, while wild type males bear this structure only on the first pair of legs. It was later demonstrated that this was caused by a mutation in a *trans*-acting regulator of the *Hox* genes and, due to the resulting phenotype, this regulator was called Polycomb. Similar phenotypes were obtained by mutating other proteins that were all included in the Polycomb group (PcG). It was demonstrated that the ordered pattern of the anterior-posterior axis needs transcriptional repression of the *Hox* genes after they carried out their function in a time and space specific manner. PcG proteins are the effectors of this silencing.

PcG gene products are organized in two different multimeric protein complexes, which, based on their function, are referred to as Polycomb Repressive Complex (PRC) 1 and 2 [Kuzmichev et al. 2002]. The two complexes are required for consecutive steps to achieve gene silencing: PRC2

initiates the repression, while PRC1 is involved in maintaining the repressed state once established by PRC2. The core proteins that constitute the PRC2 in mammals (homologs of the same genes found in Drosophila) are: EED, EZH2 (that might be substituted by EZH1), SUZ12 and RBAP46/48 [Muller et al. 2002]. The way in which PRC2 mediates gene silencing is through modification of the chromatin; indeed, the complex has a histone methyl transferase activity that methylates lysine 27 of histone H3, and to a lesser extent, lysine 9 [Cao et al. 2002, Czermin et al. 2002, Kuzmichev et al. 2002, Muller et al. 2002]. Many other proteins associate with PcG complexes, like transcription factors (GAGA factors [Mulholland et al. 2003], MYC [Ogawa et al. 2002], E2F [Trimarchi et al. 2001, Ogawa et al. 2002]), histone deacetylase (HDACs) [van der Vlag et al. 1999] and other histone methyl transferases [Sewalt et al. 2002]. When PRC2 is initially recruited at the target sites, it associates with HDACs to remove acetylation on histone H3, a well-characterized mark of active genes. The H3 methyltransferase contained in the complex methylates H3 on lysine 27, and in turn acts as docking site for the N-terminal chromodomain of proteins assembled in PRC1.

3 Polycomb Repressive Complexes and cancer

For decades, it has been assumed that cancer was caused only by genetic alterations. Now the view is changed and cancer is considered both a genetic and epigenetic disease. It was demonstrated indeed that epigenetic silencing of genes through DNA or histone methylation at their promoters can be an alternative way to achieve their loss of function; in addition, DNA demethylation of constitutive heterochromatin can promote genome instability. Particularly, several findings linked PcG proteins to cancer.

3.1 PcG proteins are involved in malignant transformation

Ezh2 and *Eed* have been shown to be targets of the pRB-E2F pathway, which is often misregulated in cancer [Bracken et al. 2003]. Also *Suz12* is transcriptionally regulated by E2F [Muller et al. 2001,

Weinmann et al. 2001]. It is known that the integrity of the pRB-E2F pathway is required for proper differentiation and embryonic development, and this can be related to the control that it exerts on PRC2. Another important evidence is that overexpression of EZH2 and EED shortens G0/G1-S phase transition, conferring a proliferative advantage to the cells, which is a typical feature of many tumor types [Bracken et al. 2003].

Some PcG proteins have been shown to be overexpressed and/or amplified in variety of tumors; EZH2 is overexpressed in metastatic prostate cancer [Varambally et al. 2002], aggressive breast cancer [Kleer et al. 2003, Raaphorst et al. 2003] and bladder [Arisan et al. 2005, Raman et al. 2005] and in some cases the overexpression might be due to amplification of the gene [Bracken et al. 2003]. SUZ12 is overexpressed in breast and liver tumors [Kirmizis et al. 2003, Kirmizis et al. 2004, Kuzmichev et al. 2005]; furthermore, it is regulated by the β-catenin pathway, which is often altered in colon cancer, thus resulting in *Suz12* upregulation [Kirmizis et al. 2003]. *Bmi1*, a member of the PRC1, is amplified in haematological tumors [Bea et al. 2001] and it was shown to cooperate with RAS and MYC in cellular transformation [Jacobs et al. 1999a].

3.2 Aberrant gene regulation in cancer

The altered expression and/or activity of PcG proteins results in aberrant silencing of tumor suppressor genes and of genes that control differentiation and restrain proliferation. Pertinent examples of this, relative to this thesis, are represented by the control of the *Ink4a/Arf* locus and by the expression of the bone morphogenetic protein receptor 1b (*Bmpr1b*) gene.

Both EZH2 and BMI1 were shown to directly regulate the silencing of the *Ink4a/Arf* locus, which encodes for two proteins, $p16^{INK4A}$ and $p14^{ARF}$ [Bracken et al. 2007, Jacobs et al. 1999b]. These two proteins control cell proliferation by regulating the pRB-E2F pathway and the p53 pathway, two of the most commonly mutated pathways in all types of cancer. A more comprehensive description of p16/p14 mode of action will be presented in the next section; briefly they inhibit RB phosphorylation or MDM2-mediated ubiquitylation of p53, inducing cell cycle arrest. In tumors in

which pRB and p53 pathways are not mutated, EZH2 and BMI1 can mediate *Ink4a/Arf* silencing, leading to epigenetic inactivation of these cell cycle regulators.

The second example stems from the finding that EZH2 was demonstrated to immunoprecipitate with all the 3 human DNA methyl transferases (DNMTs) [Vire et al. 2006], suggesting that there is a cooperation between these two silencing pathways. Furthermore, PcG targets in embryonic stem cells are 12-fold more likely to undergo DNA hypermethylation in cancer and H3K27me3 (together with another repressive histone methylation, H3K9me2) seems to be the mark required to trigger DNA methylation [Ohm et al. 2007, Schlesinger et al. 2007, Widschwendter et al. 2007]. In their study, Lee et al. showed that a subset of human glioma stem cells resemble fetal neural stem cells and are prevented from differentiating by DNA methylation and H3K27 trimethylation occurring at the promoter of *Bmpr1b* gene. This gene is involved in astroglial differentiation in normal embryogenesis. Inhibition of EZH2 was sufficient to remove the superimposed DNA methylation, restoring BMPR1B expression and, consequently, the astroglial differentiation potential of the glioma stem cells [Lee et al. 2008].

This two examples highlighted two different mechanism of action through which Polycomb axis can facilitate tumor formation. On the one hand, it can cause loss of function of tumor suppressor genes by epigenetic silencing, endowing cancer cells with growth advantage. On the other, it can overturn the normal gene expression profile of the cells by bringing them back to and/or blocking them into a less differentiated developmental stage. Moreover, it can act as template for DNA methylation which has been demonstrated to occur non-randomly but rather at specific loci in an instructive manner during tumor formation.

4 Modeling glioma in the mouse

4.1 The choice of the model system

There are many murine glioma models described in literature [reviewed in Hambardzumyan et al. 2011 and Chen et al. 2012]. For my study, I decided to make use of a well-established mouse

model that relies on the orthotopic transplantation of murine transformed astrocytes [Bachoo et al. 2002]. The astrocytes are derived from a $p16^{INK4a}$/$p19^{ARF}$ double knock-out (KO) strain and are then infected in order to overexpress the mutated form of the human epidermal growth factor receptor (EGFR) known as EGFRvIII. I have chosen the Bachoo et al. model system for five main reasons. First, it accurately recapitulates the human disease, both at the genetic and histopathological level. Indeed, p16/p19 loss of function and mutation of *Egfr* are two signature lesions of GBM, associated in a discrete proportion of gliomas. Second, this study demonstrated that it is possible to establish tumors from either astrocytes or neural stem cells [Bachoo et al. 2002]. This suggests that, regardless of the cell of origin and the differentiation stage, there is likely to be a convergent mechanism that alters the gene expression profile, confirming the hypothesis of the instructive process that leads to tumor formation. Third, it bears the intrinsic possibility to analyze cell populations with different tumorigenic potential, enabling the study of their transcription profiles, of their epigenomes and how the first varies according to changes in the latter at each stage of disease progression. Fourth, the KO of p16/p19 is of primary importance in order to study cancer relevant developmental pathways of PcG aside from its known effect on INK4a/ARF. Finally, it is a fully penetrant modeling, with associated experimental advantages.

In order to unravel the role played by the Polycomb axis during gliomagenesis, I introduced in the model described above a conditional allele for Ezh2 [Su et al. 2003], the enzymatic subunit of the PRC2.

4.1.1 The Ink4a/Arf locus

The *Ink4a/Arf* locus is characterized by an uncommon, if not unique, genomic structure which to my knowledge is not found elsewhere in the mammalian genomes. This segment spans approximately 3×10^4 bp and encodes for two different proteins each of which has its own promoter. The first exon is mutually exclusive between the two products, but the second and the

third are shared. Despite this, the two proteins have no amino-acid similarity and they are not isoforms.

The first protein being discovered from this locus was a 16 KDa protein that was shown to negatively modulate the activity of cyclin dependent kinase (CDK) 4 and 6, hence the name $p16^{INK4a}$ (inhibitor of CDK4) [Serrano et al. 1993]. This was a truncated form; the full length contains 8 more amino-acids, for a total of 156. P16 is composed of 4 ankyrin repeat (AR) motifs linked together by loops of different length; the second and third AR motifs are responsible for binding the CDKs, blocking their activity. Moreover, this binding causes the shrinkage of the CDK surface on which the activator of the kinase, cyclin D, usually binds. CDK4 and 6 are known to hyperphosphorilate pRB in the late G1 phase, causing the release of E2F transcription factors, the transcription of proliferation genes and the transition to the S phase of the cell cycle. By imparing CDKs activity, p16 impedes the phosphorilation of pRB, which sequesters E2F and arrests the cell cycle in G1 phase. INK4a was also shown to be involved in cellular senescence and aging, even if the mechanisms are still not fully understood. Indeed, while p16 expression level is low during development, it increases during aging, restricting the replicative potential of self-renewing or progenitor populations, including neural progenitor cells [Molofsky et al. 2006, Krishnamurthy et al. 2006].

The second protein was discovered when an alternative first exon was demonstrated to be transcribed from the same locus, producing an alternate reading frame (ARF, also known as p19 in mouse and p14 in humans) [Quelle et al. 1995]. No structural motifs have been identified in $p19^{ARF}$; of the 169 amino-acids, more than 20% are arginine residues, thus resulting in a very basic and hydrophobic protein which is localized in the nucleolus thanks to the nucleolar localization signal located at the N-terminal portion of the protein. Expression of oncogenes upregulates p19, inducing a block in the cell cycle in response to hyperproliferative signals. This is mainly due to the inhibition of two ubiquitin ligases, MDM2 and ARF-BP1/Mule that mediate ubiquitination of p53, causing its nuclear export and proteasomal degradation. By binding to MDM2, p19 promotes the stabilization of p53, inducing cell cyle arrest. Despite p53 stabilization seems to be the main

mechanism by which ARF is able to restrain cell growth, some evidences raised the possibility of some p53-independent mechanism of action of p19. Indeed, it was shown to induce G1 arrest in cell lacking p53 [Weber et al. 2000]. Moreover, it seems to reduce ribosomal biogenesis by sequestering nucleophosmin/B23 (NPM) in the nucleolus, but this finding is still under debate [Bertwistle et al. 2004, Itahana et al. 2003]. Finally, in mouse (but not in humans) it seems to be involved in senescence [Kamijo et al. 1997].

The *Ink4a/Arf* locus therefore plays a central role in controlling cellular growth, particularly due to malignant cues. Thus, it appears even stranger that two different and powerful tumor suppressor mechanisms have been evolutionary placed in the same locus. In fact, in different tumor types deletion or mutation of the common exons often lead to the loss of both proteins.

4.1.2 Epidermal growth factor receptor vIII

Epidermal growth factor receptor belongs to the ErbB family of receptors, which comprises four members: EGFR (also known as ErbB1), ErbB2, ErbB3 and ErbB4. All the members are constituted by 3 main domains: an extracellular domain to which ligands bound, a single membrane-spanning region and an intracellular domain with multiple tyrosine residues which are phosphorylated upon ligand binding and receptor activation. This leads to activation of downstream effectors and, ultimately, to gene transcription.

Amplification of the *EGFR* gene with consequent over-expression is the most common genetic alteration in gliomas, occurring in 40-60% of *de novo* GBMs [Libermann et al. 1984, Sugawa et al. 1990, Ekstrand et al. 1992, Jaros et al. 1992, Wong et al. 1992]. During the process of amplification of the *EGFR* locus, some amplicons undergo rearrangements resulting in mutated form of the receptor (60-75% tumors with amplified *EGFR*). The most common rearrangement is the deletion of exons 2 to 7, causing the loss of 267 aminoacids that form the extracellular ligand-binding domain [Ekstrand et al. 1994]. This form of the receptor is constitutively active and even if it does not seem to confer any growth advantage to cultured cells, it does in vivo upon

transplantation [Nishikawa et al. 1994]. Several studies also demonstrated that wild-type *EGFR* amplification and over-expression alone is quite inefficient in driving tumor formation, even in presence of other mutation such as INK4a/ARF loss of function (LOF) [Lachat et al. 2004, Zhu et al. 2009]. EGFRvIII alone is not sufficient for the tumorigenic process [Zhu et al. 2009] and this particular mutation has almost always been found in association with INK4a/ARF LOF [Hayashi et al. 1997, Hegi et al. 1997], while it is mutually exclusive with p53 LOF [Watanabe et al. 1996, Fulci et al. 2000, Labuhn et al. 2001]. Amplification of *EGFR*, and so its rearrangement, seem to be restricted to grade III and grade IV astrocytomas [Lieberman et al. 1985, Wong et al. 1987, Ekstrand et al. 1992] and it is less prevalent in secondary GBM. The mutation is not observed in normal tissue [Garcia de Palazzo et al. 1993, Moscatello et al. 1995, Wikstrand et al. 1998] and its presence is associated with poor prognosis [Furnari et al. 2007].

Many studies have attempted to identify the function of EGFRvIII, but a clear conclusion is still missing. Several studies demonstrated that it can activate the PI3K pathway [Moscatello et al. 1998, Choe et al. 2003, Huang et al. 2007], which is associated with cell proliferation, survival and migration. Moreover, it was shown to upregulate *Bcl-xl*, a potent inhibitor of apoptosis [Nagane et al. 1998], and to cause over-expression of tissue factor that can induce over-expression of angiogenic factors like IL-8 and VEGF, leading to an angiogenic phenotype [Magnus et al. 2010].

4.1.3 Enhancer of Zeste homolog 2 (Ezh2)

EZH2 is the active subunit responsible for the methylation of K27 in PRC2. This protein is constituted of 746 amino-acid residues and maps to chromosome 6 in mouse [Laible et al. 1999], close to *Hox* gene cluster.

EZH2 contains a SET domain (named after SU(VAR)3-9, E(z)and trx, the founding members of the SET domain containing proteins) which is able to catalyze the addition of methyl groups to the ε-amino group of lysine side chain. Several studies demonstrated that complex containing E(z) in Drosophila or EZH2 in humans were able to trimethylate H3K27 and, to a lesser extent, H3K9 [Cao

et al. 2002, Czermin et al. 2002, Kuzmichev et al. 2002, Muller et al. 2002]. EZH2 alone is not able to exert its histone methyl transferase activity; it requires to be in a complex with at least EED and SUZ12, which are thought to behave as modulating proteins [Cao et al. 2004, Pasini et al. 2004, Ketel et al. 2005, Montgomery et al. 2005, Nekrasov et al. 2005]. Moreover, another protein, PHF1 seems to interact with PRC2 and favor the recruitment to target genes and the methyltransferase activity of EZH2 needed to achieve the H3K27me3 [Sarma et al. 2008].

Initial studies on embryonic stem cells (ESCs) showed that Polycomb Complexes are required to maintain stem cell identity [Boyer et al. 2006, Lee et al. 2006]. However further investigation showed that this process is much more elaborated and the current model proposes that PcG proteins function dynamically during development and differentiation to lock off the expression of alternative cell fate regulators in any particular lineage [Mohn et al. 2008, Ezhkova et al. 2009]. The work by Mohn and colleagues has highlighted that, at least in the context of neural differentiation, PRC2 plays a role in the whole process. In particular this work shows that PRC2 reaches new targets not only in the precocious phase between ES cells and neural precursors, but also at a later stage between neural precursors and terminal differentiated neurons [Mohn et al. 2008]. Therefore even if EZH2 is expressed at low level in late stages of neural differentiation [Sher et al. 2008, Pereira et al. 2010], it appears still important in the repression of new targets. Hence it is clear that EZH2 and PRC2 complex have a function at all stages of neural differentiation. The role of EZH2 in neural differentiation has also been studied *in vivo* with contrasting results. In one study, *Ezh2* was deleted in the cortex at E14.5, causing impairment in the neurogenic to astrogenic fate switch that leads to a delay in the appearance of astrocytic markers and to an increased numbers of neurons in knockout cortices [Hirabayashi et al. 2009]. Instead, in a second work by Pereira et al. the deletion of *Ezh2* is achieved before the onset of neurogenesis at E12. This precocious absence of EZH2 leads to an accelerated differentiation with an earlier appearance of both neurons and astrocytes at the expenses of neural precursors that results in a thinner cortex at birth [Pereira et al. 2010].

Given the importance of EZH2 during the CNS development, it appears evident how its disregulation that often happens in cancer can affect gene expression and the "differentiation" state of a cell.

Homozygous inactivating mutations of *Ezh2* are embryonically lethal, most probably due to the lack of regulation of *Hox* genes expression and body patterning formation. In order to study EZH2 effects at specific developmental stages or cell lineages, two laboratories produced two slightly different conditional KO alleles [Su et al. 2003, Shen et al. 2008]. A conditional KO allele can be inactivated in time and/or space specific manner. Both the alleles rely on the CRE/loxP technology (see material and Methods, section 1.2 for further details) and upon exposure to CRE recombinase the SET domain is deleted. In order to study the role of Ezh2 during gliomagenesis, I made use of the conditional allele generated in the lab of Dr. Tarakhovsky.

Materials and Methods

1 Mouse strains

1.1 p16/p19$^{-/-}$ strain

INK4a (p16) and ARF (p19) are two proteins encoded by the same locus, exploiting two different promoters and two mutually exclusive exons 1, namely exon 1α for p16 and exon 1β for p19. The other two exons are instead shared between the two proteins. Mice deficient for both p16 and p19 were generated by replacing exons 2 and 3 of the *Ink4a/Arf* locus with a cassette encoding for neomycin resistance, leading to the loss of both proteins [Serrano et al. 1996].

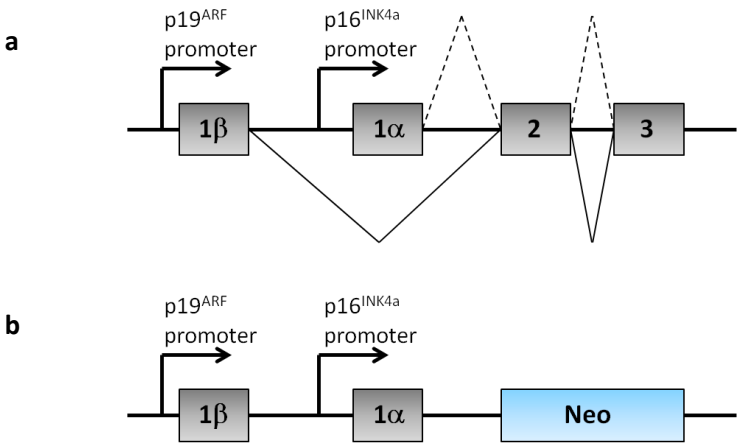

Figure 1. The *Ink4a/Arf*$^{-/-}$ locus. Schematic representation of the *Ink4a/Arf* locus (not to scale). a. *Ink4a* and *Arf* are transcribed from two independent promoters. Each transcript has a specific 5' exon (exon 1α for p16 and exon 1β for p19) which is spliced into common exons 2 and 3, as shown by black (p19^{ARF}) or dashed (p16^{INK4a}) lines. b. Mice deficient for both p16^{INK4a} and p19^{ARF} were generated by replacing exons 2 and 3 with a cassette encoding for neomycin resistance.

1.2 *Ezh2* conditional KO strain

Ezh2$^{-/-}$ mice are early embryonic lethal. To study the function of this protein I employed a conditional deletion strategy. Two loxP sites were inserted in the *Ezh2* locus. Each loxP site is a Cre (cyclization recombination) recognition element of 34 base pairs (bp), composed of two 13 bp inverted repeats flanking an asymmetrical sequence called spacer. If the spacers of two loxP sites are oriented in the same direction, the reaction will result in the excision of the interposed sequence. In the *Ezh2* conditional KO allele, two loxP sites flank the 4 exons encoding for the

catalytic domain of the protein (also known as SET domain) [Su et al. 2003]. Since the loxP sites are located in introns, the locus behaves like wild-type, but, upon administration of Cre recombinase, this portion of the gene is removed. It has been shown that the locus encodes for a truncated form of the protein, which is anyway catalytically inactive.

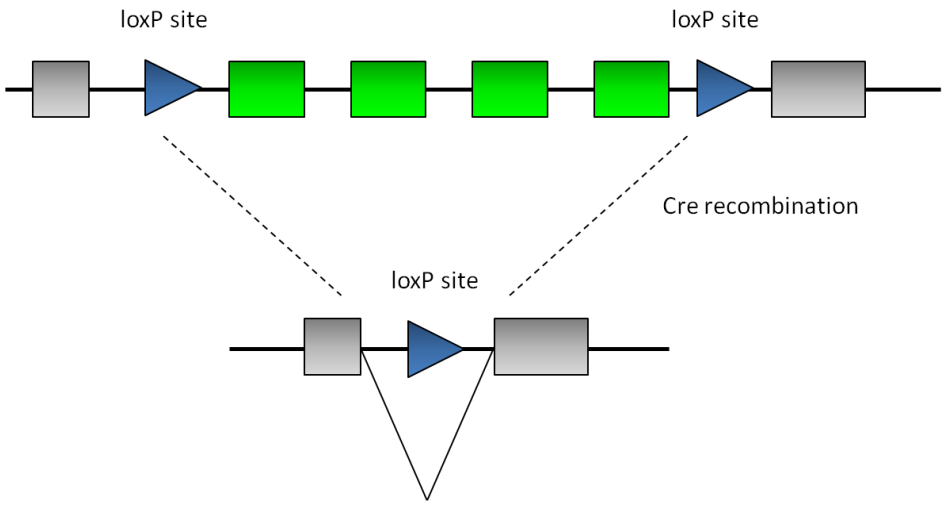

Figure 2. The *Ezh2* conditional allele. Part of the *Ezh2* conditional allele is shown before and after Cre recombination. Green boxes indicate the exons encoding the SET domain, blue triangles represent the loxP sites. Following Cre recombination, the four exons of the SET domain are deleted and the exon located upstream can be spliced into the exon located downstream, leading to the formation of a truncated protein.

1.3 Rosa26Cre-ERT2 transgenic strain

For conditional activation of Cre we made use of a tamoxifen inducible fusion protein. The transgene encoding Cre-ERT2 protein was inserted in the *Rosa26* locus from where it is ubiquitously expressed [Seibler et al. 2003]. The protein encoded by the transgene is a Cre recombinase fused to a mutated form of the estrogen binding domain (EBD). In the absence of the ligand (Tamoxifen), the EBD interacts with the ubiquitous heat shock protein 90 (HSP90), keeping Cre inactive in the cytoplasm. Upon administration of Tamoxifen, HPS90 is released and Cre is able to translocate to the nucleus where it catalyses the recombination reaction between the loxP sites.

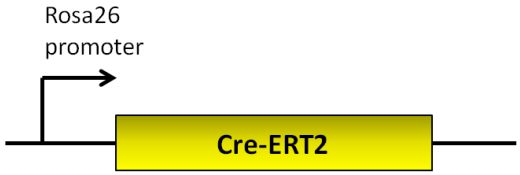

Figure 3. The *Rosa26Cre-ERT2* transgenic locus.

1.4 Breeding strategies

1.4.1.Generation of the p16/p19 KO – Ezh2 conditional KO strain

Homozygous mice from both parental strains (p16/p19$^{-/-}$ and *Ezh2* conditional KO) were bred; all the pups in F1 were then heterozygous for both mutations. By breeding two mice from F1, it was possible to obtain in F2 mice that were double homozygous or homozygous for the p16/p19$^{-/-}$ allele and heterozygous for *Ezh2* conditional allele, that were bred together. In this way, all the mice in F3 were homozygous for the p16/p19$^{-/-}$ allele and either homozygous or heterozygous for *Ezh2* conditional allele. Those mice were used to derive astrocytes (see paragraph 2.1); double homozygous cells were used as experimental population, while cells heterozygous for the *Ezh2* conditional allele were used as controls to monitor Cre toxicity.

1.4.2.Generation of the Rosa26CreERT2 – Ezh2 conditional KO strain

Homozygous mice from both parental strains (Rosa26CreERT2 and Ezh2 conditional KO) were bred; all the pups in F1 were then heterozygous for both mutations. By breeding two mice from F1, it was possible to obtain mice in F2 that were double homozygous. Those mice were bred to heterozygous mice from the parental Ezh2 conditional KO strain; the embryos coming from those breedings were all heterozygous for the Rosa26CreERT2 allele and either homozygous or heterozygous for the Ezh2 conditional allele. Those embryos were used to derive neural precursors (NPs, see paragraph 2.2); cells homozygous for the *Ezh2* allele were used as experimental population, while heterozygous were used as controls to monitor Cre toxicity.

2 Cell culture

2.1 Preparation of astrocytes

I made use of a classic protocol first described in 1980 [McCarthy et al. 1980]. Five to 7-day-old pups were sacrificed, each brain was removed from the head and transferred into a cell culture dish containing sterile Dulbecco's phosphate buffered saline (DPBS – Lonza). Cerebral cortices were isolated from the rest of the brain and placed into a 1.5 ml tube containing 0.5 ml of astrocyte culture medium (15% North American fetal bovine serum, NAFBS – Gibco, 2 mM glutamine – Lonza, 0.6% glucose – Sigma, 100 U/ml potassium penicillin + 100 U/ml streptomycin sulfate – Lonza in Basal Medium Eagle's, BME – Gibco). They were mechanically dissociated with scissors and through pipetting; the dissociated tissue was plated into a T-25 flask and incubated in a humidified incubator at 37°C with 5% CO_2. Medium was changed 3 days after plating and every other day thereafter until day 9 of culture. Cells were passaged every 4 to 5 days starting from day 10 of culture at a 1:2 or 1:3 ratio. The cells were rinsed with DPBS and detached in trypsin-versene 1:10 (Sigma and Lonza, respectively); after detachment, the dissociation into single cells was achieved through pipetting. Trypsin was inactivated by adding fetal bovine serum (FBS – Gibco, 1 volume per volume of trypsin) and cells were collected by centrifugation at 280 × *g* for 5 minutes. The cell pellet was resuspended in fresh astrocyte medium and cells were plated in 6-well-plates.

For the cryopreservation of astrocytes, cells were harvested from a confluent well of a 6-well-plate as described for passaging. The cell pellet was resuspended in freezing medium containing 50% of astrocyte culture medium, 40% of FBS and 10% of dimethylsulfoxide (DMSO – Merck) and put in one cryovial. After 2 days at -80°C in a cryobox, the vials were transferred to liquid nitrogen for long term storage.

One vial of frozen cells was quickly thawed at 37°C in a circulating water bath; cells were transferred to a new 15 ml tube and centrifuged at 280 × *g* for 5 minutes. The cell pellet was resuspended in astrocyte culture medium and plated in 1 well of a 6-well-plate. The following day medium was replaced with fresh medium.

2.2 Preparation of neural precursors (NPs)

Pregnant mice were sacrificed between day 13.5 and 14.5 of embryonic development. The uterus was extracted and washed in Hank's balanced salt solution (HBSS – Gibco). The brain from each embryo was isolated and the telencephalic vesicles transferred to a 1.5 ml tube containing 10% FBS in high glucose Dulbecco's Modified Eagle's medium with GlutaMax™ (hgDMEM – Gibco). Tissue was mechanically dissociated with glass pipettes rinsed with medium and centrifuged at 280 × *g* for 5 minutes. Supernatant was removed and the cell pellet washed in fresh media. After centrifugation (280 × *g* for 5 minutes), cells were resuspended in B27 complete medium consisting of 1x B27 (Gibco), 10 ng/ml of both murine epidermal growth factor (EGF – Peprotech), murine fibroblast growth factor 2 (FGF-2 – Peprotech) and 100 U/ml potassium penicillin + 100 U/ml streptomycin sulfate (Lonza) in hgDMEM with GlutaMax™:F-12 1:1 (both from Gibco) and plated on poly-D-lysine (PDL – Sigma) precoated wells of 24-well plates at the density of 5×10^5 cells per well. Plates were placed in a humidified incubator 37°C with 5% CO_2. At a confluence of about 90%, cells were passaged. They were detached with trypsin-EDTA (Lonza) and dispersed into single cells by gentle pipetting. Trypsin was inactivated by adding hgDMEM with 10% FBS (3 volumes per volume of trypsin) and cells were collected by centrifugation at 280 × *g* for 5 minutes. The cell pellet was resuspended in fresh B27 complete medium and cells were plated as described above.

