Chapter One: Thesis Introduction

Cancer

Cancer is a disease in which abnormal cells divide in an uncontrolled way and have the ability to invade surrounding tissues and metastasise to other regions of the body. Diverse cancers have been proposed to display several 'hallmarks', and it has been suggested that as normal cells evolve towards a cancerous state they necessarily acquire these hallmarks in a stepwise fashion [1]. It is thought that a predominant mechanism enabling acquisition of these hallmarks is somatic mutation in the preneoplastic and neoplastic cells [2], although of course there are also important contributions from epigenetic changes, germline genetic variants, and effects from the tumor microenvironment. Certain cell types are more likely to become cancerous than others, and the cell-of-origin can have important consequences for a tumor. These processes are actively being explored and still not fully understood.

Hallmarks of Cancer

A key characteristic of all cancers is sustained cellular proliferation. In order to achieve this, there must be activation of cellular signalling that induces a cell to enter and progress through the cell cycle. A common mechanism to achieve this is through growth factor receptor signalling (with activation of intracellular kinase domains) that turn on proliferative signalling cascades. Several possible ways cancer cells can do this are through: production of excessive growth factors by the cancer cells themselves for autocrine signalling; production of signalling molecules by cancer cells that induce the surrounding stromal tissue to produce growth factors that reciprocally drive cancer cell proliferation [3]; increased growth factor receptor protein on the cancer cell surface; structural changes in the receptor proteins that allow signalling independent of ligand-binding. These many ways of activating cellular signalling often converge on similar pathways, such as the Ras-MAPK and PI3K-AKT pathways that activate transcriptional programs that stimulate entry and progression through the cell cycle. For example, around 40% of melanomas have activating mutations in B-raf giving increased

signalling through the Raf to MAPK pathway, and about 40% of gliomas have activating mutations in EGFR (epidermal growth factor receptor) that turn on the PI3K-AKT pathway.

Normal cells possess several negative feedback loops that act to regulate cell proliferation and suppress excessive cell divisions by diminishing activation of signalling circuitry [4]. An example of this is the PTEN phosphatase which negatively regulates PI3K signalling through degradation of PIP3; many cancers often show loss of function mutations or deletions of the *PTEN* gene leading to increased proliferative PI3K signalling [5]. Typically, early in cancer formation the cells acquire mutations in oncogenes that activate proliferative pathways such as Ras and also in tumor suppressor genes (TSGs) that normally suppress these pathways, giving synergistic effects to maximise proliferative stimuli on the cells. However, many studies suggest that excessive oncogenic signalling in normal cells may trigger processes to counteract this as a protective mechanism, processes including senescence and apoptosis [6]. It is thought that cancer cells have overcome these defence mechanisms by disabling senescence or apoptosis processes. An important example of this is the *TP53* gene, which can activate cell death in response to oncogenic signals to protect against cancer formation; p53 was one of the earliest tumor suppressor genes to be discovered. In a similar fashion, *CDKN2A* is a tumor suppressor gene that activates upon expression of oncogenic signalling to trigger senescence and help prevent cancer formation. CDKN2A functions through the Rb (retinoblastoma-associated) pathway. The Rb protein is itself a critical tumor suppressor, loss of which is found a variety of cancers. Rb is thought to transduce signals external to the cell that are growth inhibitory and acts as a cell-cycle progression gatekeeper; in contrast, TP53 senses damage from within the cell such as to the genome, and if such DNA damage is beyond a critical level then TP53 can trigger apoptosis. Knockout mice lacking *Trp53* have normal development of tissues and organs but develop cancers later on including sarcomas and leukaemias, reflecting increased cellular proliferation due to lack of this important tumor suppressor [7].

Cancers have acquired a special property called immortalisation, whereby they can undergo an unlimited number of cell divisions unlike normal cells which can only a divide a limited number of times before entering a crisis state and undergoing cell death through senescence.

A critical step in immortalisation is overcoming the natural shortening of telomeres that occurs with successive cell divisions. Telomeres are multiple tandem hexanucleotide repeats that protect chromosomal ends, and as they are destroyed cells enter a crisis state. Cancers acquire the ability to stop telomerase erosion through expression of telomerase, a DNA polymerase that can regenerate telomere DNA. This enzyme is usually absent in normal tissue but is expressed in 90% of immortalised cells, highlighting it is an important early requirement for cellular immortalisation and therefore carcinogenesis [8].

Tumors require oxygen and nutrients to survive, even more so than normal tissue because tumors are continuously trying to grow and expand. The ability to form capillaries and other blood vessels is tightly regulated in development. However, cancers have a unique ability to break this regulation and turn on an 'angiogenic switch', whereby the balance between angiogenic stimulators and inhibitors is turned heavily in favour of activation. As such, cancers continuously produce new blood vessels and microvessels to support their growth, although these vessels tend to be abnormal themselves as evidenced by their leakiness and haemorrhage which is so characteristic of cancer. It is thought that angiogenesis is a property acquired early in cancer formation, since these abnormal blood vessels are seen in preneoplastic lesions. Common examples of angiogenic factors are VEGF (vascular endothelial growth factor) and FGF (fibroblast growth factor); VEGF overexpression is often seen in cancer [9].

Immune surveillance programs are an important checkpoint to halt early neoplastic formation – preneoplastic cells express novel antigens on their cell surface that can potentially be recognised by immune cells, which can then potentially kill them before they progress. Cancer formation thus requires the ability to escape detection by the immune system and / or mitigate destruction by it. This hypothesis is supported by experimental work in mice showing tumor incidence is increased when tumors are transplanted into immunocompromised mice compared with immunocompetent ones; in particular, mice with low numbers of natural killer (NK) cells, CD8+ T cells or CD4+ T-cells generate more tumors in this fashion suggesting these cells are important for tumor immunity [10]. Clinical observations also support these ideas given that, for example, patients with colon cancer and high numbers of T cells in their tumor

tend to have a better prognosis most likely because there is some degree of tumor immunity associated with infiltration of these T cells [11]. PD-L1 (programmed death ligand 1) inhibitors have proven successful in targeting interactions between cancer cells and T-cells. PD-L1 on some tumor cells interacts with PD-1 (programmed cell death protein 1) on T cells to suppress T cells from destroying cancer cells. PD-L1 inhibitors have proven successful in activating T cell immune destruction of tumor cells, improving patient survival in certain cases of advanced melanoma, renal cell carcinoma and non-small cell lung cancer [12, 13].

As implied from the descriptions above, a combination of mutations is required for cancer initiation and progression. When a cell acquires a mutation driving cellular proliferation, this mutation becomes selected for as the cell population carrying it expands as part of a Darwinian natural selection process. A subpopulation of these cells may then acquire an additional mutation in an oncogene or tumor suppressor gene that gives an extra growth advantage, generating further clonal selection for cells carrying this combination of mutations. This process of cancer gene mutation and clonal selection can occur repeatedly, and this is the basis of the multistep tumor progression model. Analysis of cancer genomes reveals gains and losses of many chromosomal regions; this was noticed even before modern next generation sequencing (NGS) technologies were available, for example with the use of comparative genome hybridisation (CGH). Now, with NGS it is even more apparent that cancer genomes have widespread mutations and copy number changes. These findings all imply that an elevated mutational rate is important in carcinogenesis, since this will allow preneoplastic cells to gain the required mutations in necessary cancer genes [14, 15]. Defects in cellular DNA repair machinery and / or in the mechanisms that normally protect the cell from the damaging effects of external DNA mutagens can lead to the undesirable effect of increased mutational rate, predisposing to cancer formation [16]. Once again, TP53 is an example of such a molecule, which can sense DNA damage and initiate repair mechanisms as a ‘guardian of the genome’. Similarly, caretaker genes that participate in DNA mismatch repair, for instance *MSH2*, are often tumor suppressor genes.

