Chapter Three: Evolution of *EGFRvIII*-induced Gliomas in Mice

Abstract

EGFR is one of the most frequently mutated and amplified genes in gliomas, which represent the commonest type of intrinsic brain tumor. It is unclear at what stage of gliomagenesis these alterations are acquired and the effect they have on the genetic evolution of cancer. Gliomas also occur in the spinal cord, but the genetics of these tumors and how they compare with their brain counterparts are poorly understood. Here, we set out to determine how a common activating *EGFR* mutation (*EGFRvIII*) influences glioma genetic evolution by leveraging mouse genomics. We expressed *EGFRvIII* in the mouse central nervous system, and found that this is sufficient to initiate glioma formation both in the brain and spinal cord with long latency. Whole-exome sequencing of resultant tumors revealed the secondary molecular alterations spontaneously acquired after *EGFRvIII*-tumor intiation, including amplification of *EGFRvIII*, deletions of *Cdkn2a* and *Nlrp1b*, and mutations of *Trp53*, *Tead2*, *Sub1* and *Nt5c2*. Transcriptomic profiling through RNA-sequencing of these tumors revealed enrichment for gene sets in multiple pathways, including Wnt, MAPK, p53, JAK-STAT and stem cell related pathways. Comparative analysis of these data with human glioma sequencing data demonstrates recurrent deletions in *TEAD2* and *NT5C2*, as well as methylation of *Sub1* and *Nlrp1* implicating these genes as putative contributors to human gliomagenesis too. This chapter presents data showing *EGFRvIII* can initiate gliomagenesis *in vivo* and the subsequent genetic alterations somatically acquired in tumorigenesis.

Introduction

EGFR as a glioma driver

EGFR is mutated in up to 60% of *IDH1*-wild type GBMs, and *EGFRvIII* is one of the commonest types of *EGFR* mutation in these tumors. As one would expect, previous analysis of TCGA data showed that *EGFR* mutation and / or amplification is associated with a significant increase in EGFR expression suggesting these genetic changes have functional consequences on the tumor [27]. In fact, this observation that increased *EGFR* expression is associated with its amplification in GBM was first described in 1987 [127]. Previous studies have suggested that *EGFRvIII* (variant III) mutations only cause gliomas in mice in the presence of other predisposing mutations in genes such as *Pten* and *Cdkn2a* [92, 96, 128]. These studies typically expressed the mutation(s) in specific locations, such as the basal ganglia or cerebral cortex, reflecting the location of cre injection. Studies have since implicated a role for the subventricular zone (SVZ) in the earliest phase of glioma formation, although this may be context-dependent on the genetic background; indeed the majority of studies that demonstrate SVZ as a glioma site of origin have used *Trp53* as a predisposing mutation. Another issue is that the observation times of these earlier studies were generally short (eg 12 weeks in the case of [29]), meaning tumors that arise after long latency were not detected.

In the TCGA cohort, GBMs have a somatic mutation rate in *EGFR* of 32.4% [27]. A review of the locations of the point mutations within the *EGFR* gene demonstrates that they are clustered either within the extracellular receptor domains or in the intracellular tyrosine kinase domains, Fig 3.1. This is consistent with the idea that (at least some of) these mutations are activating mutations, that either switch on the tyrosine kinase domain or the ligand-binding domain.

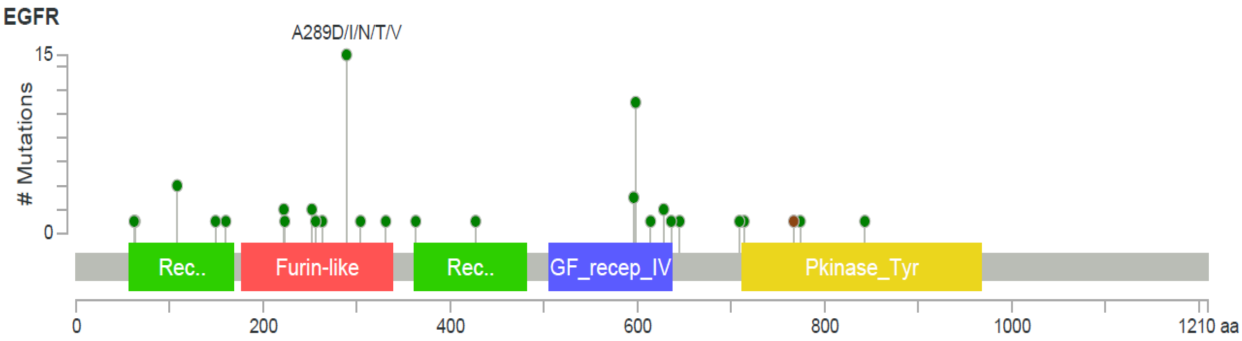


Figure 3.1. Genetic alterations in *EGFR* in human GBMs, data from TCGA. Bottom panel details the location in the gene where mutations are located; the height of the pin represents the number of tumors with a mutation in that locus. Data extracted using the publicly available software, CBioportal (see Materials and Methods). Green block = receptor ligand domain; red block = furin-like cysteine rich region; blue block = growth factor receptor domain IV; yellow block = protein tyrosine kinase.

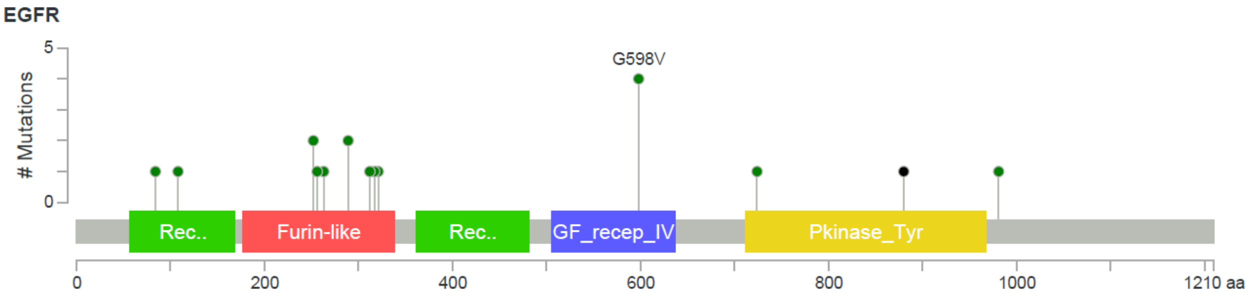


Figure 3.2. Analysis of genetic alterations of *EGFR* in human low grade gliomas from the TCGA cohort. Bottom panel displays the location of mutations within *EGFR*, the majority of which occur in the furin-like domain although there is a prominent mutation (G598V) in the growth factor receptor domain.

Analysis of TCGA data from 530 human low grade gliomas identifies a difference in *EGFR* mutational rates compared with GBMs: the *EGFR* somatic mutation rate in LGGs was 5.3% in this cohort [93]. Similar to GBMs, the vast majority of the mutations occurred either in the extracellular ligand-binding domains or the tyrosine kinase domain, Fig 3.2. However, these data do not include extrachromosomal *EGFR* mutations and mutations which are known to be present (therefore increasing the prevalence of *EGFR* driver mutations) in both LGGs and GBMs [129].

Mechanistically, *EGFRvIII* is a constitutively active form of *EGFR* and signals predominantly via the PI3K-Akt-mTOR pathway, whose ultimate effects include increased cellular proliferation [130, 131]; thus, this can explain why this mutation would predispose to tumor formation on its own, particularly if the activation of the PI3K-Akt pathway were very strong. However, a counter argument is that constitutive activation of EGFR may trigger senescence, thereby avoiding cancer formation[92]. Intriguingly, previous work has demonstrated that *EGFRvIII* may not activate the Ras-MAPK pathway as strongly as PI3K-Akt signalling, unlike its wildtype counterpart which can activate both PI3K-Akt and Ras-MAPK pathways strongly in the presence of its ligand epidermal growth factor (EGF). Glioma cells can express both wild-type *EGFR* and *EGFRvIII*, in which case EGFR phosphorylates EGFRvIII and this leads to enhanced STAT3 signalling which may increase tumorigenicity even further [132]. EGFRvIII can activate the c-MET receptor tyrosine kinase that can stimulate cell proliferation too, such that inhibiting both EGFR and c-MET together may have a strong anti-proliferative effect [133]. It has been reported that the oncogenic mutant EGFRvIII protein can be transferred between glioma cells via extracellular vesicles, thus enabling oncogenic pathway activation in neighbouring cells [134].

Human GBMs, although they frequently carry the *EGFRvIII* mutation, display substantial intra- and inter-tumor heterogeneity, with many cells not expressing the mutant protein although the tumor as a whole carries the mutation[135]. This makes it challenging to determine whether *EGFRvIII* is an initiating tumorigenic event or whether it emerges late in gliomagenesis and then establishes itself as a dominant genetic driver. However, studies have demonstrated that mutant *EGFR* provides a selective growth advantage for glioma cells *in vivo*, specifically in maintaining glioma growth following tumor initiation [136]. Whether *EGFRvIII* can initiate glioma formation and the subsequent genetic alterations driving tumor formation are still unknown.

EGFR targeted therapies for gliomas

There are a number of targeted therapies against the epidermal growth factor receptor (EGFR, either already in clinical use or in development. Thus, it is worth considering key issues surrounding these approaches here (and these are also discussed further in the Discussion Chapter). EGFR inhibitors have been demonstrated to be beneficial in certain cancers carrying an *EGFR* mutation, such as lung and colon cancers where there is a clear survival benefit associated with this treatment in subsets of patients [137]. The original EGFR inhibitors like erlotinib and gefitinib are classed as reversible, in that mutations in the gene can easily lead to tumor resistance [138-141]; as a result, newer irreversible inhibitors, such as afatinib, have been developed that bind the receptor more strongly. Early clinical trials using EGFR inhibitors in patients with gliomas demonstrated evidence of tumor regression with this treatment, particularly in tumors also containing *PTEN* co-expression [142]. However, larger trials have not demonstrated an improvement in survival with this treatment [143]. Potential reasons for this are that gliomas carrying *EGFR* mutations are no longer dependent on these for their growth and instead are dependent on other acquired mutations, and / or that these drugs do not completely block EGFR signalling [144]. Indeed, a recent study in mice demonstrated that genetic ablation of *Egfr* gave a stronger tumor inhibitory effect than EGFR inhibitors, suggesting complete suppression of the receptor is important for slowing tumor growth [128]. Given the potential for resistance to this form of therapy, studies are investigating the role of additional therapies to complement EGFR inhibitors.

Immunotherapy is a promising area within cancer research, which has revolutionized treatment of malignant melanoma for example. A 13-amino acid peptide vaccine, rindopepimut, based on the EGFRvIII protein has been generated and is being studied in randomised controlled trials after showing benefit in mouse models and in small glioblastoma patient cohorts [145]. This method relies upon the body's immune system to trigger a response to the foreign EGFRvIII peptide, which contains a novel surface glycine residue that is not normally present on the wild-type EGFR and is therefore an immunogenic epitope. The precise component of the immune system (whether antibodies or T-cells) that is responsible for removal of the tumor cells is still unknown however. In some pre-clinical and clinical studies, rindopepimut is also being injected subcutaneously with GM-CSF to enhance the immunogenic response. Antibodies against EGFR have been described in experimental models of glioma, with some promising results [146].

The *EGFRvIII* mutation has been implicated in resistance to radiotherapy in GBM, potentially through conferring an increased rate of double strand break repair compared with normal brain which mitigates the effect of radiotherapy on tumor cell killing [147, 148]. Resistance to EGFR tyrosine kinase inhibitors (TKIs) may occur through multiple mechanisms. A possible such mechanism is maintenance of mutant *EGFR* on extrachromosomal DNA, and following treatment with *EGFR* TKIs there is elimination of mutant *EGFR* from extrachromosomal DNA to give treatment resistance; after withdrawal of treatment, mutant *EGFR* may reappear on extrachromosomal DNA to drive tumor re-expansion [149]. Amplified copies of *EGFR* may also be contained on double-minute chromosomes [150]. Alternatively, *EGFRvIII* may transcriptionally suppress alternative receptors that activate similar pathways, such as PDGFRβ; treatment with EGFR TKIs may then increase transcription of *PDGFRβ* that can then drive tumor growth [151]. Other EGFR-related members of the ERBB protein family may similarly become activated in glioma stem cells if EGFR is inhibited ([152]. *EGFRvIII* itself may drive resistance to erlotinib when used to treat tumors with an *EGFR* amplification: in this situation, *EGFRvIII* becomes upregulated resulting in an increase in PI3K signalling via increased expression of PI3Kp110∂ (a regulatory subunit of PI3K) [153]. There are important effects of *EGFRvIII* on transcriptional programs as well, for example acting via transcription factors SOX9 and FOXG1[154]. *EGFR* mutants are also likely to enhance tumor invasion into normal brain, as it has been demonstrated to upregulate enzymatic effectors of invasion such as metalloproteases and serine proteases [155].

Although these novel therapeutics are generating much interest in EGFR as a clinical target, this molecule has thus far proven to be an unsuccessful target in the treatment of glioma patients. Potential reasons for this are an incomplete understanding of the biology of this gene, and particularly of its interactions with other molecules in a tumor such as cooperativity between EGFRvIII and other proteins.

Aims of Study

Whole-genome sequencing studies of human brain gliomas have shed much light on the genetic and epigenetic landscapes of these tumor types[28, 30, 93, 156]'[157]. In addition to mutations, driver genes may be altered through transcriptional, methylation or large copy number changes, and these are more difficult to identify as cancer drivers. Another complicating issue is that mutations in individual tumors can occur in different combinations, which can affect prognosis and response to therapy[1, 158], but this makes it more difficult to confidently identify which genes are truly collaborating with one another. Given the extensive intra-tumor heterogeneity of end-stage tumors, the timing of key mutations acquired during the natural history of gliomas and how these driver mutations influence tumor genetic evolution cannot easily be inferred from human genomic studies.

Activating mutations in the epidermal growth factor receptor (*EGFR*) occur in up to 60% of *IDH1*-wild-type GBMs[28] of which *EGFRvIII* is the most common (an in-frame deletion of exon 2 to 7 in the extracellular domain leading to constitutive receptor activation[96, 133]). Frequent mutations and amplifications of *EGFR* (including extrachromosomal ones) have recently been detected in *IDH1*-wild-type, histologically low-grade gliomas (LGGs)[129, 159], highlighting a need for integrated molecular diagnosis. In addition to brain tumors, studies on small cohorts of patients have identified *EGFR* amplification and expression in spinal gliomas, particularly in leptomeningeal-disseminated paediatric LGGs [160, 161], suggesting increased EGFR signalling may promote tumorigenesis in a subset of spinal gliomas. However, the timing of *EGFR* mutations in gliomagenesis, their role in spinal gliomas and their cooperative genetic lesions remain largely unknown. In particular, the genetic drivers of spinal gliomas and how they compare with their brain counterparts are obscure[162]. Understanding the functional genomic landscapes of gliomas is therefore of the utmost importance and will help us decipher human patient glioma genomes.

Here, we aimed to address these challenges by investigating the genetics of gliomas from mice expressing a constitutively-active *EGFR* mutation (*EGFRvIII*) in the central nervous system under control of the nestin-promoter. We show that *EGFRvIII* is sufficient to initiate gliomagenesis from the normal mouse brain and spinal cord with long latency in this model. By combining whole-exome sequencing, transcriptomics, and genome-wide *piggyBac* transposon mutagenesis (discussed further in Chapter Four), we identified recurrent mutations in known and novel putative glioma genes and characterized the functional genomic landscapes of *EGFR*-mutant brain and spinal gliomas in mice.

Results

Here we aimed to study the role of *EGFRvIII* in gliomagenesis. For this, we generated double heterozygous mice carrying a conditional human *EGFRvIII* transgene (integrated in the Col1a1 locus, chromosome 11) [96] and expressing cre under the control of the Nestin promoter[87] (*nes*-cre), which activates *EGFRvIII* expression primarily in the central nervous system. It has been previously shown that mice with the nestin-cre allele express cre from embryonic day 13 throughout most of the central nervous system, eye and also the kidneys – this was demonstrated by Dubois and colleagues who showed almost complete cre-mediated recombination in these tissues by embryonic day 15.5 using LacZ based reporters [88], except in early embryonic ventricular zone neural progenitors and neural stem cells in which recombination is complete by embryonic day 17.5) [89]. The outline of the experiment for our study is shown in Fig 3.3.