For the cryopreservation of neural precursors, cells were harvested from a confluent well of a 24-well-plate as described for passaging. The cell pellet was resuspended in freezing medium containing 10% DMSO in B27 complete medium and put in one cryovial. After 2 days at -80°C in a cryobox, the vials were transferred to liquid nitrogen for long term storage.

One vial of frozen cells was quickly thawed at 37°C in a circulating water bath; cells were transferred to a new 15 ml tube and centrifuged at 280 × *g* for 5 minutes. The cell pellet was resuspended in B27 complete medium and plated in 1 well of a 24-well-plate. The following day medium was replaced with fresh medium.

2.3 Viral infection of cell cultures

Epidermal growth factor receptor variant III (EGFRvIII) expressing retroviral particles for astrocyte infection and platelet derived growth factor beta (PDGFB) expressing retroviral particles for neural precursor infection were generated using the retrovirus producer cell line Phoenix-Eco. Phoenix-Eco cells were plated one day before transfection at a density of 1.5×10^6 cells per 10 cm dish. The day of transfection, cells were treated with 25 μM chloroquine (Sigma). In a 15 ml tube, 10 ng of retroviral vector (either EGFRvIII or PDGFB) were mixed to a 250 mM calcium chloride ($CaCl_2$, VWR) solution; this final solution was added dropwise to bubbling 2x HEPES buffered saline pH 7.05 (HBS; 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, HEPES, 10 mM potassium chloride, KCl, 12 mM dextrose, 280 mM sodium chloride, NaCl, 1.5 mM sodium phosphate dibasic heptahydrate, $Na_2HPO_4 \cdot 7H_2O$ – all from VWR). After 15 minutes of incubation at room temperature, the solution was added to the cells and mixed by rocking the plate. Supernatants were collected 48 hours after transfection and every 12 hours thereafter for 4-5 times.

Early-passage astrocyte cultures were infected using the viral supernatant filtered through a 0.45 μm filter. Astrocytes were trypsinized as described before; 5×10^5 cells were resuspended in 3 ml of viral supernatant supplemented with 8 μg/ml Polybrene® (Sigma) and plated in a well of a 6-well-plate. The supernatant was removed and replaced with fresh supernatant at each new collection.

Neural precursors were infected using ultracentrifugated viral supernatant. After the last collection, all the supernatant was filtered through a 0.45 μm filter, placed in ultracentrifuge tubes and ultracentrifugated at 20200 × *g* for 2 hours. After centrifugation, the medium was removed by inverting the tubes and the residual few microliters were used to resuspend the viral particles pelleted at the bottom for 2 hours on ice. Neural precursors were mixed with 50 μl of ultracentrifugated viral preparation during their derivation, immediately before plating.

2.4 Derivation of glioma propagating cells (GPCs)

2.4.1 Derivation of GPCs – astrocytoma model

I made use of a protocol that allows GPC culture in adherent conditions [Pollard et al. 2009]. Mice injected with tumorigenic astrocytes were sacrificed at the onset of neurological symptoms. The tumor mass was dissected from the brain and a small piece was put in a 1.5 ml tube containing 0.5 ml of Accutase (Sigma). The tissue was mechanically dissociated with scissors and through pipetting, passed through a 70 μm cell-strainer and single cell suspension was collected in a 50 ml tube. After centrifugation (300 × *g* for 5 minutes), cells were resuspended in glioma propagating cell medium (1x B27, 1x N2 – Gibco, 20 ng/ml of both EGF and murine FGF-2, 100 U/ml potassium penicillin + 100 U/ml streptomycin sulfate in Neurobasal Medium – Gibco) and plated in 1 or 2 wells of a 6-well-plate, coated with laminin (Roche). Plates were kept in a humidified incubator 37°C with 5% CO_2; medium was changed 2 days after plating and every other day thereafter. Cells were passaged at 80-90% confluence at a ratio ranging from 1:3 to 1:6. They were detached with Accutase and dispersed into single cells by gentle pipetting; after centrifugation at 300 × *g* for 5 minutes, cells were resuspended in glioma propagating cell medium and plated as described above.

For cryopreservation, cells were harvested from a 80-90% confluent well of a 6-well-plate as described for passaging. The cell pellet was resuspended in freezing medium containing 10% DMSO in glioma propagating cell medium and put in one cryovial. After 2 days at -80°C in a cryobox, the vials were transferred to liquid nitrogen for long term storage.

One vial of frozen cells was quickly thawed at 37°C in a circulating water bath; cells were transferred to a new 15 ml tube and centrifuged at 300 × *g* for 5 minutes. The cell pellet was resuspended in glioma propagating cell culture medium and plated in 1 well of a 6-well-plate. The following day medium was replaced with fresh medium.

2.4.2 Derivation of GPCs – oligodendroglioma model

Mice injected with tumorigenic neural precursors were sacrificed at the onset of neurological symptoms. The tumor mass was dissected from the brain under a fluorescence steromicroscope (Olympus SZX16 stereomicroscope, equipped with a Nikon Digital Sight DS-5Mc RGB 8-bit CCD camera,pixel size 7.5 μm; images were acquired with NIS Element software by Nikon and analyzed with ImageJ software) partially dissociated with forceps and put in a 15 ml tube containing 10 mM HEPES in HBSS. After centrifugation (280 × *g* for 5 minutes), supernatant was removed, trypsin-EDTA was added to the tube and incubated for 15 minutes at 37°C. Trypsin was blocked by adding hgDMEM with 10% FBS and the tumor mass was mechanically dissociated with Pasteur's glass pipettes rinsed with medium. The single cell suspension was centrifuged at 280 × *g* for 5 minutes; the cell pellet was resuspended in B27 complete medium and plated on Matrigel™ matrix (BD Biosciences) precoated wells of 6-well plates at the density of 1-1.5 × 10^6 cells per well. Plates were kept in a humidified incubator 37°C with 5% CO_2; medium was changed 2 days after plating and every other day thereafter. Cells were passaged at 80-90% confluency at a ratio ranging from 1 to 3 to 1 to 6. They were detached with Accutase and dispersed into single cells by gentle pipetting; after centrifugation at 280 × *g* for 5 minutes, cells were resuspended in B27 complete medium and plated as described above.

For cryopreservation, cells were harvested from a 80-90% confluent well of a 6-well-plate as described for passaging. The cell pellet was resuspended in freezing medium containing 10% DMSO and 10% FBS in B27 complete medium and put in one cryovial. After 2 days at -80°C in a cryobox, the vials were transferred to liquid nitrogen for long term storage.

One vial of frozen cells was quickly thawed at 37°C in a circulating water bath; cells were transferred to a new 15 ml tube and centrifuged at 280 × *g* for 5 minutes. The cell pellet was resuspended in B27 complete medium and plated in 1 well of a 6-well-plate. The following day medium was replaced with fresh medium.

2.5 Cell treatment with Cre recombinase

2.5.1 Cre-mediated recombination in astrocytes and in corresponding GPCs

To induce Cre-mediated recombination in astrocytes carrying the conditional allele for Ezh2, the cells were treated with a Cre recombinase fused with Human Immunodeficiency Virus (HIV) Transactivator of Transcription (TAT) protein [Peitz et al. 2002]. Astrocyte or GPC cultures were washed 3 times with DPBS. For transduction, purified TAT-Cre protein was diluted in 1 ml of HyClone animal derived component free (ADCF)-MAb™ medium (Thermo Scientific). After 3 hours, medium was replaced with fresh astrocyte culture medium or glioma propagating cell culture medium. The treatment was repeated every day for 5 to 7 consecutive days. To confirm that recombination had occurred, genomic DNA was extracted from cells and tested both by polymerase chain reaction (PCR) and copy-number TaqMan® assay (assay ID: Mm00178344_cn – Applied Biosystems).

2.5.2 Cre-mediated recombination in neural precursors and in corresponding GPCs

Neural precursors were derived from a mouse strain carrying both the conditional allele for Ezh2 and the Cre-ERT2 transgene under the control of the Rosa26 promoter. To activate the Cre recombinase either in the NPs or in the corresponding GPCs, 1 μM 4-hydroxytamoxifen (4-OHT, Sigma) was added to the culture medium every day for 4 to 5 consecutive days. To confirm that recombination had occurred, genomic DNA was extracted from cells and tested both by PCR and copy-number TaqMan® assay (assay ID: Mm00178344_cn – Applied Biosystems).

2.6 Protein staining and analysis

2.6.1 Immunofluorescence (IF) on cultured cells

To perform immunofluorescence on cultured cells, the cells were plated either on glass coverslips placed at the bottom of the culture dish or on plastic or glass Chamber Slides™ (Lab-Tek™ –

Nunc). After removing the culture medium, cells were fixed with 4% paraformaldehyde (PFA) for 15 minutes at room temperature, washed 3 times with DPBS, incubated in permeabilization/blocking buffer (2% bovine serum albumin, BSA – Sigma, 0.1% Triton X-100 – VWR in DPBS) for 15 minutes at room temperature and washed again 3 times in DPBS. Primary antibodies were diluted in 2% BSA DPBS according to manufacturer indications and incubated in a humidified chamber at 4°C overnight. Cells were washed 3 times in DPBS, incubated with specie-specific secondary antibodies conjugated to FITC and Cy3 (Jackson ImmunoReserch), washed again 3 times in DPBS and finally incubated with DAPI:water 1:5000. Slides were mounted with Mowiol® (Calbiochem). The cells were imaged with an Olympus AX70 upright microscope equipped with a Photometrics Coolsnap EZ b/w 12-bit CCD camera (pixel size 6.45 μm) with the software MetaVue 7.5.6.0 (Molecular Devices). Image analysis was performed with ImageJ software.

Antibody	Working dilution	Supplier	Species	Clonality	Code
GFAP	1:500	DAKO	Rabbit	Polyclonal	Z 0334
Nestin	1:500	Millipore	Mouse	Monoclonal	MAB353
OLIG2	1:150	Millipore	Rabbit	Polyclonal	AB9610

Table 2. Antibodies used in immunofluorescence

2.6.2 Fluorescence Activated Cell Sorting (FACS)

Cells were harvested as described for passaging. Cells that were infected with fluorescent-protein expressing viruses were fixed in 4% PFA for 15 minutes at room temperature and directly analyzed; in the case of sorting, the fixation step was omitted. To stain surface antigens, cells were washed in DPBS, pelleted by centrifugation (280 × *g* for 5 minutes) and resuspended in 2% BSA in DPBS. Primary purified antibodies were added according to manufacturer instructions and allowed to bind for 1-2 hours on ice. Cells were then pelleted by centrifugation, washed 3 times in DPBS (each step at 280 × *g* for 5 minutes) and resuspended in 2% BSA in DPBS. They were incubated with secondary antibodies conjugated to FITC and Cy3, washed again 3 times in DPBS and resuspended in 2% BSA in DPBS to be analyzed. FACSCantoII (Becton Dickinson) was used for

population analysis, while FACSAria (Becton Dickinson) was used to sort different populations of cells. Data were analyzed with FlowJo® software (Tree Star inc.).

Antibody	Working dilution	Supplier	Species	Clonality	Code
A2B5	1:200	Millipore	Mouse	Monoclonal	MAB312R
CD133	1:100	eBioscience	Rat	Monoclonal	13-1331

Table 3. Antibodies used for FACS

2.6.3 Immunoblotting

The cells were harvested as described for passaging. After centrifugation, the supernatant was removed, cell pellet was lysed in an appropriate volume of radio-immuno-precipitation assay buffer (RIPA buffer, 150 mM NaCl, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris pH 8) supplemented with protease inhibitors (2 μg/ml aprotin, 5-10 μg/ml leupeptin, 1 mM PMSF) and then sonicated in a 4°C water-bath sonicator (Bioruptor® Diagenode, high power, 3 cycles of 30 s pulse and 1 minute pause each) to achieve complete lysis of the cells and to shear genomic DNA. Protein concentration was determined by a Bradford assay using BSA as a protein standard; 2 μl of lysate were diluted in 800 μl of water plus 200 μl of Bio-Rad Protein Assay (Biorad) and the absorbance at 595 nm was measured with the spectrophotometer. Proteins were separated according to their molecular weight by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Fifty ng of proteins mixed with 6x Laemmli buffer (375 mM Tris HCl pH 6.8, 9% SDS, 50% glycerol, 9% beta-mercaptoethanol , 0.03% bromophenol blue) were boiled at 95°C for 5 minutes to achieve denaturation and loaded on a gel; run was performed in a standard migration buffer (25 mM Tris base, 190 mM glycine, 0.1% SDS). Proteins were transferred to a nitrocellulose membrane using a standard transfer buffer (25 mM Tris base, 190 mM glycine, 20% methanol). Efficiency of the transfer was checked by staining the membrane with Ponceau red (VWR); the membrane was blocked in 5% (w/v) skimmed milk powder in Tris buffered saline with 0.1% Tween 20 (TBS-T, 25 mM Tris, 150 mM NaCl, 2 mM KCl) for one hour at room temperature. Primary antibodies were diluted at the proper concentration in a fresh 5%

milk solution and allowed to bind overnight at 4°C under agitation. The day after, the membrane was washed 3 times for 5 minutes in TBS-T; the secondary antibody was then added in a fresh 5% milk solution and allowed to bind for 1 or 2 hours at room temperature. The membrane was washed again 3 times for 5 minutes in TBS-T and bands were detected with ECL (GE Amersham). Radiographic films were used to image the bands.

Antibody	Working dilution	Supplier	Species	Clonality	Code
EGFR	1:1000	Home made	Rabbit	Polyclonal	-
EZH2	1:1000	Home made	Mouse	Monoclonal	-
Vinculin	1:5000	Sigma	Mouse	Monoclonal	V 9131

Table 4. Antibodies used in immunoblotting

3 DNA methods

3.1 DNA extraction

3.1.1 Isolation of genomic DNA from mouse tails

I made use of a classic protocol described in 1991 [Laird et al. 1991]. Tail biopsies were taken from pups at weaning age by the IFOM-IEO Animal House facility staff. They were incubated in 0.5 ml of lysis buffer (100 mM Tris-HCl, pH 8.5 – VWR, 5 mM EDTA – VWR, 0.2% SDS – BDH, 200 mM NaCl – VWR and 100 μg/ml proteinase K – Sigma) in a thermomixer at 37°C, 800 rpm for 3 to 6 hours. Lysate was centrifuged at 16000 × g for 1 minute in order to remove non-digested tissue; half of the lysate was put in a new tube and one volume of isopropanol (Sigma) was added to precipitate DNA. After centrifugation at 16000 × g for 10 minutes, the DNA pellet was washed with 70% ethanol (Sigma) and centrifuged again. Finally, the dried DNA pellet was resuspended in an appropriate volume of water or 1x Tris-EDTA buffer (TE, 10 mM Tris-HCl pH 8.0, 1 mM EDTA). The remaining half of the lysate was put at 94°C for 10 minutes to heat-inactivate the proteinase K. 1 μl either of the crude lysate diluted 1:10 or of the precipitated DNA was used for genotyping.

3.1.2 DNA purification using phenol-chloroform extraction

To remove protein contaminants and to obtain high-molecular weight molecules, DNA samples were purified by means of phenol-chloroform extraction. An equal volume of phenol was added to the lysate (generally obtained from cell pellets, using the same lysis buffer described before) and the phases were mixed by vortexing. After centrifugation at 16000 × g for 5 minutes, the aqueous upper layer containing DNA was transferred to a new tube. An equal volume of phenol:chloroform 1:1 was added, mixed by vortexing and centrifugated as in the previous step. The aqueous phase was transferred to a new tube and one volume of chloroform:isoamyl alcohol 24:1 was added to remove any trace of phenol. After mixing and centrifugation, the aqueous phase was transferred to a new tube and DNA was precipitated by ethanol precipitation. NaCl was added to the aqueous solution to a final concentration of 200 mM; 2 volumes of cold 100% ethanol were added and mixed by inversion. After centrifugation at 16000 × g for 10 minutes at 4°C, the DNA pellet was washed with 70% ethanol and centrifuged again. The DNA pellet was resuspended in an appropriate volume of water or TE buffer and quantified using a Nanodrop (Thermo Scientific).

3.1.3 DNA purification using Qiagen DNeasy® blood and tissue kit

Samples were first lysed using proteinase K. Buffering conditions were adjusted to provide optimal DNA binding conditions and the lysate was loaded onto the column. During centrifugation, DNA was selectively bound to the membrane in the column while contaminants passed through. Remaining contaminants and enzyme inhibitors were removed in two efficient wash steps and DNA was then eluted in water or buffer, ready for use. (Adapted from Qiagen DNeasy blood and tissue kit handbook, 07/2006)

3.2 Polymerase chain reaction (PCR)

3.2.1 Genotyping PCR

PCR reactions were usually carried out in a total volume of 25 μl with 0.8 mM dNTPs (0.2 mM each), 2 μM of each primer, 0.02 U/μl DNA polymerase, DNA template (in different amount depending on the source) and the appropriate buffer provided by the supplier. The reaction conditions varied among the experiments, but generally followed this cycling steps: 3 minutes of initial denaturation at 95°C, 35 cycles each of which consisting of 30 seconds denaturation at 95°C, 30 seconds of annealing (temperature optimized for each primer combination) and variable extension time according to amplicon length at 72°C and, after the last cycle, a final extension step of 7 minutes at 72°C.

PCR products were run on agarose gels containing 0.5 μg/ml ethidium bromide in Tris-acetate-EDTA (TAE) buffer.

Gene/Locus	Primer name	Primer sequence (5'-3')
Ezh2	Ezh2F	TTATTCATAGAGCCACCTGG
	Ezh2R	CTGCTCTGAATGGCAACTCC
Rosa26CreERT2	RosawtF	AAAGTCGCTCTGAGTTGTTAT
	RosawtR	GGAGCGGGAGAAATGGATATG
	RosacreR	CCTGATCCTGGCAATTTCG
Ink4a/Arf	19C10	TTAACAGCGGAGCTTCGTACA
	20C10	CTGCACCGTAGTTGAGCAGAA
	2B2	GCCTACCCGCTTCCATTGCT

Table 5. Primers used for genotyping

3.2.2 Quantitative PCR for copy number assay

10 ng of genomic DNA was amplified in a reaction volume of 15 μl containing the following reagents: 7.5 μl of TaqMan® PCR Mastermix 2x No UNG (Applied Biosystems), 0.75 μl of TaqMan® Copy Number Assay 20x (Applied Biosystems) and 0.75 μl of TaqMan® Copy Number Reference Tert 20x (Applied Biosystems). Real-time PCR was carried out on the ABI/Prism 7900 HT Sequence Detector System (Applied Biosystems), using a pre-PCR step of 10 min at 95°C, followed by 40

cycles of 15 seconds at 95°C and 60 seconds at 60°C. Data were analyzed with CopyCaller 1.0 software (Applied Biosystems).

3.3 Plasmid preparation

3.3.1 Bacterial transformation by heat shock

Chemically competent bacteria (DH5-alpha *E. Coli*) were thawed on ice. 25 ng of DNA were added to the cells and the tube was incubated on ice for 30 minutes. The tube was then transferred to a preheated 42°C circulating water bath, incubated for exactly 90 seconds and rapidly moved to an ice bath, allowed to cool for 1-2 minutes. Eight hundreds μl of Luria Broth (LB) medium were added to the tube and the culture was then kept at 37°C for 45 minutes under gentle agitation (225 rpm). Hundred μl of transformed cells were then plated on agar-LB containing the appropriate antibiotic; plates were inverted and incubated over night at 37°C.

3.3.2 Plasmid preparation using Qiagen QIAprep® Spin Miniprep kit (miniprep)

The QIAprep miniprep procedure uses the modified alkaline lysis method of Birnboim and Doly [Birnboim et al. 1979]. Bacteria were lysed under alkaline conditions, and the lysate was subsequently neutralized and adjusted to high-salt binding conditions. The columns use a silica membrane for selective adsorption of plasmid DNA in high-salt buffer and elution in low-salt buffer. Endonucleases and salts were efficiently removed by a brief wash step. High-quality plasmid DNA was then eluted from the column with appropriate buffer or water. (Adapted from Qiagen QIAprep Miniprep kit handbook, 12/2006)

3.3.3 Plasmid preparation using QIAGEN® Plasmid Purification kit (maxiprep)

Plasmid purification protocol is based on a modified alkaline lysis procedure as described for minipreps, followed by binding of plasmid DNA to anion-exchange resin under appropriate low-

salt and pH conditions. RNA, proteins, dyes, and low-molecular-weight impurities were removed by a medium-salt wash. Plasmid DNA was eluted in a high-salt buffer and then concentrated and desalted by isopropanol precipitation. (Adapted from Qiagen Plasmid Purification kit handbook, 11/2005)

4 RNA methods

4.1 RNA extraction

4.1.1 Total RNA extraction with TRIzol® Reagent

Cell pellets were lysed by adding 1 ml of TRIzol® Reagent (Life Technologies) to 5-10 x 10^6 cells and incubated for 5 minutes at room temperature. 0,2 ml of chloroform were added per ml of TRIzol® Reagent, mixed vigorously and the samples were incubated at room temperature for 2 to 3 minutes. After centrifugation at 12000 × g for 10 minutes at 4 °C, the upper phase was transferred to a new tube. 0.5 ml of isopropanol per ml of TRIzol® Reagent were added, samples were incubated at room temperature for 5 minutes and then centrifuged at 12000 × g for 10 minutes at 4 °C. Supernatant was discarded and the RNA pellet was washed with 75% ethanol; after centrifugation at 7500 × g for 5 minutes at 4 °C, ethanol was discarded and the pellet dried for 2-3 minutes. The RNA pellet was resuspended in 30 μl of RNAse-free water.

4.1.2 Total RNA extraction with Qiagen RNeasy® Plus Mini kit

Biological samples were first lysed and homogenized in a highly denaturing guanidine-isothiocyanate–containing buffer, which immediately inactivates RNases to ensure isolation of intact RNA. The lysate was then passed through a genomic DNA eliminator column. This column, in combination with the optimized high-salt buffer, allowed efficient removal of genomic DNA. Ethanol was added to the flow-through to provide appropriate binding conditions for RNA, and the sample was then applied to a column, where total RNA binds to the membrane and

contaminants are efficiently washed away. High-quality RNA was then eluted in water. With this procedure, all RNA molecules longer than 200 nucleotides are isolated. The procedure provides an enrichment for mRNA, since most RNAs <200 nucleotides (such as 5.8S rRNA, 5S rRNA, and tRNAs, which together comprise 15–20% of total RNA) are selectively excluded. (Adapted from Qiagen RNeasy® Plus Mini kit handbook, 09/2010)

4.2 cDNA synthesis

cDNA synthesis was performed according to SuperScript® VILO™ cDNA Syntesis kit according to the instructions of the manufacturer. Briefly, the following components were combined in a tube: 4 μl of 5X VILO™ reaction mix (which includes random primers, MgCl2 and dNTPs in a buffer formulation optimized for qRT-PCR), 2 μl of 10X SuperScript® enzyme mix (which includes SuperScript® III reverse transcriptase, RNaseOUT™ recombinant ribonuclease inhibitor and an helper protein), RNA up to 2.5 μg (generally 1 μg) and water up to 20 μl. Tubes were incubated at 25°C for 10 minutes, at 42°C for 60 minutes and finally at 85°C for 5 minutes to terminate the reaction. cDNA can be used diluted or undiluted in qPCR analysis.

4.3 Quantitative real-time PCR (qRT-PCR)

4.3.1 qRT-PCR with TaqMan® assay

cDNA was diluted to achieve a concentration equivalent of starting with 100 ng of RNA in the reverse transcription reaction. Five ng of cDNA was amplified (in triplicate) in a reaction volume of 15 μl containing the following reagents: 7.5 μl of TaqMan® PCR Mastermix 2x No UNG (Applied Biosystems), 0.75 μl of TaqMan® Gene Expression Assay 20x (Applied Biosystems). Real-time PCR was carried out on the ABI/Prism 7900 HT Sequence Detector System (Applied Biosystems), using a pre-PCR step of 10 min at 95°C, followed by 40 cycles of 15 seconds at 95°C and 60 seconds at 60°C. Preparations with RNA template without reverse transcriptase were used as negative

controls. Samples were amplified with specific primers for each gene and TATA binding protein (*Tbp*) as housekeeping gene. Each sample was analyzed in triplicate and normalized to *Tbp*. Expression level for each gene is calculated using the comparative cycle threshold method (ΔCt) and relative mRNA amount is reported as $2^{-\Delta Ct}$.

Gene/Locus	TaqMan® assay ID
Aldh1l1	Mm00550947_m1
Aqp4	Mm00802131_m1
Ezh2	Mm00468464_m1
Gfap	Mm01253033_m1
Mbp	Mm01266402_m1
Nestin	Mm00450205_m1
Pdgfra	Mm00440701_m1
Slc12a5	Mm00803929_m1
Slc1a3	Mm00600697_m1
Snap25	Mm00456921_m1
Sox10	Mm01300162_m1
Syt1	Mm00436858_m1
Tbp	Mm00446973_m1

Table 6. TaqMan® assays used in qRT-PCR

5 Hi-throughput experiments

5.1 Chromatin immunoprecipitation followed by deep-sequencing (ChIP-seq)

5.1.1 ChIP-seq

Cells were cross-linked with 1% formaldehyde in DPBS for 10 minutes at room temperature. Fixation was quenched by addition of glycine to a final concentration of 0.125 M for 5 minutes at room temperature. Cells were washed 2 times in DPBS and then harvested in SDS buffer (0.5% SDS, 50 mM Tris HCl pH8.1, 100 mM NaCl, 5 mM EDTA pH 8, 0.02% NaN_3) with protease inhibitors (2 μg/ml aprotin, 5-10 μg/ml leupeptin, 1 mM PMSF). After centrifugation at 2000 rpm for 6 minutes, cells were resuspended in 3 ml of SDS buffer:Triton Dilution buffer 1:0.5 (Triton Dilution buffer: 5% Triton X-100, 100 mM Tris HCl pH 8.6, 100 mM NaCl, 5 mM EDTA pH 8, 0.02% NaN_3) and sonicated on ice with a large tip in a Branson Sonifier (20-30 seconds with 2-3 minutes rest in between, power set at 30%). Sonicated chromatin is centrifuged at 20000 × g for 30 minutes at

4°C and protein content is quantified using Bradford assay (see paragraph 2.6.3). Hundred μg are diluted in 1 ml of SDS buffer:Triton Dilution buffer 1:0.5; 10 μl are taken to be used as total control (1% input).

Primary antibodies were added according to manufacturer instructions and incubated overnight at 4°C on a rotating wheel. ChIP complexes were collected by incubating with 40 μl of protein A sepharose beads (GE Healthcare) for 2-4 hours at 4°C on a rotating wheel; ChIP-bead complexes were pelleted 2000 rpm and washed 3 times in 150 mM NaCl Wash buffer (150 mM NaCl, 1% Triton X-100, 20 mM Tris HCl pH 8.0, 0.1% SDS, 2 mM EDTA pH 8) and 1 time in 500 mM NaCl Wash buffer (500 mM NaCl, 1% Triton X-100, 20 mM Tris HCl pH 8.0, 0.1% SDS, 2 mM EDTA pH 8). ChIP-bead complexes and the 1% input were incubated in 120 μl of 1% SDS, 0.1 M $NaHCO_3$ at 65°C overnight to reverse crosslink of proteins to DNA. DNA was purified with the QIAquick PCR purification kit (Qiagen) following manufacturer instructions and eluted in 100 μl of provided buffer EB.

Antibody	Working dilution	Supplier	Species	Code
H3K27me3	2,7 μg / 100 μg of chromatin	Cell Signaling	Rabbit (M)	9733
H3	2 μg / 100 μg of chromatin	Abcam	Rabbit (P)	ab1791
IgG	1 μg / 100 μg of chromatin	Sigma	Rabbit	I8140

Table 7. Antibodies used in chromatin immunoprecipitation (M: monoclonal, P: polyclonal).