Over the last few decades, it has been increasingly recognised that there is a great deal of cellular intratumor heterogeneity, with different cell subpopulations in a tumor having

different phenotypes and genetics. Moreover, these different cell populations can have varying degrees of tumorigenicity – not all cancer cells are born equal. A prominent hypothesis that has emerged from these observations is the cancer stem cell (CSC) hypothesis which states that cancers consist of a hierarchy of cells with varying degrees of differentiation, with cancer stem cells sitting at the top of this hierarchy and able to generate all of the more differentiated progeny [17]. These cancer stem cells often express the same markers of stemness that are used to distinguish their normal stem cell counterparts, and their transcriptomic profiles often overlap with normal stem cells from the same tissue of origin. Importantly, cancer stem cells have the ability to efficiently initiate tumor formation in mice, in contrast to more differentiated cancer cells which are much less efficient at doing so. These findings have prompted many to speculate that cancers themselves originate from normal stem cells that acquire mutations in oncogenes or TSGs that endow these cells with proliferative abilities and thus cancer stem cell properties. However, it is still unclear whether CSCs originate from aberrations in normal stem cells or in progenitor cells, or more radically yet through mutations in differentiated cells that then acquire stem cell-like characteristics. Indeed, it has been recently shown that even in fully formed cancers there is plasticity in phenotype of differentiated and cancer stem cells, with some differentiated cancer cells having the ability to form CSCs in the right conditions[18]. Given the clinical findings that large numbers of CSCs are associated with shorter time to recurrence, poorer prognosis and resistance to chemo- and radiotherapy [19], it will be essential to improve our understanding of these cells in order to design better targeted therapeutics.

Sequencing Human Tumors

With the advent of next-generation sequencing technologies and the continually diminishing costs of sequencing genomes, large-scale studies involving sequencing of hundreds of tumor genomes or exomes have become increasingly attractive and feasible. The Cancer Genome Atlas (TCGA) has published several major studies sequencing several hundred human low-grade gliomas and glioblastomas, which have revealed the mutational landscapes of these

tumors. These are primarily observational studies looking for associations between certain types of cancer and specific mutations. With increasingly advanced statistical methods however, the ability of such studies to differentiate true cancer 'driver' mutations (those that confer a selective growth advantage on a tumor) from the background 'passenger' ones (mutations that do not confer a growth advantage) is continually improving. In order to prove that these putative drivers are indeed real drivers and to understand the underlying biology, it is necessary to functionally validate them in model systems, for example by inducing the same mutations in mice to determine if they generate the intended cancers. This move to model organisms from human sequencing data will also more readily enable investigation of the molecular mechanisms underpinning how genetic mutations drive cancer forward.

Gliomas

There is a diverse spectrum of brain tumors, reflecting the presence of different tissues in the central nervous system (CNS). With the exception of metastases from primary sites other than the brain, gliomas and meningiomas are the commonest types of brain tumor. Meningiomas are typically benign tumors that arise from the meningeal coverings of the brain. Other types of tumor include medulloblastomas (which are commoner in children), ependymomas, pilocytic astrocytomas and primitive neuroectodermal tumors.

Low grade gliomas (LGGs, World Health Organisation, grade II) are a heterogeneous population of intrinsic brain tumors whose natural history is to evolve to higher grade tumors. These tumors histologically contain cells with similar appearances to glial cells (neuronal support cells), including astrocytes and oligodendrocytes. LGGs constitute 15% of all adult brain tumors, and they most commonly present with seizures (in 80% of cases) [20]. A model for the natural history of gliomas posits four phases: 1) the occult stage, in which tumor-initiating cells proliferate but there is no detectable tumor on MRI; 2) the clinically silent stage, in which tumor mass becomes apparent on MRI but the patient does not have any symptoms (incidental glioma)[21]; 3) the symptomatic stage, in which the tumor elicits symptoms such as seizures or weakness; 4) malignant transformation, in which the low grade glioma switches to a more biologically aggressive high grade glioma [22, 23]. Upon malignant transformation, the tumor is termed a secondary glioblastoma, Fig 1.

Glioblastoma multiforme (GBM), or high-grade glioma, is the most common intrinsic brain tumor, and characteristically invades surrounding brain aggressively, making complete surgical resection unachievable. It tends to affect middle-aged to elderly people, and can either arise de novo (primary glioblastoma) or by transformation from an LGG (secondary glioblastoma). Its prognosis is therefore poor, with a median survival of only 14 months despite maximal therapy with surgery and chemo- or radiotherapy [24]. These survival times have not substantially changed over the last few decades despite improvements in treatment, and clearly further understanding of the biology of this cancer is needed before significant advances in treating it are made. The current chemotherapeutic standard of care for GBM is

temozolomide, a DNA alkylating agent that improves prognosis by around 2 months when given with radiotherapy [24, 25]. Given the limited impact this has on survival, there is a strong need for more molecularly targeted treatments for this cancer that will improve prognosis further. A recent example of such a therapy that has entered clinical practice for recurrent GBM is bevacizumab (Avastin): a monoclonal antibody targeting vascular endothelial growth factor (VEGF) and thus aims to reduce angiogenesis within tumors. Although bevacizumab slows down glioma growth, it does not improve overall survival in GBM; there is a benefit on progression-free survival of around 2 – 3 months in a randomised controlled study on patients with newly diagnosed GBM [26]. Nevertheless, all patients succumb to the disease after developing resistance to treatment. Surgery, although beneficial, is not curative because glioma cells tend to invade well beyond the visible margins of resection, and such cells (possibly glioma stem cells) trigger disease recurrence. A number of genes are implicated in resistance to temozolomide, most are involved in DNA repair, for example *MGMT*, *MSH1* and *MSH2*. However, much is still to be learnt in this field because tumors even in those patients without such resistance mutations become resistant to chemotherapy, suggesting other unknown mechanisms are involved.

Recent genome-wide sequencing studies of GBMs have provided insight into common genetic drivers of this tumor and have highlighted genetic differences between primary and secondary GBMs. Primary GBMs usually have one or more mutations in three main molecular pathways: Ras/RTK pathway, p53 pathway, and the Rb pathway [27, 28]. Within the Ras/RTK pathway, *EGFR* (30-50% of tumors) and *PTEN* (30%) are the most commonly mutated in GBM, although mutations in *NF1* and *RAS* have also been documented. Mutations within this pathway tend to enhance cellular proliferation. Of the p53 pathway, *TP53* (25%) itself is most commonly mutated in GBM. The *TP53* gene is normally activated following DNA damage to cells, inducing transcription of genes whose ultimate effects include apoptosis. Mutations in *TP53* are thought to have effects such as inhibition of apoptosis, stimulation of cell proliferation and neovascularisation, which are hallmarks of cancer [1]. Although a mutation in the Rb pathway is present in most GBMs, the *RB* gene itself is infrequently mutated and instead mutations in *CDKN2A* (50% of tumors) are particularly common. *CDKN2A* is the locus

for two tumor suppressor genes – *INK4A* and *P19-ARF*. *In vitro* and *in vivo* models have validated a number of such mutations as driving tumor growth and invasion. Mouse models have been particularly helpful in demonstrating how mutations in multiple pathways can co-operate together to accelerate tumorigenesis [29].