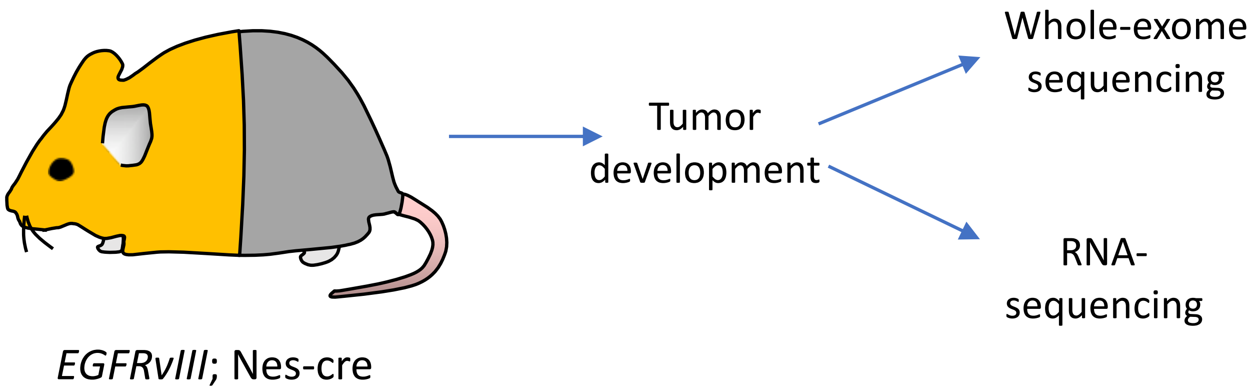


Fig 3.3. Outline of the experiment: *EGFRvIII* was conditionally expressed throughout the central nervous system using *nes-cre*, and resulting tumors were subjected to whole-exome sequencing and RNA-sequencing.

Eye lesions in *EGFRvIII*; *nes*-cre mice

EGFRvIII; *nes*-cre mice started developing lesions within the eye that were clinically apparent from around 7 weeks of age. These lesions typically eventually affect both eyes with 100% penetrance. The complex eye phenotype presented with a number of features, including one or more of: cataract, hyphema, secondary glaucoma and proptosis, Fig 3.4. Histology revealed abnormal neovascularisation in the retina associated with chronic micro-haemorrhages and occasionally larger bleeds (n=24 eyes). In all cases, there was either partial or complete degeneration of the lens. There was also a proliferation within the ciliary body in many cases; although this could be mistaken for an ocular melanoma, the cytology of the proliferation does not match that of a melanoma. It is likely therefore that the ciliary body proliferation is a secondary reaction to chronic haemorrhages within the eye, rather than presenting a tumor or tumor precursor. No such lesions were observed histologically in age-matched control eyes from wild-type mice (with EGFRvIII but lacking nestin-cre; n=10 eyes). However, these lesions were not the focus of this project so these data are not presented here.

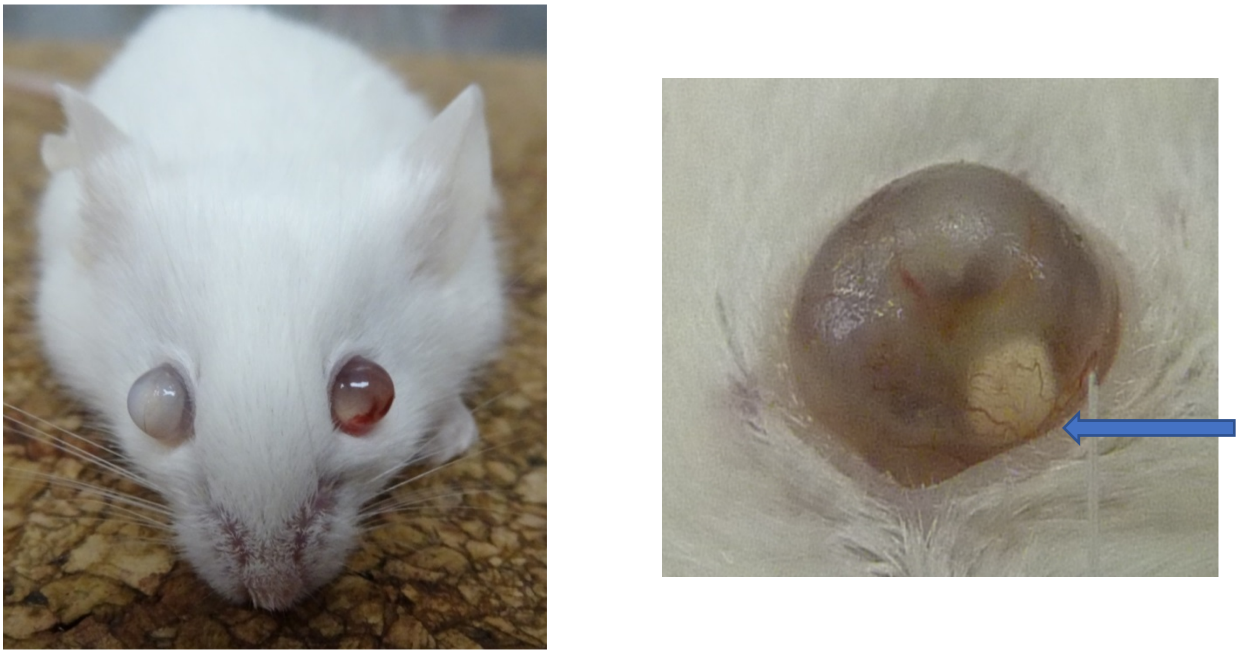


Figure 3.4. Eye lesions in *EGFRvIII*; *nes*-cre mice. Left panel shows proptosis of eyes and small haemorrhages in left eye. Right panel shows apoptotic lens, confirmed with histology (histology not shown).

Clinical Phenotypes of Mice

EGFRvIII; *nes*-cre mice started developing signs of neurological disease from 14 weeks onwards. These include progressive macrocephaly (enlargement of the head, Fig 3.5), lateral leaning behaviours or a head tilt, circling, weakness of limbs, and seizures. At later stages, the mice displayed a combination of these signs, and the vast majority also display the eye phenotype with a degenerate lens. Mice required culling when the phenotype restricted their basic functioning, such as inability to mobilise and therefore to feed. Records were kept of the age at which these mice were culled and of their clinical phenotype. Control mice with *nes*-cre but not *EGFRvIII* did not display any signs of neurological disease after one year of observation.

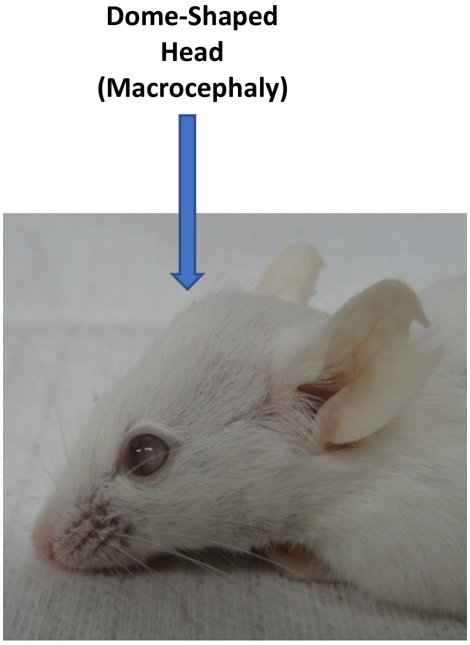


Figure 3.5. Clinical phenotypes of mice expressing the *EGFRvIII* allele in nestin-expressing tissues. The phenotypes are primarily neurological; left panel shows a typical example of macrocephaly due to hydrocephalus.

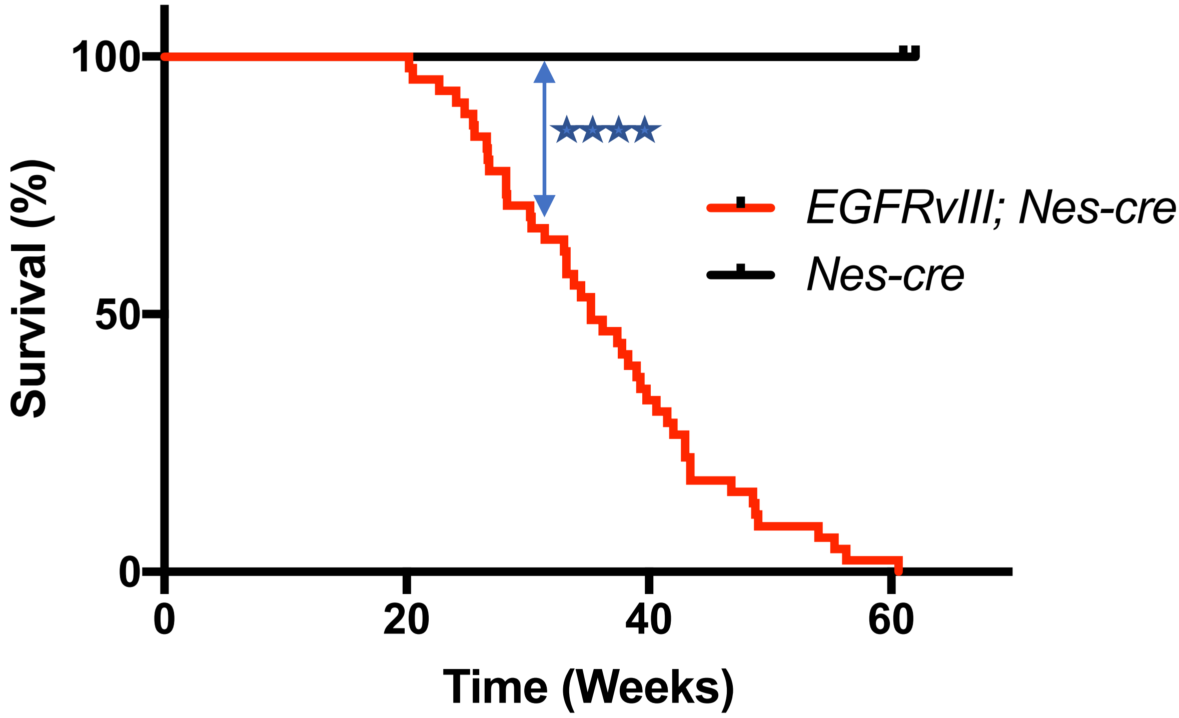


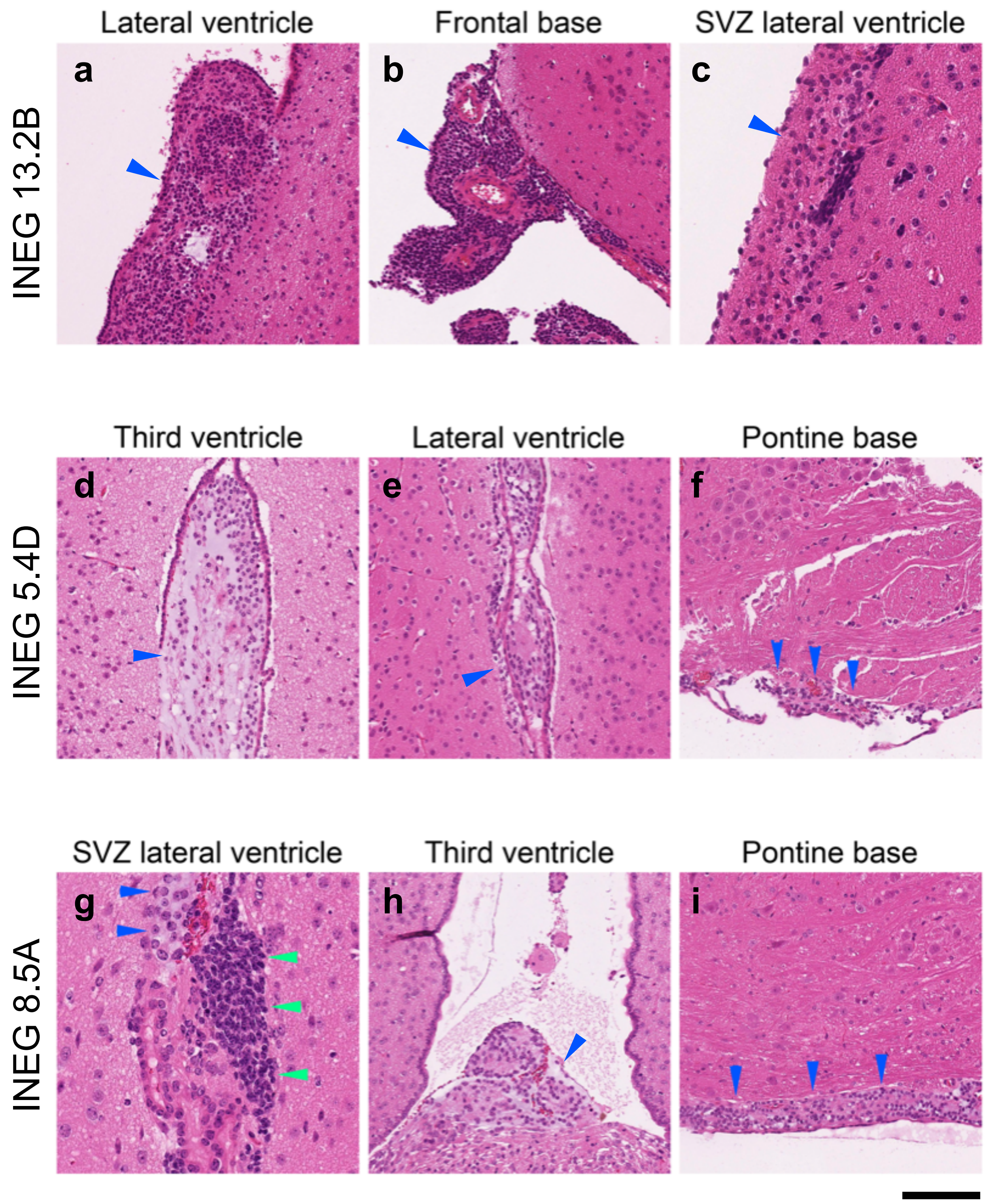
Fig 3.6. Kaplan-Meier plot of *EGFRvIII ; nes-cre* mice and control (*nes-cre*) mice (**** denotes p < 0.0001, log-rank test, n=31 and n=10 mice respectively). Control mice with *nes*-cre but lacking *EGFRvIII* did not display signs of neurological disease after one year of observation.

***EGFRvIII* initiates gliomagenesis**

By 60 weeks of age, 100% of mice had succumbed to brain and/or spinal tumors (n=31), Fig 3.6. Spinal tumors will be described in the next sub-section.

Pathological examination of brains prior to clinically overt disease (mice aged 12-28-weeks) revealed small glioma precursor lesions with proliferative activity (as indicated by immunohistochemical staining with Ki67) – these lesions are also described as 'microneoplasias' (13/13 mice) and have also been reported in mice from different genetic contexts before[97, 163]. The size of each microneoplasia was between 100 and 200µm. Multiple microneoplasias were detected bilaterally protruding into the lateral ventricles, third ventricle and from the brain surface, and they had subpopulations of proliferating cells as detected by Ki67 immunostaining, Fig 3.7 and Fig 3.9. In contrast, no such lesions were seen in 5 control mice carrying only the *nes*-cre allele (age 10-30 weeks; p = 0.0001, two-sided Fisher's exact test), Fig 3.8. The control and experimental brain samples were examined in exactly the same way by our Consultant Neuropathologist, with four standardised cut sections from each brain, as described in the Materials and Methods. Importantly, Professor Sebastian Brandner (Consultant Neuropathologist) provided histological diagnoses for all samples in this study. Using immunohistochemical staining, we demonstrated these microneoplasias expressed protein markers of neural stem cells and transit-amplifying cells, specifically Sox2, Nestin, PDGFRa, GFAP and Olig2, Fig 3.10 and 3.11.

Next, we examined mice that were culled following development of clinical signs of underlying disease. *EGFRvIII*; *nes*-cre mice displayed neurological signs due to one or multiple gliomas within the lateral ventricles and / or brain surface with histological evidence of subarachnoid involvement (26/31 mice had brain gliomas; mean survival 36.2 weeks), Fig 3.12. In order to confirm *EGFRvIII* recombination had occurred specifically in tumor cells from microneoplasias and gliomas, we performed immunostaining for human EGFR and EGFRvIII in these mouse brains which demonstrated strong EGFR and EGFRvIII expression specifically in gliomas and their precursors but not in normal brain (5/5 tumors positive for EGFR immunostaining; 4/4 tumors positive for EGFRvIII immunostaining), Fig 3.13 and 3.14. Histopathological analysis by a Consultant Neuropathologist revealed these tumors had histological features and expressed protein markers comparable to those of human gliomas, Fig 3.15. The cells of these tumors have relatively monomorphic round tumor cell nuclei of a glial nature. The tumors displayed small lakes of myxoid matrix, similar to those observed in human astrocytomas or oligodendrogliomas, and range in size from 200μm upwards. Although the majority were histologically LGGs, a small proportion also displayed necrosis and microvascular proliferation (endothelial hyperplasia) that is characteristic of human GBMs, Fig 3.16. The grading system we used to define glioma grades is described in the Materials and Methods. Tables 3.1 and3.2 show all mice in this study and their associated pathologies.