For the sequencing, samples were quantified and given to the Illumina facility of the IFOM-IEO Campus which performed the following part of the protocol; the starting amount of DNA was 10 ng. The overhangs of the DNA fragments were converted into phosphorylated blunt ends, using T4 DNA polymerase, E. coli DNA polymerase I large fragment (Klenow polymerase), and T4 polynucleotide kinase. The 3' to 5' exonuclease activity of these enzymes removed 3' overhangs and the polymerase activity filled in the 5' overhangs. A single 'A' nucleotide was added to the to the 3' end of the blunt phosphorylated DNA fragments, using the polymerase activity of Klenow fragment (3' to 5' exo minus). This prepared the DNA fragments for ligation to the adapters, which have a single 'T' base overhang at their 3' end. Adapters were ligated to the ends of the DNA

fragments, preparing them to be hybridized to a flow cell. DNA was run on a TAE 2% agarose gel to remove excess adaptors and selects a size range of templates; a gel slice containing the material in the 200±25 bp range was cut from the gel and purified with QIAquick Gel Extraction Kit (Qiagen) according to manufacturer instructions. The adapter-modified DNA fragments were enriched by PCR amplification using the gel-extracted DNA. (Adapted from ChIP-seq Sample Preparation Guide 2007, Illumina). Sequencing was performed on Illumina Genome Analyzer II, using 36 bp paired end reads.

5.1.2 Bioinformatic analysis of ChIP-seq

Raw data from the sequencing experiments were aligned to the reference mouse genomic sequence (NCBI37/mm9, July 2007) with BowTie v.0.12.7 [Langmead et al. 2009] allowing up to two mismatches per read and discarding multiply-aligning reads. To identify enriched domains we used MACS v.1.4.0 [Zhang et al. 2008] which estimates the enrichment of the specific signal against a noise signal (input). Enriched regions were assigned to genes by determining, for each RefSeq transcript, whether an enriched domain overlapped with a +/-5kb region around the transcription start site (TSS).

5.2 Whole Transcriptome Shotgun Sequencing

5.2.1 RNA-seq

Total RNA was extracted with the Qiagen RNeasy® Plus Mini kit (see paragraph 4.1.2).

The following part of the protocol was performed by the Illumina facility of the IFOM-IEO Campus; the starting amount of total RNA was 4 μg. From the total RNA, the poly-A containing mRNA molecules were purified using poly-T oligo attached magnetic beads using two rounds of purification. During the second elution of the poly-A RNA, the RNA was also fragmented using divalent cations under elevated temperature and primed for cDNA synthesis with random

hexamers. The primed and cleaved RNA fragments were reverse into first strand cDNA using reverse transcriptase and random primers. The RNA template was removed with RNase H and a replacement strand was synthesized with DNA Polymerase I to generate double-strand (ds) cDNA. Beads (AMPure XP beads) are used to separate the ds cDNA from the second strand reaction mix. Overhangs resulting from fragmentation were converted into blunt ends using an End Repair Mix: the 3' to 5' exonuclease activity of this mix removes the 3' overhangs and the polymerase activity fills in the 5' overhangs. A single 'A' nucleotide was added to the 3' ends of the blunt fragments to prevent them from ligating to one another during the adapter ligation reaction. A corresponding single 'T' nucleotide on the 3' end of the adapter provided a complementary overhang for ligating the adapter to the fragment. This strategy ensures a low rate of chimera (concatenated template) formation. Multiple indexing adapters were ligated to the ends of the ds cDNA, preparing them for hybridization onto a flow cell. PCR was used to selectively enrich those DNA fragments that have adapter molecules on both ends and to amplify the amount of DNA in the library. Fragments with only one or no adapters on their ends are by-products of inefficiencies in the ligation reaction. Neither species can be used to make clusters, as fragments without any adapters cannot hybridize to surface-bound primers in the flow cell, and fragments with an adapter on only one end can hybridize to surface bound primers but cannot form clusters. The PCR was performed with a PCR primer cocktail that anneals to the ends of the adapters. (Adapted from TruSeq® RNA Sample Preparation v2 Guide 05/2012, Illumina). Sequencing was performed on Illumina I-seq 2K, using 50 bp paired end reads.

5.2.2 Bioinformatic analysis of RNA-seq

Raw data from the sequencing experiments were aligned to the reference mouse genomic sequence (NCBI37/mm9, July 2007) with TopHat-CuffLink software pipeline [Trapnell et al. 2012]. TopHat is a program that aligns RNA-Seq reads to a genome in order to identify exon-exon splice junctions. TopHat finds splice junctions without a reference annotation. By first mapping RNA-Seq

reads to the genome, TopHat identifies potential exons, since many RNA-Seq reads will contiguously align to the genome. Using this initial mapping information, TopHat builds a database of possible splice junctions, and then maps the reads against these junctions to confirm them. Short read sequencing machines can currently produce reads 100bp or longer, but many exons are shorter than this, and so would be missed in the initial mapping. TopHat solves this problem by splitting all input reads into smaller segments, and then mapping them independently. The segment alignments are "glued" back together in a final step of the program to produce the end-to-end read alignments.

The alignment is then revised by the software Cufflink, which generated enrichment maps used to quantify the abundance of the transcripts. Cufflinks assembles transcripts, estimates their abundances, and tests for differential expression and regulation in RNA-Seq samples. It accepts aligned RNA-Seq reads and assembles the alignments into a parsimonious set of transcripts. Cufflinks then estimates the relative abundances of these transcripts based on how many reads support each one, taking into account biases in library preparation protocols. Replicates (e.g.: all the data from different samples of astrocytes) were analyzed together and a p-value was associated at each transcript; the lower the p-value, the more expression values were similar among the samples. After this first step, through the CuffDiff plug-in, different group of samples were compared (e.g.: astrocytes versus gliomas) and a second p-value was generated; the lower the p-value, the more expression levels were different between the groups. Data on differential expression were then validated by qRT-PCR.

5.3 Functional Annotation Analysis (FAA)

5.3.1 Gene Ontology analysis

Cytoscape software and the BiNGO plugin [Maere et al. 2005] were used to analyze Gene Ontology terms, particularly biological processes and molecular functions, significantly associated to the various gene sets. The enrichment for each term was tested using hypergeometric test and

P values were corrected using FDR procedure. All terms with a FDR <0.05 were considered associated.

5.3.2 Analysis of KEGG pathway enrichment

DAVID online analysis tools [Huang et al. 2009] were used to interrogate KEGG database in order to find annotated pathways significantly associated to the various gene sets. The enrichment for each term was tested using hypergeometric test and *P* values were corrected using FDR procedure. All the pathways with a FDR <0.05 were considered associated.

6 Stereotaxic injection

Tumor initiation assays were carried out using stereotaxic apparatus to inject cells into the brain of immunocompromised recipient mice (CD1® nude HO – Charles River). Cells were harvested as described for passaging. Astrocytes and GICs were counted with a hemocytometer, pelleted by centrifugation and resuspended in an appropriate volume of DPBS in order to have 2-2.5 × 10^4 cells per microliter. NPs were counted with an hemocytometer under a fluorescence microscope (Olympus IX81 inverted microscope, equipped with a Photometrics Coolsnap EZ b/w 12-bit CCD camera, pixel size 6.45 μm; images were acquired with MetaMorph 7.5.6.0 software by Molecular Device and analyzed with ImageJ software) in order to determine the percentage of infected cells; after centrifugation, cells were resuspended in an appropriate volume of DPBS in order to have 5 × 10^3 infected cells per microliter. Each animal was anaesthetized by intraperitoneal administration of tri-bromo-ethanol (avertin) at the concentration of 0.25-0.5 mg/g. The mouse was then placed into a stereotaxic frame to determine the exact coordinates for the site of injection on the 3 axis (x,y,z), as represented in Figure 4. The skin of the head was cut with a scalpel and the needle tip (gauge 26s, point style 2) of a Hamilton syringe was aligned to the bregma; the x and y coordinates were measured and in that point a hole was drilled in the skull with a dentist drill. The syringe was loaded with cells resuspended in DPBS (3 to 5 μl) and the needle was introduced in the skull hole according to the z coordinate. Cells were injected in two

steps (half of the volume at each step), with a 1 minute pause in between. One minute after the second step, the needle was pulled up by half of the z coordinate and after 3 more minutes it was pulled out from the skull. The wound of the skin was sealed with a surgical clip.

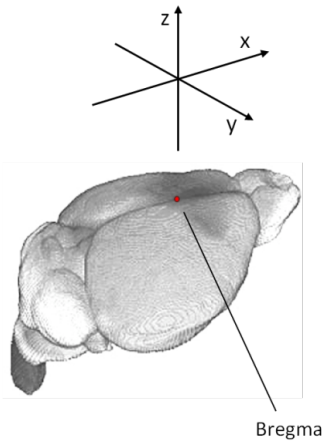

Figure 4. Determination of stereotaxic coordinates.

7 Histopathology

7.1 Brain processing

7.1.1 Paraffin-embedding of adult brain

Mice were sacrificed at the onset of frank neurological symptoms. The brain was taken, rinsed quickly in PBS to remove blood and then put overnight in formalin for fixation. The following day, it was cut in 1 mm- thick slices which were in turn placed in histological cassettes. The tissue was processed in an automatized tissue processor. After processing, the slices were embedded in paraffin and stored at room temperature; paraffin blocks were cut with a microtome (Leica RM 1450 S) in 4 μm thick sections that were attached onto glass slides for further analysis.

7.1.2 OCT- embedded brain

Mice were sacrificed at the onset of frank neurological symptoms. The brain was taken and rinsed quickly in DPBS to remove blood and then put overnight in PFA for fixation. The following day, the brain was put in a solution of 30% (w/v) sucrose (Sigma) until sank (generally overnight). It was

then transferred into a 15 ml tube and incubated in a 1:1 mixture of OCT and 30% (w/v) sucrose on a rotating wheel at room temperature for one hour. The brain was then embedded and frozen in OCT on dry-ice and placed at -80 °C for long term storage. The night before cutting the brain, it was moved to -20 °C to equilibrate the OCT at the cutting temperature. 12 μm thick sections were cut with a Leica cryostat and placed onto SuperFrost® Microscope Slides (Thermo Scientific) for further analysis.

7.2 Staining

7.2.1 Hematoxylin – eosin staining

Hematoxylin – eosin staining was performed in the Anatomo-pathology Unit of the S. Raffaele Institute using an automated colorator (Leica ST 5020). Briefly, glass slides with paraffin sections were deparaffinized, re-hydrated, coloured, de-hydrated and finally mounted with a toluemic-based non-aqueous glue. Glass slides were washed for 3 minutes in Bioclear (Bio Optica) for 4 times to achieve deparaffinization; they were then soaked in a descending alcohol scale, 2 times in 99% ethanol, 2 times in 95% ethanol and 1 time in 70% ethanol, 1.5 minutes for each step followed by a longer step of 3 minutes in 70% ethanol. Slides were moved into hematoxylin for 1.5 minutes, allowed to change colour for 5 minutes in water and stained with eosin for 3 minutes. De-hydration was performed by soaking the slides once in 95% ethanol and 2 times in 99% ethanol, each step lasting 1.5 minutes. To clarify the staining, 2 steps of 1.5 minutes were performed in Bioclear. After the staining, slides are mounted with Mowiol by an automated glass-mounter Leica CV 5030.

7.2.2 Immunohistochemistry

Paraffin-embedded sections were cut as previously described. Sections were washed in Xylol (Sigma) for 30 minutes to remove paraffin and hydrated in a descending alcoholic scale with 99%,

95% and 70% ethanol (10 minutes for each step). The antigen is unmasked in Tris-EDTA pH 9 in a water-bath at 97°C for 30 minutes; endogenous peroxidase is inhibited with 3% hydrogen peroxide (H_2O_2) for 5 minutes. Sections were then incubated in blocking buffer for 10 minutes; primary antibodies were diluted in 2% BSA in DPBS and allowed to bind for 1 hour at room temperature. After washing, sections are incubated with the secondary antibodies for half an hour at room temperature. Staining is contrasted by hematoxylin staining (10 s); sections are then de-hydrated (95% and 99% ethanol, 10 minutes each) and glass are mounted with 90% glycerol in DPBS.

Antibody	Working dilution	Supplier
Ezh2	1:500	Novocastra
Ki-67	1:500	DAKO

Table 8. Antibodies used in immunohistochemistry

7.2.3 Immunofluorescence

Paraffin-embedded sections were cut as previously described. Sections were washed in Xylol (Sigma) for 30 minutes to remove paraffin and hydrated in a descending alcoholic scale with 99%, 95% and 70% ethanol (10 minutes for each step). The antigen is unmasked in Tris-EDTA pH 9 in a water-bath at 97°C for 30 minutes. Sections were then incubated in blocking buffer for 10 minutes; primary antibodies were diluted in 2% BSA in DPBS and allowed to bind for 1 hour at room temperature. Sections were washed and incubated with specie-specific secondary antibodies conjugated to Alexa fluor 555 or 488. Nuclei are stained with DAPI and glasses are mounted with 90% glycerol in DPBS. Imaging was done as described for immunofluorescence (see paragraph 2.6.1).

Antibody	Working dilution	Supplier
Ezh2	1:500	Novocastra

Table 9. Antibodies used in immunofluorescence

7.3 Histopathological classification

Tumors were classified according to the World Health Organization (WHO) classification following the guidelines described in “WHO classification of tumours of the central nervous system” - Lyon, 2007.

Results

1. The *Ink4a/Arf*$^{-/-}$/EGFRvIII overexpression tumor model

The first model I exploited to pursue the objectives of my project is a well established model of gliomagenesis, that relies on the loss of *Ink4a/Arf* combined with the over expression of EGFRvIII [Bachoo et al. 2002]. The EGFRvIII is vital, as it has been demonstrated that p16^{INK4a}/p19$^{ARF-/-}$ mice develop spontaneous sarcomas and lymphomas, but this mutation alone is not sufficient to drive glioma [Serrano et al. 1996]. This model recapitulates well the human counterpart, both at the genetic and histological level, since the association of those two mutations is frequent in human gliomas [von Deimling et al. 1992, Hayashi et al. 1997, Hegi et al. 1997, TCGAN 2008]. Moreover, the combination of these mutations was shown to provoke high-grade gliomas with 100% penetrance.

1.1 Characterization of astrocyte cultures

I started off by validating the set up of the experimental system. For this purpose p16/p19$^{-/-}$ cortical astrocytes were derived from pups 5 days after birth following the same protocol used in the original study by Bachoo et al. (see Material and Methods for detailed description) [McCarthy et al. 1980]. Cultures were characterized by checking the expression at the mRNA level of several markers that were shown to be specific for different central nervous system cell types [Cahoy et al. 2008]. I assessed expression of four astrocyte-specific (*Gfap*, *Aqp4*, *Aldh1l1*, *Slc1a3*), three neuronal-specific (*Syt1*, *Slc12a5*, *Snap25*), three oligodendroglial-specific markers (*Pdgfra*, *Sox10*, *Mbp*) and *Nestin* (Figure 5). Astrocyte markers are highly expressed indicating that the majority of the cells in culture belong to the astroglial lineage, but also *Pdgfra* (oligodendrocyte precursor markers) and *Nestin* are expressed though at a low level.

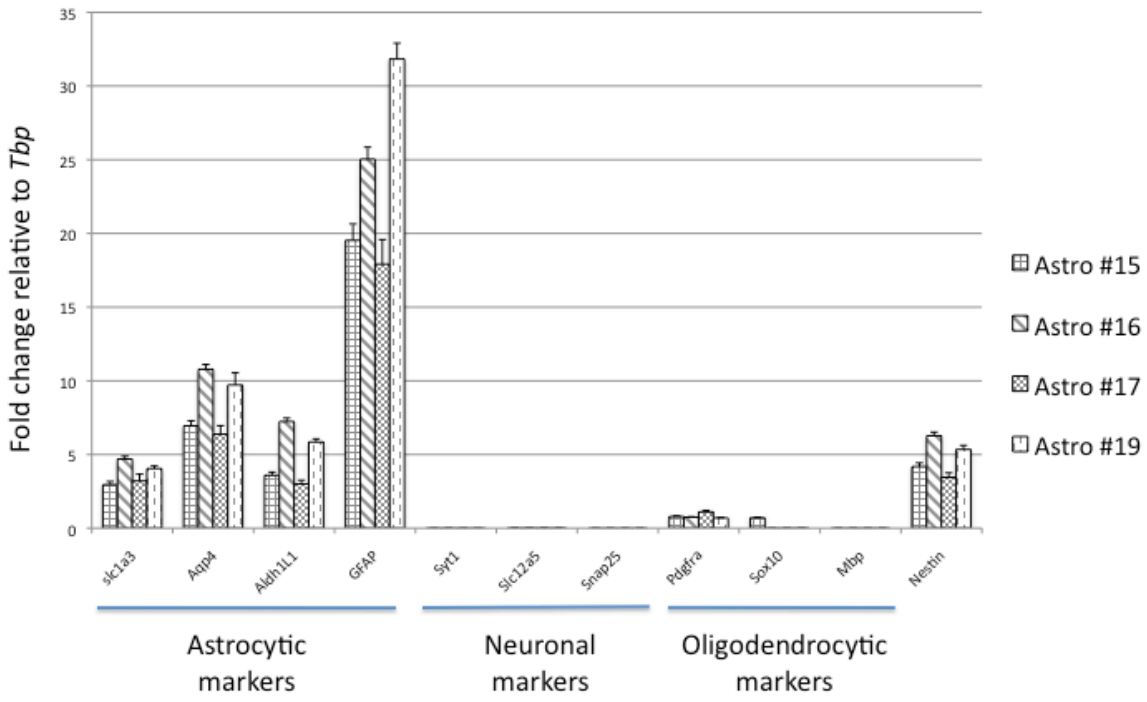

Figure 5. Astrocyte specific markers are highly expressed by the cells in culture. TaqMan assay on 4 different astrocytic cell lines revealed that cells in culture are expressing high levels of astrocyte-specific genes (*Gfap*, *Aqp4*, *Aldh1l1*, *Slc1a3*) and low levels of *Pdgfra* (marker of oligodendrocyte precursors) and *Nestin*. Other oligodendrocytic markers corresponding to a more differentiated stage (i.e. myelinating oligodendrocytes) and neuronal markers are not expressed. mRNA levels are calculated for each gene using the comparative cycle threshold method (ΔCt) and relative mRNA amount compared to *Tbp* is reported as $2^{-\Delta Ct}$. Bars represent the positive standard deviation (S.D.).

Astrocyte cultures showed a mild grade of heterogeneity in the morphology of the cells (figure 6, a, c, e). This was not completely unexpected since maturation of astrocytes in vivo is achieved towards the 4th week after birth [Baumann et al. 2001] and, with this protocol, I derived cells within the 1st week after birth. In order to characterize my cultures, I performed immunocytochemistry for glial fibrillary acidic protein (GFAP), nestin and OLIG2 (figure 6, b, d, f). OLIG2 is a transcription factor required for oligodendrocyte differentiation that is expressed in oligodendrocyte precursors [Lu et al. 2000, Zhou et al. 2000]. Nestin is a member of type VI intermediate filament (IF) protein family [Lendhal et al. 1990]. It starts being expressed in neural stem cells throughout the developing neuraxis during embryonic development and it is down-regulated postnatally, when it is replaced by cell-type specific IFs. Specifically, during gliogenesis it is gradually replaced by GFAP, a type III IF protein. Nestin is currently used as a neural stem cell or neural progenitor marker, while GFAP identifies astroglial and radial glial cells. In my cultures a large proportion of the cells were GFAP^{+}, but it was nevertheless possible to find clusters of

nestin^{+} or OLIG2^{+} cells. It was previously shown that nestin^{+} neural progenitors are found in astrocyte cultures established from newborn rodent cortices [Sergent-Tanguy et al. 2006]. In addition, OLIG2 is expressed in immature cortical astrocytes during early postnatal stages [Cai et al. 2007].

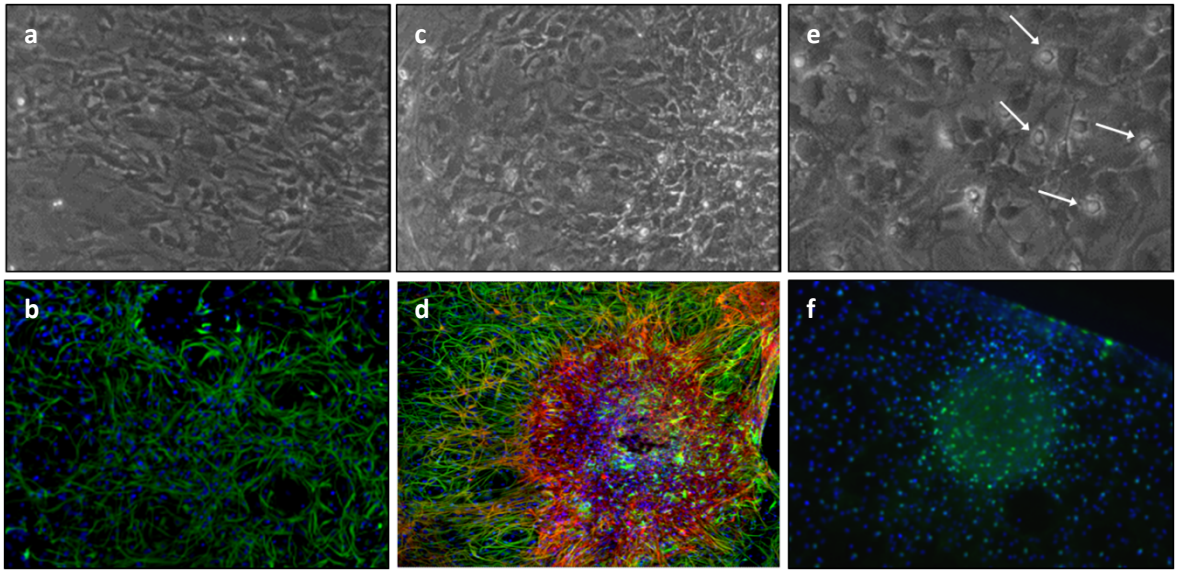

Figure 6. The astrocyte culture is mildly heterogeneous. Morphologically, the astrocyte culture appears to own a minor level of heterogeneity. Astrocytes (**a**), stained for GFAP (**b**), represent the most abundant population. Sparse clusters of immature cells from which more differentiated cells are spreading out (**c**) are identified through nestin staining (**d**; nestin: red, GFAP: green). Oligodendrocyte precursors (**e**, arrows) and OLIG2^{+} cells (**f**) are also found. (10x in all the panels except for panel **e**, 20x)

1.2 Orthotopic transplantation of astrocytes into immunocompromised mice and tumor formation

In order to introduce the second alteration upon the constitutive p16/p19 knockout to fully transform astrocytes, cells were infected with a retroviral vector encoding the mutant form of human EGFR, namely hEGFRvIII and expression of this was confirmed by immunoblotting (Figure 7).

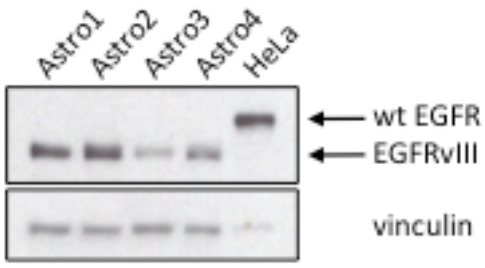

Figure 7. EGFRvIII is highly expressed in astrocyte cultures upon infection. Western Blot analysis on 4 different astrocytic cell lines (Astro 1-4) showed that hEGFRvIII is expressed upon infection (HeLa extract is used as positive control for the detection of human EGFR).

In order to assess the tumorigenicity of the transformed astrocytes, I transplanted them into the brain of recipient immunocompromised mice. These animals lack the thymus and therefore fail to generate mature T-cells. Being unable to mount a proper immune response, they are prevented from rejecting xenografts. The day of injection, cells were harvested, counted with an hemocytometer and resuspended in DPBS in order to have a concentration of 2-2.5 × 10^4 cells per μl. Nude mice were anaesthetized, placed in the stereotaxic frame and injected with 5 × 10^4 to 1 × 10^5 transformed astrocytes using stereotaxic coordinates corresponding to a posterior portion of the caudate (0.7 mm back, 3 mm left and 3.5 mm deep from the bregma). At the onset of neurological symptoms, mice were sacrificed and brains were isolated from the head for further analyses. I confirmed that this model was fully penetrant as I obtained lethal tumors from all the mice I injected (n=29). Tumors were obvious and had glioma like appearance at the macroscopic level as shown in figure 8.

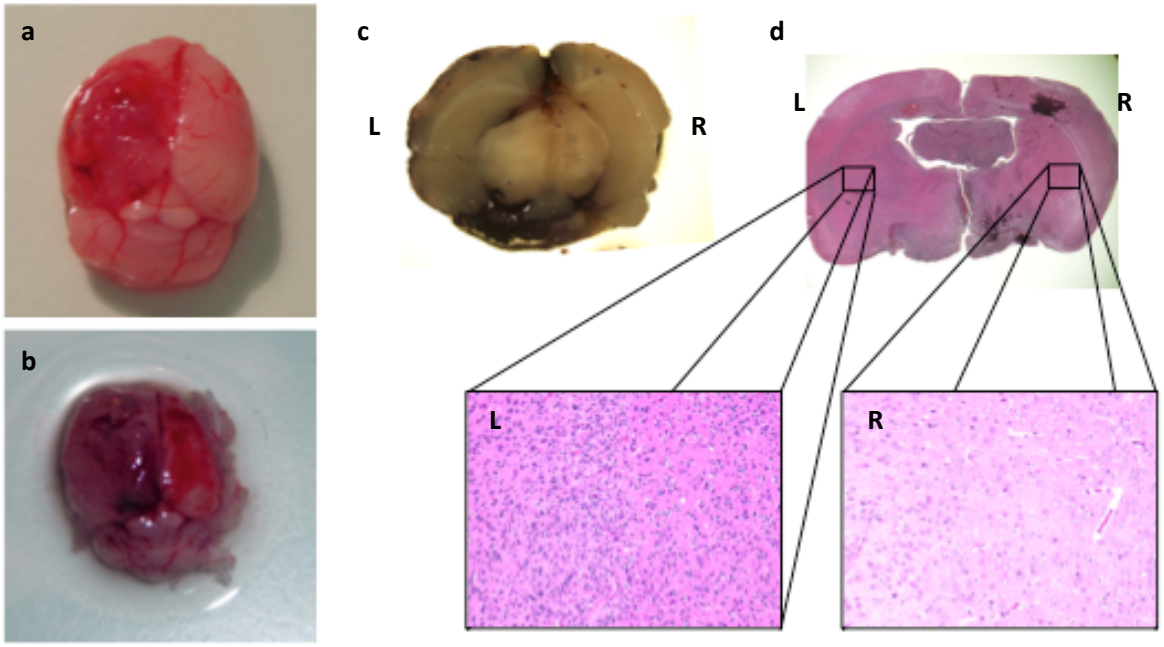

Figure 8. Tumors are identified at macroscopic level. **a, b**. Swollen hemorrhagic lesion typical of high grade gliomas. The injected hemisphere (left, L) appears enlarged compared to the right (R) counterpart, both in the freshly isolated brain (**c**) and in the fixed brain (**d**, hematoxylin-eosin staining). In the close-ups (10x) a comparison is made between the left and the right parenchyma, the first one showing hypercellularity and a complete invasion by tumor cells.

1.3 Histopathological analysis and classification of the tumors

Tumors were processed according to standard procedures used for histopathological analysis. They were independently classified following the World Health Organization guidelines, in collaboration with Dr. Doglioni and Dr. Terreni from the Anatomo-pathology Unit of the San Raffaele Hospital in Milan. Some of the tumors clearly showed glioblastoma multiforme (GBM) features such as pseudopalisading necrosis, neoangiogenesis and a high mitotic index (Figure 9, a-b). In addition, a thorough histological classification pointed out distinct variants of these gliomas, such as oligodendrogliomas, oligoastrocytomas and gliosarcomas (Figure 9, d-e; Table 10, a) were also generated. Nevertheless, the majority of tumors showed a high malignancy grade (Table 10, b), in accordance with the phenotype described for this model. In certain cases the subcutaneous layer was also infiltrated (Figure 9, c). The heterogeneity in tumor types can be explained by the heterogeneity of the starting astrocyte cultures. These results are in line with those obtained by

Bruggeman et al. [Bruggeman et al. 2007] in their study: using the same model, they could also generate different types of gliomas but all characterized by a high grade of malignancy.