The *IDH1* (isocitrate dehydrogenase 1) mutation is characteristically found more commonly in secondary GBMs and also in LGGs [30], and although the mechanism by which this mutation contributes to carcinogenesis is still unclear it is thought to act epigenetically through abnormal methylation of DNA, Fig 1.1. Although *IDH1* is mutated in the majority of LGGs and secondary GBMs, it is still not clear how and at what stage this mutation may contribute to carcinogenesis: *IDH1* mutations (of which the R132H mutation is the most frequent) predict a better prognosis, and the mutation in conditional transgenic mice has not been found to accelerate established gliomas although there is some evidence suggesting it initiates gliomagenesis and also induces acute myeloid leukaemia in mice [31-33]. The recent pathological classification of gliomas has been changed to take into account key genetic changes, including not only *IDH1* which predicts a better prognosis, but also the presence or absence of 1p/19q co-deletions and *TERT* promoter mutations. Gliomas can thus be classified into five types depending on the presence or absence of these three genetic alterations, with *TERT* promoter mutations alone signifying the worst prognosis [34].

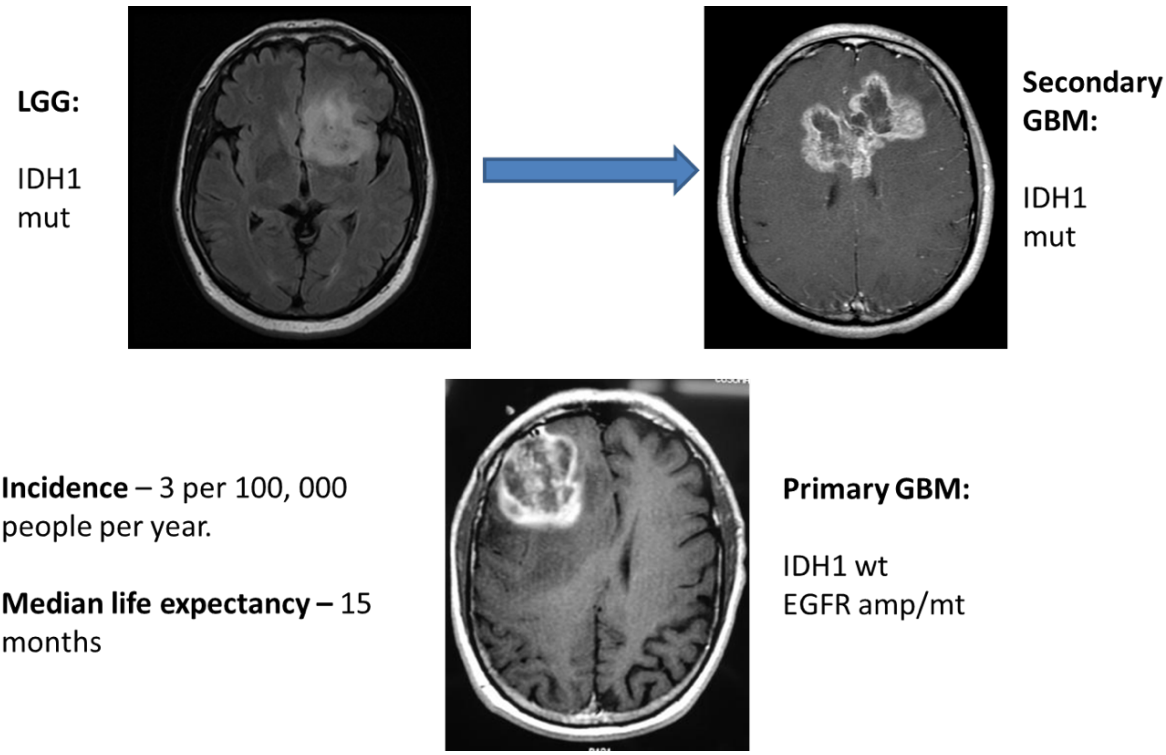

Figure 1.1. Glioma is classified into low-grade glioma (LGG) and high-grade glioma or glioblastoma (GBM). These subtypes have different clinical properties, genetics and prognoses. Only the key mutations described from the literature are illustrated here, although the reality is that these tumors have complex genomes.

Spinal Gliomas

Spinal tumors that arise from the spinal cord itself, so called intramedullary spinal cord tumors (IMSTs), are rare tumors – they make up 2% of all tumors in the CNS [35]. Of these IMSTs, spinal ependymomas are the commonest in adults, whereas spinal astrocytomas are the commonest in children and adolescents. Although these are rare tumors, spinal gliomas are a significant clinical problem – it is often difficult or not possible to achieve complete surgical resection of the tumor, and adjuvant chemotherapy and radiotherapy often only have minimal benefit but significant adverse effects; therefore, novel treatments based on improved understanding of the molecular biology and genetics of these tumors are in great need.

Spinal astrocytomas account for 60% of spinal tumors in children and adolescents. They often present with progressive back pain, often waking patients at night due to severity; due to their proximity to the spinal cord, expansion of these tumors gives symptoms of numbness or tingling as well as motor weakness, which can be misconstrued as clumsiness in children [36]. Most spinal astrocytomas are low grade (grade I or II), although around 20% are high grade (III or IV) and these have a poor prognosis associated with increased invasiveness of the tumor: the mean survival is 15.5 months [37]. It is typically difficult to achieve complete surgical resection of spinal astrocytomas because these tumors have cells that infiltrate into the spinal cord, and the surgeon is unable to resect beyond the margin of the spinal cord for concerns over inducing very disabling paralysis. Nevertheless, leaving behind groups of tumor cells in the spinal cord is likely to lead to recurrence of the tumor, which is often the case. Following recurrence, the treatment options are limited – radiotherapy is controversial for these tumors in children as it can cause radiation necrosis, radiation myelopathy and reduced spinal growth [38]; chemotherapy is therefore favoured but itself is of limited benefit. Given that temozolomide has been showed to provide a limited survival benefit for brain gliomas in adult patients, it can be used as a chemotherapeutic agent in spinal gliomas but has associated systemic toxicities such as neutropenia and lymphopenia [39].

The genetic basis of spinal astrocytomas is poorly understood. The main reason behind this, apart from the rarity of the tumor, is the lack of fresh tissue available for performing next generation sequencing. As explained, it is challenging to achieve complete resection of the tumor due to its dangerous location, so the amount of tumor resected is limited and used for pathological purposes only. Further efforts are needed to prospectively collect suitable material for deep sequencing to yield some insight into the biology of spinal astrocytomas; appropriate mouse models are also required to complement observations in genetic sequencing of these tumors in patients with functional validation and molecular characterisation.