EGFRvIII-induced brain microneoplasias

Lateral ventricle

Frontal base

SVZ lateral ventricle

Third ventricle

Lateral ventricle

Pontine base

SVZ lateral ventricle

Third ventricle

Pontine base

INEG 13.2B
INEG 5.4D
INEG 8.5A
a
b
c
d
e
f
g
h
i

Fig 3.7. Microneoplasias in EGFRvIII mouse brains. Examples of the formation of small tumors in the ventricular system and subarachnoid space. A, B, C, tumor growth in the lateral ventricle, the base of the frontal brain and the subventricular zone (SVZ) adjacent to the lateral ventricle. D, E, F, formation of the hypercellular myxoid intrinsic tumor in the third ventricle (D) the lateral ventricle (E) and the base of the pons (F, arrows). G, H, I, hypercellular cluster (dark nuclei of expanded SVZ stem/progenitor cells (green arrows) and adjacent a small glial neoplasm (blue arrows). H, small glioma protruding from the floor of the 3rd ventricle and I, subarachnoid spread of a glial neoplasm on the base of the pons, in a “sugarcoat” fashion (arrows). Lettering on sides of panels reflect mouse IDs from which these tumor originated. All histology in this Chapter and this Thesis was reviewed by Consultant Neuropathologist, Professor Sebastian Brandner (University College London, Department of Neuropathology).

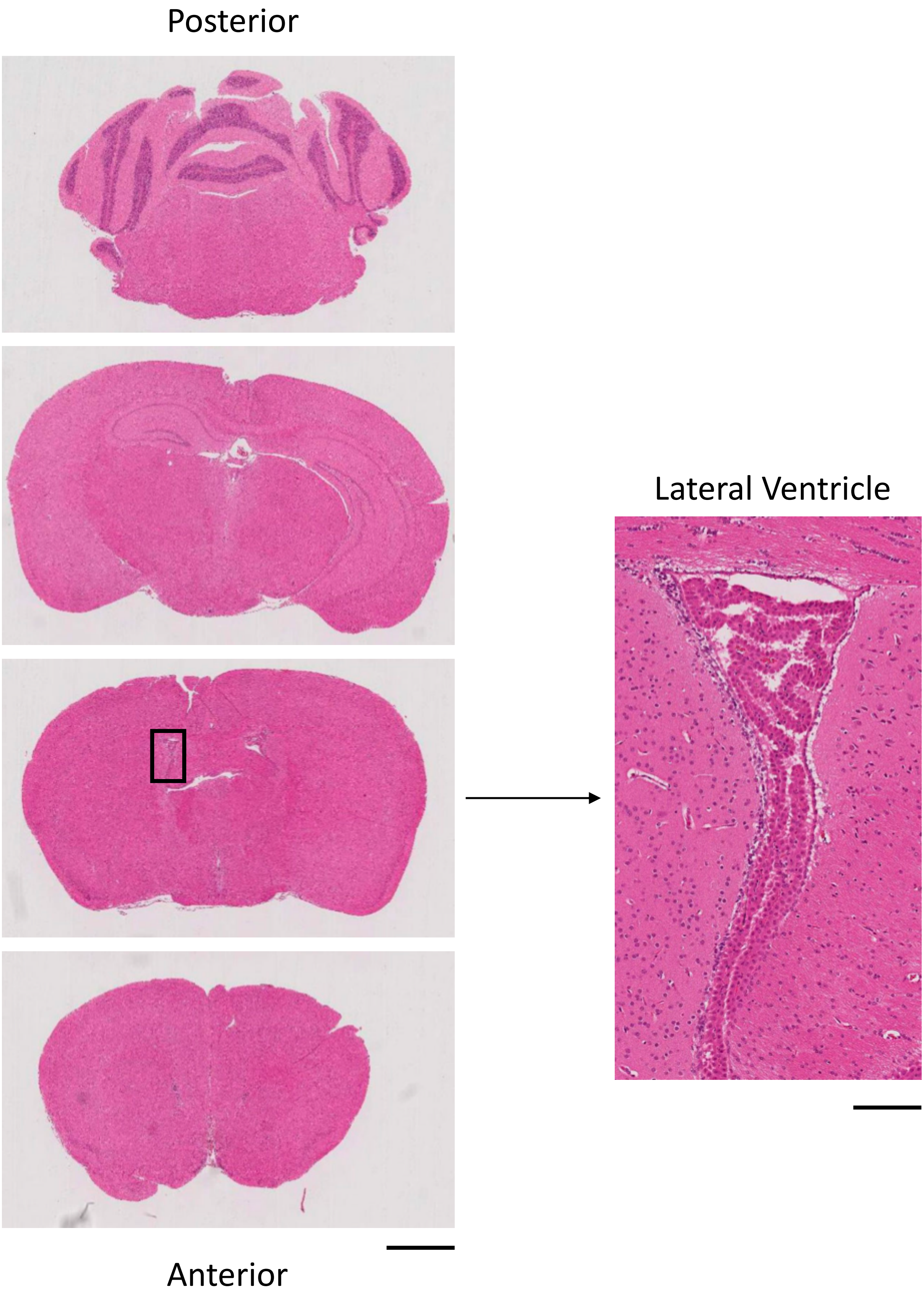

Posterior
Lateral Ventricle
Anterior

Fig 3.8. H&E stain of a typical control (nes-cre) mouse brain, showing no evidence of a microneoplasia or glioma in any location from 4 independent coronal sections. 5/5 such mice were examined with the same result. Scale bar represents 1mm for left panels and 100μm for the right panel.

Mouse ID	Age (weeks)	Pathology of brain
13.2B	12.9	hydrocephalus, hyperplasia of SVZ with multple cell clusters. Microneoplasia - Budding SVZ growth, small tumor SVZ derived
13.2f	22.1	hydrocephalus, small ventricular glioma
8.2e	24.1	bilateral intraventricular glial neoplasm/ glioma (small - microneoplasia), mild hydrocephalus, skull base microneoplasia
8.3g	21.8	Microneoplasia
31.1h	17.8	Focal clusters in SVZ, microneoplasia
5.5d	20.2	small Glioma / microneoplasia on brain surface, base of brain
5.4d	28.2	small Glioma/ microneoplasia in lateral ventricle and 3rd ventricle
25.2e	18.9	Subventricular cell cluster, small tumors on base of brain and SVZ (microneoplasias).
30.2e	16.2	Microneoplasia on base of brain and cerebellar flocculus
36.1j	21.1	Brain surface microneoplasia
8.3B	28	hydrocephalus, tumor cells (microneoplasia) in corner of lateral ventricle, 3rd ventricle and brain base
20.1e	27.6	Microneoplasia in lateral ventricle and 3rd ventricle, base of the brain.
20.3b	20.2	Microneoplasia in lateral ventricle and 3rd ventricle, base of the brain.

Table 3.1. Clinical and pathological details of *EGFRvIII* ; *nes*-cre mice with brain microneoplasias.

Ki67 Proliferation index of microneoplasias in a brain of an *EGFRvIII* ; Nes-cre mouse

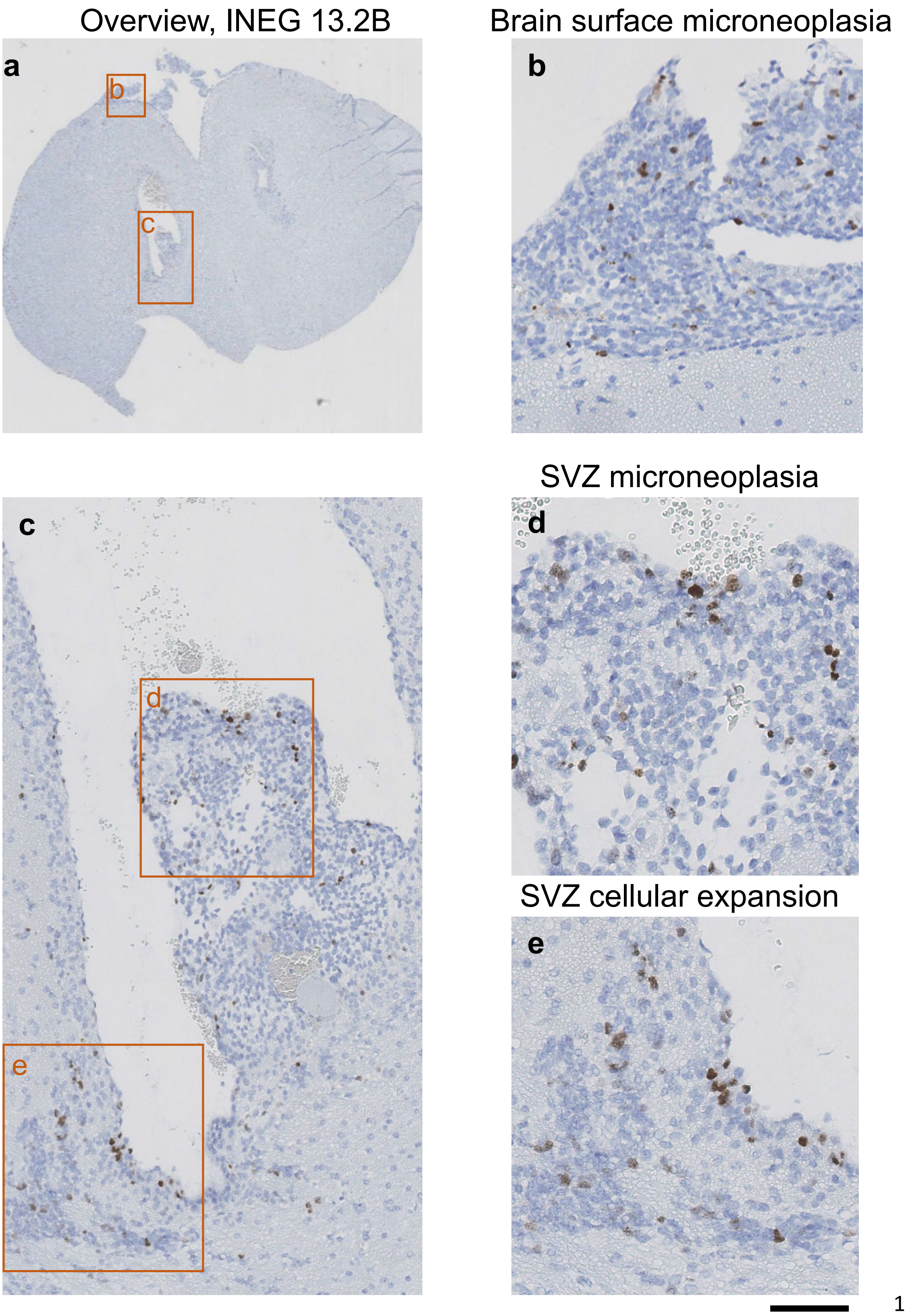

Overview, INEG 13.2B
a
b
c
Brain surface microneoplasia
b
c
d
e
SVZ microneoplasia
d
SVZ cellular expansion
e

Fig 3.9. Ki67 immunohistochemical staining of early gliomas from *EGFRvIII*-conditional mouse. Positivity for the stain is observed exclusively in a minority of cells in this early glioma and in a nearby SVZ cellular expansion. Ki67, a proliferative marker, stained a small proportion of these cells suggesting these lesions are not characterised by brisk mitotic activity. Scale bar represents 1mm for a, 50 µm for b, d, e, and 200µm for c . 3/3 mice had early tumors showing similar positivity for Ki67.

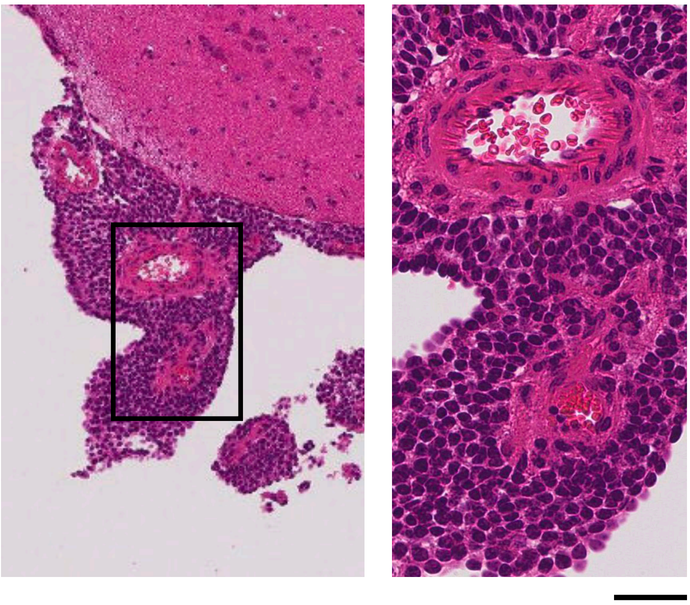


Fig 3.10. Low and high power views of a small glioma (microneoplasia) protruding from the cortical surface of the brain. Scale bar corresponds to 90µm for left panel, and 25µm for right panel.

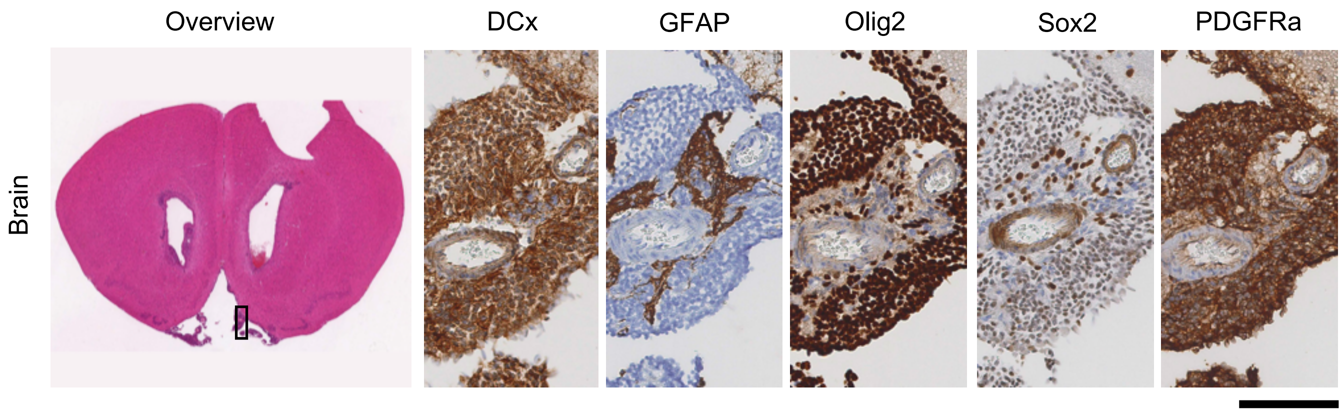


Fig 3.11. Histopathology of brain microneoplasia (left to right): low power view of H&E stain of a brain with a typical microneoplasia (same as in Fig 3.11), and high power view of immunostains of this neoplasm showing positivity for neural lineage markers double-cortin (DCx), GFAP (reflecting reactive astrocytes between tumor cells), Olig2, Sox2 and PDGFRa (n=5/5). Scale bar corresponds to 1mm for left H&E panel, 70µm for immunostain panels.

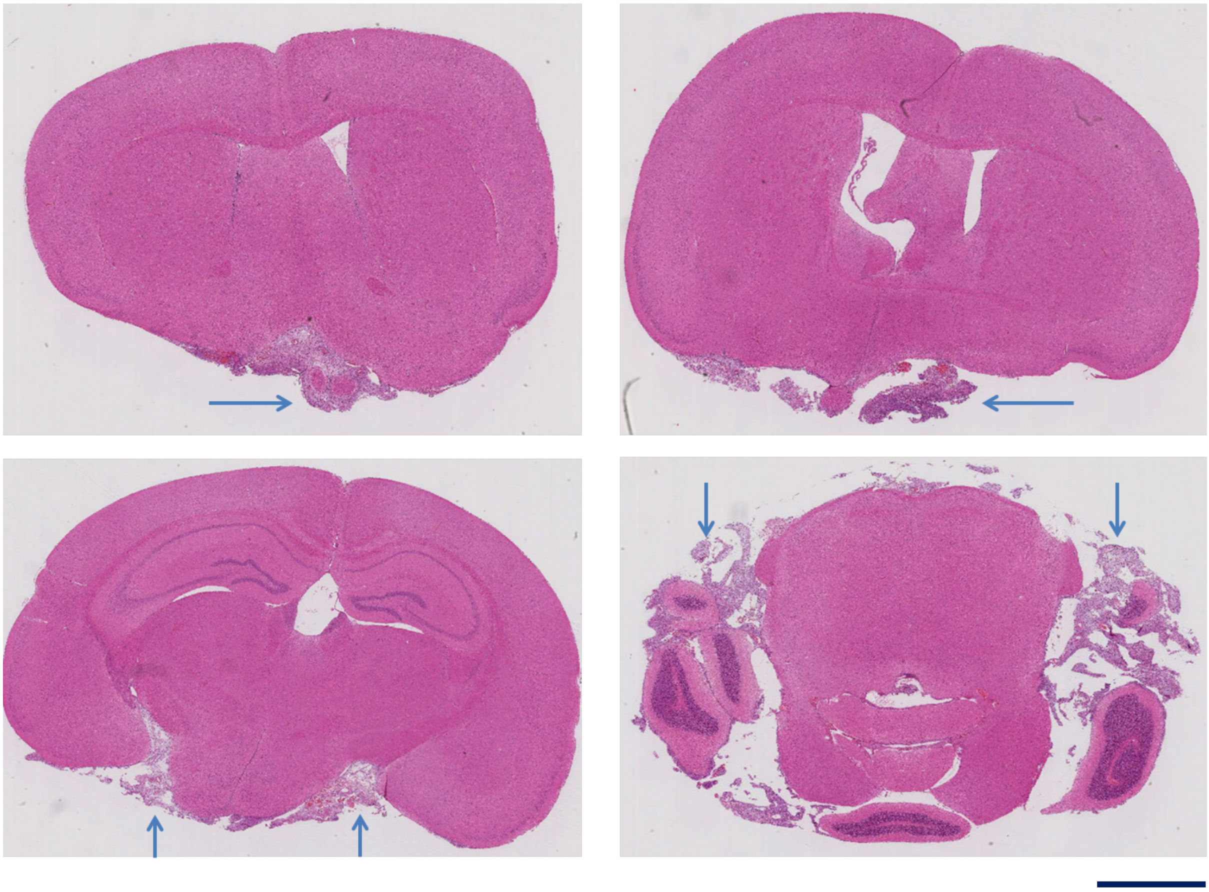


Fig 3.12. Serial sections showing subarachnoid spread of gliomas in *EGFRvIII* mice. Serial slices of a whole brain from one mouse showing dissemination of a glioma on the brain surface (highlighted by arrows); these are H&E stained sections. Scale bar represents 1.2mm.