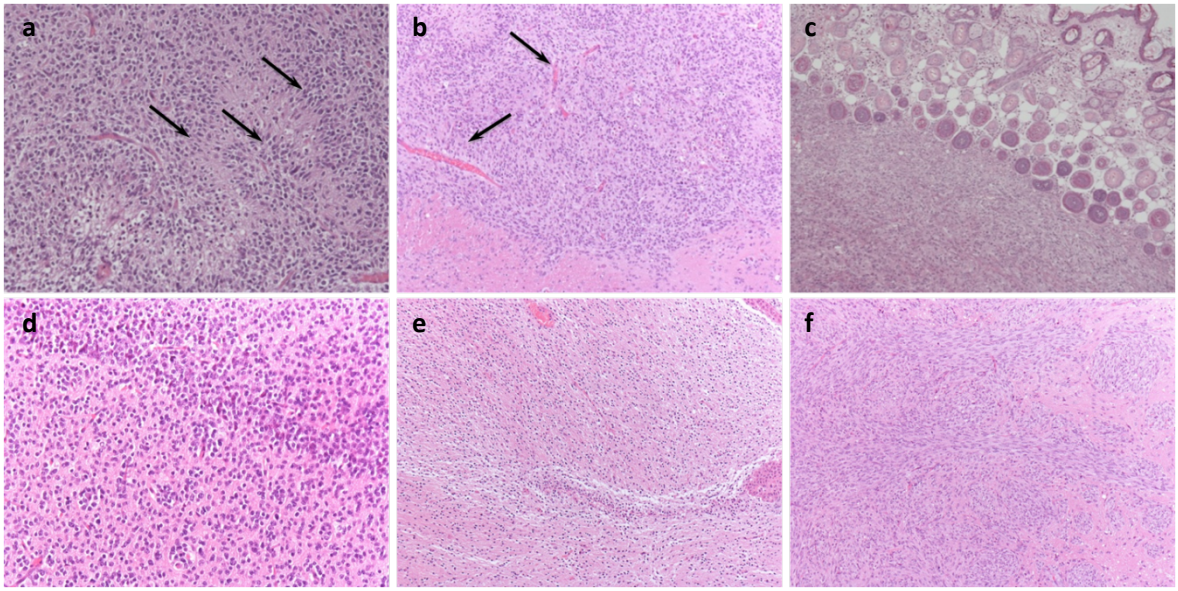

Figure 9. High malignancy features are identified in histopathological analysis of tumors. a. Pseudopalisade necrosis (indicated by arrows) **b**. New vessels formed during tumor development (indicated by arrows). **c**. Subcutaneous layer invasion. Representative examples of different tumor types are shown; **d**: oligodendroglioma, **e**: oligoastrocytoma, **f**: gliosarcoma. (Hematoxylin-eosin staining in all the panels, 10x).

a

Classification	N.
GBM	2
Oligoastrocytoma	3
Oligodendroglioma	8
Gliosarcoma	4
Other	7
NA	5

b

Malignancy grade	N.
I	0
II	3
III	8
IV	6

Table 10. Histopathological classification of the tumors according to the World Health Organization system. **a**. Histolgical classification of tumors. "NA": not assigned. **b**. Malignancy grade were defined for all categories except "Others" and "NA".

1.4 Derivation and characterization of glioma propagating cells (GPCs)

The cancer stem cell model proposes the existence of a pool of cells that is able both to self-renew and to maintain and expand the bulk of the tumor. In order to confirm the presence of such cells in my model system, a portion of the tumor mass was isolated from each brain and processed to obtain GPC cell lines. I derived GPCs from the tumors formed in primary hosts with

an average efficiency of 50%. Traditionally, GPCs were cultured as non-adherent spheroids, also known as neurospheres [Galli et al. 2004, Singh et al. 2004]. However, the condensed structure of the neurosphere hampers the diffusion of growth factors at the innermost cells, causing cell death and differentiation [Woolard et al. 2009]. Therefore, I decided to grow GPCs on laminin-coated dishes as adherent cultures [Pollard et al. 2009]; in these conditions, all the cells can equally access growth factors and the culture can be expanded homogeneously without spontaneous differentiation. The morphology of GPCs varied among different samples, as shown in Figure 10. Despite grown in adherent conditions, one line never managed to attach to the plate and formed floating neurospheres. No correlation was found between the tumor type and the morphology of the GPCs.

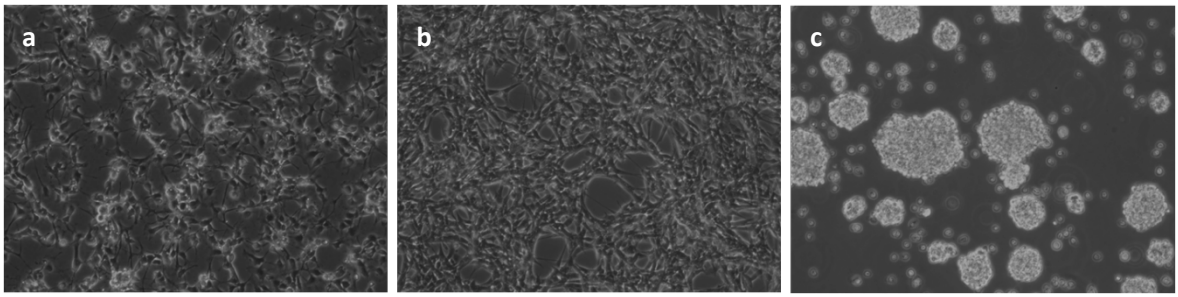

Figure 10. GPC lines showed different morphology. **a-c**: phase contrast image of different GPC lines (10x). As shown in the right panel, one line never attached to the plate and grew forming floating neurospheres.

GPCs derived from human brain tumors express markers associated with neural stem and progenitor cells, like CD133 and A2B5 [Singh et al. 2004, Ogden et al. 2008, Tchoghandjian et al. 2010]. Flow cytometry analysis showed that GPCs obtained in my experiments were $A2B5^{+}$, but $CD133^{-}$ (Figure 11).

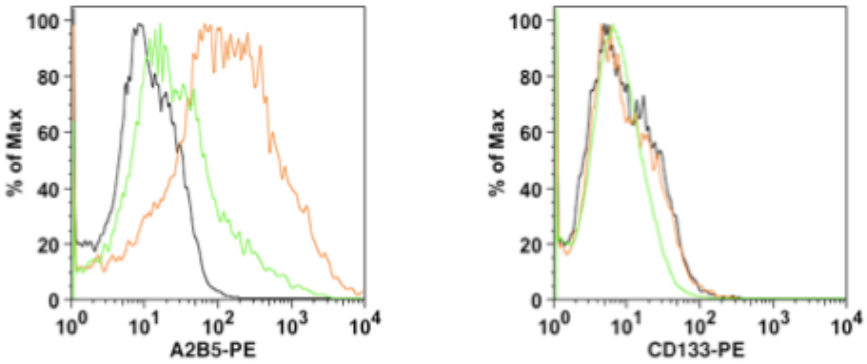

Figure 11. Glioma initiating cells express A2B5 but not CD133. Flow cytometry analysis of GPCs showed a high positivity for A2B5, while CD133 was almost absent. In each graph, red and green curves represent two different samples, while the black curve represent the negative control.

Since serial orthotopic transplantations represent the best functional test to evaluate self-renewal and tumor propagation, I re-injected 5×10^4 GPCs per mouse as described for astrocytes into secondary hosts (n=10). I obtained secondary tumors with 100% penetrance. Moreover, the development of these tumors was faster compared to primaries, with an average survival of 32.5 days (Figure 12).

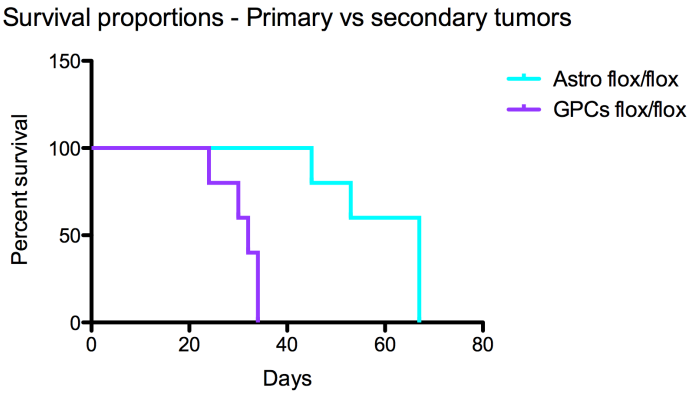

Figure 12. Glioma initiating cells generate secondary tumors with a shorter latency. Kaplan-Meyer curves showing mice survival rate for primary tumors (n=5, average = 53.2 days, median = 67 days) and secondary tumors (n=5, average = 32.5 days, median = 32 days). *P* value = 0.0027 (Log-rank test)

2. Analysis of EZH2 function in gliomagenesis

In order to assess the role of EZH2 during gliomagenesis, I chose a conditional mutagenesis approach. Indeed, straight *Ezh2* KO causes embryonic lethality. By employing a conditional allele, the mice develop normally and the allele can be deleted afterwards, in a time- and tissue- specific

manner. The astrocytes used for this experiment carried a conditional allele of *Ezh2* (hereafter referred to as *Ezh2*^{flox}), in which the SET domain can be removed by Cre recombinase administration, causing the loss and/or truncation of EZH2 [Su et al. 2003]. Despite the possible formation of a truncated form of the protein, the lack of the catalytic domain allows to consider this as a *bona fide* KO. As Polycomb-mediated silencing of the *Ink4a/Arf* locus can occur in cancer [Bracken et al. 2003, Sparmann et al. 2006, Bracken et al. 2007], the model I chose bypasses the problem of a potential restored expression of the same locus when *Ezh2* is deleted. With this strategy, it is possible to study pathways that are relevant for glioma development, that are regulated by PcG but that are independent of its effect on this locus.

2.1 Assessing the role of EZH2 in glioma formation

2.1.1 Ezh2 expression during gliomagenesis

I measured the expression of *Ezh2* in the astrocytes and in cultured GPCs samples (primary and secondary) by qRT-PCR. *Ezh2* is increased in tumors compared to astrocytes (Figure 13).

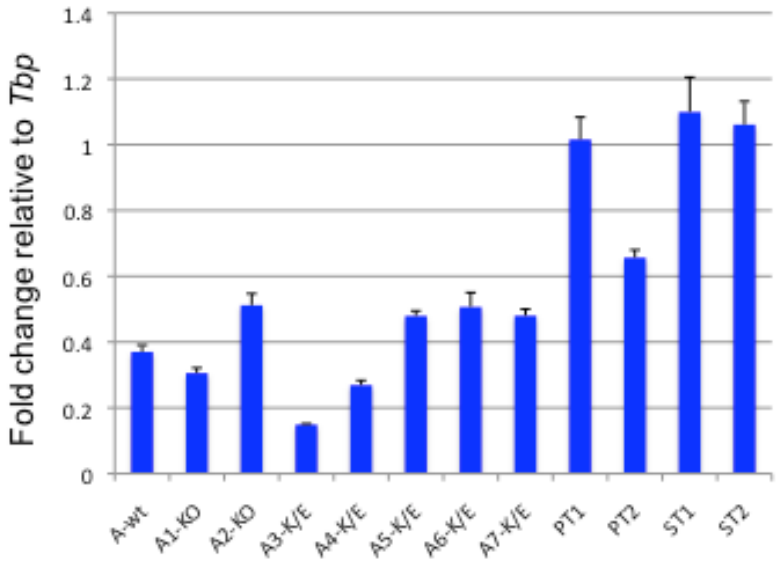

Figure 13. Ezh2 is increased in tumors. **a**. TaqMan assay on different cell populations. Astrocytes (A) with different tumorigenic potential (one wild type - wt, two p16/p19$^{-/-}$ - KO, 5 p16/p19$^{-/-}$ overexpressing EGFRvIII – K/E), two primary GPCs (PT) and two secondary GPCs (ST) revealed that *Ezh2* is more expressed in tumors compared to astrocytes. mRNA levels are calculated using the comparative cycle threshold method (ΔCt) and relative mRNA amount compared to *Tbp* is reported as $2^{-\Delta Ct}$. The bars represent the positive standard deviation (S.D.).

Since it is known that gliomas are highly proliferative, I combined the analysis of EZH2 with Ki-67 by immunohistochemical staining on subsequent slices of tumor samples. Ki-67 is a well defined marker that identifies proliferating cells. It is expressed during G1, S, G2 and M phases of the cell cycle but not in G0. My data revealed a spatial co-localization between Ki-67 and EZH2, while the two proteins were almost absent in the normal brain parenchyma (Figure 14). Thus, EZH2 appears to be selectively over-expressed in the tumor regions characterized by a higher proliferative index.

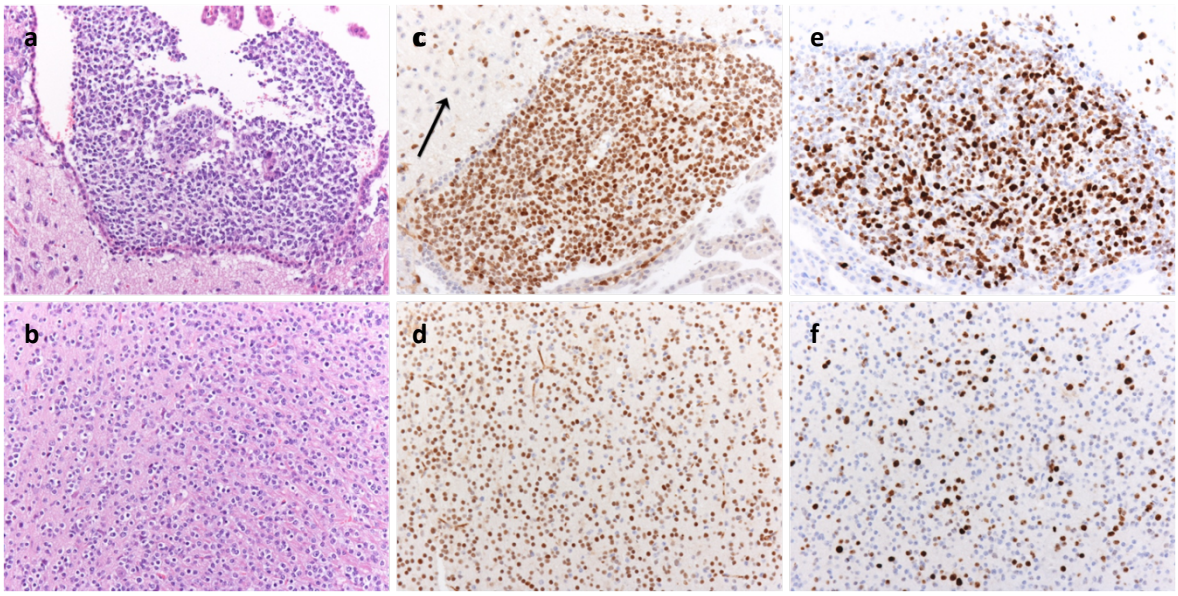

Figure 14. Co-localization of EZH2 and Ki-67 in the tumor mass. **a-b**. Haematoxylin/eosin staining. **c-d**. EZH2 immunostaining; in the normal parenchyma EZH2 is almost absent (panel c, arrow), while it is highly expressed in the tumor mass. **e-f**. Ki-67 immunostaining. (10x magnification in all the panels).

2.1.2 Deletion of Ezh2 by administration of Cre recombinase

After demonstrating that *Ezh2* is expressed at high levels both primary and secondary gliomas, I proceeded to inactivate the *Ezh2* gene by administration of the cell-permeable Cre protein (Cre fused to the HIV-derived TAT peptide). In order to check for Cre-mediated toxicity in $Ezh2^{flox/flox}$ cells (hereafter called "experimental" or "$Astro^{del/del}$"), I used as controls $Ezh2^{flox/+}$ cells (hereafter called "controls" or "$Astro^{del/+}$"). This control is useful for assessing Cre toxicity while having the recombinase activity readout of the deletion of the floxed allele. The deleted allele will be referred to as $Ezh2^{del}$. I experimented different concentrations and incubation times in order to

define the optimal conditions of TAT-Cre treatment; I tried the following 3 different regimens in parallel:

i. 50 μg/ml of TAT-Cre with overnight incubation in 2 ml of medium, one pulse;

ii. 100 μg/ml of TAT-Cre with 3 hours incubation in 2 ml of medium, one pulse;

iii. 15 μg/ml of TAT-Cre with 3 hours incubation in 2 ml of medium, one pulse per day for 3 consecutive days;

The first two protocols showed a high degree of acute toxicity with massive cell death. In particular, the death in the controls treated according to the first protocol was so high that it was impossible to rescue the cell line. The experimental counterpart was also extremely affected by the treatment, but I could retrieve enough cells to perform injections in mice (n=3). Despite those cells had an *in vivo* performance in line with the other 2 experimental sets of cells treated according to the second and third protocol, the results are never included in the subsequent analysis because of the lack of the proper control. I assessed the level of deletion in the cells treated according to the second and third regimens of TAT-Cre by genomic qPCR with a TaqMan® assay designed on the SET domain. In both cases and for both genotypes the efficiency was low (Figure 15).

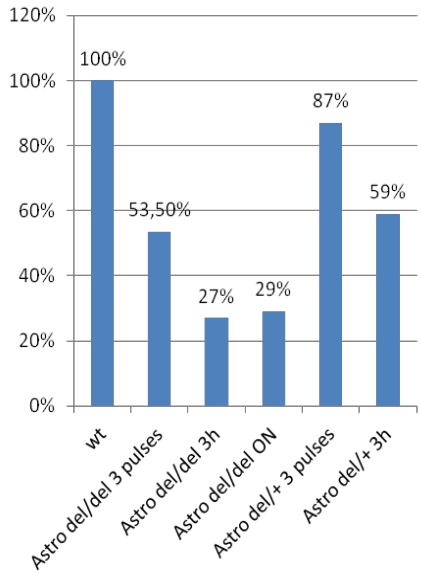

Figure 15. *Ezh2* SET domain deletion in astrocytes was inefficient. qPCR analysis of the Cre-mediated *Ezh2* deletion; the efficiency was sub-optimal in both genotypes (expected results: 50% and 0% for Astro$^{del/+}$ and Astro$^{del/del}$, respectively). Results are expressed as percentages compared to wild-type, that represents 100%.

2.1.3 Injection of TAT-Cre treated astrocytes in nude mice and tumor formation

After the TAT-Cre treatment, Astro$^{del/del}$ and Astro$^{del/+}$ were injected in nude mice to assess their tumorigenic capacity as indicated in the table below (Table 11).

Cell type	TAT-Cre tratment	N. of injected animals
Astro$^{del/del}$	100 μg/ml, 3 hours	3
Astro$^{del/del}$	15 μg/ml, 3 pulses	5
Astro$^{del/+}$	100 μg/ml, 3 hours	3
Astro$^{del/+}$	15 μg/ml, 3 pulses	5

Table 11. Panel of injections of tumorigenic astrocytes .

The two cohorts showed a nearly twofold difference in survival, with a mean of 42.4 days for control animals and of 73.1 days for the experimental animals (Figure 16).

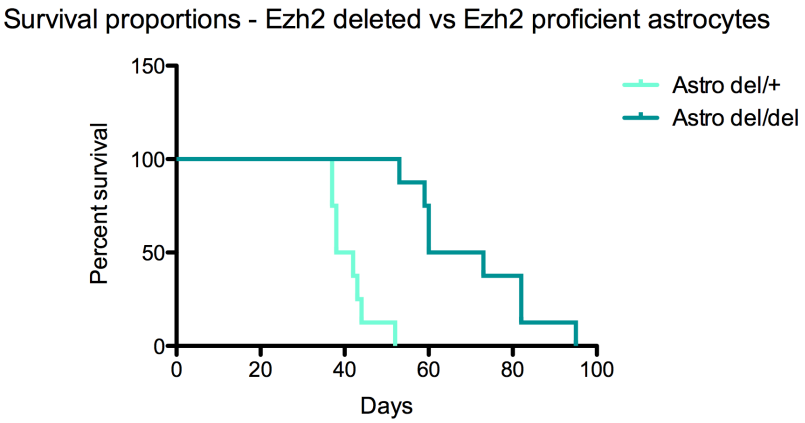

Figure 16. EZH2 loss in tumorigenic astrocytes increases survival of recipient mice. Kaplan-Meyer curves showing mice survival rate. Animals receiving Astro$^{del/del}$ (experimental, n=8, average = 73.1 days, median = 66.5 days) survived longer than animals receiving Astro$^{del/+}$ (controls, n=8, average = 42.4 days, median = 40 days). *P* value < 0.0001 (Log-rank test).

Although tumor development was severely delayed, the full penetrance of tumorigenesis even in mice injected with the Astro$^{del/del}$ was at odds with the hypothesis that EZH2 is required for glioma initiation and suggested two alternative scenarios: whether EZH2 loss is negatively selected during gliomagenesis or if tumorigenic compensatory mechanisms set in once EZH2 is lost. To distinguish

between these two hypotheses, I performed a copy number assay on genomic DNA purified from from GPCs derived from both control and experimental tumors. Though TAT-Cre treatment showed a certain level of deletion on both populations before injection, results on GPCs DNA showed that experimental gliomas were composed only of cells that had escaped Cre-deletion, thus indicating a massive counterselection for the loss of *Ezh2* during primary gliomagenesis (Figure 17). This *in vivo* competition experiment suggests that cells retaining *Ezh2* might have a proliferative advantage over cells that lost it; this could explain why mice eventually succumbed to tumors.

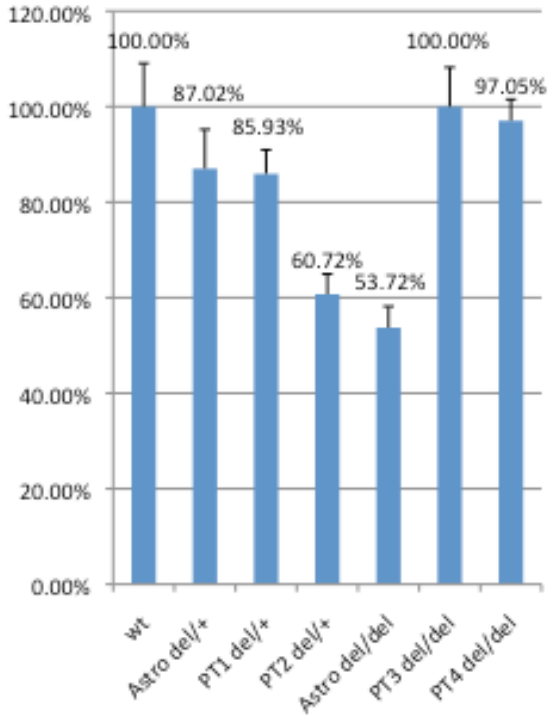

Figure 17. GPCs from primary tumors are *Ezh2* proficient. PTs are primary glioma initiating cells derived from tumors generated with the respective astrocyte populations. In tumors derived from Astro$^{del/del}$ (PT3 and PT4) there is a strong counter selection for cells that lost *Ezh2*. Astro$^{del/+}$ and Astro$^{del/del}$ data representing the deletion in transformed astrocytes treated with 15 μg/ml of TAT-Cre with 3 hours incubation for 3 consecutive days are reported for clarity. Results are expressed as percentages compared to wild-type, that represents 100%.

2.2 Assessing the role of EZH2 in tumor maintenance

2.2.1 Generation of Ezh2-proficient GPCs

To test whether EZH2 is required for the maintenance of the tumor, I generated primary tumors using astrocytes homozygous for the *Ezh2^{flox}* allele without treating them with Cre before

injection. This allowed me to obtain primary GPC cell lines that were still proficient for *Ezh2* but in which the allele could be deleted afterwards. GPCs were treated in culture with 15 ng/ml TAT-Cre for 3 hours for 3 consecutive days to remove *Ezh2* SET domain. Secondary hosts receiving experimental primary GPCs treated with Cre recombinase ($GPC^{del/del}$; 5×10^4 cells/mouse) didn't show any difference in the survival rate from those injected with untreated primary GPCs ($GPC^{flox/flox}$; 5×10^4 cells/mouse), as shown in Figure 18.

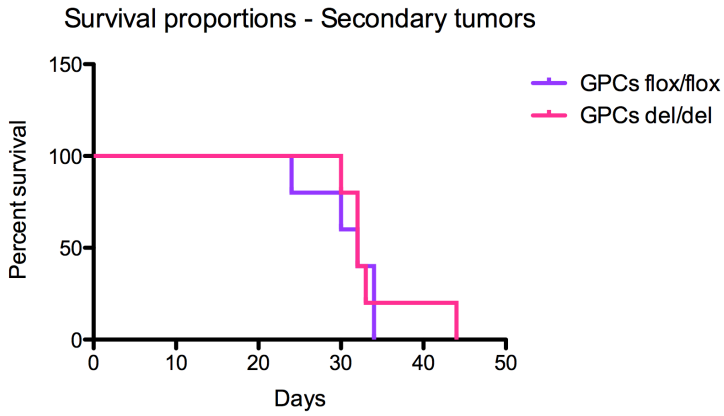

Figure 18. Secondary tumors are not affected by Ezh2 loss. In secondary tumor formation, there is no difference between the survival of experimental (n=5, average = 34.2 days, median = 32 days) and control (n=5, average = 30.8 days, median = 32 days) mice. *P* value = 0.7012 (Log-rank test).

I checked for *Ezh2* deletion using the copy number assay, but in this case no counter selection was seen in GPCs derived from secondary tumors (Figure 19). $GPC^{del/del}$ were able to form the tumor to the same extent as the $GPC^{flox/flox}$. Together, these experiments suggest that while EZH2 plays a central role at the beginning of tumor formation, it is dispensable for secondary tumorigenesis.

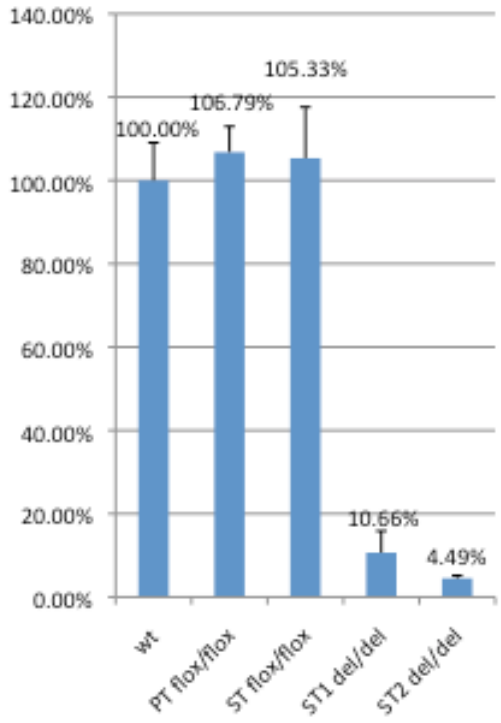

Figure 19. GPCs derived from secondary tumors are deleted for *Ezh2*. Copy number assay on genomic DNA show no positive selection for cells *Ezh2* proficient in secondary tumors (ST). Results are expressed as percentages compared to wild-type, that represents 100%.

3. Validation of the results in the model of PDGFB-dependent gliomagenesis

As discussed earlier, gliomas are associated with a range of distinct genetic lesions. For this reason, I wanted to validate my results in a distinct model of this tumor. This was done in order to determine if the role of EZH2 is consistent among different tumor driving mutations and signaling pathways or if it is specific for the one I used previously. The second model I adopted relies on the overexpression of PDGFB in cortical neural progenitor cultures prepared from embryonic telencephalic explants [Appolloni et al. 2009].

3.1 Setup of the PDGFB glioma model

Also the second model I decided to exploit is a well eastablished model with 100% penetrance. In the classical version, the transformed neural progenitors (NPs) obtained from C57 mice are infected in order to overexpress PDGFB and then transplanted into singenic C57 mice. The neural progenitors that I was using in my experiment were also derived from mice of the C57

background; nevertheless, they were transgenic for the Cre-ERT2 gene, ubiquitously expressed through the Rosa26 promoter. I first wanted to assess if this transgenic protein could have been recognized as non-self, causing immune rejection in recipient C57 mice and thus preventing tumor formation. I performed a pilot experiment injecting transformed NPs in C57 mice, nude mice and SCID beige mouse (n=3 for each strain). All mice developed tumors within a similar timeframe; thus I decided to perform all following experiments in C57 mice.