In vitro Models of Cancer

An alternative model for studying cancer is *in vitro* cancer cell lines, which are available for human glioblastomas. Such cell lines have advantages in that they are generally derived from patient tumors, they proliferate well *in vitro* and can therefore be expanded relatively quickly, they are amenable to genetic manipulation such as knockout of individual genes, and the efficacy of drugs as potential therapeutic agents can be tested on these lines. However, there are also several disadvantages for using cell lines as models of human cancer in comparison to using mice for *in vivo* study. It is known that human GBM cell lines acquire multiple additional genetic lesions in culture that were not originally present in the parental tumor, including further chromosomal copies and translocations in addition to mutations. In this way, the genetics of these cell lines do not necessarily reflect that of the patient tumor, which can lead to false conclusions drawn from genetic experiments performed in such cell lines alone. One method to help overcome this problem is to culture cells derived from patient GBMs in neural stem cell conditions either as 'gliomaspheres' or as adherent cell lines, which has been shown to help maintain the original tumor genetics [40], at least in short-term culture.

Another limitation of cancer cell lines is that they do not model more complex features of carcinogenesis such as the tumor microenvironment. As discussed previously, rather than simply being collections of cancer cells, malignant tumors contain complex interactions between cancer cells and other cells such as macrophages and T cells, and these interactions are very difficult to model accurately *in vitro*. In comparison, the use of transgenic mice that develop cancer has the potential to model such interactions between the tumor and its microenvironment. Other features of cancer, including invasion, angiogenesis and metastasis, are also more readily demonstrable *in vivo* than *in vitro*. Furthermore, cancer cell lines are derived from fully formed tumors which are typically late in cancer evolution; this precludes study of the processes initiating tumor development that may differ greatly from those driving progression. Mouse models on the other hand have the potential for investigation of both early and late evolutionary forces in cancer.

The Mouse as a Model Organism

There are several reasons why the mouse is a useful model organism for studying human diseases, particularly with genetic approaches. Firstly, being a mammal the mouse shares many similarities with humans in its organ systems, allowing a diverse range of diseases affecting different organs to be studied. In particular, the anatomy of its central nervous system is not too dissimilar from humans, containing major features of the brain and spinal cord that are found in man, making the investigation of certain neurological disorders possible in mice. Mice have relatively short breeding times, with pregnancy only lasting around two weeks, and typical litters contain 5 – 10 mice, which allow for large numbers of mice to be generated within a short period of time. Given their small size, it is also practical to house even up to thousands of mice in a single room. These latter features are particularly important for testing whether experimental results gained from a few mice can be replicated in larger cohorts and of course for performing large-scale *in vivo* genetic screens.

In order to carry out genetic studies in a particular organism, its genome must be mapped and well-understood. The mouse genome was fully sequenced and published not long after that of the human[41]. This revealed a great deal of homology in gene sequences between mouse and human genomes, a crucial feature for enabling investigation of human genetic diseases in mice.

Reverse Genetics in Mice

Reverse genetics involves inserting a known mutation into mice (or any other model organism) and observing the induced phenotype. In the context of cancer genes, this entails selecting a gene that has been associated with a specific cancer, for example from a cancer sequencing study, and then determining whether it can produce the intended cancer phenotype in mice. Of course, this ideally (but not necessarily) requires some hypothesis about the gene itself. Examples of this technique are that introducing *Tp53* or *Pten* mutations in mice lead to generation of cancers, as will be discussed in depth later [42, 43].

Forward Genetics in Mice

Forward genetic screening aims to identify gene mutations that give rise to a specific phenotype, such as cancer. In this approach, many different mutants are generated (that are not previously known to give rise to the phenotype of interest) and then these are observed for development of the phenotype. The mutants that generate such a phenotype are then further evaluated to identify the underlying mutations. This can be a powerful approach for cancer gene discovery, for example, and can be effectively applied *in vivo* in mice. In the case of cancer, there is often a multitude of genetic and epigenetic aberrations and intratumoral heterogeneity, complicating efforts for identifying the true drivers. This problem is particularly true for human GBMs. Forward genetic screening in mice can provide a complementary approach to sequencing human tumors for the identification of these true cancer driver genes, which may be mutated or epigenetically altered in human tumors.

Inbred Mice

Genetic diversity is introduced during gamete formation in mice by recombination of chromosomes in meiosis, similar to most other sexually reproducing organisms. This can lead to confounding effects in experiments performed with different mouse strains. Therefore, pure-bred mouse strains that are homozygous at all genomic loci have been invaluable for genetic experiments. These are generated by many generations of sibling matings: 98.7% of the genome is homozygous after 20 generations of these matings [44]. In order to study a specific mutation from a different genetic background, a mouse carrying the mutation of interest can be crossed with an inbred mouse, forming a congenic strain with the mutation of interest on an invariant genetic background that mitigates confounding effects of comparing results gathered from different strains of mice. This also helps ensure that the introduced mutations are the direct cause of any phenotype observed, aiding in establishing putative cancer genes as true oncogenes or tumor suppressor genes.

Different forward genetic approaches

There are several different ways for conducting forward genetic screens in mice *in vivo*, and the most important of these will be briefly discussed here. Mutagenesis can either be conducted in the germline or somatically in tissues of interest.

Irradiation

Irradiation with gamma rays introduces double strand breaks into DNA, and these can be inaccurately repaired leading to chromosomal imbalances such as deletions that effectively knockout a gene. This makes irradiation potentially useful for forward genetic screening. However, deletions caused by irradiation can be large enough to span many genes, making identification of the gene causing the phenotype more difficult and requiring detailed sequencing approaches such as whole-exome or whole-genome sequencing, and extensive validation of identified targets to confirm their true nature as genes causing the phenotype.

ENU Mutagenesis

ENU (N-ethyl-N-nitrosourea) is a very potent mutagen, acting as a DNA alkylating agent. It can induce point mutations in genes, and can therefore be used in genome wide screens by sequencing for point mutations in tissue with a phenotype of interest, such as cancer. However, given the high mutagenic rate of ENU, it can be more difficult to identify the underlying driver genes from this approach than with others such as transposons, requiring deep sequencing and more complex bioinformatics analyses.

Insertional Mutagenesis

Insertional mutagens are mobile DNA elements that insert into the genome to act as a mutagen. These have the advantage that they effectively tag the mutated locus with their

known sequence, and linker-based PCR methods can amplify the adjacent genomic regions. Insertional mutagenesis within an exon (or even just within the gene) is likely to lead to disruption of the function of the gene.

Viral Mutagenesis

Viruses can be used as cellular mutagens for cancer screening in mice. Insertion of a proviral retrovirus into the mouse genome can lead to increased expression or disruption of a gene depending on the location of the insertion within a gene. The proviral genome contains long terminal repeats (LTRs) to control transcription not only of the retrovirus but also of the host cell genes if they are downstream of these viral LTRs. Sequencing of resulting tumors formed from insertion of the provirus into cancer genes can help identify these underlying genes by determining common insertion sites of the virus [45]. An example of such a transforming virus that can be used in screens is the murine leukaemia virus (MLV). A disadvantage of this approach is that viral mutagens have insertional biases that prevent this being a truly unbiased genome-wide screening approach, and the more time-consuming and challenging nature of insertion site cloning and mapping for viruses makes this approach less favourable than others such as transposons. Retroviral screens have led to early discoveries of genes such as Myc and some of its cooperative partners in cancer [46, 47].