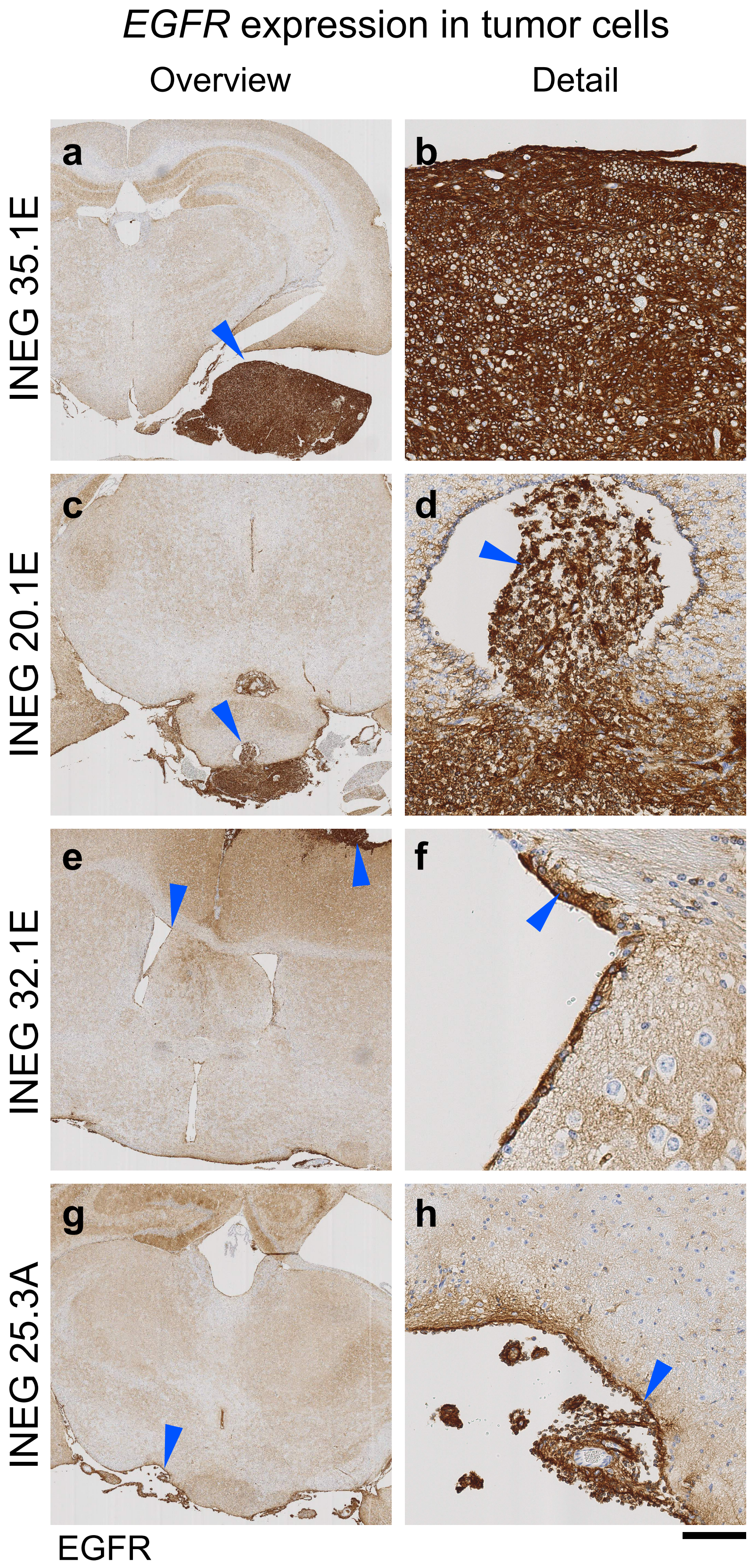

EGFR expression in tumor cells
Overview
Detail
INEG 35.1E
INEG 20.1E
INEG 32.1E
INEG 25.3A
a
b
c
d
e
f
g
h
EGFR

Fig 3.13. Expression of human EGFR, as detected by immunostaining, is limited to tumor cells in *EGFRvIII*; *nes*-cre mouse brains. 5/5 tumors from these mice stained positive for EGFR. Overview (left) and detail (right) of tumors and microneoplasias of different sizes and locations. A, B, medium-sized circumscribed, extra-parenchymal growing tumor attached to the temporal lobe. B, detail showing strong and diffuse EGFR expression specifically in the tumor. C, D, small circumscribed tumor growing on the floor of the third ventricle and expanding towards the optic tract. E, F, Likely transformed cells, with possibly incipient formation of microneoplasia in the left lateral ventricle. G, H, scattered small neoplastic lesion on the floor of the midbrain. Scale bar corresponds to 0.7mm for A, 200μm for B, 0.4mm for C, 200μm for D, 0.7mm for E, 100μm for F, 0.5mm for G, 200μm for H.

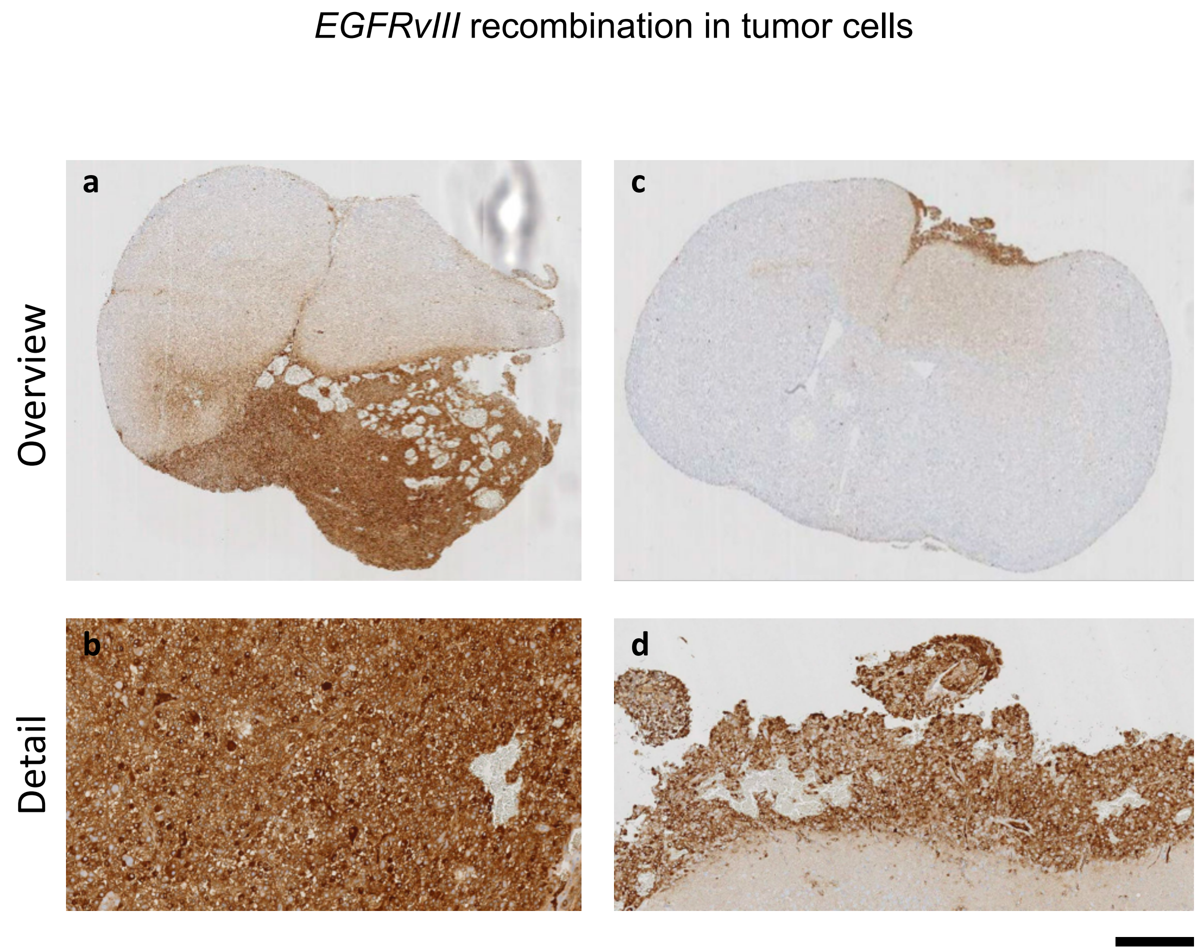

EGFRvIII recombination in tumor cells
Overview
Detail
a
c
b
d

Fig 3.14. Expression of human EGFRvIII is limited to tumor cells. A, B, overview and detail images demonstrating EGFRvIII immunostaining is positive across glioma cells but not normal mouse brain in *EGFRvIII*; *nes*-cre mice (n=4). C, D, overview and detail images demonstrating *EGFRvIII* is expressed in smaller glioma nests (precursors to larger tumors) in these mice. Scale bar corresponds to 1mm for A and C, and 100μm for B and D.

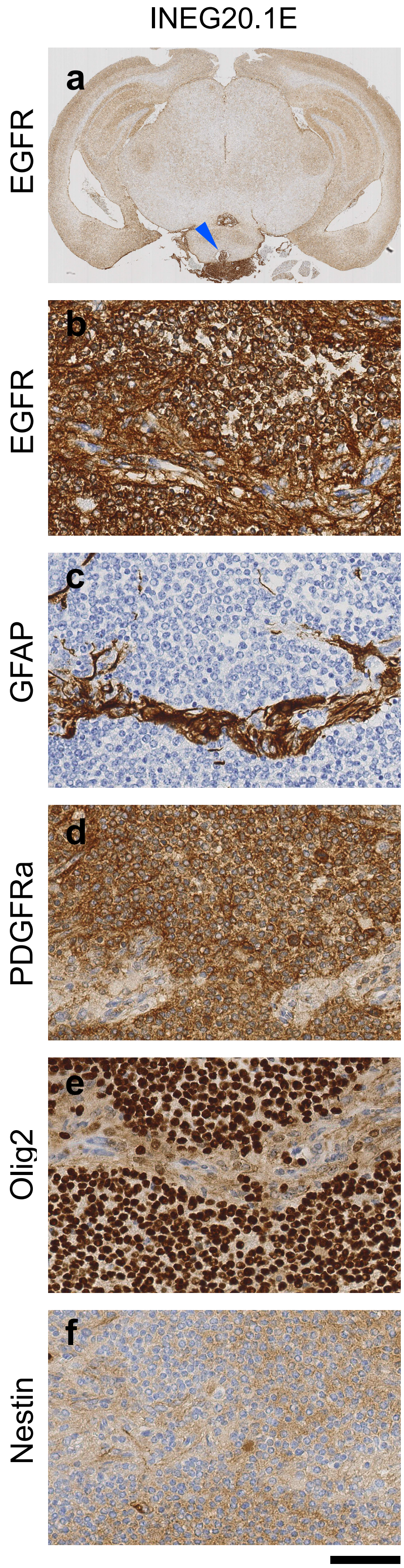

INEG20.1E
a
EGFR
b
EGFR
c
GFAP
d
PDGFRa
e
Olig2
f
Nestin

Fig 3.15. Histopathology of a typical small glioma from an *EGFRvIII*; *nes*-cre mouse. a, overview of the coronal brain section with a circumscribed extracerebral intrinsic tumor, highlighted with an immunostaining for EGFR. b, detail of the tumor, stained for EGFR. c, immunostain for GFAP shows negative tumor cells enclosing a strand of reactive glial tissue. All tumor cells strongly express PDGFRa (d), Olig2 (e), and Nestin (f), which are typical markers expressed by human gliomas. Scale bar corresponds to 50μm for b, c, d, e, f; 1.3mm for a.

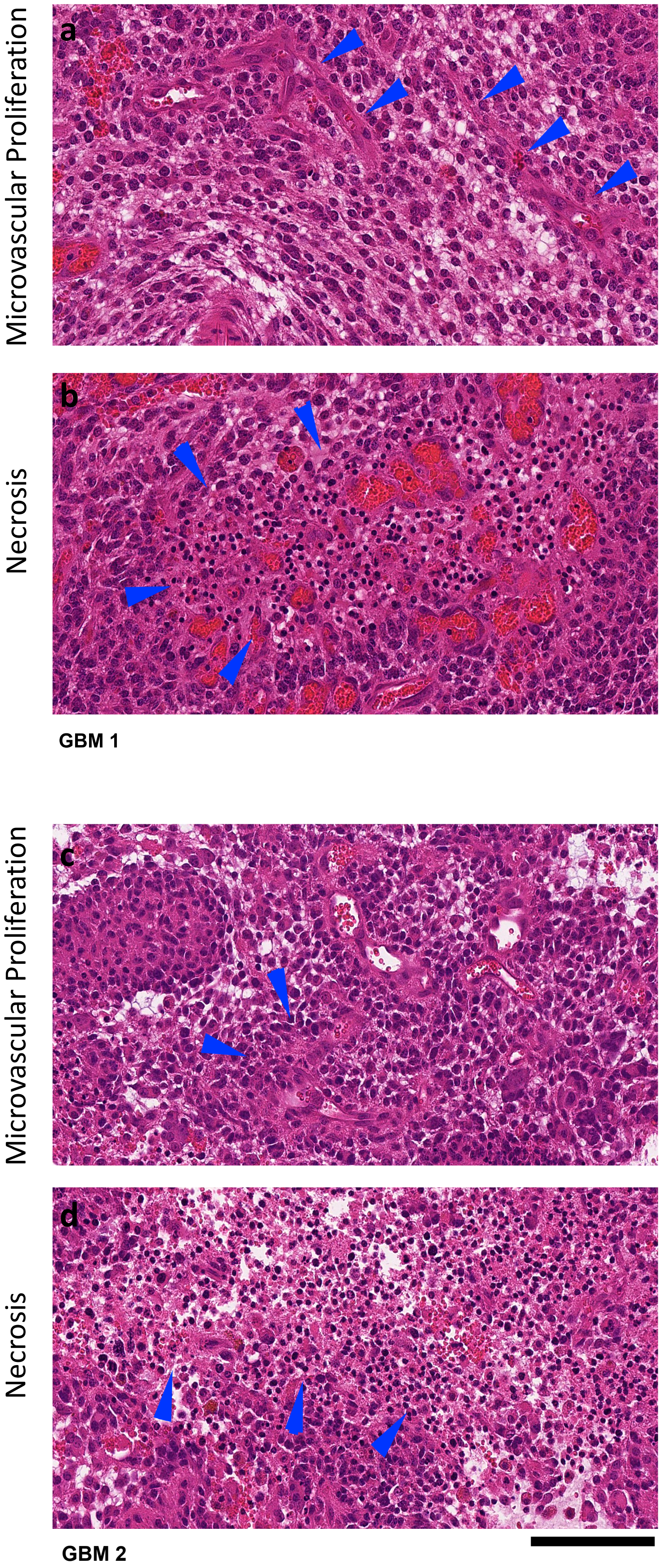

Microvascular Proliferation
a
Necrosis
b
GBM 1
Microvascular Proliferation
c
Necrosis
d
GBM 2

Fig 3.16. Histological features of GBM in *EGFRvIII*-mice. Typical examples showing defining histological features of glioblastoma in two mouse tumors – A, B show microvascular proliferation and necrosis in one GBM, and C, D show these features in another. Scale bar corresponds to 50μm.

Spinal Cord Gliomas

We next examined the spinal cords of *EGFRvIII*; *nes*-cre mice. Aside from brain tumors, *EGFRvIII*; *nes-cre* mice also developed multiple and widespread spinal tumors with 100% penetrance (31/31 mice), which account for the apparent neurological deficits including focal limb weakness. The tumors were located on the spinal cord surface, with evidence of local invasion into the surrounding nerve roots, soft tissue and cranial nerve ganglia. In contrast, most tumors did not show infiltration into the cord: only two *EGFRvIII*; *nes*-cre spinal tumors showed histopathological evidence of spinal cord parenchymal invasion by tumor cells, reminiscent of intramedullary spread of spinal astrocytomas in humans. The spinal tumors were present throughout the leptomeningeal space (in cervical, thoracic and lumbar spine) indicating leptomeningeal-spread, which is a poor prognostic indicator in human patients [164].