3.2 Setting up the experiment

The retroviral vector used to overexpress PDGFB encodes also for a fluorescent protein, either dsRED or GFP. I decided to use both vectors to infect *Ezh2*$^{flox/flox}$ and *Ezh2*$^{flox/+}$ cells with two different colors in order to perform an *in vivo* competition experiment. Cells heterozygous for the *Ezh2* conditional allele were again used to monitor the Cre-mediated toxicity as in the previous model. To check for any possible toxicity caused by the fluorescent proteins, I decided to perform infections in order to get all the possible combinations of genotypes and colors, as listed in Table 12.

Genotype	Fluorescent protein expressed
NP$^{flox/flox}$	dsRed
NP$^{flox/flox}$	GFP
NP$^{flox/+}$	dsRed
NP$^{flox/+}$	GFP

Table 12. Panel of genotype/infection combinations in NPs.

3.3 Assessing the role of EZH2 in tumor formation

3.3.1 Deletion of Ezh2 by administration of 4-hydroxytamoxifen

The cells used for this experiment carried the conditional allele of *Ezh2* (Ezh2^{flox}) described earlier and the Cre-ERT2 transgene expressed under the control of the Rosa26 promoter. The Cre-ERT2

recombinase is activated upon Tamoxifen administration. Particularly, its active metabolite 4-Hydroxytamoxifen has to be administered in cell cultures. I treated the cells with 1 μM 4-OHT for 4 days starting from the day after derivation. The copy number assay demonstrated that the treatment was effective and the deletion complete in both genotypes (Figure 20).

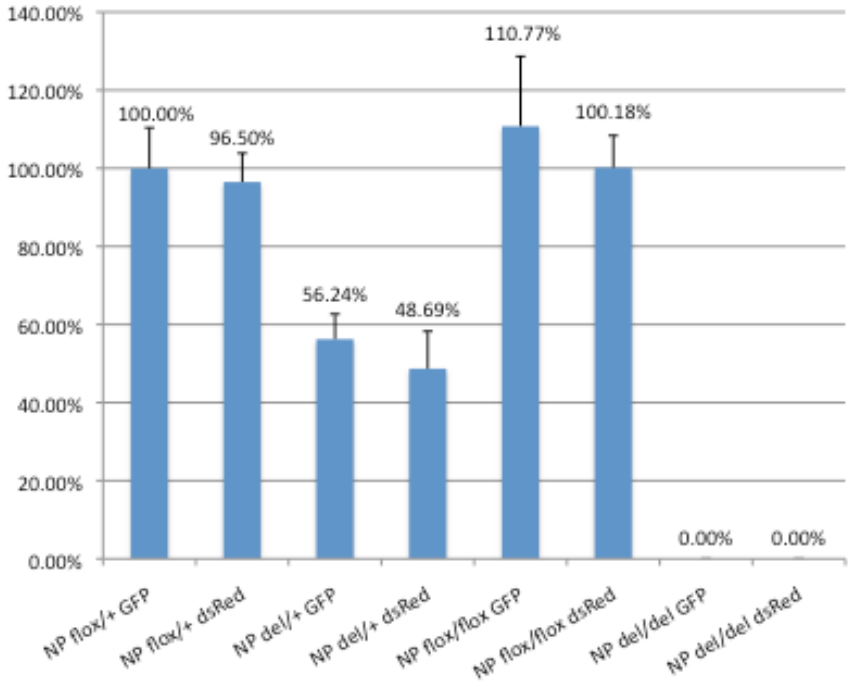

Figure 20. 4-OHT treatment is effective in NPs. The 4-OHT treatment completely deleted *Ezh2* SET domain in all the NP populations (expected results: 50% and 0% for $NP^{del/+}$ and $NP^{del/del}$, respectively). Values for the untreated samples ($NP^{flox/+}$ and $NP^{flox/flox}$) are reported for clarity.

3.3.2 Transplantation of NPs into C57 mice and assessment of glioma formation

The set of transplantation experiments performed is indicated in the table below (Table 13).

Genotype	4-OHT	N. of injected animals
$NP^{flox/flox}$	-	5
$NP^{flox/flox}$	+	5
$NP^{flox/flox}$	-	5
$NP^{flox/flox}$	+	5
$NP^{flox/+}$	-	5
$NP^{flox/+}$	+	5
$NP^{flox/+}$	-	5
$NP^{flox/+}$	+	5

Table 13. Panel of injections of NPs.

On the day of injection, cells were harvested and counted with a hemocytometer both in transmitted light and with the appropriate fluorescent light in order to determine the percentage of infected cells in each sample. The percentage of infection was further validated by flow cytometry analysis; estimated percentages varied among 5-7%. Mice were injected with 1.5×10^4 infected cells; for the *in vivo* competition experiment, mice were injected with 1.5×10^4 green cells + 1.5×10^4 red cells. Stereotaxic coordinates used for this experiment correspond to a more anterior part of the caudate, compared to the ones used for previous experiment (1 mm anterior, 1.5 mm left and 2.5 mm under the bregma).

At the onset of neurological symptoms, mice were sacrificed and all the brains were imaged with a fluorescent stereomicroscope (Figure 21). The results were fully consistent with the ones obtained with the previous model. Cells proficient for EZH2 ($NPs^{flox/+}$, $NPs^{flox/flox}$ and $NPs^{del/+}$) were able to form the tumor regardless of the fluorescent reporter and all mice injected with those cells died, except one. This is most probably due to a technical problem during injection (most likely the syringe was not loaded with cells). By contrast, all mice receiving $Ezh2^{del/del}$ NPs are all currently alive, 250 days after injection (Figure 22). In the *in vivo* competition this finding was further corroborated: the tumor mass was consistently formed only by $Ezh2^{del/+}$ cells, with no

trace of any *Ezh2*$^{del/del}$ cell in both combinations of reporters (Figure 21). Therefore, I conclude that los of EZH2 in two distinct glioma mouse models eliminates their ability to generate tumors.

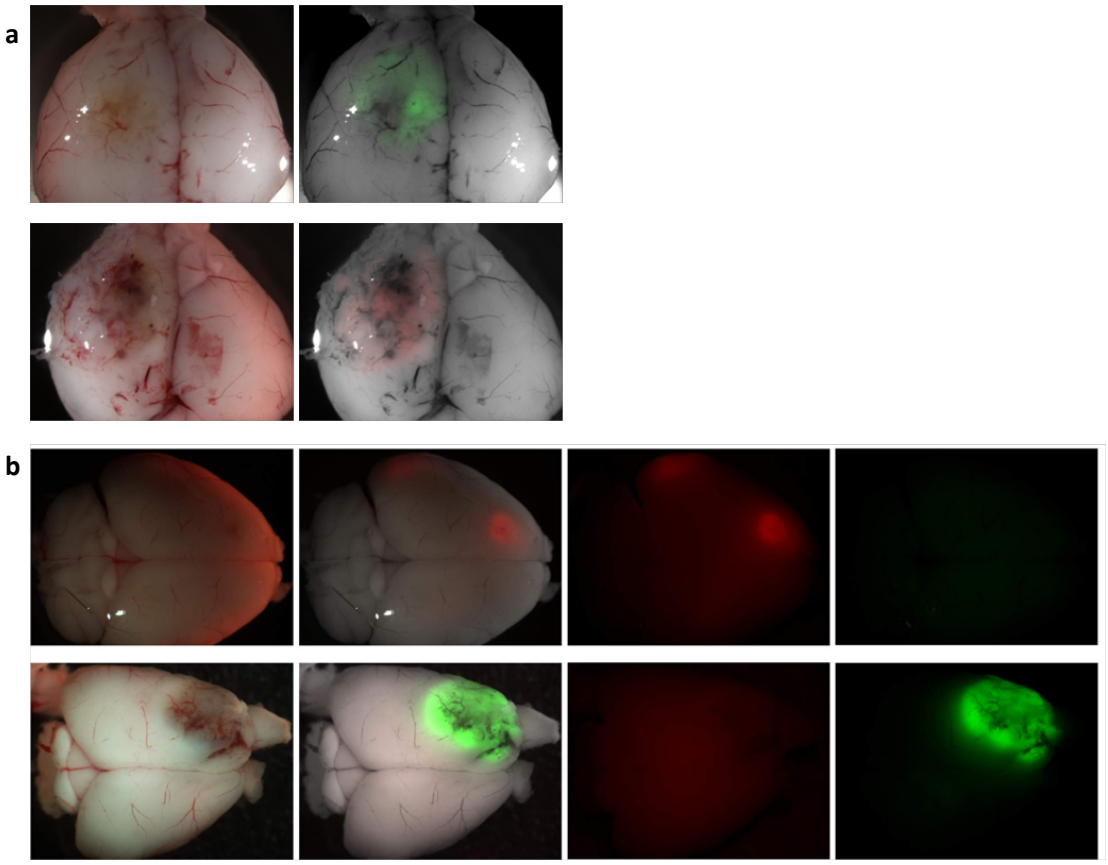

Figure 21. Tumor is formed exclusively by EZH2 proficient cells. **a-b**. Brain were imaged under a fluorescent stereomicroscope in order to visualize the tumor mass. **b**. In the in vivo competition experiment, the mass is made only by the EZH2-proficent population (NP$^{del/+}$, either dsRED^{+} or GFP^{+}).

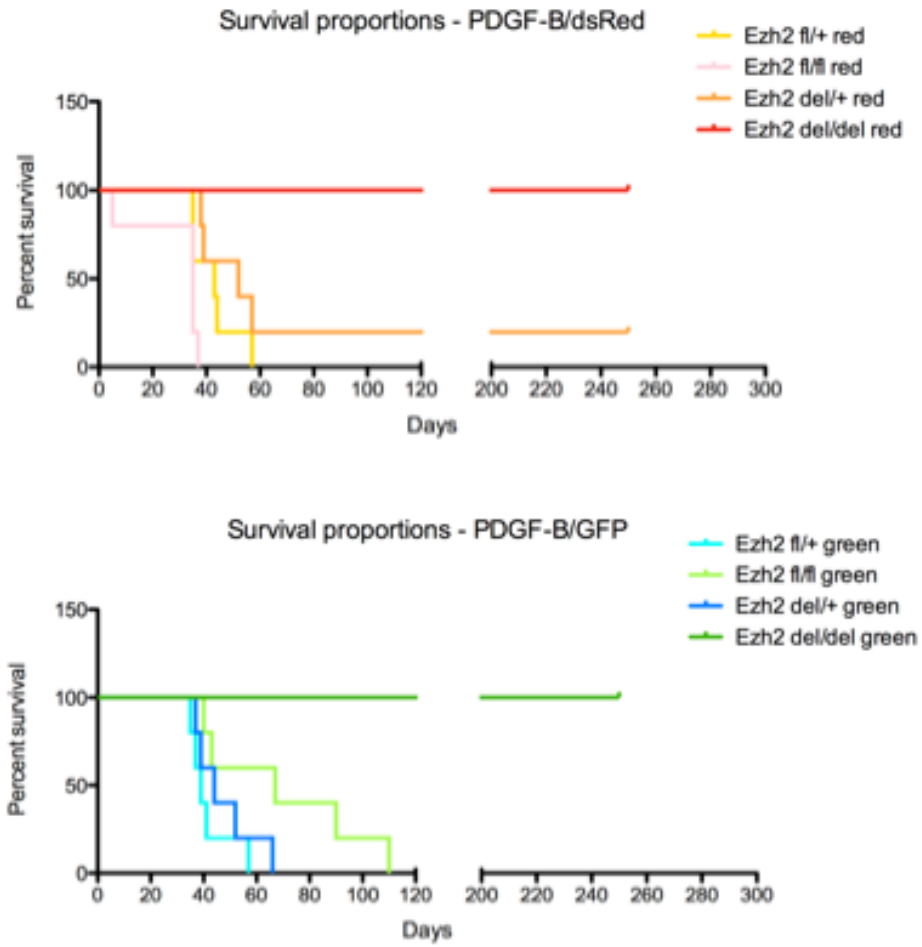

Figure 22. *Ezh2* deleted cells are prevented of forming a tumor. Survival analysis of mice injected with tumorigenic NPs shows that all the mice receiving $NP^{del/del}$ are still alive.

I next sought to determine the fate of the $NP^{del/del}$ cells. I dissected carefully tissues around the injection site to try to identify the injected cells, i.e. if the $NP^{del/del}$ cells are able to engraft but not to survive in the host microenvironment or if they are completely prevented from engrafting the host brain. Three new mice were injected with $NP^{flox/flox}$ and 3 mice with $NP^{del/del}$, both the cell populations being $dsRED^{+}$. I sacrificed one mouse per group 1 week after injection and the remaining 2 mice per group 2 weeks after injections. Brains were isolated, processed according to normal histological procedures and imaged under a fluorescent microscope in order to see how the injected cells were performing. It was impossible to find any $dsRED^{+}$ cell in brains taken 1 week after injection – most likely due both to limited microscope sensitivity, but also to the loss of cells during the stressful injection procedure. However, in brains taken 2 weeks after injection it was already possible to see a very small mass formed by $dsRED^{+}$ $NPs^{flox/flox}$ cells. Moreover, cells

were clearly invading the surrounding tissue, particularly egressing along the corpus callosum (Figure 23, a-b). Importantly, for those mice injected with $NP^{del/del}$ it was impossible to identify any $dsRED^{+}$ cells at the same stage of tumor formation (Figure 23, c-d). This result indicates that EZH2 proficient cells are immediately enabled to form a tumor mass upon transplantation whereas those lacking fail to engraft and establish masses early on. I cannot rule out that EZH2 depleted cells were present and engrafting as higher resolution microscopy would be needed.

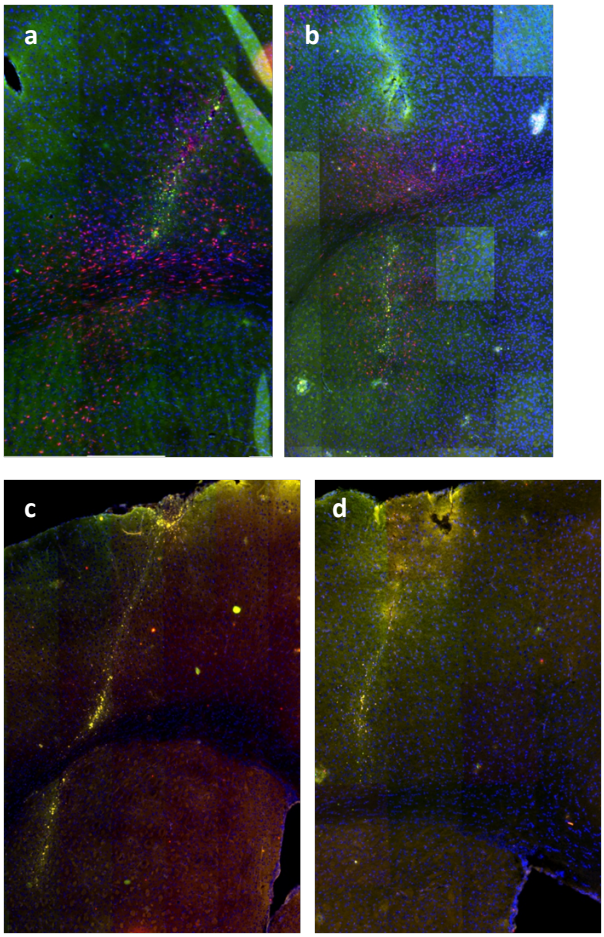

Figure 23. EZH2 proficient cells successfully engrafted and start invading the host brain 2 weeks after injection. The channel of injection is clearly traceable in all the brains but $dsRED^{+}$ cells are found only in the $NP^{flox//flox}$ injected animals (**a,b**) and not in the $NP^{del/del}$ injected ones (**c,d**).

3.2 Assessing the role of EZH2 in tumor maintenance

3.2.1 Generation of Ezh2-proficient GPCs

Also in this second model it is possible to isolate GPCs from the tumor mass. The primary tumors generated with *Ezh2*$^{flox/flox}$ cells were dissected and the allele was deleted in the resulting GPCs before injecting them for secondary tumorigenesis assay, following the same scheme described in the previous experiment. GPCs were treated while in culture with 1 μM 4-OHT for 9 days to remove the catalytic domain of Ezh2. All the secondary hosts receiving both experimental (GPC$^{del/del}$; 1.5 × 10^4 cells/mouse) and control GPCs (GPC$^{flox/flox}$; 5 × 10^4 cells/mouse) died (Figure 24).

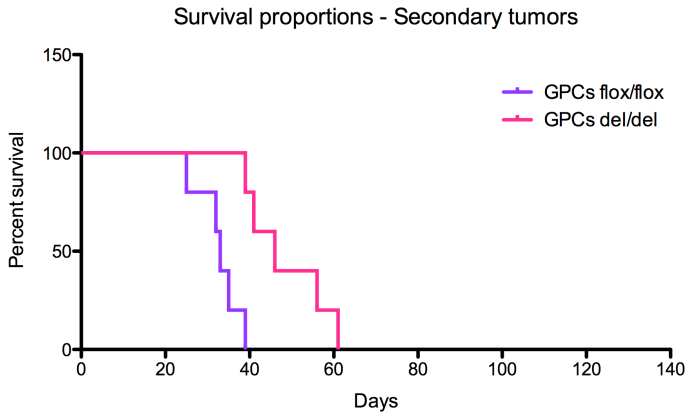

Figure 24. Secondary tumors are formed without EZH2. All the mice transplanted with primary GPCs developed lethal tumors, both in the presence (GPC$^{flox/flox}$) or absence (GPC$^{del/del}$) of EZH2.

I checked for *Ezh2* deletion using the copy number assay and as reported for the previous model, no counter selection was seen in GPCs derived from secondary tumors (Figure 25). GPC$^{del/del}$ were able to form the tumor to the as the GPC$^{flox/flox}$. These results confirmed that EZH2 is dispensable for secondary tumorigenesis, thus corroborating my previous findings and indicating that EZH2 function is independent from the driving mutations in the tumorigenic cells.

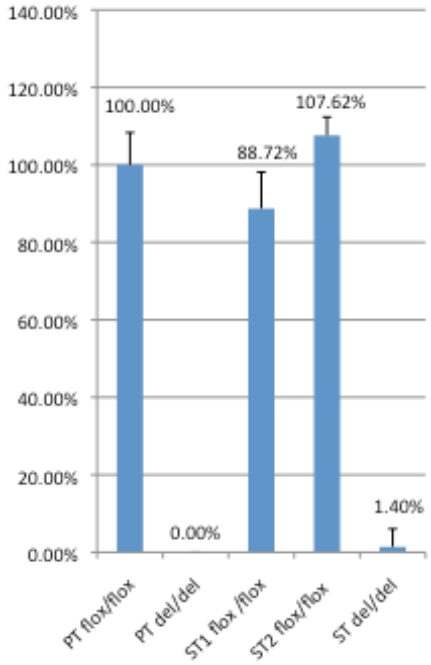

Figure 25. Secondary GPCs are *Ezh2*-deleted. As in the previous experiment, copy number assay on secondary GPCs showed no counter selection for *Ezh2*-deleted cells.

4. Exploring the transcriptome of the tumor – relation with H3K27me3 mark

I took advantage of the intrinsic possibility of my model system to analyze various cell populations with different tumorigenic potential (primary wild-type astrocytes, primary p16/p19$^{-/-}$ astrocytes, transformed astrocytes, GPCs derived from primary tumors, GPCs derived from secondary tumors). This allows me to investigate at the molecular level how the H3K27me3 mark is relocated during tumorigenesis, how the transcriptome changes and which is the possible correlation between these two processes. So far, one sample of transformed astrocytes and two samples of primary GPCs have been analyzed in terms of transcriptome and genome-wide H3K27me3 distribution.

4.1 Analysis of differentially expressed genes (DEGs) between tumorigenic astrocytes and primary tumors

In order to unravel the transcriptomic changes that occur during the transition from tumorigenic astrocytes (i.e., astrocytes p16/p19$^{-/-}$ and overexpressing EGFRvIII) to primary tumors, we defined

the gene expression profile of one tumorigenic astrocyte line ("Astro") and of two different samples of primary GPCs (Glio1 and Glio2; together they will be referred to as "Glio") by RNA-seq. Through the direct comparison of the expression level determined for each gene in each sample, we found that the 2 GPC samples are globally more similar to each other (correlation index r=0.97) than each of them compared to "Astro" (r=0.83 for Glio1 and r=0.85 for Glio2). This suggests that changes in the transcriptional profile during tumorigenesis are not stochastic, but rather they tend to converge towards a common signature.

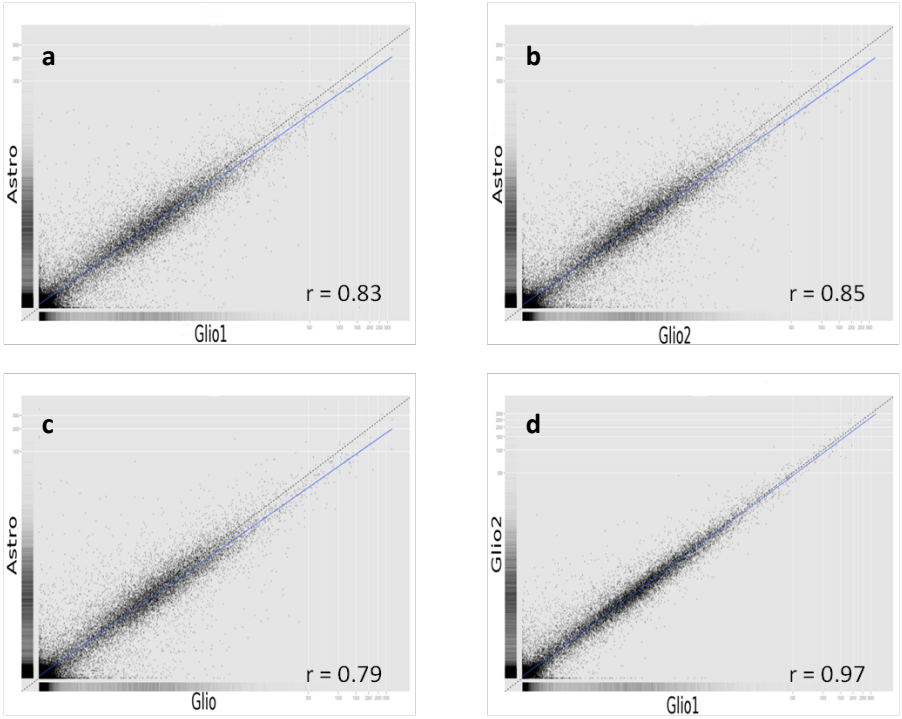

Figure 26. Different tumor samples show a high correlation according to the transcriptional profile. Scatterplots representing the direct comparison of expression levels for each gene. Astro were compared to Glio1 (**a**), Glio2 (**b**) and to the two Glio samples considered together (**c**). When comparing the two glioma samples to each other, the correlation index increased substantially (**d**, r=0.97).

We determined the number of differentially expressed genes (DEGs) comparing "Astro" to each "Glio" and we found 2396 DEGs for Glio1 and 2626 DEGs for Glio2 (FDR<0.05). Particularly, 882 and 1235 genes were upregulated in Glio1 and Glio2 respectively, while 1514 and 1391 genes were downregulated in Glio1 and Glio2 respectively. We performed a Gene Ontology (GO) analysis to determine which are the biological processes enriched in the DEGs for each "Glio" sample. We found that the most significantly associated biological processes are related to

classical Polycomb functions, like development and differentiation. Intersecting the 2 sets of DEGs, we found that about 70% of the genes (1738) are consistently up- or down-regulated in the two samples. If we restrict the GO analysis on this subset of genes, we could find a stronger enrichment with classical Polycomb functions, suggesting that this epigenetic axis is actively involved and plays an important role in gliomagenesis.

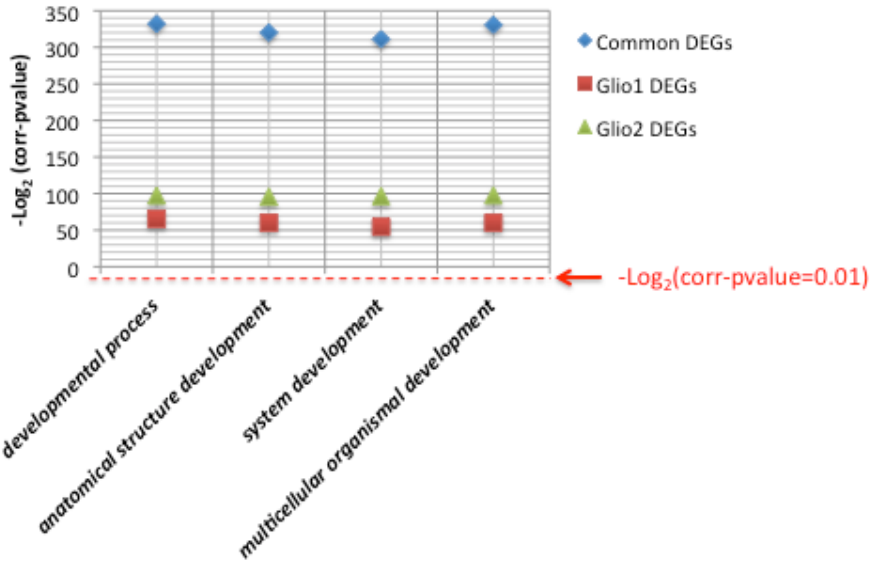

Figure 27. DEGs showed an enrichment for classical Polycomb functions. GO analysis of the biological processes associated to DEGs shows an enrichment for classical Polycomb functions, like development and differentiation. In the common DEGs this association becomes even stronger.

The common 1738 DEGs were then divided into up-regulated or down-regulated (i.e., genes that are up-regulated moving from astrocytes towards glioma or down-regulated following the same direction). We found 643 up-regulated genes and 1098 down-regulated genes. Biological processes enriched in downregulated genes were involved in tissue development and differentiation while those associated to upregulated genes seemed to be involved in cell metabolism. We analyzed also the KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways enriched in each gene set. Interestingly, pathways associated to downregulated genes are involved in cell adhesion while pathways associated to upregulated genes are again related to metabolism. Taken together, these findings fit with typical features described for GPCs. Indeed, metabolism is known to be quite active in this cell population. Moreover GPCs are less

differentiated compared to astrocytes; they resemble neural progenitors or neural stem cells and it has already been shown that they are able to differentiate in 3 cell types, namely neurons, astrocytes and oligodendrocytes. In addition, even if gliomas don't form metastasis in different organs, cells are extremely invasive and able to migrate diffusely in the brain.

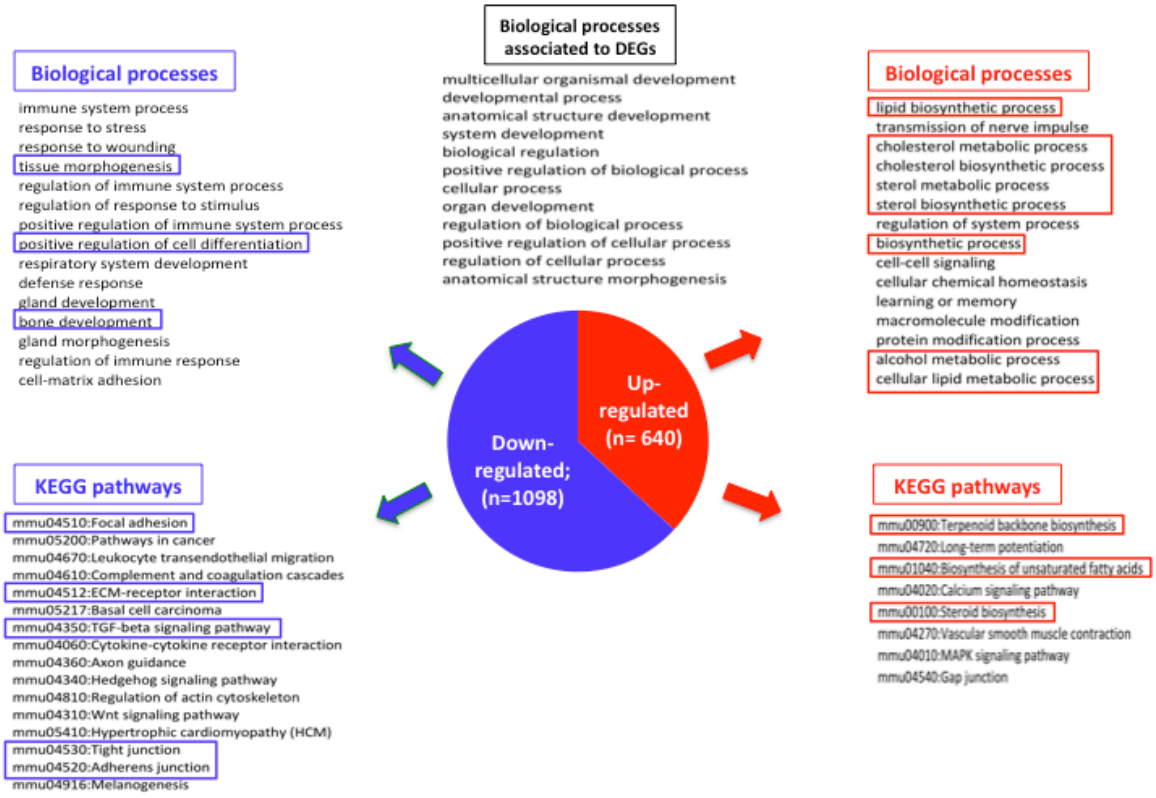

Figure 28. Up-regulated genes are involved in metabolism. For each set of genes, we defined the enriched biological processes and KEGG pathways. Up-regulated genes are related to metabolism and particularly to lipid metabolism. Down-regulated genes are involved in differentiation and development and associated to adhesion pathways. The associations with biological processes and KEGG pathways have been calculated considering a FDR<0.05.