Sleeping Beauty Transposons

Sleeping beauty (SB) transposons are from the Tc1-Mariner family and are naturally found in salmonid fish. This transposon was reconstructed from phylogenetic data and the investigators thus gave the name '*Sleeping beauty*' to reflect that fact that it was 'awakened from a long evolutionary sleep.' Ivics et al discovered the sequence of the ancestral transposon from this class of fish, demonstrated to be two 250bp terminal DNA sequences containing inverted repeats that flank an open reading frame that codes for the transposase enzyme [48]. The entire sequence is 1.6kbp. *Sleeping beauty* is therefore a two-component

system composed of the transposon vector and the transposase enzyme. When these are present in the same cell, the transposase recognises the inverted repeats of the transposon and excises it from the donor locus. The transposon can then insert itself at a TA dinucleotide region elsewhere in the genome. In this way, the transposase catalyses a 'cut and paste' reaction of the transposon. Modifications to the inverted repeats were made early on in order to improve transposition efficiency [49]. Site-directed mutagenesis of the SB transposase produced alternative versions of the enzyme with different transposition efficiencies [50, 51]. The first SB transposase was SB10, and modified versions were numerically labelled (SB11 and so on).

A consequence of SB transposition is the creation of a 'footprint' mutation. The reason for this is that when SB inserts into genome it duplicates its target site (a TA dinucleotide) and excision generates 3-nucleotide overhangs. A potential advantage of this footprint mutation is that insertions of likely biological relevance can still be identified through genetic sequencing even if the transposon has subsequently mobilised into another site. On the other hand, a disadvantage of a footprint mutation is that some gene disruption may remain in the transposon insertion site, which may not be of relevance for the phenotype of interest. Another feature of the SB transposon is its tendency for so-called 'local hopping' wherein SB preferentially inserts into DNA near to the original site of the transposon, typically within 2 – 10 Mbp on the same chromosome. This bias in its insertion sites must be considered when analysing the data from a SB forward genetic screen, in practice meaning that investigators usually exclude all common insertion sites in the transposon donor locus.

The first sleeping beauty screens in mice for cancer used a constitutively active SB transposase and a transposon line (T2/Onc or T2/Onc2), in which the transposon was mobilising in all tissues [52]. This transposon can induce a gain-of-function or a loss-of-function of a gene, depending on where in the gene it inserts. T2/Onc contains a murine stem cell virus (MSCV) long terminal repeat (LTR) promoter with artificial exon with splice donor (SD). The LTR and SD can lead to fusion transcripts through splicing, and if T2/Onc has inserted upstream in a gene these fusion transcripts will be overexpressed because of the T2/Onc promoter. This is equivalent to activation of proto-oncogenes. The transposon also contains splice acceptors

and a bidirectional polyadenylation signal to cause termination of transcripts that arise when the transposon inserts into an intron of a gene. This is important for inactivating tumor suppressor genes and thus promoting tumorigenesis.

Several SB insertional mutagenesis screens have been conducted in mice, successfully contributing to driver gene discovery for many different cancers [53-63]. One constitutive SB screen in mice generated gliomas in the brain, although the incidence was low [64, 65]. To increase the incidence of these tumors, the authors crossed the mice with a *P19Arf* allele which is a known tumor suppressor gene in gliomas. Although this did indeed increase the incidence of brain tumors, the difference was small. A mixture of anaplastic astrocytomas (grade III) and glioblastomas (grade IV) were produced. Sequencing of the resulting 21 gliomas from this screen yielded 887 common integration sites (CIS), and identified *Csf1* as a recurrently hit gene. Immunohistochemistry in human tumors demonstrated overexpression of CSF1 in high-grade astrocytomas, providing some support for a role of this gene in supporting malignant glioma formation in humans as well as mice. Importantly however, the CIS genes did not include a number of well-established glioma genes such as *Egfr*, *Pdgfr* and *Tp53*, and there were only single insertions found in other important tumor-specific genes including *Pten* and *Akt*. This may be explained by certain insertion site preference biases of SB or that SB-induced gliomas represent only a subset of gliomas that is not driven by the major cancer drivers in most human gliomas. It is therefore crucial that transposon-driven gliomas are studied in other genetically-predisposed backgrounds in order to gain a more complete understanding of the cancer drivers, ideally with an alternative transposon system such as *piggyBac* as well. A systematic comparison of the insertion sites from their study and human glioma sequencing data was not performed.

To enable screens to be performed for tissue-specific cancers, a conditional SB transposase allele was also developed that is active in the presence of cre expression and has been used for screens of many cancer types in mice. One study which employed this conditional SB transposon system for investigating gliomas used a nestin-cre allele on a *Trp53*-mutant background to drive expression of the SB transposase in mouse neural stem cells, although these did not directly generate tumors *in vivo*. *In vitro* culturing of embryonic neural stem

cells derived from the subventricular zone of these mice demonstrated that these cells can be immortalised by mobilisation of SB: immortalisation occurred significantly more frequently in cell lines with both Trp*53*^{R172H} and mobilising SB than those with *Trp53*^{R172H} alone, and not at all in lines with neither *Trp53*^{R172H} nor mobilising SB. When these immortalised cells were subcutaneously transplanted into SCID mice, they generated tumors with a latency of two – four months. The authors identified 106 CIS genes in the immortalised cell lines and 114 in the tumors, of which 34 CIS genes were present in both cohorts. Comparing the CIS from the immortalised cells with those of the tumors in mice showed that a further round of transposon mobilisation for *in vivo* tumor establishment was needed in addition to the insertions present in immortalised cells alone [66]. The authors therefore were able to categorise SB insertions according to whether they drove cellular proliferation *in vitro* or *in vivo* tumor growth or both, as part of a two-step process of cancer evolution. Amongst the CIS in the immortalised cell lines were a few known glioma genes, including *Pten*, and similarly amongst the tumor CIS were genes such as *Pdgfrb* and *Nf1*. The study also identified cancer genes that were not previously linked with gliomas, such as *Met* and *Klf3* which were amongst their top-ranking tumor CIS. The CIS genes clustered into biological pathways that are thought to underlie gliomagenesis, in particular the Ras-MAPK and PI3K-Akt pathways, confirming that these major pathways that are mutated in human tumors can promote neural stem cell immortalisation *in vitro* and subcutaneous glioma formation. A systematic comparison with genetic and epigenetic alterations in human gliomas and glioma cell lines was not performed however, and neither was there an analysis for network interactions between the CIS genes.

More recently, work from the Bradley group has demonstrated the usefulness of a single-copy sleeping beauty transposon for cancer forward genetic screens in mice. The study employed a transposon linked with the *Pten* gene, such that mobilisation of this single copy transposon also led to heterozygous loss of *Pten* (an important cancer driver). The resulting developed tumors in multiple organs, particularly in the prostate. The advantage of this approach is that a single cell will simultaneously have transposition and *Pten* loss, helpful in establishing cooperativity between *Pten* and transposon insertion genes in tumorigenesis. Additionally, the study found that the common insertion sites from this screen gave the

strongest candidate driver genes from an otherwise longer list of genes produced from a multiple-copy transposon screen [67].

PiggyBac Transposon

The *piggyBac* (PB) transposon is naturally found in the cabbage-looper moth. *PiggyBac* has some key advantages over SB as a transposon: it has a larger cargo capacity than SB, and also can mobilise within the genome without leaving a footprint unlike SB. Given these advantages, efforts were made to adapt PB for mammalian systems [68], and specifically for forward genetic screens of cancer both *in vitro* and *in vivo*. There are several mouse PB transgenic lines, including both high-copy and low-copy PB number. The low-copy number PB transposon mice were demonstrated to be better for modelling solid cancers, since high copy mutagenesis has a tendency for embryonic lethality.