In 5/31 mice without established brain tumors (but with microneoplasia), there were still spinal tumors present at all levels of the spines examined (cervical, thoracic and lumbar). Histological examination of the spinal tumors classified them as grade II glioma in all cases (Fig 3.17), even in the presence of grade IV intracranial gliomas, suggesting these are primary spinal gliomas, most likely arising independently. These spinal tumors expressed classical glioma markers, such as GFAP, Sox2, Olig2 and PDGFRa, Fig 3.18.

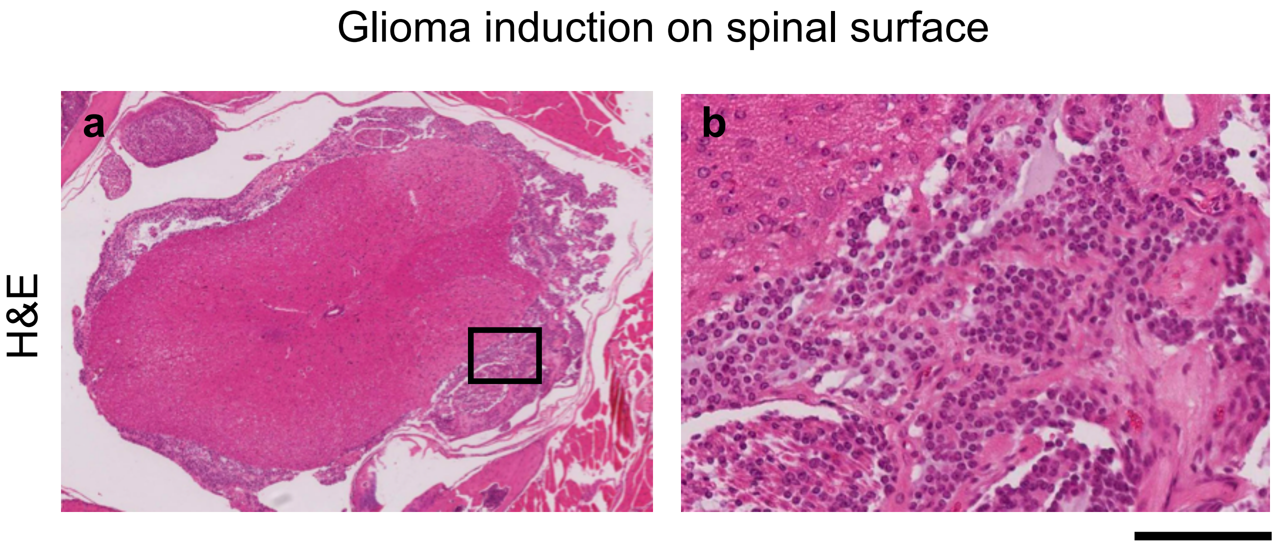

Glioma induction on spinal surface
H&E
a
b

Figure 3.17. Spinal cord glioma initiated by *EGFRvIII*. Haematoxylin and eosin stains. High powered view is presented on the right. Scale bar represents 1mm for a and 100μm for b.

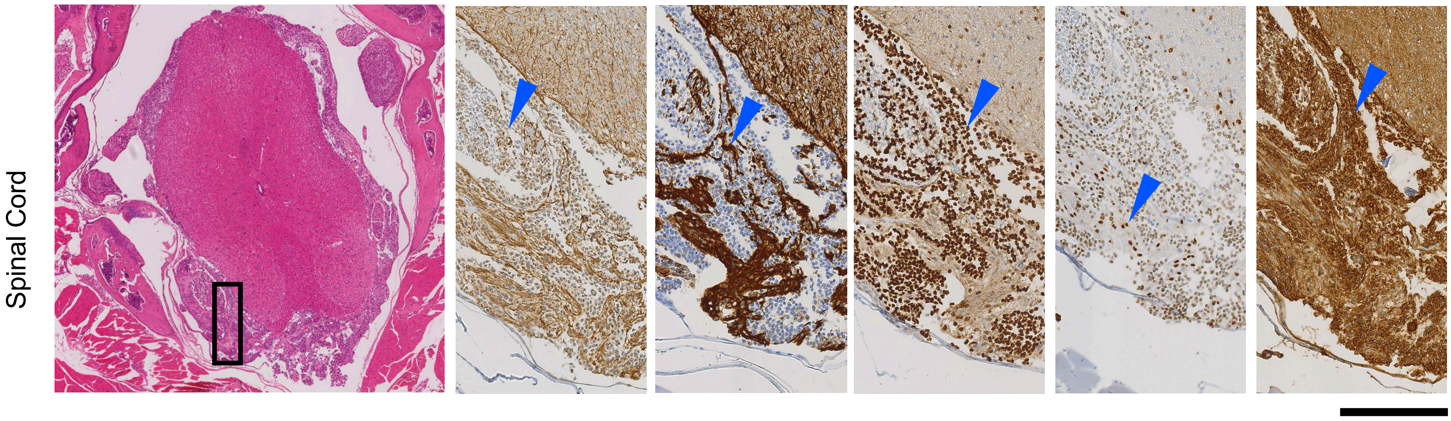


Fig 3.18. Histopathology of a typical spinal glioma in an *EGFRvIII*; *nes*-cre mouse (left to right): low power view of spinal cord with an encasing glioma, and high power views of immunostains of this tumor for neural lineage markers – tumor cells are negative for DCx, reactive astrocytes are positive for GFAP, and tumor cells are positive for Olig2, Sox2 and PDGFRa. Scale bar corresponds to 1mm for left H&E panel, 70µm for upper immunostain panels; 0.7mm for lower H&E panel, 140µm for lower immunostain panels.

Mouse ID	**Age (weeks)**	**Brain Pathology**	**Spine Pathology**
50.1h	34	tumor on lateral ventricle and subarachnoid spread, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
35.2e	23.8	mild hydrocephalus, subarachnoid tumor, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
36.1D	40	SVZ thickening and tumor spread, subarachnoid widespread, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
51.1f	22.7	brain base tumor, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
36.2c	24.8	SVZ thickening and tumor spread, subarachnoid widespread, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
8.2g	41.5	intraventricular and subarachnoid tumor, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
36.2b	37.8	SVZ thickening, widespread thick tumor growth, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
8.5d	33.2	Tumor cells basal and possibly also ventricular, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
30.1b	33	normal SVZ, subarachnoid tumor spread, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
35.1a	33.8	widespread subarachnoidal growth, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
50.1f	26.6	intraventricular tumor growth grade 2/3 glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
35.2a	28.2	extraventricular subarachnoid tumor growth grade 2/3 glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
51.1j	26.7	large extra cerebral tumor grade 2/3 glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
30.1d	34.4	glioblastoma, large, lateral grade 4	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
30.2d	31.4	tumor spread grade 2/3 glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
5.3g	43.4	svz thickening, extra cerebral subdural spread grade 2/3 glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II

50.1a	28.3	grade 4, MVP (microvascular proliferation), necrosis, no infiltration	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
5.4c	42	grade 2, ventricular tumor, bilateral at the base of the brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
Mouse ID	Age (weeks)	Brain Pathology	Spine Pathology
5.5e	39.3	grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
33.1a	37.8	grade 2, Ventricular growth, base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
51.1d	30.2	grade 2, Widespread ventricular and subarachnoidal spread wth base accentuate	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
36.1g	38.3	grade 2, Ventricular growth, base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
5.3d	46.8	grade 2, Ventricular growth, base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
33.1b	50.6	grade 2, Ventricular growth, base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
31.1j	43	grade 3, large intracerebral glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
30.1h	42.2	microneoplasia, small tumor nest at the base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
51.1a	35.2	microneoplasia, small tumor islands at the base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
51.1c	35.2	microneoplasia, intraventricular tumor, small tumor islands at the base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
5.4H	48.8	microneoplasia, small ventricular tumor nests and base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
5.4a	49	microneoplasia, small intraventricular tumor, small tumor islands at the base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
35.2b	39	grade II, ventricular tumor and islands at the base of brain.	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II

Table 3.2. Clinical and pathological details of all EGFRvIII-only mice. Although brain tumors displayed heterogeneity, spinal tumors were pathologically very homogenous.

Primary Cultures

Human glioma cells tend to grow as gliomaspheres (spheres of tumors cells) in neural stem cell media [165]. *EGFRvIII*-expressing mouse tumor cells placed into neural stem cell culture media led to the rapid production (within 1 – 2 days) of gliomaspheres: clusters of tumor cells in suspension, akin to the growth of normal neural stem cells in these culture conditions, Fig 3.19. The cells continued to proliferate (measured up to 8 passages) and the cultures required splitting approximately twice weekly. Gliomasphere cultures were produced for 8 / 8 mouse tumors.

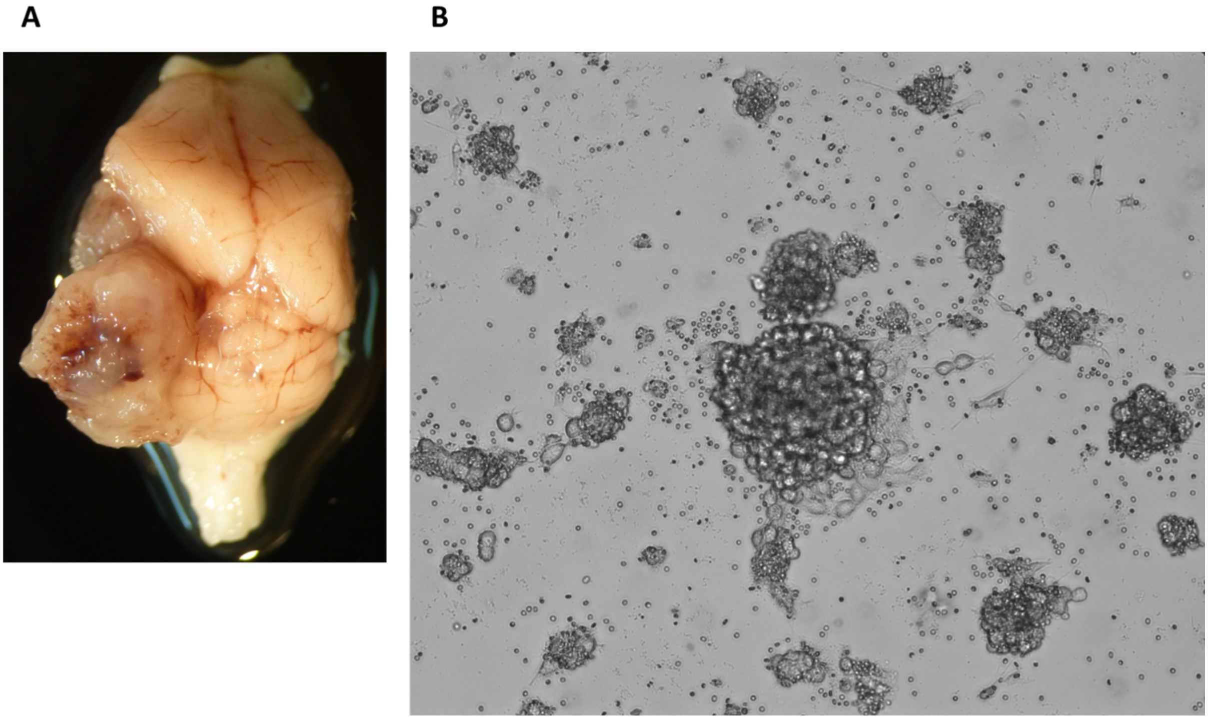


Fig 3.19. Establishing primary cultures from *EGFRvIII*-expressing mouse gliomas. A shows a large glioma on brain surface, which was cultured in neural stem cell media to yield gliomaspheres as shown in B.

Cytogenetic Analysis

Although there was some heterogeneity between the samples, in general these tumors were diploid but with extensive polyploid components. M-FISH karyotyping was performed for 3 *EGFRvIII*-only tumors. Cytogenetic analysis is shown for two representative tumor samples, Fig 3.20 and 3.21.

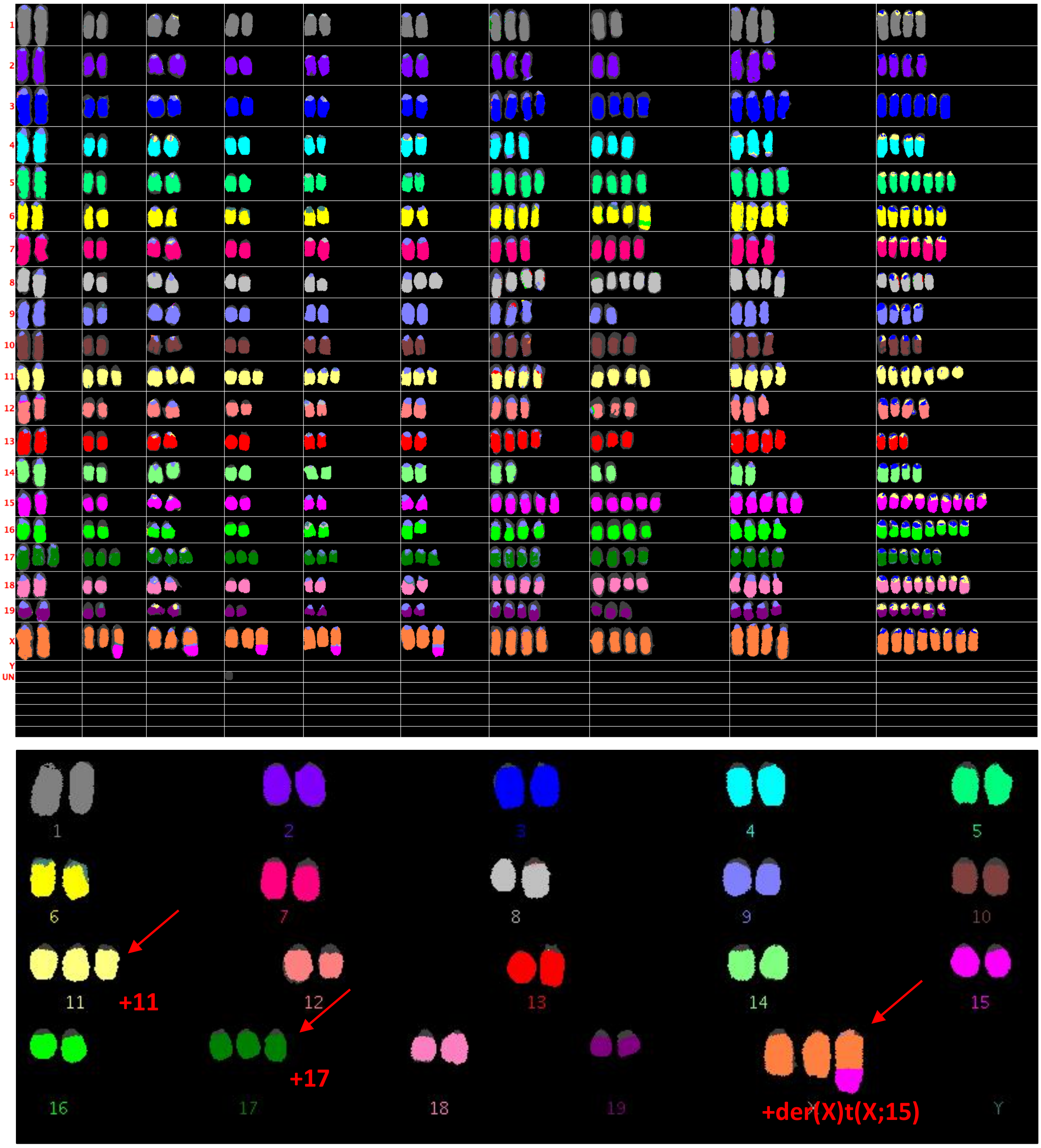


Figure 3.20. Cytogenetic analysis of an *EGFRvIII*-expressing mouse glioma. Top panel shows 10 sets of mitotic chromosomes in metaphase spreads from a single tumor; a metaphase is shown in a column and each row is a chromosome. Note the high degree of polyploidy. Bottom panel shows the dominant subclone (40% of cells had this chromosome profile), with an extra copy of chromosome 11 and 17, as well as an amplification of X and chromosome 15 (with a translocation between them)