Among the KEGG pathways enriched in the up-regulated gene set, we found the calcium-signaling and the MAPK-signaling pathways, strongly associated with the activation of EGFR signaling pathway. Indeed, we found that EGFR is more highly expressed in "Glio" compared to "Astro".

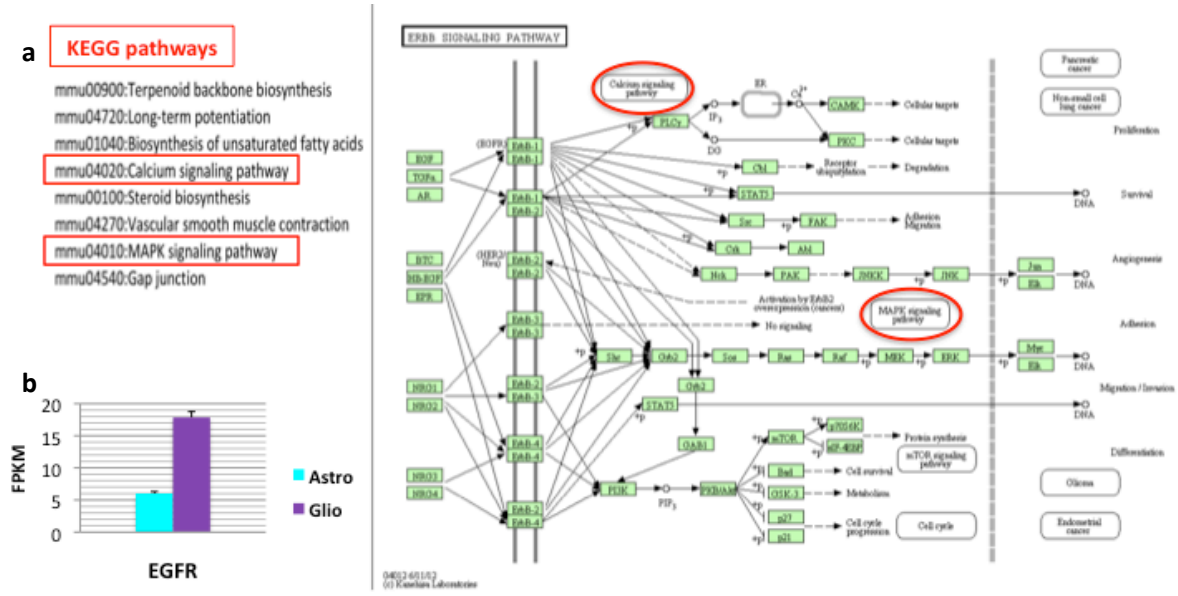

Figure 29. Up-regulated genes are related to EGFR pathway. The functional annotation analysis revealed that up-regulated genes are enriched for the calcium signaling and the MAPK signaling pathways, which are both involved in the EGFR pathway (**a**). Indeed, EGFR is overexpressed in Glio compared to Astro (**b**).

4.2 Analysis of the H3K27me3 distribution

In the same samples used to analyze the transcriptome, we also examined the distribution of H3K27me3 by ChIP-seq. We divided the genome in discrete windows of 1 Mbp and we counted how many reads aligned in each window for each sample. By comparing the read counts in two samples, we were able to define the similarity in the H3K27me3 distribution, expressed as a correlation index. The two tumor samples have a correlation index of 0.97, while when comparing each Glio to Astro the correlation index decreases (r=0.92 for Glio1 and r=0.90 for Glio2), indicating that the deposition of H3K27me3 along the genome is almost identical between the 2 "Glio". Taken together these results seem to indicate that the relocation of K27 trimethylation is not random but rather happens at very specific sites, common across different glioma samples. Moreover, the redistribution moving from transformed astrocytes to gliomas is not massive but is targeted to a specific subset of sites within the genome.

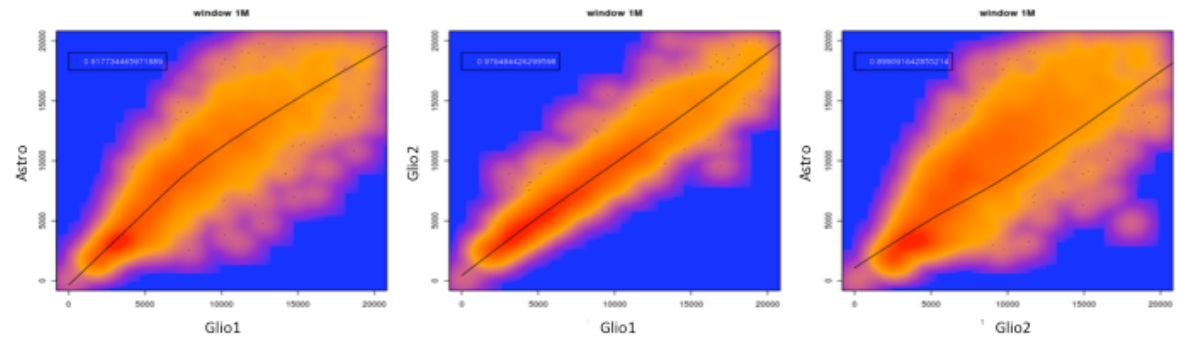

Figure 30. Different tumor samples show an high correlation according to the H3K27me3 distribution. Scatterplots representing the direct comparison of read counts in 1 Mbp windows on the genome. The correlation between the two Glio is higher (r=0.97), while it decreases when each of the Glio is compared to Astro (r=0.92 for Astro/Glio1 comparison and r=0.90 for Astro/Glio2 comparison).

We determined the number of genes carrying the trimethylation of lysine 27 for each sample and we found an increase in both tumors (3636 for Glio1 and 3585 for Glio2) compared to astrocytes (2515). The overlap of the trimethylated genes between Glio1 and Glio2 is more than 80% (3010). We analyze the trend of this histone mark moving from tumorigenic astrocytes to gliomas and we discovered that of the 3010 trimethylated genes found in the tumors, 1049 acquired the mark during the transition, while 1961 were already trimethylated in the tumorigenic astrocytes. Interestingly, 554 genes that were trimethylated in astrocytes lost the mark in the tumors (Figure 31). Overall, the relocation of the H3K27me3 involved almost 1600 genes.

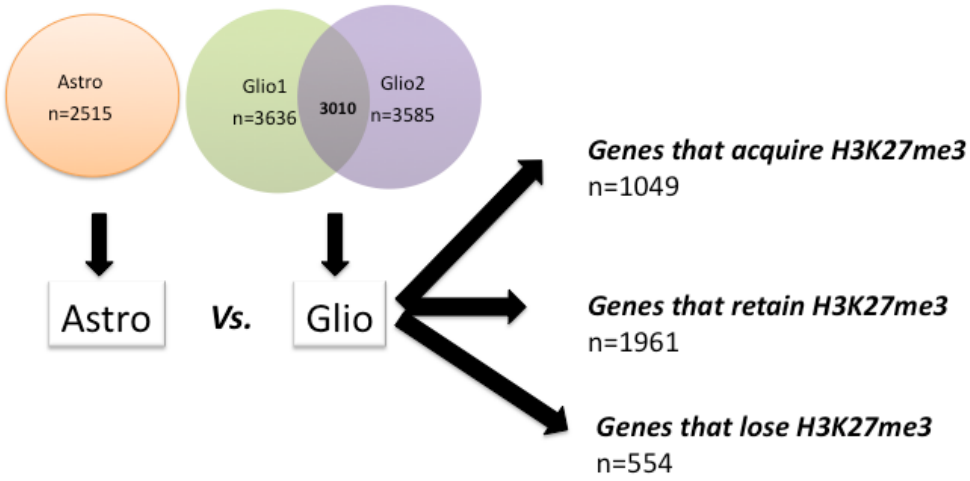

Figure 31. Relocation of H3K27me3 during tumorigenesis involves almost 1600 genes. Analysis of the H3K27me3 patterns revealed that the two Glio samples shared 80% of trimethylated genes. Moreover, of these genes, only 1049 were *de novo* methylated during gliomagenesis.

We went then back to our common DEG list and we defined how many up-regulated gene and down-regulated genes retained, gained or lost the H3K27me3. Analyzing the KEGG pathways enriched in the set of genes that acquired the trimethylation and that were also down-regulated, we could confirm that all of them were included in the pathways enriched in the broader down-regulated gene set.

4.3 Comparison of expression profile data with known human molecular signatures

We wanted to compare our transcriptional profiles with the human molecular signatures previously defined by 3 different groups [Phillips et al. 2006, Carro et al. 2010, Verhaak et al. 2010] in order to determine if there was any overlap between our data and their classification. In their study, Phillips et al. defined a list of 254, 217 and 170 overexpressed genes that characterize the proneural (PN), proliferative (PR) and mesenchymal (MES) signatures, respectively. In our set of up-regulated genes, we found 33 genes shared with the proneural signature. More interestingly, among the biological processes enriched in this group of 33 genes we found "neurogenesis" and "Notch signaling pathway", which are typically features of the proneural tumors according to the same work [Phillips et al. 2006]. In our set of down-regulated genes, 22 belong to the mesenchymal signature; also in this case one biological process enriched in these genes, namely "angiogenesis", is reported to be characteristic of the mesenchymal type (Figure 32).

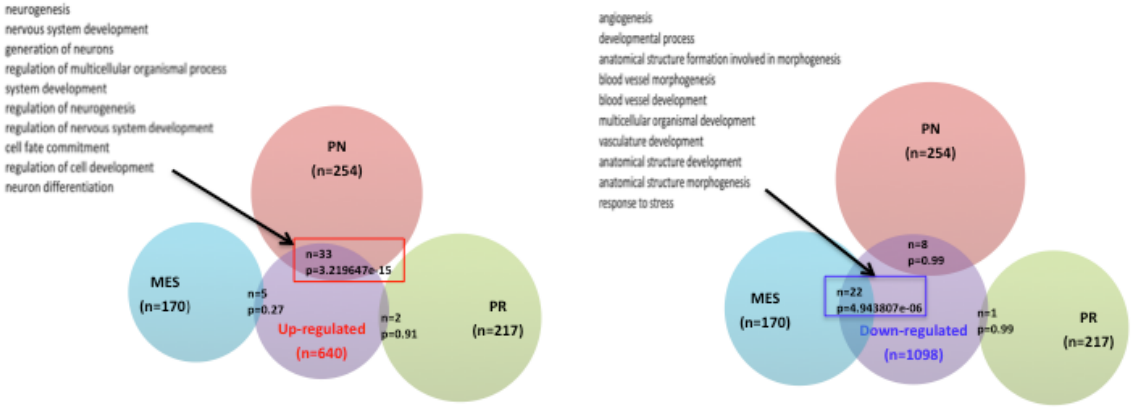

Figure 32. Up-regulated genes are enriched for Phillips' proneural signature genes. Comparing our set of genes with the molecular subtypes defined by Phillips et al. in their work, we found that up-regulated genes were enriched for genes belonging to the proneural signature, while down-regulated genes were enriched for mesenchymal genes.

Carro et al. re-defined the classification proposed by Phillips et al. analyzing a broader group of samples (including the one used by Phillips et al.). They reduced the number of genes associated to each signature, indicating 244, 182 and 82 genes for the proneural, proliferative and mesenchymal signature respectively. In our up-regulated set of genes, 35 overlapped again with the proneural signature and 17 of the down-regulated overlapped with the mesenchymal, confirming previous results. Verhaak et al. introduced a different molecular classification based on 4 signatures. They maintained the proneural and mesenchymal subtypes, but they introduced two new classes, called neural and classical. When we compared our transcriptional profiles to their results, our common up-regulated genes were mostly associated with the proneural subtype (25 genes), but also with the classical (13 genes) and with the neural (11 genes). Down-regulated genes instead overlapped again with the mesenchymal signature (43 genes). Taking all these results together, we can then conclude that our tumors closely resemble the human proneural subtype.

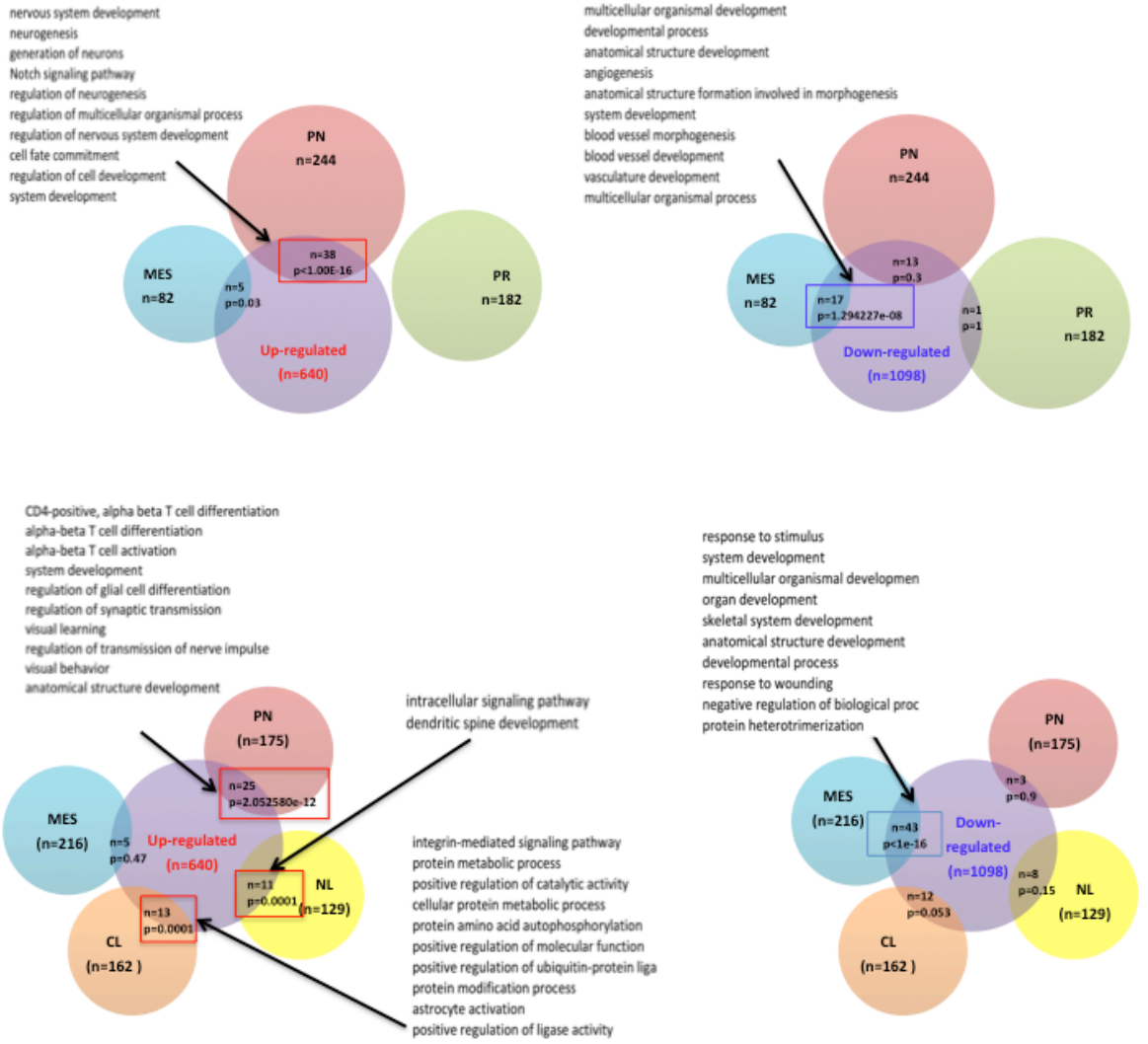

Figure 33. Up-regulated genes are enriched for Carro's and Verhaak's proneural signature genes. Comparison of our down- or up-regulated genes with Carro's (upper graphs) and Verhaak's (lower graphs) signatures. In both cases our up-regulated genes are enriched for genes belonging to the proneural signature.

4.4 Transcription factor binding site overrepresentation analysis

We examined our DEGs in order to define which were the transcription factors (TFs) consensus binding sequences present in their promoters. We defined overrepresented TFs in order to find a possible master regulator for gliomagenesis using the R package pcaGoPromoter [Hansen et al. 2012]. In developmental biology, a master regulator is defined as a gene (usually a transcription factor) that is necessary and sufficient to coordinate the activity of a broader group of genes in order to activate a developmental program specific for a certain tissue or organ. Since the process of tumorigenesis can be seen as an aberrant developmental program, it could be interesting to

pick out one or few genes that carry out the same function in the disease process. Possibly, the expression of our master regulator should have been controlled by H3K27me3.

With those parameters we found one TF, ZFP423, that is down-regulated and acquires the trimethylation of lysine 27 in Glio. Based on the consensus binding site, ZFP423 seems to control 219 genes in our DEG set, 128 of which are down-regulated and 81 up-regulated. This transcription factor is known to activate the transcription of BMP target genes by associating with SMADs in response to BMP2/4/7 [reviewed in Liu et al. 2005]. Moreover it can act as repressor of the olfactory neurogenesis by binding to EBF1 and preventing the transcription of EBF1 olfactory-specific genes [Roby et al. 2012]. Interestingly, another putative target of ZFP423 is *Dll3*, which seems involved in the block of astrogenesis (GO:0048712, negative regulation of astrocyte differentiation) and that in our data turned out to be up-regulated.

If validated, this could represent an interesting molecular mechanism through which one or few genes regulated by H3K27me3 can in turn regulate several molecular processes, leading to the final tumorigenic phenotype (e.g., persistence of an undifferentiated state of the cell).

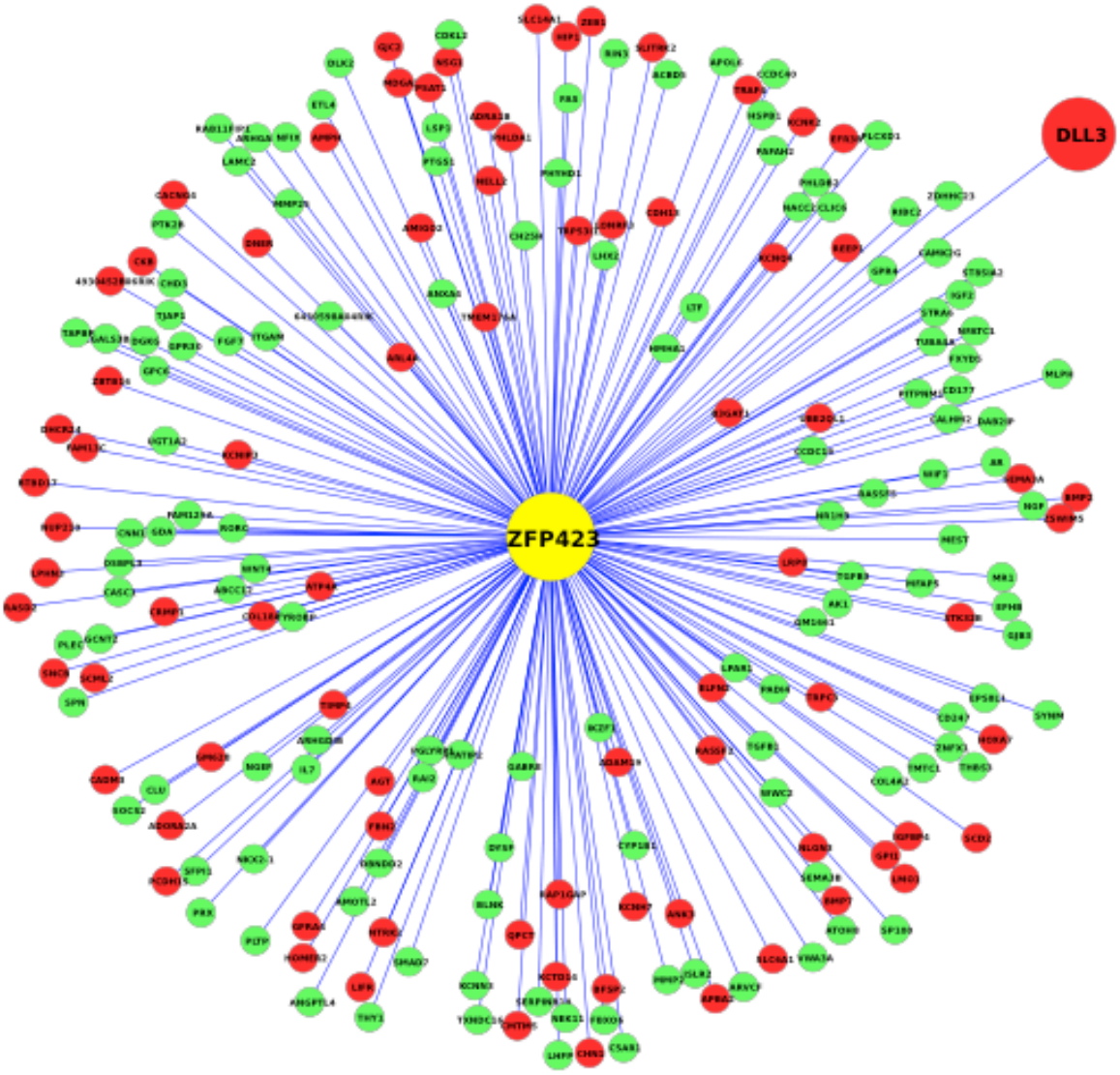

Figure 34. ZFP423 controls more than 200 DEGs. The transcription factor ZFP423 is *de novo* H3K27 trimethylated during gliomagenesis, is down-regulated and is overrepresented in our DEG list, controlling 219 genes (128 down-regulated, green circles and 81 up-regulated, red circles)

Discussion

The work presented in this thesis aimed at investigating the functional role of H3K27me3 in gliomagenesis.

In the last decade, several works contributed to change the general view of cancer as a genetic disease, demonstrating that epigenetics is deeply involved in the transforming process of cells and that genetics and epigenetics cooperate during cancer development. Particularly, several studies defined Polycomb group (PcG) proteins as key players in this process.

Historically, PcG were shown to exert a transcriptional repressive function that determines the patterned expression of the homeotic genes in Drosophila. Homeotic genes (*Hox* genes in vertebrates) are responsible for the specification of characteristic structures in each body segment along the antero-posterior axis and their expression needs to be tightly controlled both spatially and temporally ("homeo" means similar and homeotic mutants generally show a phenotype in which a body structure is replaced by another). Mutants for PcG proteins showed homeotic transformation of body segments due to an ectopic expression of homeotic genes. From this evidence, it was demonstrated that they exert a transcriptional repressive function. PcG proteins are assembled in two different multiproteic complexes, namely Polycomb Repressive Complex (PRC) 1 and 2. The current model proposes that PRC2 is responsible for the initiation of repression while PRC1 is involved in its maintenance. PRC2 is evolutionary conserved in invertebrates, vertebrates and plants and in mammals it contains four core members, namely EED, SUZ12, RBAP46/48 and EZH2. The proposed model suggests that PRC2 recruits histone deacetylases (HDAC) at target genes, causing deacetylation of histones. Through its active subunit, EZH2, it catalyzes the di- and tri-methylation of H3K27 which acts as a docking site for PRC1, that in turn ubiquitilates lysine 119 of histone H2A. Moreover, members of the PRCs have been shown to interact with the DNA methyl transferases [Vire et al. 2006, Mohammad et al. 2009], suggesting that they can actively control this further layer of epigenetic silencing. The epigenetic regulation mediated by PcG is inherited during cell division, and this cell memory is

necessary for the proper organism development by controlling cell growth and differentiation. Both these processes are deregulated in cancer and several studies demonstrated that this is linked with the aberrant expression of PcG proteins.

In my work I wanted to determine whether EZH2 plays a fundamental role in the set up of gliomas. To assess my hypothesis, I used two different mouse model systems. The first is mainly considered as an astrocytoma model [Bachoo et al. 2002], even if several evidences from mine and previous studies [Bachoo et al. 2002, Bruggeman et al. 2007] revealed that different histotypes can be originated with the same model (oligodendogliomas, oligoastrocytomas and gliosarcomas). The second is an oligodendrocytoma model [Appolloni et al. 2009]. Regardless of the histological outcome, these two models were chosen to assess if the role of EZH2 could vary according to different tumor-driving genetic lesions; indeed, while the first model relies on the loss of the *Ink4a/Arf* locus combined with the over-expression of a constitutively active mutant form of EGFR, EGFRvIII, the second is driven by the over-expression of PDGFB.

With both models, I was able to show that EZH2 depleted cells are counterselected in the first steps of glioma formation, indicating that the genetic lesions *per se* are not sufficient to get the tumor. Instead, they have to be accompanied by a re-setting of the epigenome, mediated by PRC2. Moreover, this appears to be relevant only in the earliest stages of cancer development.

It has been shown that epigenetic silencing is another way to achieve loss of function of a gene and sometimes this accounts for the “second hit”, the first being a genetic mutation. In colon cancer and in other cancer types it was shown that some genes are already silenced in pre-invasive stages of the disease [reviewed in Baylin and Ohm 2006]. These genes are defined as “epigenetic gatekeepers”, since their normal epigenetic modulation activates them during development and differentiation to prevent immortalization of stem or progenitor cells. The silencing of these genes leads to extended survival and proliferation and to a block in the differentiation of these cell populations, which are considered to be preliminary steps in

malignant transformation. Interestingly, among the genes undergoing epigenetic silencing there is HIC1, which is a transcription factor that negatively modulates the transcription of sirtuin 1 (SIRT1), a NAD-dependent histone deacetylase belonging to class III HDACs. SIRT1 is often found in association with a different Polycomb complex, called PRC4, in which EED isoform 1 is replaced by isoform 2 [Kuzmichev et al. 2005]. SIRT1 deacetylates histones, thus favouring PRC4-mediated methylation. Interestingly, PRC4 is able to methylate lysine 26 of histone H1 [Kuzmichev et al. 2004], which in turn can recruit heterochromatin binding protein 1 (HP1) causing a remodeling in higher order chromatin structure [Daujat et al. 2005]. Accumulation of EZH2 caused by over-expression often occurs in cancer as shown in prostate [Varambally et al. 2002], breast [Kleer et al. 2003] and bladder [Arisan et al. 2005, Raman et al. 2005, Weikert et al. 2005], but also in my own study. The different stoichiometry of the PRC members can facilitate the formation of this alternative PRC complex [Kuzmichev et al. 2005], adding another layer in the regulation of gene expression. Epigenetic changes happening in the early stages of glioma development may favor the expansion of cells harbouring genetic lesions, providing a substrate for accumulation of further genomic or epigenomic alterations. Furthermore, this aberrant system is maintained once established because H3K27me3, as well as genetic mutations, is mitotically inheritable [Margueron et al. 2009].