A constitutive PB screen in mice generated a variety of solid and haematological cancers [69]. In this work, there were a number of different transposon lines: it was demonstrated that lines with a transposon driven by the CAG promoter (ATP1) were more efficient at generating solid tumors, likely because of superior transposon mobilisation in the underlying tissues. Conversely, lines with the transposon under control of the MSCV promoter (ATP2) were better at producing leukaemias and lymphomas. With these models, the authors were able to identify known and novel cancer genes from the sequenced transposon common insertion sites. To enable screening in specific tissues, a conditional *piggyBac* transposase mouse line was engineered; a screen on pancreatic cancers employed this conditional *piggyBac* system, identifying new cancer genes such as *Foxp2* in the specific tissue of interest [70]. This approach for cancer screening is discussed in further detail in the relevant chapter of this thesis.

CRISPR / cas9 screening

Technologies for genetic manipulation of mammalian genomes based on engineered nucleases have evolved rapidly over the last few years. Zinc finger nucleases (ZFNs) were the first to be widely used in mammalian cells [71]. Their use established the basic paradigm for using double strand DNA breaks in the genome to stimulate error-prone or template targeted repair mechanisms to effect a genetic change. The use of ZFNs for this purpose quickly declined with the development of transcription-activator like effector nucleases (TALENs) primarily because the reliability of the nucleotide recognition encoded by the TAL repeats supported reliable synthesis of highly specific molecules [72]. The use of TALENs for mammalian genome engineering was also relatively short lived, as these were rapidly superseded by Clustered regularly interspaced short palindromic repeats-Cas9 (CRISPR-Cas9) because of the simplicity of deploying the cas9 nuclease to any nucleotide in highly complex genomes.

CRISPRs were identified in *E.coli* in 1987 and in other bacteria and archaea a decade later [73, 74]. The phage origin of these repeats and the identification of genes with putative nucleases associated with these repeats (CRISPR-associated) *cas*-genes led to the hypothesis and subsequent demonstration that the CRISPR-cas system had a role in microbial adaptive immunity [75]. This is achieved by directing the cas-nuclease to the incoming phage DNA by a guide RNA transcribed from the clustered repeats [76].

In contrast to the ZFN and TALEN systems in which specificity is achieved by complex protein-nucleic acid interactions, the cas-nuclease is directed to a genetic target by nucleic acid base pairing determined by a unique 20 nucleotide region of short guide RNA (sgRNA). Experimentally this sgRNA sequence can be adjusted to guide the nuclease to virtually any site in a complex genome [77]. The efficiency of the cas-nuclease coupled with the simplicity with which it can be directed has resulted in its rapid adoption. The CRISPR-cas9 system has been shown to be effective for manipulating genes in a variety of cell types from different organisms. When used as a nuclease, cleaved DNA is re-joined by an error-prone end-joining process resulting in small insertions and deletions (‘indels’) at the target site and concomitant

loss of the gene's function. Larger genetic alterations such as deletions and inversions can also be generated. In other applications the break generated by the nuclease will catalyse a process of homology directed repair if a suitable vector is also provided resulting in replacement of one sequence (for instance a defective copy) with a normal one provided by the vector, so called gene-editing, Fig 1.2. Studies have provided a cautionary note of potential off-target effects (unintended modifications at other sites in the genome) with this platform [78].

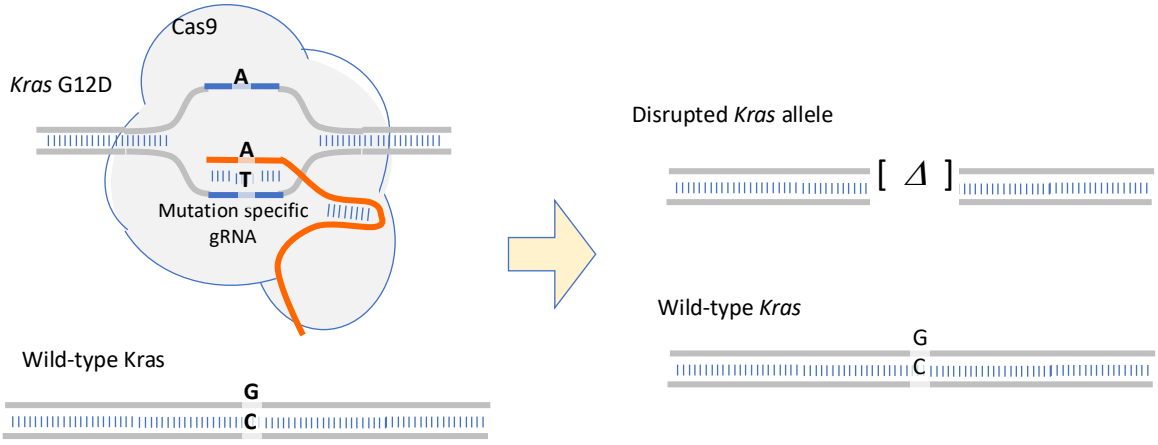

Figure 1.2. CRISPR/cas9 mediated introduction of DNA indels. Example is given of the Kras G12D oncogene which can be targeted with an sgRNA that leaved the wild-type allele untouched.

Over the last few years, a number of studies have demonstrated the usefulness of CRISPR-cas9 in both positive and negative selection genome-wide screens, including in human cancer cells [79] [80-83]. Such screens employ large lentiviral libraries with multiple sgRNAs per gene, and consequently require large starting populations of cells.

Glioma Mouse Models

The majority of glioma mouse models have employed cre / LoxP technology for specifically targeting cancer genes in certain neural tissues of interest. I will therefore describe this technology before discussing examples of glioma mouse models in more detail.

Cre / LoxP, Flp/FRT, RCAS Technology

Site-specific recombination allows for the generation of genetic alterations such as deletions, point mutations, duplications and inversions. The flippase / flippase recognition target (Flp/FRT) system was the first one to achieve site-specific recombination in multicellular organisms, and this was originally performed in Drosophila [84]; in this system, the flp recombinase mediates recombination between FRT sites in the genome. In the mouse however, the commonest method for recombination is the use of the cre/LoxP system in which the cre (cyclization recombinase) mediates recombination between two LoxP sites. The LoxP sites are 34-base pair consensus sequences, each with a central 8-bp core spacer sequence that determines the orientation of the LoxP site, and two inverted 13-bp flanking sequences that bind cre. The cre/LoxP system was first implemented in mice in the early 1990s, and since then has been widely used for generating conditional genetic alterations *in vivo* in a variety of specific tissues, including the brain. Indeed, many cre transgenic mouse lines have been created in the last few decades to allow study of organ or tissue-specific physiology and pathology [85].

An alternative system for introducing targeted mutations to cells of interest is the RCAS (replication-competent ASLV-long terminal repeat with a splice acceptor) vector system. These vectors derive from the Rous sarcoma virus, which belongs to the avian sarcoma – leukosis virus (ASLV) family, and they contain the src (oncogene) splice site and express an inserted gene (such as an oncogene) via a spliced message. This system is limited by the small size of the insert (2.5 kb) and the low number of cells that are typically infected and express the gene of interest [86].