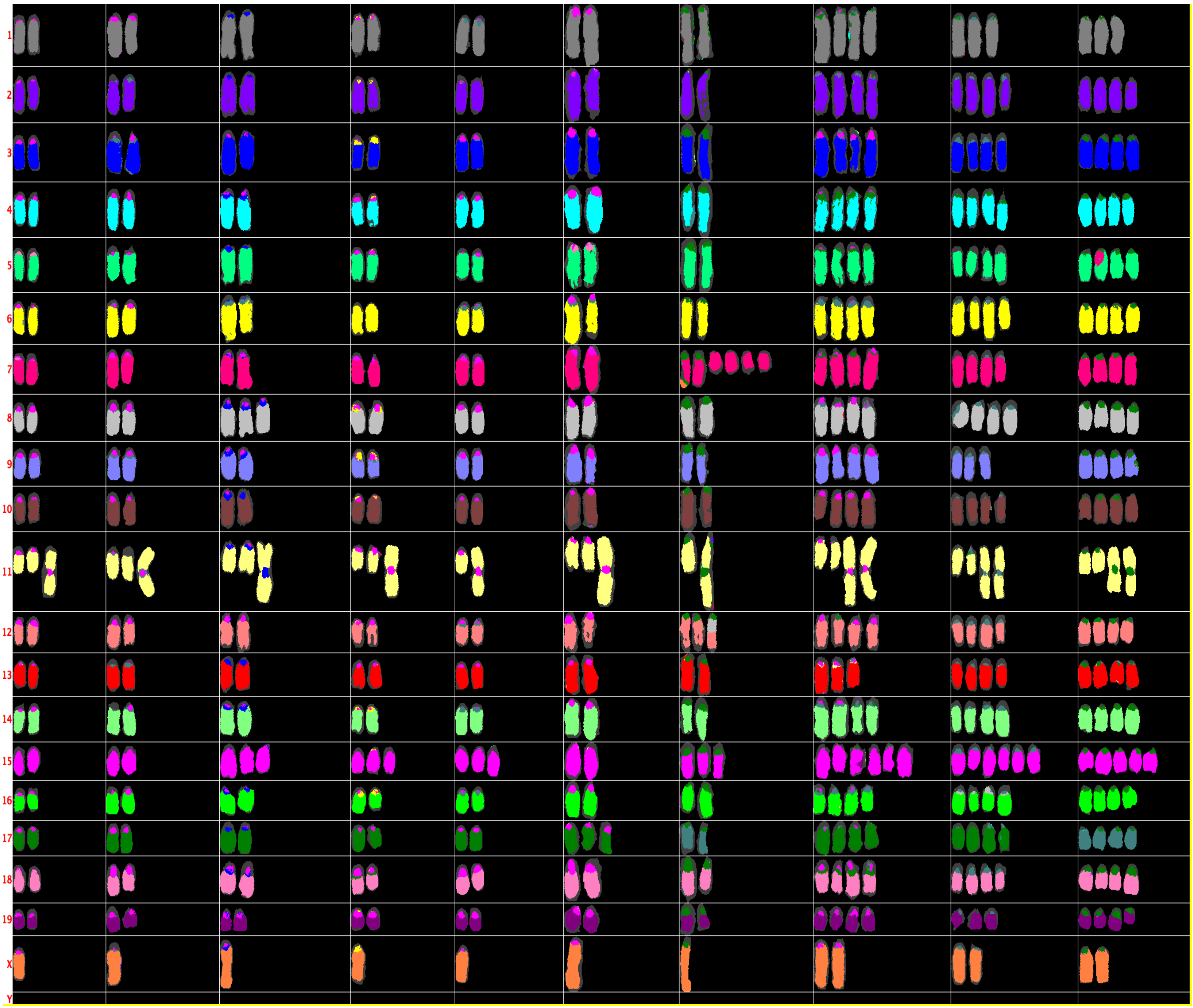

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
X
Y

Fig 3.21. Cytogenetic profile for a second *EGFRvIII* mouse glioma, highlighting chromosome 11 amplification (in this case with a Robertsonian translocation) in all cells. Note also in this case that 4 out of 7 diploids have trisomy for chromosome 15.

RNA-Sequencing

We sequenced the RNA from *EGFRvIII*-expressing mouse gliomas (primary tumors) in order to determine the pathways activated in these tumors to provide insight into the relevant tumorigenic mechanisms. Whole-transcriptomic profiling of 11 brain tumors revealed a distinctive expression pattern for *EGFRvIII*-gliomas. There were 2000 genes that were significantly upregulated in *EGFRvIII*-expressing brain gliomas compared with control (wild-type mouse brains, n=6) with a log-fold change of more than 2 and a Benjamini-Hochberg adjusted (for multiple testing) $p < 0.01$, Wald's test (Supplementary Table 2). Gene ontology analysis for pathway enrichment of the 300 most upregulated genes in these brain tumors showed there was significant enrichment for multiple pathways, particularly those related to the cell cycle and mitosis, cell differentiation, central nervous system development and neurogenesis (FDR < 0.001 in all cases), Table 3.3. Downregulated genes showed enrichment for pathways such as neuron differentiation and migration (FDR < 0.001; downregulated genes are shown in Supplementary Table 3).

The most "upregulated" gene was the *EGFRvIII* transgene but as this human transgene is not present in normal mouse tissue, fold-change is not meaningful. The endogenous mouse *Egfr* gene was also upregulated (mean $\log_2$-fold-change = 3.71) in both brain and spinal tumors, suggesting both mutant-*EGFR* and wild-type *Egfr* expression are advantageous to tumor growth (Fig 3.22; human EGFRvIII and mouse Egfr could be differentiated based on sequence differences between the species, see Materials and Methods). In brain and spinal tumors, we confirmed that the majority of the top mutated genes are also expressed, including *Sub1*, *Trp53*, *Tead2, Nlrp1b, Nt5c2*, *Prex2*, *Uimc1* and *Itga6*.

Hox (homeobox) genes have been implicated in escape from apoptosis, epithelial-mesenchymal transition, and angiogenesis in other cancers[166]. Nineteen of the 30 most strongly upregulated genes in the brain tumors compared with wild-type brains were homeobox (*Hox*) genes (Benjamini-Hochberg adjusted $p < 1 \times 10^{-12}$, Wald's test, Fig 3.23).

Gene set enrichment analysis (GSEA) is a type of analysis used to delineate sets of genes that are over-represented in a large set of genes and that therefore may be linked to a disease phenotype, in this case gliomas. GSEA of differentially expressed genes in *EGFRvIII*- brain gliomas showed significantly enriched gene sets (FDR < 0.01, Kolmogorov-Smirnov test) including the p53 pathway, Wnt and Jak-Stat pathways, Rb pathway, and stem cell-related pathways, implicating these oncogenic pathways in gliomagenesis by cooperating with EGFR signalling, Fig 3.24.

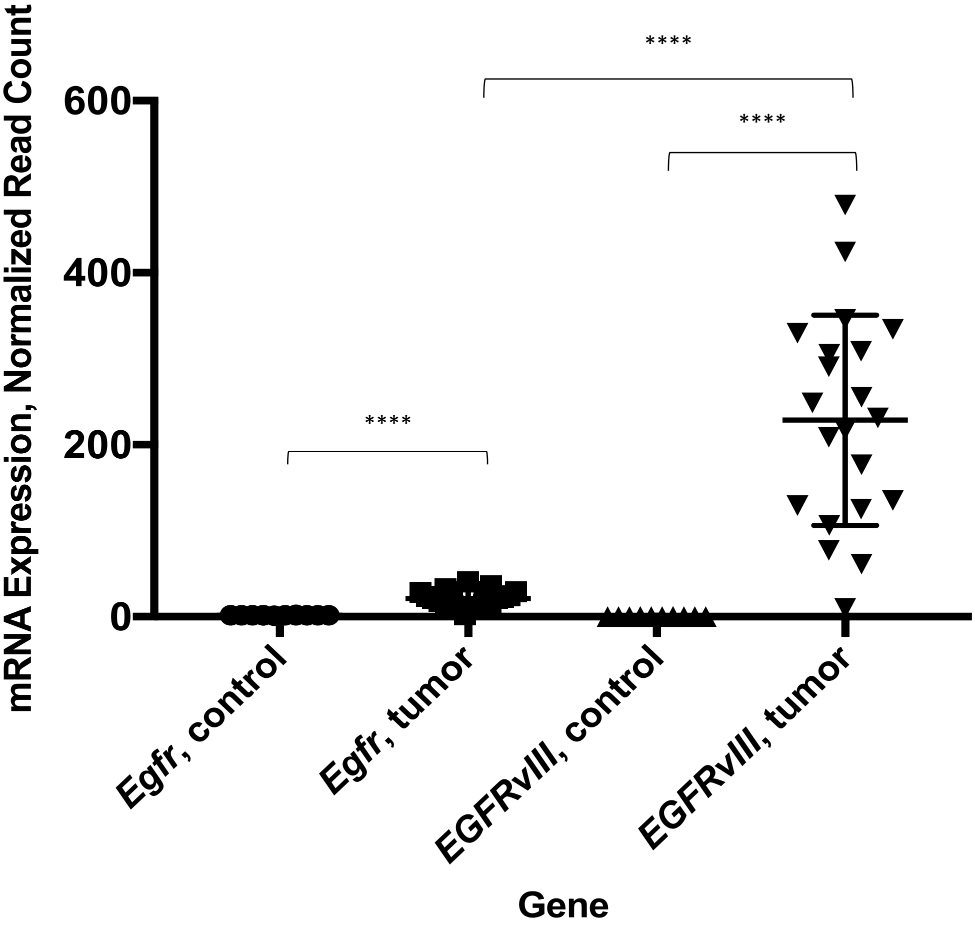


Figure 3.22. Plot showing stronger upregulation of EGFRvIII mRNA expression (from RNA-sequencing) compared with wild-type Egfr in tumors, highlighting the former is the more prominent driver (**** denotes p < 0.0001, paired t-test; n=11 brain tumors, n=10 spinal

tumors, relative to wild-type brain, n=6, and spinal cord, n=6). Mean expression and standard deviation values (error bars) are plotted.

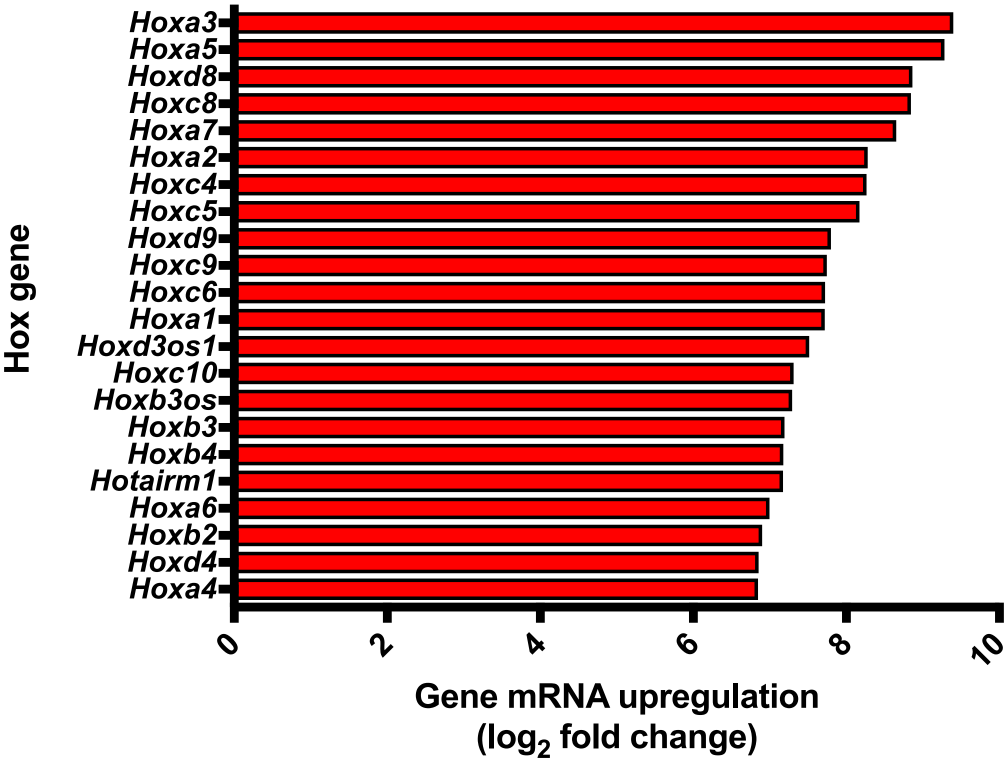


Fig 3.23. *Hox* gene upregulation in EGFRvIII gliomas arising in the mouse brain. Genes are ranked according to $\log_2$-fold change compared to wild-type mouse brain, Benjamini-Hochberg adjusted $p < 1 \times 10^{-12}$ for each gene (Wald's test).

	Biological Process (GO)		
Pathway ID	**Pathway Description**	**Count in Gene Set**	**False Discovery Rate**
GO:0048598	embryonic morphogenesis	43	2.00E-20
GO:0048513	organ development	79	2.06E-15
GO:0051301	cell division	30	5.96E-13
GO:0000278	mitotic cell cycle	31	5.32E-12
GO:0030154	cell differentiation	75	1.30E-10
GO:0007067	mitotic nuclear division	23	2.84E-10
GO:0080090	regulation of primary metabolic process	95	6.02E-10
GO:0000280	nuclear division	25	1.90E-09
GO:0048523	negative regulation of cellular process	80	1.93E-09
GO:0022402	cell cycle process	33	2.47E-09
GO:0031323	regulation of cellular metabolic process	95	3.51E-09
GO:0007049	cell cycle	38	3.66E-09
GO:0006351	transcription, DNA-templated	55	3.96E-09
GO:0048519	negative regulation of biological process	83	4.37E-09
GO:0019219	regulation of nucleobase-containing compound metabolic process	71	1.24E-08
GO:0050794	regulation of cellular process	129	1.25E-08
GO:0051171	regulation of nitrogen compound metabolic process	74	1.54E-08
GO:0060255	regulation of macromolecule metabolic process	91	1.56E-08
GO:0031326	regulation of cellular biosynthetic process	73	1.94E-08
GO:1901362	organic cyclic compound biosynthetic process	60	2.38E-08
GO:0051252	regulation of RNA metabolic process	65	9.47E-08
GO:0007059	chromosome segregation	16	1.21E-07
GO:2001141	regulation of RNA biosynthetic process	63	1.80E-07
GO:0045595	regulation of cell differentiation	42	1.82E-07
GO:0035270	endocrine system development	13	2.66E-07
GO:0009059	macromolecule biosynthetic process	63	3.06E-07
GO:0006355	regulation of transcription, DNA-templated	62	3.81E-07
GO:0048522	positive regulation of cellular process	78	1.97E-06
GO:0022008	Neurogenesis	37	3.32E-06
GO:0021546	rhombomere development	5	3.53E-06
GO:0090304	nucleic acid metabolic process	64	3.89E-06
GO:0007399	nervous system development	45	4.46E-06
GO:0031324	negative regulation of cellular metabolic process	48	7.19E-06
GO:0000122	negative regulation of transcription from RNA polymerase II promoter	26	7.84E-06
GO:0006357	regulation of transcription from RNA polymerase II promoter	41	1.14E-05
GO:0048518	positive regulation of biological process	83	1.37E-05
GO:0010605	negative regulation of macromolecule metabolic process	47	1.58E-05
GO:0048699	generation of neurons	34	1.58E-05
GO:0007417	central nervous system development	27	1.82E-05
GO:1902679	negative regulation of RNA biosynthetic process	32	1.91E-05

Table 3.3. Gene ontology (DAVID) analysis for enriched pathways in transcriptomic profile of *EGFRvIII*-induced brain tumors. Fisher's exact test with FDR multiple testing correction is used for significance testing. This list was generated using publicly available DAVID analysis software, https://david.ncifcrf.gov/.

Comparison of the RNA-sequencing profile of 11 *EGFRvIII*-induced brain gliomas with 10 spinal gliomas revealed that there were significantly differentially expressed genes between these two groups of tumors. Differentially expressed genes were ranked according to the log-fold change from one tumor type to another. With a cut-off of log-fold change of 2 and a p-value less than 0.05 (Benjamini-Hochberg corrected for multiple testing, Wald's test), there were 68 upregulated genes in the brain tumor cohort compared with spinal tumors and 228 significantly downregulated genes. Analysis of the upregulated genes in brain tumors using STRING demonstrated that there was enrichment for interactions between the proteins represented by the genes, suggesting that these genes are connected as part of a network (p-value = 2.6×10^{-7}, Hypergeometric test), Fig 3.25. Gene ontology (GO) analysis of this gene set by DAVID (see Methods) showed significant enrichment for pathways involved in brain and head development, and specifically in forebrain development, Table 3.4. GO analysis of genes upregulated in spinal compared with brain tumors demonstrated significant enrichment for GO pathways involved in sensory and motor processes, such as detection of pain and thermal stimuli, response to external stimuli and regulation of muscle contraction, Table 3.5 and Fig 3.26. Collectively, these pathways are representative of the major functions of the spinal cord and spinal nerves, reflecting the location (and likely independent origin) of spinal tumors.

RNA-sequencing of spinal tumors showed gene set enrichment for the same pathways in tumors compared with wild-type spinal cord (FDR q-value < 0.01 in all cases, Kolmogorov-Smirnov test), including for example p53, Wnt and MAPK pathways, Fig 3.28. Supplementary Table 4 shows upregulated genes in spinal tumors and Supplementary Table 5 shows downregulated genes in spinal tumors.