While I showed that EZH2 is necessary in the early stage of glioma formation, I demonstrated that it is dispensable in later phases, represented here by secondary tumorigenesis experiments. Another important epigenetic mark that is modified in early stages of cancer is DNA methylation. It was demonstrated that PcG proteins, and particularly EZH2, are able to directly control the deposition of this mark by interacting with the DNA methyl transferases (DNMTs) [Vire et al. 2006]. Generally, in normal tissues pericentromeric hetrochromatin and retrotransposons elements are hypermethylated, while CpG islands at gene promoters are hypomethylated. In cancer this paradigm is subverted and characterized by genome-wide hypomethylation causing genomic instability, while at some specific loci, CpG islands found in gene promoters are hypermethylated, causing transcriptional repression [reviewed in Robertson 2005]. Interestingly,

the majority of genes that undergo *de novo* DNA hypermethylation in cancer bear H3K27me3 in normal tissue [Ohm et al. 2007, Schlesinger et al. 2007, Widschwendter et al. 2007] and PcG targets genes are preferentially methylated in poorly differentiated tumors [Ben-Porath et al. 2008]. In their studies Widshwendter et al. and Ohm et al. postulated also that the DNA methylation pattern found in cancer might reflect the chromatin status of the cells from which the tumor developed, cells in which a reversible epigenetic repressive mark as H3K27me3 is replaced by a permanent silencing, locked in by DNA methylation. Also DNA methylation is mitotically inheritable [Hansen et al. 2008] and once established it can be passed over subsequent cell generations. My study raises the possibility that once driven by H3K27me3 and permanently established by DNMTs, DNA methylation could be sufficient to control the glioma gene expression pattern and thus no longer require PRC2 to be maintained.

The requirement of Ezh2 in the early period of gliomagenesis and its dispensability at later points was confirmed in both the systems chosen for this study, indicating that the critical role of Ezh2 is independent of the glioma-driving mutations. This finding sets the stage for further investigations aimed at identifying the existence of a common mechanism that regulates gliomagenesis. It is possible that one or few molecular pathways are universally altered at the earliest stages of different gliomas, either through genetic mutations or by epigenetic modification (e.g. epigenetic silencing, which is an alternative mechanism underlying gene loss of function) through an instructive mechanism. Once the common axis has been established, other alterations can contribute to the final phenotype, in terms of histotype, gene expression profile, prognosis and response to therapy. Those alterations might be due to the genomic instability or to other mechanisms that either in an adaptive or in a random manner, selected afterwards in a evolutionary fashion, hit different loci in the genome.

Taking advantage of the intrinsic possibility of my model systems to isolate homogeneous cell populations that have different tumorigenic potential, I explored the relocation of H3K27me3 and the gene expression profile modifications in tumorigenic astrocytes and primary tumor glioma propagating cells (GPCs) obtained in the astrocytoma model. Concerning the transcriptome

analysis, I identified about 1700 genes that were differentially expressed (i.e. consistently up- or down- regulated in two different samples) between tumorigenic astrocytes and primary GPCs. The transcriptomes of poorly differentiated gliomas show what has been defined as “embrionyc stem cell (ESC)-like signature”, in which PcG targets are under-expressed [Ben-Porath et al. 2008]; consistent with this finding, biological processes associated to our down-regulated gene set concern typical functions orchestrated by PcG, namely development and differentiation [Bracken et al. 2006]. The analysis of the redistribution of H3K27me3 revealed that almost 1600 genes are affected by the relocalization. Of these, roughly 1000 acquired the K27 trimethylation while about 550 lost it during glioma formation. Interestingly, comparing the H3K27me3 occupancy data with the transcriptome profiling, we found that only 100 of the 1098 down-regulated genes acquired the mark, while the vast majority genes belonging to this subset were unmethylated in astrocytes and retain this status also in primary GPCs. Strikingly, a small portion of genes that are up-regulated in gliomas acquired the H3K27me3; from previous studies of PcG target sites in both *Drosophila* and mammals it has already been shown that the presence of the H3K27me3 mark at gene promoters is compatible with transcriptional activity [Ringrose et al. 2004, Schwartz et al. 2006, Pasini et al. 2007]. All these results have several implication for the biology of gliomas. First, GPCs acquire and maintain a more undifferentiated profile, compared to tumorigenic astrocytes by inhibiting pathways related to development and differentiation; this is consistent with two major findings in the field that demonstrated how GPCs have aberrant developmental blocks making them similar to neural stem cells [Lee et al. 2008, Abdouh et al. 2009] and how poorly differentiated tumors restore an ESC-like molecular signature through the down-regulation of ESC Polycomb targets [Ben-Porath et al. 2008]. Second, the redistribution of the H3K27me3 mark happens at about 6-7% of genes in the genome; roughly two thirds (1049 out of 1603) of the involved genes acquire the mark, but only one tenth (100 out 1049) are down-regulated. This strengthens the hypothesis that PcG mediated regulation requires other effectors to be unfold, such as transcriptional activator or repressors, and that histone modifications can have different outcomes. The combination of different histone modifications on the same nucleosome [Voigt et

al. 2012], as for example the one described for bivalent domains in which H3K27me3 is coupled with H3K4me3 and results in low expression levels [Bernstein et al. 2006], or changes in the "reading machinery" of the histone code can affect the transcriptional activity. Third, in my comparative analysis I found that the vast majority of up- or down-regulated genes were H3K27 unmethylated in astrocytes and remained so in GPCs, suggesting that the regulation of few genes through epigenetic modifications can lead to a cascade of events that initiate circuitry effect on the entire transcriptome. This is quite well exemplified by the case of ZFP423, a transcription factor that is found to be H3K27me3 and downregulated in my GPCs samples. Moreover, the transcription factor binding site overrepresentation analysis showed that more than 10% of the genes among our DEGs (219 out of 1738) has a consensus sequence for ZFP423. This TF has both activating and repressive activity. Among its 219 targets in our DEG set, I identified the Delta-like protein 3 (DLL3) which is thought to block astrocyte differentiation. Both DLL3 and ZFP423 are involved in the Notch signaling pathway. When DLL3 binds to Notch receptor, the intracellular domain of the receptor is cleaved [Brou et al. 2000, Mumm et al. 2000] and induces the transcription of target genes including *Hes1* and *Hes5*, two basic helix-loop-helix (bHLH) transcription factors that are known to inhibit neuronal differentiation [Ohtsuka et al. 1999, Hatakeyama et al. 2004]. ZFP423 interacts with the intracellular domain of the Notch receptor inducing the transcription of *Hes5*; moreover since it is known to interact with SMAD proteins in response to BMP signaling, and that BMP can induce *Hes5* expression in neuroepithelial cells [Takizawa et al. 2003], it has been postulated that ZFP423 can coordinate the activity of the Notch and BMP pathways by regulating the expression of *Hes5* [Masserdotti et al. 2010]. If validated, this could be a clear example of a cascade mechanism by which the methylation of a single or at most few genes, particularly transcription factor with pleiotropic effects on the transcriptional activity of target genes, entailed multiple effects, including a differentiation block that helps GPCs maintain a NSC-like state.

The two glioma samples analyzed so far appear nearly identical at the transcriptional level but also in the distribution of H3K27me3, supporting the hypothesis of an instructive mechanism that

drives tumorigenesis, at least in the early phase. It will be very interesting now to examine the same properties in the second model that was chosen to probe the results obtained in the astrocyte-based model. If the trimethylation of H3K27 will hit loci that have already been identified as methylated in this preliminary analysis, it will be a further step in order to find a common basic mechanism that sets in to drive glioma formation. Also the comparison of the transcriptomes can reveal alterations that are taking place and whether there is a common underlying signature or pathway in different gliomas. Indeed, much of the work present in the literature of the last years is aimed at identifying differences, rather than common mechanisms, with the main focus of tailoring the therapy according to molecular alterations. The caveat in these studies is that they are performed at later stages of the disease, when the genomic instability and other processes could have further perturbed the initial (common) landscape.

The identification of a molecular signature in order to define molecular subclasses of gliomas has been strongly pursued in the last years; several studies tried to solve this issue, reporting successful identification of glioma profiles, but a unified picture is still missing [Phillips et al. 2006, Li et al. 2009, Verhaak et al 2010]. This can be due to different biases, as intrinsic biological variability, abnormal data distribution (e.g., skewed distribution that was thought to be Gaussian) in the tumor samples compared to normal tissue and analytical factors due to commonly employed algorithms [Marko et al. 2011]. I testified the lack of consistency among different data sets by comparing gene expression profiles published in three different studies [Phillips et al 2006, Carro et al. 2010, Verhaak et al. 2010]; despite the limited overlap among the reported signatures, up-regulated genes in my DEG dataset were always enriched for genes belonging to the proneural subtypes. The overlap of the three proneural signatures defined by Phillips, Carro and Verhaak is reduced to 21 genes, but 9 out of these 21 are also found to be upregulated in my experiment. I can then argue that, despite the weakness of the human molecular classification, my data seem to be invariably correlated with the human proneural profile. Two are the possible scenarios at this point. In the first, there is a correspondence between the mouse model and the human counterpart, and there is a minimal but significant group of genes that are overexpressed in the

two systems and that form a strong “core” for the proneural signature. The second instead is that, regardless of the driving mutation, modeling glioma in the mouse will always give rise to tumors belonging to the proneural signature. It has already been published that glioma developed using the $p53^{-/-}$/$Pten^{-/-}$/PDGF model correspond to this particular gene expression profile [Lei et al. 2011]. To partially settle this controversy, it is now of primary importance to extend my analysis on a larger cohort of samples and see if the relation with the human counterpart still holds. Even more important will be to repeat the same screening on the second glioma model I adopted.

In conclusion, results presented in this work revealed the relevance of Ezh2 activity during the early stages of gliomagenesis, which on the contrary is no longer required for the maintenance of the tumor. This suggests that there is a narrow time window in which the Polycomb axis is required and that probably this happens before the establishment of other molecular mechanisms that enable the GPC pool to escape from this dependency. Furthermore, H3K27me3 relocation upon tumorigenesis seems to happen consistently in different glioma samples, highlighting the instructive properties of this process and causing almost identical effects on the transcriptome, suggesting the existence of a “starting” PcG-dependent molecular signature that drives glioma formation.

References

Abdouh, M., Facchino, S., Chatoo, W., Balasingam, V., Ferreira, J., and Bernier, G. (2009). BMI1 sustains human glioblastoma multiforme stem cell renewal. J Neurosci *29*, 8884-8896.

Appolloni, I., Calzolari, F., Tutucci, E., Caviglia, S., Terrile, M., Corte, G., and Malatesta, P. (2009). PDGF-B induces a homogeneous class of oligodendrogliomas from embryonic neural progenitors. Int J Cancer *124*, 2251-2259.

Arisan, S., Buyuktuncer, E.D., Palavan-Unsal, N., Caskurlu, T., Cakir, O.O., and Ergenekon, E. (2005). Increased expression of EZH2, a polycomb group protein, in bladder carcinoma. Urol Int *75*, 252-257.

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H.F., John, R.M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M.*, et al.* (2006). Chromatin signatures of pluripotent cell lines. Nat Cell Biol *8*, 532-538.

Bachoo, R.M., Maher, E.A., Ligon, K.L., Sharpless, N.E., Chan, S.S., You, M.J., Tang, Y., DeFrances, J., Stover, E., Weissleder, R.*, et al.* (2002). Epidermal growth factor receptor and Ink4a/Arf: convergent mechanisms governing terminal differentiation and transformation along the neural stem cell to astrocyte axis. Cancer Cell *1*, 269-277.

Baumann, N., and Pham-Dinh, D. (2001). Biology of oligodendrocyte and myelin in the mammalian central nervous system. Physiol Rev *81*, 871-927.

Baylin, S.B., and Ohm, J.E. (2006). Epigenetic gene silencing in cancer - a mechanism for early oncogenic pathway addiction? Nat Rev Cancer *6*, 107-116.

Bea, S., Tort, F., Pinyol, M., Puig, X., Hernandez, L., Hernandez, S., Fernandez, P.L., van Lohuizen, M., Colomer, D., and Campo, E. (2001). BMI-1 gene amplification and overexpression in hematological malignancies occur mainly in mantle cell lymphomas. Cancer Res *61*, 2409-2412.

Ben-Porath, I., Thomson, M.W., Carey, V.J., Ge, R., Bell, G.W., Regev, A., and Weinberg, R.A. (2008). An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. Nat Genet *40*, 499-507.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K.*, et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.

Bertwistle, D., Sugimoto, M., and Sherr, C.J. (2004). Physical and functional interactions of the Arf tumor suppressor protein with nucleophosmin/B23. Mol Cell Biol *24*, 985-996.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev *16*, 6-21.

Birnboim, H.C., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res *7*, 1513-1523.

Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K.*, et al.* (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature *441*, 349-353.

Bracken, A.P., Dietrich, N., Pasini, D., Hansen, K.H., and Helin, K. (2006). Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes Dev *20*, 1123-1136.

Bracken, A.P., Kleine-Kohlbrecher, D., Dietrich, N., Pasini, D., Gargiulo, G., Beekman, C., Theilgaard-Monch, K., Minucci, S., Porse, B.T., Marine, J.C.*, et al.* (2007). The Polycomb group

proteins bind throughout the INK4A-ARF locus and are disassociated in senescent cells. Genes Dev *21*, 525-530.

Bracken, A.P., Pasini, D., Capra, M., Prosperini, E., Colli, E., and Helin, K. (2003). EZH2 is downstream of the pRB-E2F pathway, essential for proliferation and amplified in cancer. EMBO J *22*, 5323-5335.

Brou, C., Logeat, F., Gupta, N., Bessia, C., LeBail, O., Doedens, J.R., Cumano, A., Roux, P., Black, R.A., and Israel, A. (2000). A novel proteolytic cleavage involved in Notch signaling: the role of the disintegrin-metalloprotease TACE. Mol Cell *5*, 207-216.

Bruggeman, S.W., Hulsman, D., Tanger, E., Buckle, T., Blom, M., Zevenhoven, J., van Tellingen, O., and van Lohuizen, M. (2007). Bmi1 controls tumor development in an Ink4a/Arf-independent manner in a mouse model for glioma. Cancer Cell *12*, 328-341.

Cahoy, J.D., Emery, B., Kaushal, A., Foo, L.C., Zamanian, J.L., Christopherson, K.S., Xing, Y., Lubischer, J.L., Krieg, P.A., Krupenko, S.A*., et al.* (2008). A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. J Neurosci *28*, 264-278.

Cai, J., Chen, Y., Cai, W.H., Hurlock, E.C., Wu, H., Kernie, S.G., Parada, L.F., and Lu, Q.R. (2007). A crucial role for Olig2 in white matter astrocyte development. Development *134*, 1887-1899.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science *298*, 1039-1043.

Cao, R., and Zhang, Y. (2004). SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. Mol Cell *15*, 57-67.

Carro, M.S., Lim, W.K., Alvarez, M.J., Bollo, R.J., Zhao, X., Snyder, E.Y., Sulman, E.P., Anne, S.L., Doetsch, F., Colman, H*., et al.* The transcriptional network for mesenchymal transformation of brain tumours. Nature *463*, 318-325.

Chen, J., McKay, R.M., and Parada, L.F. (2012). Malignant glioma: lessons from genomics, mouse models, and stem cells. Cell *149*, 36-47.

Choe, G., Horvath, S., Cloughesy, T.F., Crosby, K., Seligson, D., Palotie, A., Inge, L., Smith, B.L., Sawyers, C.L., and Mischel, P.S. (2003). Analysis of the phosphatidylinositol 3'-kinase signaling pathway in glioblastoma patients in vivo. Cancer Res *63*, 2742-2746.

Clarke, M.F., Dick, J.E., Dirks, P.B., Eaves, C.J., Jamieson, C.H., Jones, D.L., Visvader, J., Weissman, I.L., and Wahl, G.M. (2006). Cancer stem cells--perspectives on current status and future directions: AACR Workshop on cancer stem cells. Cancer Res *66*, 9339-9344.

Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A., and Pirrotta, V. (2002). Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. Cell *111*, 185-196.

Daujat, S., Zeissler, U., Waldmann, T., Happel, N., and Schneider, R. (2005). HP1 binds specifically to Lys26-methylated histone H1.4, whereas simultaneous Ser27 phosphorylation blocks HP1 binding. J Biol Chem *280*, 38090-38095.

Doetsch, F., Caille, I., Lim, D.A., Garcia-Verdugo, J.M., and Alvarez-Buylla, A. (1999). Subventricular zone astrocytes are neural stem cells in the adult mammalian brain. Cell *97*, 703-716.

Ekstrand, A.J., Longo, N., Hamid, M.L., Olson, J.J., Liu, L., Collins, V.P., and James, C.D. (1994). Functional characterization of an EGF receptor with a truncated extracellular domain expressed in glioblastomas with EGFR gene amplification. Oncogene *9*, 2313-2320.

Ekstrand, A.J., Sugawa, N., James, C.D., and Collins, V.P. (1992). Amplified and rearranged epidermal growth factor receptor genes in human glioblastomas reveal deletions of sequences encoding portions of the N- and/or C-terminal tails. Proc Natl Acad Sci U S A *89*, 4309-4313.

Ezhkova, E., Pasolli, H.A., Parker, J.S., Stokes, N., Su, I.H., Hannon, G., Tarakhovsky, A., and Fuchs, E. (2009). Ezh2 orchestrates gene expression for the stepwise differentiation of tissue-specific stem cells. Cell *136*, 1122-1135.

Freije, W.A., Castro-Vargas, F.E., Fang, Z., Horvath, S., Cloughesy, T., Liau, L.M., Mischel, P.S., and Nelson, S.F. (2004). Gene expression profiling of gliomas strongly predicts survival. Cancer Res *64*, 6503-6510.

Friedmann-Morvinski, D., Bushong, E.A., Ke, E., Soda, Y., Marumoto, T., Singer, O., Ellisman, M.H., and Verma, I.M. (2012). Dedifferentiation of Neurons and Astrocytes by Oncogenes Can Induce Gliomas in Mice. Science.

Fulci, G., Labuhn, M., Maier, D., Lachat, Y., Hausmann, O., Hegi, M.E., Janzer, R.C., Merlo, A., and Van Meir, E.G. (2000). p53 gene mutation and ink4a-arf deletion appear to be two mutually exclusive events in human glioblastoma. Oncogene *19*, 3816-3822.

Furnari, F.B., Fenton, T., Bachoo, R.M., Mukasa, A., Stommel, J.M., Stegh, A., Hahn, W.C., Ligon, K.L., Louis, D.N., Brennan, C.*, et al.* (2007). Malignant astrocytic glioma: genetics, biology, and paths to treatment. Genes Dev *21*, 2683-2710.

Galli, R., Binda, E., Orfanelli, U., Cipelletti, B., Gritti, A., De Vitis, S., Fiocco, R., Foroni, C., Dimeco, F., and Vescovi, A. (2004). Isolation and characterization of tumorigenic, stem-like neural precursors from human glioblastoma. Cancer Res *64*, 7011-7021.

Garcia de Palazzo, I.E., Adams, G.P., Sundareshan, P., Wong, A.J., Testa, J.R., Bigner, D.D., and Weiner, L.M. (1993). Expression of mutated epidermal growth factor receptor by non-small cell lung carcinomas. Cancer Res *53*, 3217-3220.

Garcia-Marques, J., and Lopez-Mascaraque, L. (2012). Clonal Identity Determines Astrocyte Cortical Heterogeneity. Cereb Cortex [Epub ahead of print]

Godard, S., Getz, G., Delorenzi, M., Farmer, P., Kobayashi, H., Desbaillets, I., Nozaki, M., Diserens, A.C., Hamou, M.F., Dietrich, P.Y.*, et al.* (2003). Classification of human astrocytic gliomas on the basis of gene expression: a correlated group of genes with angiogenic activity emerges as a strong predictor of subtypes. Cancer Res *63*, 6613-6625.

Gurdon, J. (1962). Adult frogs derived from the nuclei of single somatic cells. Developmental biology *4*, 256-273.

Hambardzumyan, D., Parada, L.F., Holland, E.C., and Charest, A. (2011). Genetic modeling of gliomas in mice: new tools to tackle old problems. Glia *59*, 1155-1168.

Hansen, K.H., Bracken, A.P., Pasini, D., Dietrich, N., Gehani, S.S., Monrad, A., Rappsilber, J., Lerdrup, M., and Helin, K. (2008). A model for transmission of the H3K27me3 epigenetic mark. Nat Cell Biol *10*, 1291-1300.

Hansen, M., Gerds, T.A., Nielsen, O.H., Seidelin, J.B., Troelsen, J.T., and Olsen, J. (2012). pcaGoPromoter--an R package for biological and regulatory interpretation of principal components in genome-wide gene expression data. PLoS One *7*, e32394.

Hatakeyama, J., Bessho, Y., Katoh, K., Ookawara, S., Fujioka, M., Guillemot, F., and Kageyama, R. (2004). Hes genes regulate size, shape and histogenesis of the nervous system by control of the timing of neural stem cell differentiation. Development *131*, 5539-5550.

Hayashi, Y., Ueki, K., Waha, A., Wiestler, O.D., Louis, D.N., and von Deimling, A. (1997). Association of EGFR gene amplification and CDKN2 (p16/MTS1) gene deletion in glioblastoma multiforme. Brain Pathol *7*, 871-875.

Hegi, M.E., zur Hausen, A., Ruedi, D., Malin, G., and Kleihues, P. (1997). Hemizygous or homozygous deletion of the chromosomal region containing the p16INK4a gene is associated with amplification of the EGF receptor gene in glioblastomas. Int J Cancer *73*, 57-63.

Hemmati, H.D., Nakano, I., Lazareff, J.A., Masterman-Smith, M., Geschwind, D.H., Bronner-Fraser, M., and Kornblum, H.I. (2003). Cancerous stem cells can arise from pediatric brain tumors. Proc Natl Acad Sci U S A *100*, 15178-15183.

Hirabayashi, Y., Suzki, N., Tsuboi, M., Endo, T.A., Toyoda, T., Shinga, J., Koseki, H., Vidal, M., and Gotoh, Y. (2009). Polycomb limits the neurogenic competence of neural precursor cells to promote astrogenic fate transition. Neuron *63*, 600-613.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

Huang, P.H., Mukasa, A., Bonavia, R., Flynn, R.A., Brewer, Z.E., Cavenee, W.K., Furnari, F.B., and White, F.M. (2007). Quantitative analysis of EGFRvIII cellular signaling networks reveals a combinatorial therapeutic strategy for glioblastoma. Proc Natl Acad Sci U S A *104*, 12867-12872.

Ignatova, T.N., Kukekov, V.G., Laywell, E.D., Suslov, O.N., Vrionis, F.D., and Steindler, D.A. (2002). Human cortical glial tumors contain neural stem-like cells expressing astroglial and neuronal markers in vitro. Glia *39*, 193-206.

Itahana, K., Bhat, K.P., Jin, A., Itahana, Y., Hawke, D., Kobayashi, R., and Zhang, Y. (2003). Tumor suppressor ARF degrades B23, a nucleolar protein involved in ribosome biogenesis and cell proliferation. Mol Cell *12*, 1151-1164.

Jackson, E.L., Garcia-Verdugo, J.M., Gil-Perotin, S., Roy, M., Quinones-Hinojosa, A., VandenBerg, S., and Alvarez-Buylla, A. (2006). PDGFR alpha-positive B cells are neural stem cells in the adult SVZ that form glioma-like growths in response to increased PDGF signaling. Neuron *51*, 187-199.

Jacobs, J.J., Kieboom, K., Marino, S., DePinho, R.A., and van Lohuizen, M. (1999b). The oncogene and Polycomb-group gene bmi-1 regulates cell proliferation and senescence through the ink4a locus. Nature *397*, 164-168.

Jacobs, J.J., Scheijen, B., Voncken, J.W., Kieboom, K., Berns, A., and van Lohuizen, M. (1999a). Bmi-1 collaborates with c-Myc in tumorigenesis by inhibiting c-Myc-induced apoptosis via INK4a/ARF. Genes Dev *13*, 2678-2690.

James, C.D., Carlbom, E., Dumanski, J.P., Hansen, M., Nordenskjold, M., Collins, V.P., and Cavenee, W.K. (1988). Clonal genomic alterations in glioma malignancy stages. Cancer Res *48*, 5546-5551.

Jaros, E., Perry, R.H., Adam, L., Kelly, P.J., Crawford, P.J., Kalbag, R.M., Mendelow, A.D., Sengupta, R.P., and Pearson, A.D. (1992). Prognostic implications of p53 protein, epidermal growth factor receptor, and Ki-67 labelling in brain tumours. Br J Cancer *66*, 373-385.

Kamijo, T., Zindy, F., Roussel, M.F., Quelle, D.E., Downing, J.R., Ashmun, R.A., Grosveld, G., and Sherr, C.J. (1997). Tumor suppression at the mouse INK4a locus mediated by the alternative reading frame product p19ARF. Cell *91*, 649-659.

Ketel, C.S., Andersen, E.F., Vargas, M.L., Suh, J., Strome, S., and Simon, J.A. (2005). Subunit contributions to histone methyltransferase activities of fly and worm polycomb group complexes. Mol Cell Biol *25*, 6857-6868.

Kimelberg, H.K. (2004). The problem of astrocyte identity. Neurochem Int *45*, 191-202.

Kirmizis, A., Bartley, S.M., and Farnham, P.J. (2003). Identification of the polycomb group protein SU(Z)12 as a potential molecular target for human cancer therapy. Mol Cancer Ther *2*, 113-121.

Kirmizis, A., Bartley, S.M., Kuzmichev, A., Margueron, R., Reinberg, D., Green, R., and Farnham, P.J. (2004). Silencing of human polycomb target genes is associated with methylation of histone H3 Lys 27. Genes Dev *18*, 1592-1605.

Kleer, C.G., Cao, Q., Varambally, S., Shen, R., Ota, I., Tomlins, S.A., Ghosh, D., Sewalt, R.G., Otte, A.P., Hayes, D.F*., et al.* (2003). EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. Proc Natl Acad Sci U S A *100*, 11606-11611.

Kondo, T., Setoguchi, T., and Taga, T. (2004). Persistence of a small subpopulation of cancer stem-like cells in the C6 glioma cell line. Proc Natl Acad Sci U S A *101*, 781-786.

Krishnamurthy, J., Ramsey, M.R., Ligon, K.L., Torrice, C., Koh, A., Bonner-Weir, S., and Sharpless, N.E. (2006). p16INK4a induces an age-dependent decline in islet regenerative potential. Nature *443*, 453-457.

Kuzmichev, A., Jenuwein, T., Tempst, P., and Reinberg, D. (2004). Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. Mol Cell *14*, 183-193.

Kuzmichev, A., Margueron, R., Vaquero, A., Preissner, T.S., Scher, M., Kirmizis, A., Ouyang, X., Brockdorff, N., Abate-Shen, C., Farnham, P*., et al.* (2005). Composition and histone substrates of polycomb repressive group complexes change during cellular differentiation. Proc Natl Acad Sci U S A *102*, 1859-1864.

Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2002). Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes Dev *16*, 2893-2905.

Kwon, C.H., Zhao, D., Chen, J., Alcantara, S., Li, Y., Burns, D.K., Mason, R.P., Lee, E.Y., Wu, H., and Parada, L.F. (2008). Pten haploinsufficiency accelerates formation of high-grade astrocytomas. Cancer Res *68*, 3286-3294.

Labuhn, M., Jones, G., Speel, E.J., Maier, D., Zweifel, C., Gratzl, O., Van Meir, E.G., Hegi, M.E., and Merlo, A. (2001). Quantitative real-time PCR does not show selective targeting of p14(ARF) but concomitant inactivation of both p16(INK4A) and p14(ARF) in 105 human primary gliomas. Oncogene *20*, 1103-1109.

Lachat, Y., Diserens, A.C., Nozaki, M., Kobayashi, H., Hamou, M.F., Godard, S., De Tribolet, N., and Hegi, M.E. (2004). INK4a/Arf is required for suppression of EGFR/DeltaEGFR(2-7)-dependent ERK activation in mouse astrocytes and glioma. Oncogene *23*, 6854-6863.

Laible, G., Haynes, A.R., Lebersorger, A., O'Carroll, D., Mattei, M.G., Denny, P., Brown, S.D., and Jenuwein, T. (1999). The murine polycomb-group genes Ezh1 and Ezh2 map close to Hox gene clusters on mouse chromosomes 11 and 6. Mamm Genome *10*, 311-314.

Laird, P.W., Zijderveld, A., Linders, K., Rudnicki, M.A., Jaenisch, R., and Berns, A. (1991). Simplified mammalian DNA isolation procedure. Nucleic Acids Res *19*, 4293.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Lathia, J.D., Gallagher, J., Heddleston, J.M., Wang, J., Eyler, C.E., Macswords, J., Wu, Q., Vasanji, A., McLendon, R.E., Hjelmeland, A.B.*, et al.* (2010). Integrin alpha 6 regulates glioblastoma stem cells. Cell Stem Cell *6*, 421-432.