Nestin is an intermediate filament protein expressed in neural stem cells and neural progenitors. Mice with the nestin-cre allele express cre from embryonic day 13, at which stage the embryonic neural progenitors are able to undergo differentiation into many cell types including astrocytes, neurons and oligodendrocytes. Therefore, in postnatal and adult mice containing nestin-cre, cre is expressed throughout most of the central nervous system, eye and also the kidneys – this was demonstrated by Dubois and colleagues who showed virtually complete cre-mediated recombination in these tissues by embryonic day 15.5 using LacZ based reporters [87, 88]. The allele is however insufficient for driving recombination in early embryonic ventricular zone neural progenitors and neural stem cells (before embryonic day 17.5), as determined using multiple cre-dependent reporters. [89]. An alternative cre line that is frequently used in glioma mouse models is hGFAP-cre, which is also expressed from pre-natal stages and in the majority of cell types in the brain and spinal cord [90]. For studies where the timing of recombination needs to be controlled, it is possible to do so through tamoxifen injections using nestin-creERT2 and GFAP-creERT2 mouse lines, which are inducible cre lines. These are useful for studying tumor origins from adult brain cells as opposed to embryonic cells, for example; the cre expression onset is controlled by specifying age of the mice at which tamoxifen is given. In order to induce recombination in more specific groups of brain cells, alternative cre lines can be used. For example, Olig2-cre allows site-specific recombination in oligodendrocyte precursor cells and oligodendrocytes, Syn-cre gives recombination specifically in neurons, and Glast-cre is a newer alternative line for recombination in neural stem cells in the SVZ. However, an important drawback of all of these cre lines is the specificity of the regions and cell types in which recombination occurs, in that there is typically recombination in other cells than those of interest.

Key Glioma Mouse Models

One of the earliest oncogenes to be discovered in gliomas is the epidermal growth factor receptor (*EGFR*) gene [91], which is mutated and/or amplified in 50 -60% of primary glioblastomas. EGFR is a cell-surface receptor that binds epidermal growth factor as its ligand

and then signals via intracellular cascades, including the Ras-MAPK and PI3K-Akt pathways. In primary glioblastomas, the variant III mutation of *EGFR* is particularly common, and involves deletion of exons 2 – 7 of the gene (the extracellular ligand binding domain) leading to constitutive signalling the resulting receptor. An early study aimed at determining whether excessive EGFR signalling can induce gliomas *in vivo* employed the RCAS vector system to introduce an *Egfr* activating mutation (the *EgfrvIII* deletion and another deletion that removes the intracellular regulatory kinase domain) in mice expressing the avian tumor virus receptor A (TVA) under brain cell specific promoters. The vector was introduced into the frontal lobes and hippocampus. After 15 weeks, none of the mice developed gliomas. In contrast, when an *Egfr* activating mutation was introduced in the presence of *Cdkn2a* loss, gliomas arose at a high frequency particularly on the nestin-TVA (Ntv) background. The authors concluded that *Egfr* activating mutations alone are insufficient to generate gliomas, but can cooperate with predisposing mutations such as those of *Cdkn2a* to produce these tumors [92]. Given the incidence of tumors was higher in Ntv compared with glial-specific GFAP-TVA (Gtv) mice, they suggested that the presence of these mutations in a neural stem cell lineage is a likely origin for gliomas.

Another early glioma mouse model that used the RCAS vector system was that by Holland et al in 2000 [29]. In this study, a $Kras^{G12D}$ mutated gene and a constitutively active *Akt* mutant were virally transferred into the brain of mice using RCAS vectors. Each of these genes was insufficient to induce gliomas when expressed alone; however, when they were expressed in combination with each other, lesions similar to human glioblastomas were produced. Although previously it was thought that neither *Kras* or *Akt* mutations are found in human GBMs, recent large-scale sequencing efforts of human tumors have demonstrated that *Kras* is a likely genetic driver of these cancers albeit at a low frequency [93]. Moreover, the authors demonstrated elevated Ras pathway activation in all GBMs they tested and increased Akt protein phosphorylation in the majority of GBMs, suggesting that upstream mutations are likely to lead to activation of these pathways.

Given that *TP53* is mutated in ~30% of GBMs and *PTEN* in around ~40%, Zheng et al hypothesised that these two mutations cooperate with each other in gliomagenesis[94]. They

crossed a hGFAP-cre mouse with *Trp53* mutant and *Pten* knockout mice, which led to grade III and grade IV gliomas at a median latency of approximately 7 months [94]. Gliomaspheres with stem cell like properties could be generated from these tumors, and the authors demonstrated that activation of myc was crucial in driving tumorigenesis in this model. Importantly, although *TP53* and *PTEN* mutations are commonly found in low grade gliomas as well, all of the tumors generated in these mice were either grade III or IV. Work from Luis Parada's laboratory supports these findings and also demonstrate cooperation of *Trp53* and *Pten* with *Nf1* in mice [95].

Zhu et al explored the cooperation between the *EGFRvIII* mutation and other genes in gliomagenesis by using transgenic mice [96]. They generated both an *EGFRvIII* transgenic mouse, in which the mutation was overexpressed at the *Col1a* locus, and also an *EGFRwt* transgenic mouse with the human gene sequence inserted and over-expressed at the *Col1a* locus. These mice were conditional and required injections of cre into the brain for the mutations to be expressed. Cre was injected into the basal ganglia (striatum) of adult mice. Neither of these mutations was sufficient to induce gliomas alone; but when expressed in combination with homozygous loss of *Pten* and *Ink4a*, both mutations were able to produce high-grade gliomas with a short latency. However, the *EGFRwt* allele produced tumors with a low incidence and long latency in comparison with a single *EGFRvIII* allele that dramatically enhanced the incidence and reduced the latency of tumor formation. Homozygous *EGFRvIII* was more efficient in producing tumors than heterozygous *EGFRvIII*, although the difference was rather small.

Glioma Cell of Origin

An important question in glioma biology is which cell type gives rise to the tumor. This is a well-studied yet still controversial topic, and is thus worth giving some consideration to here. Although it is unclear which is the key cell type of origin, it appears that the combination of genetic alterations affects whether one particular cell type can give rise to a glioma. Jacques et al introduced combinations of *Trp53* / *Pten* and *Rb* mutations in adult subventricular zone (SVZ) neural stem cells (NSCs) and in astrocytes of mice [97]. Only SVZ stem cells produced

tumors, whereas introducing these mutations into cultures astrocytes did not; moreover, *Trp53* and *Pten* mutations together induced gliomas, whereas deletion of *Rb* in addition to *Trp53* / *Pten* led to primitive neuroectodermal tumors (PNETs). Importantly, despite containing the same mutations as those induced in the SVZ, mature astrocytes were unable to form tumors.

A related study into the cellular origin of gliomas investigated the role of *Egfr* (activation) and *Cdkn2a* (loss) mutations in different brain cell types [98]. These mutations were introduced into cultured mouse astrocytes and neural stem cells, which were then transplanted into the brain (striatum) of SCID mice. If these mutations were introduced independently of each other, the cells were unable to induce gliomas. In combination however, they led to the formation of gliomas from both astrocytes and neural stem cells, suggesting that the combination of mutations rather than the cell type was more important in driving tumor formation. *Cdkn2a* loss led to dedifferentiation of the astrocytes, which allowed the cells to later be transformed if an activating *Egfr* mutation was introduced. The authors concluded that loss of *Cdkn2a* was a critical initial step in gliomagenesis that must precede *Egfr* activation if the latter is to trigger glioma formation.