To determine if the transcriptional profiles of these mouse tumors have similarity to human cancers, we used GSEA comparing with all known datasets in MSigDB. Critically, the transcriptional profile of both brain and spinal *EGFR*-mutant gliomas significantly overlapped with the human Verhaak mesenchymal glioblastoma profile (FDR q-value < 0.01, Kolmogorov-Smirnov test) ([167]), Fig 3.27 and Fig 3.28. These data imply the transcriptome of these mouse gliomas resembles that of this particular subtype of human glioblastoma.

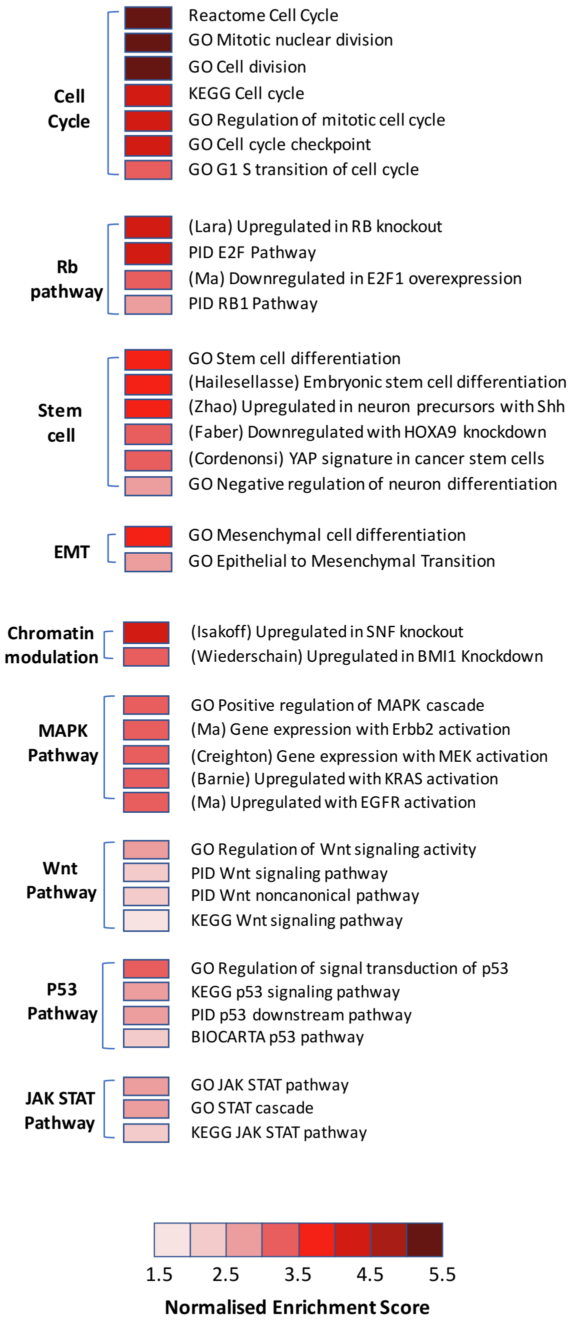


Fig 3.24. Gene set enrichment analysis reveals collaborative pathways in *EGFR*-mutant brain tumors identifies oncogenic pathways, stem-cell related genes and epithelial to mesenchymal transition (EMT) related genes. Each line identifies a transcriptomic profile with an FDR q-value < 0.01 (Kolmogorov-Smirnov test). Although not displayed here, spinal tumors enrich for the same pathways implying conserved molecular mechanisms. Gene set enrichment analysis produces an enrichment

score for a given gene set, which is then ‘normalised’ to account for the gene set size (as this can influence the enrichment score), producing a normalised enrichment score (NES), [119].

Biological Process (GO)			
Pathway ID	Pathway Description	Count in Gene Set	False Discovery Rate
GO:0060322	head development	13	4.95E-05
GO:0007417	central nervous system development	13	0.000186
GO:0007399	nervous system development	18	0.000206
GO:0035108	limb morphogenesis	7	0.00022
GO:0007420	brain development	11	0.000293
GO:0060173	limb development	7	0.000293
GO:0007389	pattern specification process	9	0.000873
GO:0030326	embryonic limb morphogenesis	6	0.000873
GO:0003002	Regionalization	8	0.0012
GO:0006366	transcription from RNA polymerase II promoter	9	0.0012
GO:0009653	anatomical structure morphogenesis	17	0.00128
GO:0030900	forebrain development	8	0.00139
GO:0060021	palate development	5	0.00141
GO:0007267	cell-cell signaling	9	0.00193
GO:0021761	limbic system development	5	0.00211
GO:0048731	system development	22	0.00211
GO:0009887	organ morphogenesis	11	0.00286
GO:0021877	forebrain neuron fate commitment	3	0.00286

Table 3.4. The gene ontology pathways most enriched in *EGFRvIII* brain gliomas compared with spinal gliomas. Analysis was performed using DAVID, with all genes upregulated in brain compared with spine tumors with $\log_2$-fold change > 2.0 and Benjamini-Hochberg adjusted p-value < 0.01 (Wald's test). Fisher's exact test with FDR multiple testing correction is used for significance testing. This list was generated using publicly available DAVID analysis software, https://david.ncifcrf.gov/.

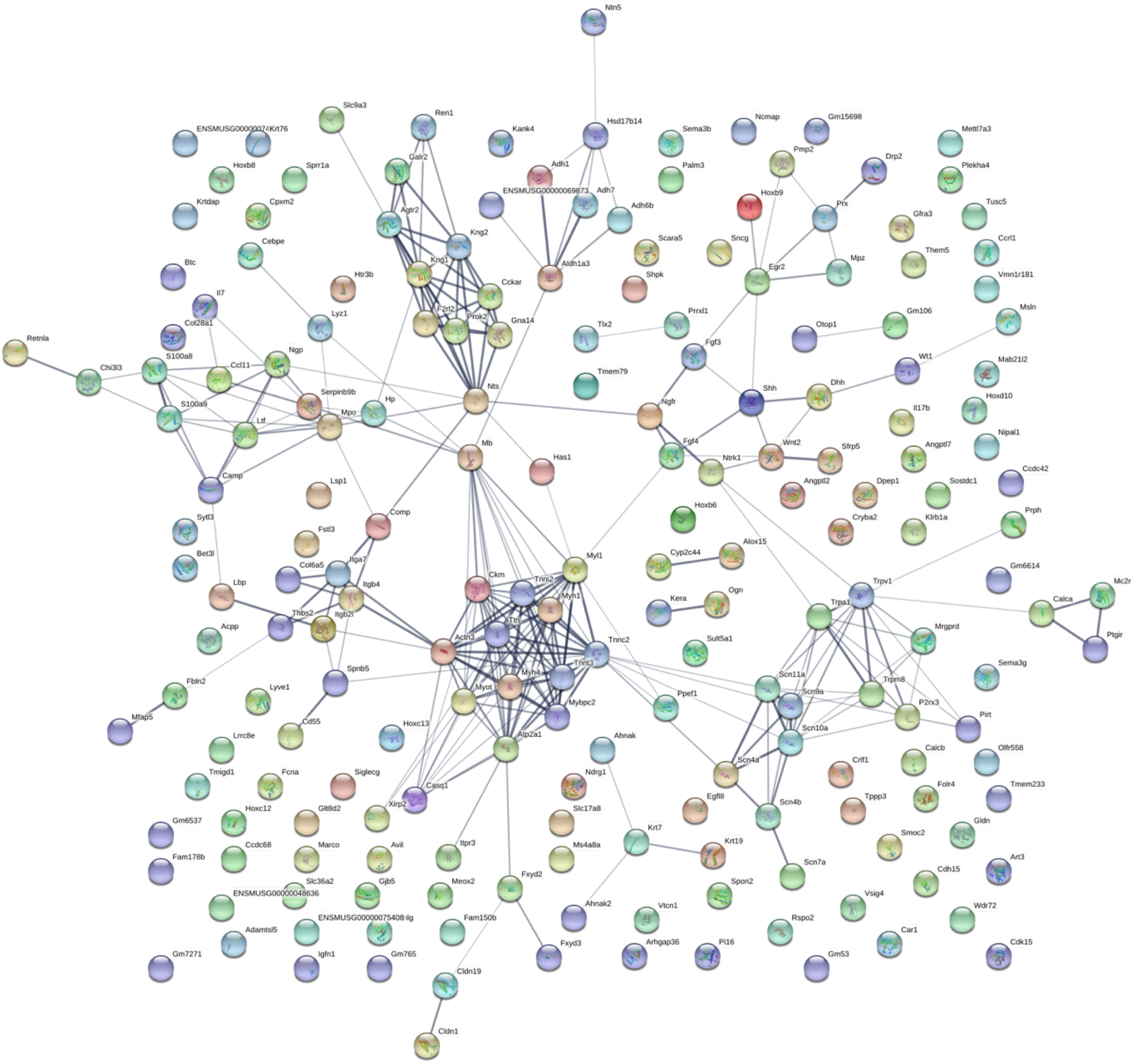

Ntn5
Slc9a3
Ren1
ENSMUSG00000074
Krt76
Kank4
Hsd17b14
Sema3b
Ncmap
Gm15698
Mettl7a3
Hoxb8
Sprr1a
Galr2
Adh1
Pmp2
Drp2
Plekha4
Krtdap
Cpxm2
ENSMUSG00000069873
Adh7
Palm3
Hoxb9
Prx
Tusc5
Agtr2
Adh6b
Gfra3
Cebpe
Kng2
Scara5
Sncg
Ccrl1
Btc
Kng1
Aldh1a3
Them5
Mpz
Htr3b
Cckar
Shpk
Egr2
Vmn1r181
Il7
Lyz1
F2rl2
Prok2
Gna14
Prrxl1
Gm106
Msln
Col28a1
Tlx2
Otop1
Retnla
S100a8
Ngp
Fgf3
Wt1
Mab21l2
Chi3l3
Ccl11
Tmem79
Shh
Dhh
Hoxd10
Serpinb9b
Nts
Hp
S100a9
Mpo
Ngfr
Il17b
Ltf
Mb
Fgf4
Nipal1
Has1
Ntrk1
Wnt2
Sfrp5
Angptl7
Camp
Lsp1
Angptl2
Dpep1
Sostdc1
Ccdc42
Sytl3
Comp
Hoxb6
Cryba2
Klrb1a
Prph
Fstl3
Alox15
Bet3l
Col6a5
Itga7
Myl1
Cyp2c44
Gm6614
Mc2r
Itgb4
Ckm
Tnni2
Trpv1
Lbp
Myh1
Ogn
Kera
Calca
Thbs2
Itgb2l
Ttn
Trpa1
Mrgprd
Acpp
Actn3
Tnnc2
Sult5a1
Ptgir
Spnb5
Myh4
Tnnt3
Scn11a
Trpm8
Sema3g
Fbln2
Myot
Lyve1
Ppef1
Scn9a
P2rx3
Mybpc2
Pirt
Mfap5
Cd55
Hoxc13
Ahnak
Scn10a
Lrrc8e
Atp2a1
Calcb
Olfr558
Tmigd1
Fcna
Siglecg
Ndrg1
Scn4a
Crlf1
Casq1
Egfl8
Tppp3
Folr4
Tmem233
Gm6537
Hoxc12
Glt8d2
Xirp2
Slc17a8
Scn4b
Gldn
Krt7
Smoc2
Fam178b
Ccdc68
Marco
Avil
Itpr3
Ms4a8a
Krt19
Art3
Scn7a
Cdh15
Slc36a2
Gjb5
Meox2
Fxyd2
Spon2
ENSMUSG00000048636
Vsig4
Ahnak2
Vtcn1
Wdr72
ENSMUSG00000075408
Adamtsl5
Fam150b
Rspo2
Car1
Arhgap36
Pi16
Cdk15
Gm7271
Igfn1
Gm765
Fxyd3
Gm53
Cldn19
Cldn1

Fig 3.25. STRING network analysis of differentially upregulated genes in spinal tumors compared with brain tumors. The figure shows significant interactions between proteins of the genes in the list (Benjamini-Hochberg adjusted $p < 0.001$, Hypergeometric test), suggesting a functional network of genes are differentially expressed in spinal gliomas. This analysis was performed using the publicly available software http://string-db.org/cgi/network.pl. Color coding: colored nodes are proteins from

CIS genes; connecting lines are known or predicted interactions between proteins; see https://string-db.org for further details.

Biological Process (GO)			
Pathway ID	Pathway Description	Count in Gene Set	False Discovery Rate
GO:0003008	system process	34	9.63E-07
GO:0003012	muscle system process	13	7.76E-06
GO:0016048	detection of temperature stimulus	6	2.15E-05
GO:0009266	response to temperature stimulus	11	2.21E-05
GO:0006936	muscle contraction	11	3.19E-05
GO:0048731	system development	59	8.22E-05
GO:0048513	organ development	49	0.000173
GO:0009408	response to heat	8	0.000529
GO:0048265	response to pain	6	0.000571
GO:0050951	sensory perception of temperature stimulus	5	0.000571
GO:0006935	Chemotaxis	15	0.000591
GO:0009653	anatomical structure morphogenesis	39	0.000591
GO:0035962	response to interleukin-13	3	0.000591
GO:0006937	regulation of muscle contraction	9	0.000628
GO:0009605	response to external stimulus	32	0.000628
GO:0090257	regulation of muscle system process	10	0.000649
GO:0019228	neuronal action potential	5	0.000977
GO:0098655	cation transmembrane transport	17	0.00106
GO:0009582	detection of abiotic stimulus	8	0.00133
GO:0009607	response to biotic stimulus	19	0.00133
GO:0019226	transmission of nerve impulse	6	0.00133
GO:0009581	detection of external stimulus	8	0.00138

Table 3.5. Gene ontology pathways upregulated in *EGFRvIII* spinal tumors compared with brain tumors. DAVID analysis for enriched pathways reveals processes related to the spinal cord are upregulated in spinal tumors, including responses to external stimuli and muscle control. This result is more consistent with spinal gliomas being primary tumors rather than metastases from the brain. Fisher's exact test with multiple test FDR correction used for significance testing (FDR < 0.05 taken as significant). This list was generated using publicly available DAVID analysis software, https://david.ncifcrf.gov/.

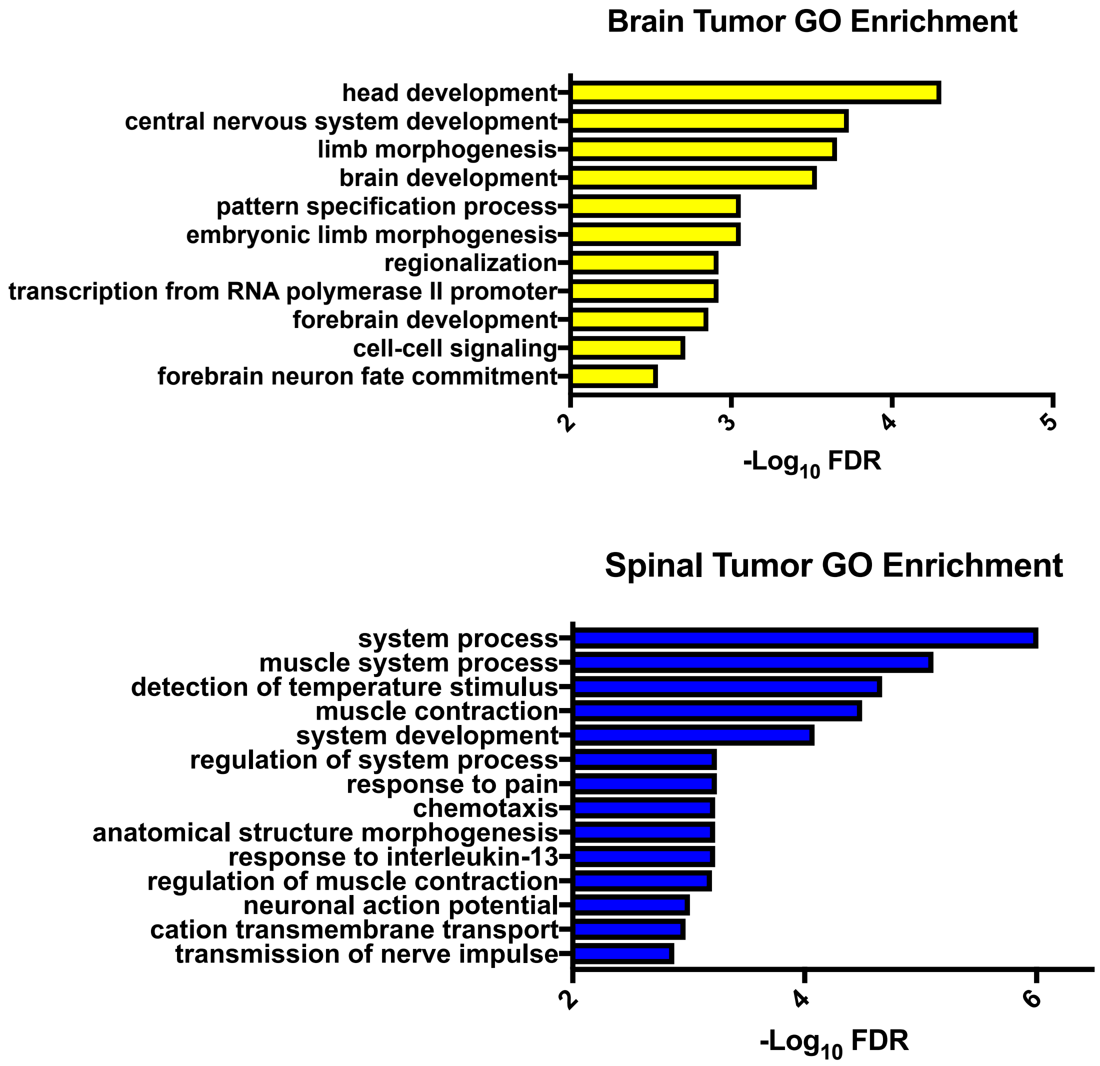
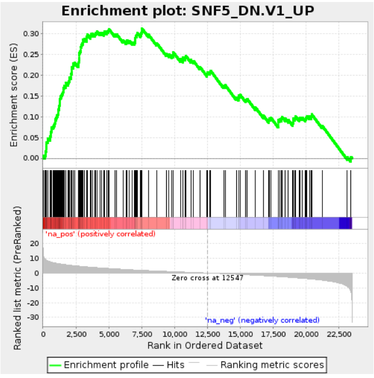
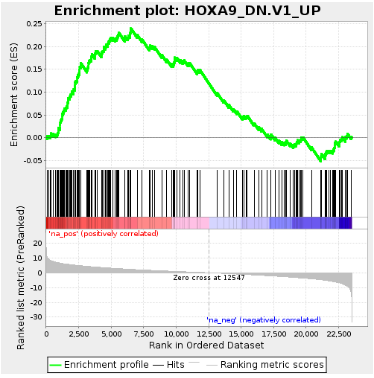


Fig 3.26. Gene ontology (DAVID) analysis of differentially expressed genes in *EGFRvIII; nes*-cre mouse brain and spinal gliomas. In brain tumors, there is significant enrichment for gene sets reflecting brain developmental processes, whereas in spinal tumors the gene sets reflect processes intrinsic to the spinal cord. These data reflect the different tissue origins of these tumors; note the absence of gene sets for oncogenic pathways here, as these are largely shared between the two types of tumor. FDR – false discovery rate. Fisher's exact test with FDR multiple testing correction is used for significance testing (FDR < 0.05 taken as significant).