Laywell, E.D., Rakic, P., Kukekov, V.G., Holland, E.C., and Steindler, D.A. (2000). Identification of a multipotent astrocytic stem cell in the immature and adult mouse brain. Proc Natl Acad Sci U S A *97*, 13883-13888.

Lee, J., Son, M.J., Woolard, K., Donin, N.M., Li, A., Cheng, C.H., Kotliarova, S., Kotliarov, Y., Walling, J., Ahn, S.*, et al.* (2008). Epigenetic-mediated dysfunction of the bone morphogenetic protein pathway inhibits differentiation of glioblastoma-initiating cells. Cancer Cell *13*, 69-80.

Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F., Isono, K.*, et al.* (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. Cell *125*, 301-313.

Lei, L., Sonabend, A.M., Guarnieri, P., Soderquist, C., Ludwig, T., Rosenfeld, S., Bruce, J.N., and Canoll, P. (2011). Glioblastoma models reveal the connection between adult glial progenitors and the proneural phenotype. PLoS One *6*, e20041.

Lendahl, U., Zimmerman, L.B., and McKay, R.D. (1990). CNS stem cells express a new class of intermediate filament protein. Cell *60*, 585-595.

Lewis, E.B. (1978). A gene complex controlling segmentation in Drosophila. Nature *276*, 565-570.

Li, A., Walling, J., Ahn, S., Kotliarov, Y., Su, Q., Quezado, M., Oberholtzer, J.C., Park, J., Zenklusen, J.C., and Fine, H.A. (2009). Unsupervised analysis of transcriptomic profiles reveals six glioma subtypes. Cancer Res *69*, 2091-2099.

Liang, Y., Diehn, M., Watson, N., Bollen, A.W., Aldape, K.D., Nicholas, M.K., Lamborn, K.R., Berger, M.S., Botstein, D., Brown, P.O.*, et al.* (2005). Gene expression profiling reveals molecularly and clinically distinct subtypes of glioblastoma multiforme. Proc Natl Acad Sci U S A *102*, 5814-5819.

Libermann, T.A., Nusbaum, H.R., Razon, N., Kris, R., Lax, I., Soreq, H., Whittle, N., Waterfield, M.D., Ullrich, A., and Schlessinger, J. (1985). Amplification, enhanced expression and possible rearrangement of EGF receptor gene in primary human brain tumours of glial origin. Nature *313*, 144-147.

Libermann, T.A., Razon, N., Bartal, A.D., Yarden, Y., Schlessinger, J., and Soreq, H. (1984). Expression of epidermal growth factor receptors in human brain tumors. Cancer Res *44*, 753-760.

Liu, A., and Niswander, L.A. (2005). Bone morphogenetic protein signalling and vertebrate nervous system development. Nat Rev Neurosci *6*, 945-954.

Liu, C., Sage, J.C., Miller, M.R., Verhaak, R.G., Hippenmeyer, S., Vogel, H., Foreman, O., Bronson, R.T., Nishiyama, A., Luo, L.*, et al.* (2011). Mosaic analysis with double markers reveals tumor cell of origin in glioma. Cell *146*, 209-221.

Louis, D.N., Ohgaki, H., Wiestler, O.D., Cavenee, W.K., Burger, P.C., Jouvet, A., Scheithauer, B.W., and Kleihues, P. (2007). The 2007 WHO classification of tumours of the central nervous system. Acta Neuropathol *114*, 97-109.

Lu, Q.R., Yuk, D., Alberta, J.A., Zhu, Z., Pawlitzky, I., Chan, J., McMahon, A.P., Stiles, C.D., and Rowitch, D.H. (2000). Sonic hedgehog--regulated oligodendrocyte lineage genes encoding bHLH proteins in the mammalian central nervous system. Neuron *25*, 317-329.

Maere, S., Heymans, K., and Kuiper, M. (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics *21*, 3448-3449.

Magnus, N., Garnier, D., and Rak, J. (2010). Oncogenic epidermal growth factor receptor up-regulates multiple elements of the tissue factor signaling pathway in human glioma cells. Blood *116*, 815-818.

Maher, E.A., Brennan, C., Wen, P.Y., Durso, L., Ligon, K.L., Richardson, A., Khatry, D., Feng, B., Sinha, R., Louis, D.N.*, et al.* (2006). Marked genomic differences characterize primary and secondary glioblastoma subtypes and identify two distinct molecular and clinical secondary glioblastoma entities. Cancer Res *66*, 11502-11513.

Margueron, R., Justin, N., Ohno, K., Sharpe, M.L., Son, J., Drury, W.J., 3rd, Voigt, P., Martin, S.R., Taylor, W.R., De Marco, V.*, et al.* (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. Nature *461*, 762-767.

Marko, N.F., Quackenbush, J., and Weil, R.J. (2011) Why is there a lack of consensus on molecular subgroups of glioblastoma? Understanding the nature of biological and statistical variability in glioblastoma expression data. PLoS One *6*, e20826.

Masserdotti, G., Badaloni, A., Green, Y.S., Croci, L., Barili, V., Bergamini, G., Vetter, M.L., and Consalez, G.G. ZFP423 coordinates Notch and bone morphogenetic protein signaling, selectively up-regulating Hes5 gene expression. J Biol Chem *285*, 30814-30824.

McClintock, C.B. (1950). The origin and behavior of mutable loci in maize. Proc Natl Acad Sci U S A *36*, 344-355.

McCarthy, K.D., and de Vellis, J. (1980). Preparation of separate astroglial and oligodendroglial cell cultures from rat cerebral tissue. J Cell Biol *85*, 890-902.

Mohammad, H.P., Cai, Y., McGarvey, K.M., Easwaran, H., Van Neste, L., Ohm, J.E., O'Hagan, H.M., and Baylin, S.B. (2009). Polycomb CBX7 promotes initiation of heritable repression of genes frequently silenced with cancer-specific DNA hypermethylation. Cancer Res *69*, 6322-6330.

Mohn, F., Weber, M., Rebhan, M., Roloff, T.C., Richter, J., Stadler, M.B., Bibel, M., and Schubeler, D. (2008). Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. Mol Cell *30*, 755-766.

Molofsky, A.V., Slutsky, S.G., Joseph, N.M., He, S., Pardal, R., Krishnamurthy, J., Sharpless, N.E., and Morrison, S.J. (2006). Increasing p16INK4a expression decreases forebrain progenitors and neurogenesis during ageing. Nature *443*, 448-452.

Montgomery, N.D., Yee, D., Chen, A., Kalantry, S., Chamberlain, S.J., Otte, A.P., and Magnuson, T. (2005). The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. Curr Biol *15*, 942-947.

Moscatello, D.K., Holgado-Madruga, M., Emlet, D.R., Montgomery, R.B., and Wong, A.J. (1998). Constitutive activation of phosphatidylinositol 3-kinase by a naturally occurring mutant epidermal growth factor receptor. J Biol Chem *273*, 200-206.

Moscatello, D.K., Holgado-Madruga, M., Godwin, A.K., Ramirez, G., Gunn, G., Zoltick, P.W., Biegel, J.A., Hayes, R.L., and Wong, A.J. (1995). Frequent expression of a mutant epidermal growth factor receptor in multiple human tumors. Cancer Res *55*, 5536-5539.

Mulholland, N.M., King, I.F., and Kingston, R.E. (2003). Regulation of Polycomb group complexes by the sequence-specific DNA binding proteins Zeste and GAGA. Genes Dev *17*, 2741-2746.

Muller, H., Bracken, A.P., Vernell, R., Moroni, M.C., Christians, F., Grassilli, E., Prosperini, E., Vigo, E., Oliner, J.D., and Helin, K. (2001). E2Fs regulate the expression of genes involved in differentiation, development, proliferation, and apoptosis. Genes Dev *15*, 267-285.

Muller, H.J., and Altenburg, E. (1930). The Frequency of Translocations Produced by X-Rays in Drosophila. Genetics *15*, 283-311.

Muller, J., Hart, C.M., Francis, N.J., Vargas, M.L., Sengupta, A., Wild, B., Miller, E.L., O'Connor, M.B., Kingston, R.E., and Simon, J.A. (2002). Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. Cell *111*, 197-208.

Mumm, J.S., Schroeter, E.H., Saxena, M.T., Griesemer, A., Tian, X., Pan, D.J., Ray, W.J., and Kopan, R. (2000). A ligand-induced extracellular cleavage regulates gamma-secretase-like proteolytic activation of Notch1. Mol Cell *5*, 197-206.

Nagane, M., Levitzki, A., Gazit, A., Cavenee, W.K., and Huang, H.J. (1998). Drug resistance of human glioblastoma cells conferred by a tumor-specific mutant epidermal growth factor receptor through modulation of Bcl-XL and caspase-3-like proteases. Proc Natl Acad Sci U S A *95*, 5724-5729.

Nekrasov, M., Wild, B., and Muller, J. (2005). Nucleosome binding and histone methyltransferase activity of Drosophila PRC2. EMBO Rep *6*, 348-353.

Nigro, J.M., Misra, A., Zhang, L., Smirnov, I., Colman, H., Griffin, C., Ozburn, N., Chen, M., Pan, E., Koul, D*., et al.* (2005). Integrated array-comparative genomic hybridization and expression array profiles identify clinically relevant molecular subtypes of glioblastoma. Cancer Res *65*, 1678-1686.

Nishikawa, R., Ji, X.D., Harmon, R.C., Lazar, C.S., Gill, G.N., Cavenee, W.K., and Huang, H.J. (1994). A mutant epidermal growth factor receptor common in human glioma confers enhanced tumorigenicity. Proc Natl Acad Sci U S A *91*, 7727-7731.

Nutt, C.L., Mani, D.R., Betensky, R.A., Tamayo, P., Cairncross, J.G., Ladd, C., Pohl, U., Hartmann, C., McLaughlin, M.E., Batchelor, T.T*., et al.* (2003). Gene expression-based classification of malignant gliomas correlates better with survival than histological classification. Cancer Res *63*, 1602-1607.

Ogawa, H., Ishiguro, K., Gaubatz, S., Livingston, D.M., and Nakatani, Y. (2002). A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells. Science *296*, 1132-1136.

Ogden, A.T., Waziri, A.E., Lochhead, R.A., Fusco, D., Lopez, K., Ellis, J.A., Kang, J., Assanah, M., McKhann, G.M., Sisti, M.B.*, et al.* (2008). Identification of A2B5+CD133- tumor-initiating cells in adult human gliomas. Neurosurgery *62*, 505-514; discussion 514-505.

Ohm, J.E., McGarvey, K.M., Yu, X., Cheng, L., Schuebel, K.E., Cope, L., Mohammad, H.P., Chen, W., Daniel, V.C., Yu, W.*, et al.* (2007). A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat Genet *39*, 237-242.

Ohtsuka, T., Ishibashi, M., Gradwohl, G., Nakanishi, S., Guillemot, F., and Kageyama, R. (1999). Hes1 and Hes5 as notch effectors in mammalian neuronal differentiation. EMBO J *18*, 2196-2207.

Parsons, D.W., Jones, S., Zhang, X., Lin, J.C., Leary, R.J., Angenendt, P., Mankoo, P., Carter, H., Siu, I.M., Gallia, G.L.*, et al.* (2008). An integrated genomic analysis of human glioblastoma multiforme. Science *321*, 1807-1812.

Pasini, D., Bracken, A.P., Hansen, J.B., Capillo, M., and Helin, K. (2007). The polycomb group protein Suz12 is required for embryonic stem cell differentiation. Mol Cell Biol *27*, 3769-3779.

Pasini, D., Bracken, A.P., Jensen, M.R., Lazzerini Denchi, E., and Helin, K. (2004). Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. EMBO J *23*, 4061-4071.

Peitz, M., Pfannkuche, K., Rajewsky, K., and Edenhofer, F. (2002). Ability of the hydrophobic FGF and basic TAT peptides to promote cellular uptake of recombinant Cre recombinase: a tool for efficient genetic engineering of mammalian genomes. Proc Natl Acad Sci U S A *99*, 4489-4494.

Pereira, J.D., Sansom, S.N., Smith, J., Dobenecker, M.W., Tarakhovsky, A., and Livesey, F.J. (2010). Ezh2, the histone methyltransferase of PRC2, regulates the balance between self-renewal and differentiation in the cerebral cortex. Proc Natl Acad Sci U S A *107*, 15957-15962.

Phillips, H.S., Kharbanda, S., Chen, R., Forrest, W.F., Soriano, R.H., Wu, T.D., Misra, A., Nigro, J.M., Colman, H., Soroceanu, L.*, et al.* (2006). Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. Cancer Cell *9*, 157-173.

Pollard, S.M., Yoshikawa, K., Clarke, I.D., Danovi, D., Stricker, S., Russell, R., Bayani, J., Head, R., Lee, M., Bernstein, M.*, et al.* (2009). Glioma stem cell lines expanded in adherent culture have tumor-specific phenotypes and are suitable for chemical and genetic screens. Cell Stem Cell *4*, 568-580.

Quelle, D.E., Zindy, F., Ashmun, R.A., and Sherr, C.J. (1995). Alternative reading frames of the INK4a tumor suppressor gene encode two unrelated proteins capable of inducing cell cycle arrest. Cell *83*, 993-1000.

Raaphorst, F.M., Meijer, C.J., Fieret, E., Blokzijl, T., Mommers, E., Buerger, H., Packeisen, J., Sewalt, R.A., Otte, A.P., and van Diest, P.J. (2003). Poorly differentiated breast carcinoma is associated with increased expression of the human polycomb group EZH2 gene. Neoplasia *5*, 481-488.

Raman, J.D., Mongan, N.P., Tickoo, S.K., Boorjian, S.A., Scherr, D.S., and Gudas, L.J. (2005). Increased expression of the polycomb group gene, EZH2, in transitional cell carcinoma of the bladder. Clin Cancer Res *11*, 8570-8576.

Ringrose, L., Ehret, H., and Paro, R. (2004). Distinct contributions of histone H3 lysine 9 and 27 methylation to locus-specific stability of polycomb complexes. Mol Cell *16*, 641-653.

Robertson, K.D. (2005). DNA methylation and human disease. Nat Rev Genet *6*, 597-610.

Roby, Y.A., Bushey, M.A., Cheng, L.E., Kulaga, H.M., Lee, S.J., and Reed, R.R. (2012). Zfp423/OAZ Mutation Reveals the Importance of Olf/EBF Transcription Activity in Olfactory Neuronal Maturation. J Neurosci *32*, 13679-13688.

Sarma, K., Margueron, R., Ivanov, A., Pirrotta, V., and Reinberg, D. (2008). Ezh2 requires PHF1 to efficiently catalyze H3 lysine 27 trimethylation in vivo. Mol Cell Biol *28*, 2718-2731.

Schlesinger, Y., Straussman, R., Keshet, I., Farkash, S., Hecht, M., Zimmerman, J., Eden, E., Yakhini, Z., Ben-Shushan, E., Reubinoff, B.E*., et al.* (2007). Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat Genet *39*, 232-236.

Schwartz, Y.B., Kahn, T.G., Nix, D.A., Li, X.Y., Bourgon, R., Biggin, M., and Pirrotta, V. (2006). Genome-wide analysis of Polycomb targets in Drosophila melanogaster. Nat Genet *38*, 700-705.

Seibler, J., Zevnik, B., Kuter-Luks, B., Andreas, S., Kern, H., Hennek, T., Rode, A., Heimann, C., Faust, N., Kauselmann, G*., et al.* (2003). Rapid generation of inducible mouse mutants. Nucleic Acids Res *31*, e12.

Sergent-Tanguy, S., Michel, D.C., Neveu, I., and Naveilhan, P. (2006). Long-lasting coexpression of nestin and glial fibrillary acidic protein in primary cultures of astroglial cells with a major participation of nestin(+)/GFAP(-) cells in cell proliferation. J Neurosci Res *83*, 1515-1524.

Serrano, M., Hannon, G.J., and Beach, D. (1993). A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. Nature *366*, 704-707.

Serrano, M., Lee, H., Chin, L., Cordon-Cardo, C., Beach, D., and DePinho, R.A. (1996). Role of the INK4a locus in tumor suppression and cell mortality. Cell *85*, 27-37.

Sewalt, R.G., Lachner, M., Vargas, M., Hamer, K.M., den Blaauwen, J.L., Hendrix, T., Melcher, M., Schweizer, D., Jenuwein, T., and Otte, A.P. (2002). Selective interactions between vertebrate polycomb homologs and the SUV39H1 histone lysine methyltransferase suggest that histone H3-K9 methylation contributes to chromosomal targeting of Polycomb group proteins. Mol Cell Biol *22*, 5539-5553.

Shai, R., Shi, T., Kremen, T.J., Horvath, S., Liau, L.M., Cloughesy, T.F., Mischel, P.S., and Nelson, S.F. (2003). Gene expression profiling identifies molecular subtypes of gliomas. Oncogene *22*, 4918-4923.

Shen, X., Liu, Y., Hsu, Y.J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G.C., and Orkin, S.H. (2008). EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. Mol Cell *32*, 491-502.

Sher, F., Rossler, R., Brouwer, N., Balasubramaniyan, V., Boddeke, E., and Copray, S. (2008). Differentiation of neural stem cells into oligodendrocytes: involvement of the polycomb group protein Ezh2. Stem Cells *26*, 2875-2883.

Singh, S.K., Clarke, I.D., Hide, T., and Dirks, P.B. (2004). Cancer stem cells in nervous system tumors. Oncogene *23*, 7267-7273.

Singh, S.K., Clarke, I.D., Terasaki, M., Bonn, V.E., Hawkins, C., Squire, J., and Dirks, P.B. (2003). Identification of a cancer stem cell in human brain tumors. Cancer Res *63*, 5821-5828.

Son, M.J., Woolard, K., Nam, D.H., Lee, J., and Fine, H.A. (2009). SSEA-1 is an enrichment marker for tumor-initiating cells in human glioblastoma. Cell Stem Cell *4*, 440-452.

Sparmann, A., and van Lohuizen, M. (2006). Polycomb silencers control cell fate, development and cancer. Nat Rev Cancer *6*, 846-856.

Sturm, D., Witt, H., Hovestadt, V., Khuong-Quang, D.A., Jones, D.T., Konermann, C., Pfaff, E., Tonjes, M., Sill, M., Bender, S.*, et al.* (2012). Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. Cancer Cell *22*, 425-437.

Su, I.H., Basavaraj, A., Krutchinsky, A.N., Hobert, O., Ullrich, A., Chait, B.T., and Tarakhovsky, A. (2003). Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. Nat Immunol *4*, 124-131.

Sugawa, N., Ekstrand, A.J., James, C.D., and Collins, V.P. (1990). Identical splicing of aberrant epidermal growth factor receptor transcripts from amplified rearranged genes in human glioblastomas. Proc Natl Acad Sci U S A *87*, 8602-8606.

Sugiarto, S., Persson, A.I., Munoz, E.G., Waldhuber, M., Lamagna, C., Andor, N., Hanecker, P., Ayers-Ringler, J., Phillips, J., Siu, J.*, et al.* (2011). Asymmetry-defective oligodendrocyte progenitors are glioma precursors. Cancer Cell *20*, 328-340.

Takizawa, T., Ochiai, W., Nakashima, K., and Taga, T. (2003). Enhanced gene activation by Notch and BMP signaling cross-talk. Nucleic Acids Res *31*, 5723-5731.

Tchoghandjian, A., Baeza, N., Colin, C., Cayre, M., Metellus, P., Beclin, C., Ouafik, L., and Figarella-Branger, D. (2010). A2B5 cells from human glioblastoma have cancer stem cell properties. Brain Pathol *20*, 211-221.

The Cancer Genome Atlas Network (2008). Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature *455*, 1061-1068.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc *7*, 562-578.

Trimarchi, J.M., Fairchild, B., Wen, J., and Lees, J.A. (2001). The E2F6 transcription factor is a component of the mammalian Bmi1-containing polycomb complex. Proc Natl Acad Sci U S A *98*, 1519-1524.

Tso, C.L., Freije, W.A., Day, A., Chen, Z., Merriman, B., Perlina, A., Lee, Y., Dia, E.Q., Yoshimoto, K., Mischel, P.S.*, et al.* (2006). Distinct transcription profiles of primary and secondary glioblastoma subgroups. Cancer Res *66*, 159-167.

Uhrbom, L., Dai, C., Celestino, J.C., Rosenblum, M.K., Fuller, G.N., and Holland, E.C. (2002). Ink4a-Arf loss cooperates with KRas activation in astrocytes and neural progenitors to generate glioblastomas of various morphologies depending on activated Akt. Cancer Res *62*, 5551-5558.

van der Vlag, J., and Otte, A.P. (1999). Transcriptional repression mediated by the human polycomb-group protein EED involves histone deacetylation. Nat Genet *23*, 474-478.

Varambally, S., Dhanasekaran, S.M., Zhou, M., Barrette, T.R., Kumar-Sinha, C., Sanda, M.G., Ghosh, D., Pienta, K.J., Sewalt, R.G., Otte, A.P.*, et al.* (2002). The polycomb group protein EZH2 is involved in progression of prostate cancer. Nature *419*, 624-629.

Verhaak, R.G., Hoadley, K.A., Purdom, E., Wang, V., Qi, Y., Wilkerson, M.D., Miller, C.R., Ding, L., Golub, T., Mesirov, J.P.*, et al.* (2010).Integrated genomic analysis identifies clinically relevant

subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell *17*, 98-110.

Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J.M.*, et al.* (2006). The Polycomb group protein EZH2 directly controls DNA methylation. Nature *439*, 871-874.

Visvader, J.E. (2011). Cells of origin in cancer. Nature *469*, 314-322.

Voigt, P., Leroy, G., Drury, W.J., 3rd, Zee, B.M., Son, J., Beck, D.B., Young, N.L., Garcia, B.A., and Reinberg, D. (2012).Asymmetrically modified nucleosomes. Cell *151*, 181-193.

von Deimling, A., Louis, D.N., von Ammon, K., Petersen, I., Wiestler, O.D., and Seizinger, B.R. (1992). Evidence for a tumor suppressor gene on chromosome 19q associated with human astrocytomas, oligodendrogliomas, and mixed gliomas. Cancer Res *52*, 4277-4279.

Waddington, C.H. (1942). The epigenotype. Int J Epidemiol *41*, 10-13.

Watanabe, K., Tachibana, O., Sata, K., Yonekawa, Y., Kleihues, P., and Ohgaki, H. (1996). Overexpression of the EGF receptor and p53 mutations are mutually exclusive in the evolution of primary and secondary glioblastomas. Brain Pathol *6*, 217-223; discussion 223-214.

Weber, J.D., Jeffers, J.R., Rehg, J.E., Randle, D.H., Lozano, G., Roussel, M.F., Sherr, C.J., and Zambetti, G.P. (2000). p53-independent functions of the p19(ARF) tumor suppressor. Genes Dev *14*, 2358-2365.

Weikert, S., Christoph, F., Kollermann, J., Muller, M., Schrader, M., Miller, K., and Krause, H. (2005). Expression levels of the EZH2 polycomb transcriptional repressor correlate with aggressiveness and invasive potential of bladder carcinomas. Int J Mol Med *16*, 349-353.

Weinmann, A.S., Bartley, S.M., Zhang, T., Zhang, M.Q., and Farnham, P.J. (2001). Use of chromatin immunoprecipitation to clone novel E2F target promoters. Mol Cell Biol *21*, 6820-6832.

Widschwendter, M., Fiegl, H., Egle, D., Mueller-Holzner, E., Spizzo, G., Marth, C., Weisenberger, D.J., Campan, M., Young, J., Jacobs, I.*, et al.* (2007). Epigenetic stem cell signature in cancer. Nat Genet *39*, 157-158.

Wikstrand, C.J., Reist, C.J., Archer, G.E., Zalutsky, M.R., and Bigner, D.D. (1998). The class III variant of the epidermal growth factor receptor (EGFRvIII): characterization and utilization as an immunotherapeutic target. J Neurovirol *4*, 148-158.

Wong, A.J., Bigner, S.H., Bigner, D.D., Kinzler, K.W., Hamilton, S.R., and Vogelstein, B. (1987). Increased expression of the epidermal growth factor receptor gene in malignant gliomas is invariably associated with gene amplification. Proc Natl Acad Sci U S A *84*, 6899-6903.

Wong, A.J., Ruppert, J.M., Bigner, S.H., Grzeschik, C.H., Humphrey, P.A., Bigner, D.S., and Vogelstein, B. (1992). Structural alterations of the epidermal growth factor receptor gene in human gliomas. Proc Natl Acad Sci U S A *89*, 2965-2969.

Woolard, K., and Fine, H.A. (2009). Glioma stem cells: better flat than round. Cell Stem Cell *4*, 466-467.

Yuan, X., Curtin, J., Xiong, Y., Liu, G., Waschsmann-Hogiu, S., Farkas, D.L., Black, K.L., and Yu, J.S. (2004). Isolation of cancer stem cells from adult glioblastoma multiforme. Oncogene *23*, 9392-9400.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W*., et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zheng, H., Ying, H., Yan, H., Kimmelman, A.C., Hiller, D.J., Chen, A.J., Perry, S.R., Tonon, G., Chu, G.C., Ding, Z*., et al.* (2008). p53 and Pten control neural and glioma stem/progenitor cell renewal and differentiation. Nature *455*, 1129-1133.

Zhou, Q., Wang, S., and Anderson, D.J. (2000). Identification of a novel family of oligodendrocyte lineage-specific basic helix-loop-helix transcription factors. Neuron *25*, 331-343.

Zhu, H., Acquaviva, J., Ramachandran, P., Boskovitz, A., Woolfenden, S., Pfannl, R., Bronson, R.T., Chen, J.W., Weissleder, R., Housman, D.E*., et al.* (2009). Oncogenic EGFR signaling cooperates with loss of tumor suppressor gene functions in gliomagenesis. Proc Natl Acad Sci U S A *106*, 2712-2716.

Zhu, X., Bergles, D.E., and Nishiyama, A. (2008). NG2 cells generate both oligodendrocytes and gray matter astrocytes. Development *135*, 145-157.

Zhu, Y., Guignard, F., Zhao, D., Liu, L., Burns, D.K., Mason, R.P., Messing, A., and Parada, L.F. (2005). Early inactivation of p53 tumor suppressor gene cooperating with NF1 loss induces malignant astrocytoma. Cancer Cell *8*, 119-130.

Books

Histological typing of tumors of the central nervous system. World Health Organization - Geneva 1979. (Zulch KJ ed.)

World Health Organization classification of tumours of the central nervous system. Louis, D.N., Ohgaki, H., Wiestler, O.D., Cavenee, W.K. - Lyon 2007. (IARC)

Aknowledgements

First, I would like to thank Dr. Giuseppe Testa for giving me the opportunity to work on this exciting project. His scientific suggestions but also his optimistic and encouraging attitude really helped me, even in the darkest moments of my PhD. With endless patience he listened and answered to all my questions, supporting me in my professional and personal growth.

I would like to thank my external co-supervisor Dr. Steven Pollard for being such a dedicated supervisor. Thanks for discussion and suggestions and for meticulous reading of reports and thesis throughout the last 4 years.

I would like to thank my internal co-supervisor Dr. Gioacchino Natoli for his scientific supervision.

I wish to thank all the members of the GT-group, past and present, that I'm going to mention in alphabetical order: Antonio Adamo, Sina Atashpaz, Giulia Barbagiovanni, Serena Buontempo, Thomas Burgold, Agnieszka Chronowska, Silvia Cristofanon, Giulia Fragola, Pierre-Luc Germain, Pasquale Laise, Pietro Lo Riso, Jacopo Sgualdino and Prem Tripathi. In particular, I want to thank Tommy, Aga, Sere and Giulius; thanks for all the hints, for sharing your experience with me, for coffees and early lunches but mainly for your friendship.

Another big thanks goes to Margherita Turco for the nice time we had in the tissue culture room.

Thanks to the facilities for their technical support, to the girls of the IFOM-IEO soccer team and to all the people I met here at the Campus.

Thanks to my family and to all my friends that followed me during this experience, even being completely outside from the scientific world.

Soprattutto, grazie ancora una volta a Matteo, per l'aiuto, la pazienza e i pomeriggi estivi passati in casa insieme a me durante la produzione di queste "sudate carte". Grazie per essermi stato vicino, per avermi ascoltato e per aver reso piu' facile questo arduo compito.