To expand on these observations, Friedmann-Morvinski and colleagues used performed lentiviral vector injections to cause p53/Nf1 knockdown or H-ras expression with p53 knockdown in neurons, astrocytes and NSCs of mice. They found that all of these cell types generated malignant gliomas in mice with these genetic alterations, and concluded that most CNS cell types undergo dedifferentiation in response to defined oncogenic mutations to NSCs or progenitors, enabling tumor initiation and maintenance [99]. Although this demonstrated these differentiated cell types can give rise to GBMs *in vivo*, this does not necessarily establish which cell type is the most likely origin.

Another group investigated whether a particular cell type in the SVZ was particularly responsive to EGF; they demonstrated in mice that infusion of EGF into the lateral ventricles caused increased proliferation of C cells (transit amplifying progenitor cells that express nestin) in the SVZ, and these cells then invaded the brain parenchyma. Although no tumors

occurred in this model, the study demonstrated that exogenous EGF can increase proliferation of neural stem cells through the wild-type *Egfr* activation [100]. It is unclear from this study alone though what the effect of the *EGFRvIII* mutation in absence of exogenous EGF would be on these cells.

Another study elegantly used mosaic analysis with double markers (MADM) in mice with *p53*/*Nf1* inactivation in NSCs. Prior to GBM establishment, MADM-based lineage tracing identified aberrant growth only in oligodendrocyte precursor cells (OPCs), but not in NSCs or other NSC-lineages. Moreover, induction of *p53*/*Nf1* mutations directly in OPCs caused glioma formation, leading the authors to conclude that OPCs are the likely origin of glioma, even if the initiating mutations occur in NSCs [101], Fig 1.3.

Very recent work using sequencing data from human patients lends support to the subventricular zone being the origin of at least some GBMs [102] – the investigators performed deep sequencing of triple matched tissues from *IDH*-wild type GBM patients, including normal SVZ, tumor tissue and normal cerebral cortex. They found that normal SVZ in 56.3% of cases contained low level driver mutations (1% of tumor mutational burden) that were also present in the GBMs; introduction of driver mutations in astrocyte-like NSCs in the SVZ in mice led to migration of these cells and formation of GBMs at distant brain regions. This evidence supports these cells as being potential origins of GBMs.

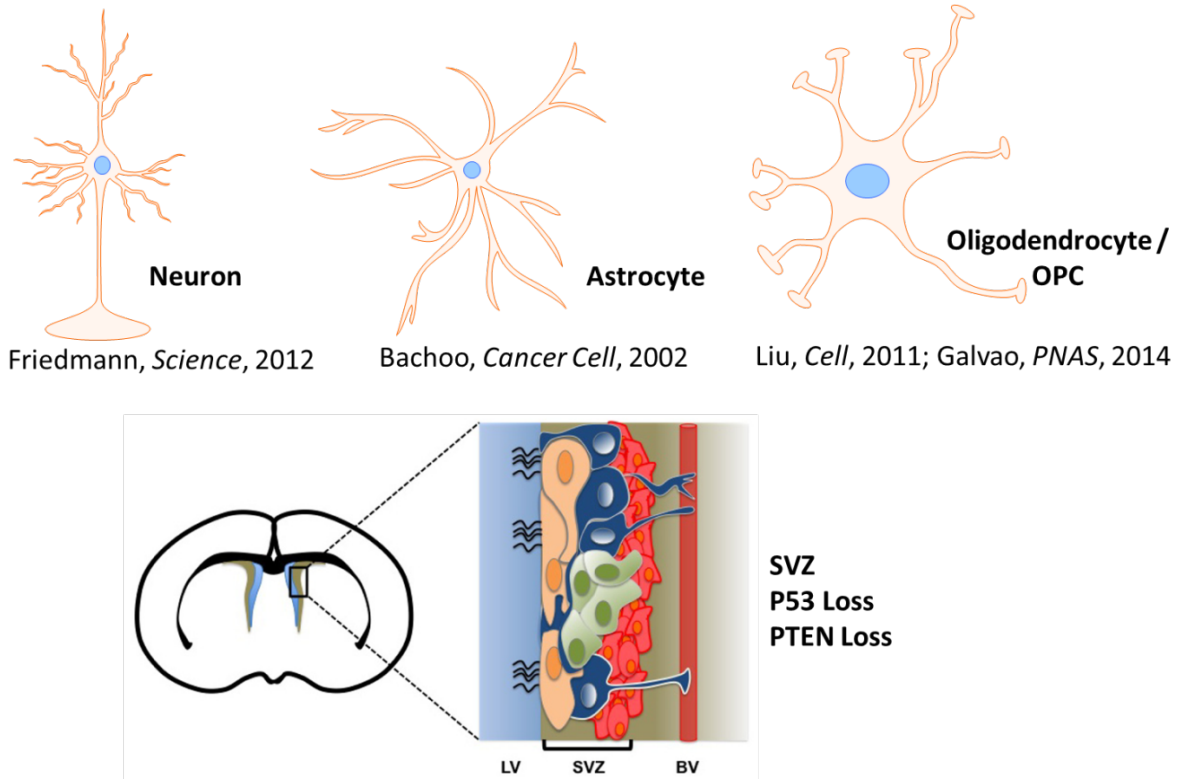

Figure 1.3. Potential sites of origin for gliomas as demonstrated in mouse models, as demonstrated in various studies suggesting these tumors may arise from neurons, astrocytes, oligodendrocyte precursor cells (OPCs) or subventricular zone (SVZ) neural stem cells [98, 99, 101, 103].

Spinal Glioma Mouse Model

Spinal gliomas cause significant morbidity such as limb paralysis, and the prognosis associated with these tumors is poor. There are very few animal models of spinal gliomas and their molecular pathology is poorly understood. Hitoshi et al used transgenic mice expressing *Pdgfβ* under the GFAP promoter using a tetracycline responsive element (TRE); they developed several mouse lines and selected a line with the highest expression of *Pdgfβ* in the spinal cord instead of the brain [104]. With this model, they demonstrated that mice developed spinal gliomas with a high incidence, and these tumors reflected a spectrum from oligodendrogliomas to astrocytomas. Loss of one copy of *Trp53* in addition to expression of

Pdgfβ led to acceleration in the time taken to develop spinal tumors; the incidence of brain gliomas is this model was relatively low (< 5%), likely due to the lower expression of *Pdgfβ* in the brain. This model provides an indication that gliomas can arise in the spinal cord with the same genetic aberration that can be used to generate brain gliomas. A strength of this model is the use of transgenic mice instead of RCAS vectors (a popular method amongst early glioma models), in which it is not possible to exclude a role for insertional mutagenesis in tumor formation and in which it is more difficult to target less accessible regions such as the spinal cord.

Summary

In summary, I have discussed the critical processes behind cancer development, introduced brain tumors and in particular gliomas, and how these tumors may be modelled and studied in mice. This will form the basis for understanding the experiments performed as part of this PhD thesis. The key aims of this Thesis, as will be described in the relevant Chapters, are to establish the role of *EGFR* in glioma initiation, and to map the cooperative mutational and functional genomic landscapes of gliomas in mice. Such knowledge will be important for deciphering complex human glioma genomes and potentially for developing new biomarkers and therapeutics.