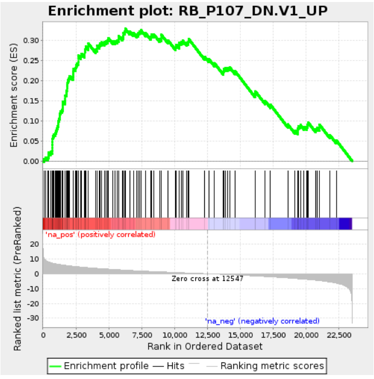
a


Enrichment plot: SNF5_DN.V1_UP
Enrichment score (ES)
Ranked list metric (PreRanked)
'na_pos' (positively correlated)
Zero cross at 12547
'na_neg' (negatively correlated)
Rank in Ordered Dataset
Enrichment profile — Hits — Ranking metric scores

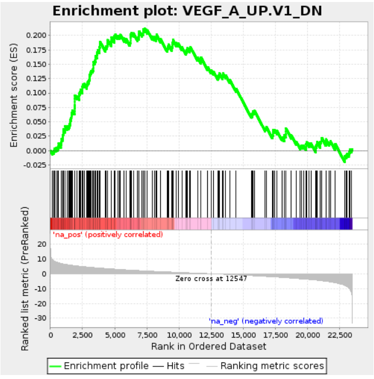
b


Enrichment plot: HOXA9_DN.V1_UP
Enrichment score (ES)
Ranked list metric (PreRanked)
'na_pos' (positively correlated)
Zero cross at 12547
'na_neg' (negatively correlated)
Rank in Ordered Dataset
Enrichment profile — Hits — Ranking metric scores

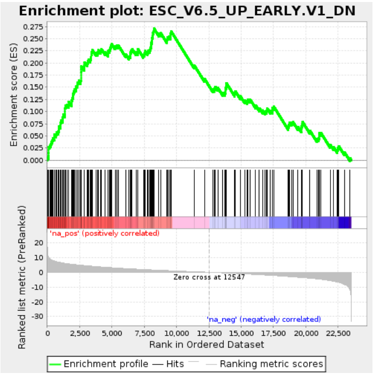
c


Enrichment plot: RB_P107_DN.V1_UP
Enrichment score (ES)
Ranked list metric (PreRanked)
'na_pos' (positively correlated)
Zero cross at 12547
'na_neg' (negatively correlated)
Rank in Ordered Dataset
Enrichment profile — Hits — Ranking metric scores

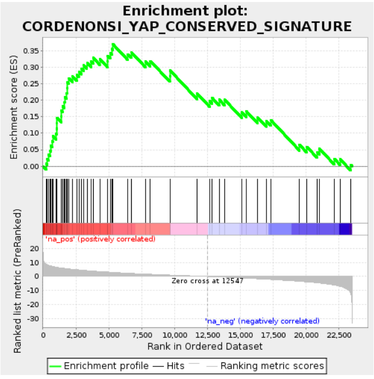
d


Enrichment plot: VEGF_A_UP.V1_DN
Enrichment score (ES)
Ranked list metric (PreRanked)
'na_pos' (positively correlated)
Zero cross at 12547
'na_neg' (negatively correlated)
Rank in Ordered Dataset
Enrichment profile — Hits — Ranking metric scores

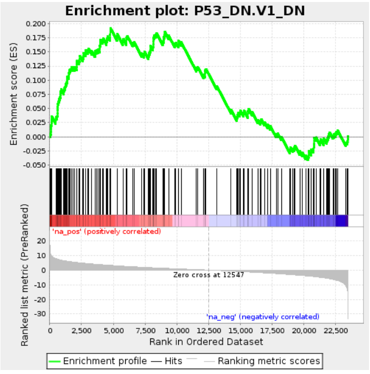
e


Enrichment plot: ESC_V6.5_UP_EARLY.V1_DN
Enrichment score (ES)
Ranked list metric (PreRanked)
'na_pos' (positively correlated)
Zero cross at 12547
'na_neg' (negatively correlated)
Rank in Ordered Dataset
Enrichment profile — Hits — Ranking metric scores

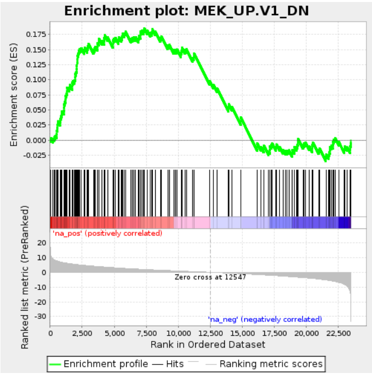
f


Enrichment plot:
CORDENONSI_YAP_CONSERVED_SIGNATURE
Enrichment score (ES)
Ranked list metric (PreRanked)
'na_pos' (positively correlated)
Zero cross at 12547
'na_neg' (negatively correlated)
Rank in Ordered Dataset
Enrichment profile — Hits — Ranking metric scores

g


Enrichment plot: P53_DN.V1_DN
Enrichment score (ES)
Ranked list metric (PreRanked)
'na_pos' (positively correlated)
Zero cross at 12547
'na_neg' (negatively correlated)
Rank in Ordered Dataset
Enrichment profile — Hits — Ranking metric scores

h


Enrichment plot: MEK_UP.V1_DN
Enrichment score (ES)
Ranked list metric (PreRanked)
'na_pos' (positively correlated)
Zero cross at 12547
'na_neg' (negatively correlated)
Rank in Ordered Dataset
Enrichment profile — Hits — Ranking metric scores

i

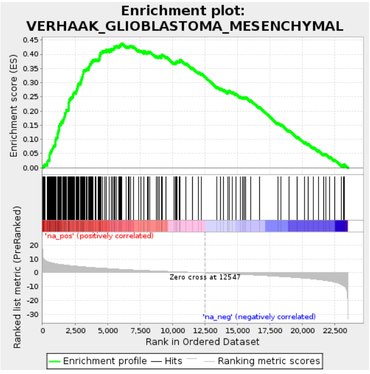

Enrichment plot:
VERHAAK_GLIOBLASTOMA_MESENCHYMAL
Enrichment score (ES)
Ranked list metric (PreRanked)
'na_pos' (positively correlated)
Zero cross at 12547
'na_neg' (negatively correlated)
Rank in Ordered Dataset
Enrichment profile — Hits — Ranking metric scores

Figure 3.27. Gene set enrichment analysis (GSEA) of *EGFRvIII; nes*-cre brain gliomas defines their key oncogenic pathways. Plots are displayed for a selection of significantly enriched gene sets, including *SNF5*, *HOXA9*, *RB*, *VEGF*, *ESC*, *YAP*, *P53* and *MEK* gene sets (FDR q-value < 0.01, Kolmogorov-Smirnov test). I. Transcriptional profile of these tumors are significantly enriched for the Verhaak human mesenchymal glioblastoma profile (FDR q-value < 0.01, Kolmogorov-Smirnov test).

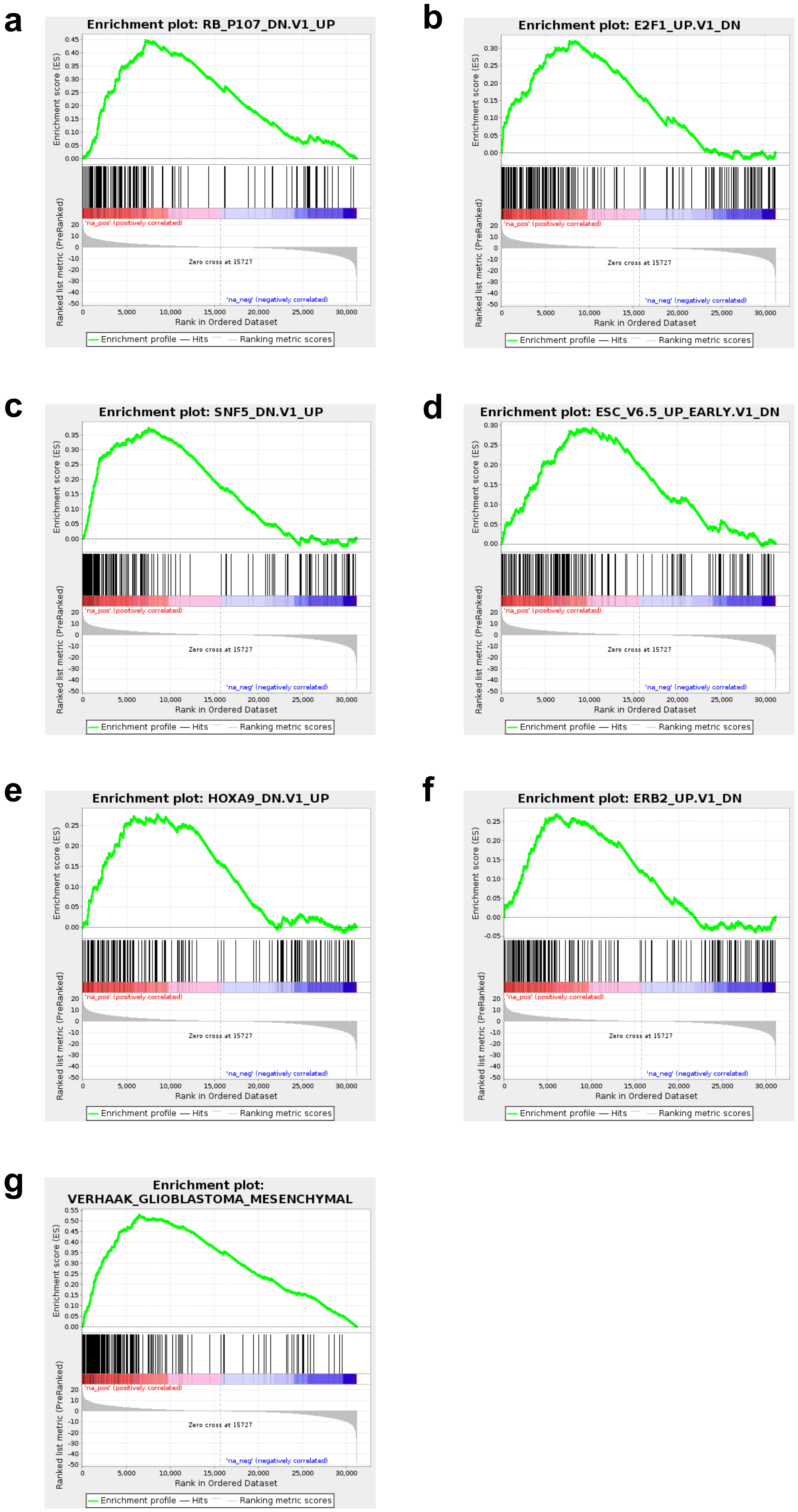

a
Enrichment plot: RB_P107_DN.V1_UP
b
Enrichment plot: E2F1_UP.V1_DN
c
Enrichment plot: SNF5_DN.V1_UP
d
Enrichment plot: ESC_V6.5_UP_EARLY.V1_DN
e
Enrichment plot: HOXA9_DN.V1_UP
f
Enrichment plot: ERB2_UP.V1_DN
g
Enrichment plot: VERHAAK_GLIOBLASTOMA_MESENCHYMAL
Enrichment score (ES)
Ranked list metric (PreRanked)
'na_pos' (positively correlated)
Zero cross at 15727
'na_neg' (negatively correlated)
Rank in Ordered Dataset
0
5,000
10,000
15,000
20,000
25,000
30,000
Enrichment profile — Hits — Ranking metric scores

Figure 3.28. Gene set enrichment analysis of *EGFRvIII; nes*-cre spinal gliomas defines their key oncogenic pathways. Plots for a selection of significantly enriched gene sets are presented here, including *Rb*, *E2F1*, *SNF5*, *ESC* (embryonic stem cell), *HOXA9*, and *ERBB2* gene sets (FDR q-value < 0.01, Kolmogorov-Smirnov test). G. Transcriptional profile of these tumors are significantly enriched for the Verhaak human mesenchymal glioblastoma profile (FDR q-value < 0.01, Kolmogorov-Smirnov test).

Whole-Exome Sequencing for Mutations and Copy Number Changes

In order to identify somatic mutations and copy number changes acquired following glioma initiation by *EGFRvIII* in our mice, we performed whole-exome sequencing (WES) on 17 tumors (9 brain and 8 spinal gliomas) with matched normal spleen controls from the same mice. Across all tumors, we found 85 significant recurrently mutated genes with mutations in two or more tumors identified by MuSiC[111] (adapted for mouse data- see Materials and Methods for details for mutation calling and determination of significance); most had single-nucleotide variants (SNVs) but some genes exhibited INDELS (insertions or deletions; Fig 3.29, Supplementary Table 1). The median number of exonic mutations per tumor was 29 of which missense mutations were the most common . Amongst the single nucleotide variants, T > C and C > T were the commonest changes. *Sub1* was the most frequently mutated gene (6 mutations in 5/17 tumors, $p=1.1 \times 10^{-16}$, FDR 2.27×10^{-12}, Likelihood ratio test, LRT) displaying INDELs and SNVs, all in splice sites suggesting loss-of-function, Fig 3.29. *Sub1* is a transcriptional regulator whose precise function is unknown, however recent work has demonstrated it is upregulated in many cancers and its knockdown reduced prostate cancer cell invasion *in vitro* and tumor growth *in vivo*, with chromatin immunoprecipitation showing that *Sub1* binds promoter regions of oncogenes *C-MYC* and *PLK1* in prostate cancer [168]. *Trp53*, a known glioma tumor suppressor[169], was the second most frequently mutated gene (5/17 tumors had a *Trp53* missense mutation, all within *Trp53*'s DNA-binding domain; $p=1.13 \times 10^{-12}$, FDR 7.75×10^{-9}, LRT), validating the application of WES to identify relevant collaborative mutations, Fig 3.30. Similarly, *Nf1,* a known genetic driver of brain glioma[170], was found to be mutated in two tumors (p=0.0010, FDR 0.17, LRT).

Other frequently mutated novel genes were *Tead2*, *Nt5c2, Ces1c, Nlrp1b*, *Prex2, Uimc1* and *Itga6*. *Tead2*, a transcription factor in the Hippo pathway, had recurrent mutations across its TEA/ATTS (DNA-binding) domain (4 mutations in 3/17 tumors; $p=2.80 \times 10^{-11}$, FDR 1.15×10^{-7}, LRT), including splice site mutations and one frameshift mutation, suggesting loss-of-function. *TEAD2* is thought to be involved in tumor suppression via interaction with the YAP oncoprotein in the Hippo signalling pathway, restricting cell proliferation and promoting apoptosis [171, 172]. *Uimc1* and *Itga6* had three mutations each ($p=1.39 \times 10^{-7}$ and FDR 1.9 x

10^{-4}, $p=2.7 \times 10^{-7}$ and FDR 3.2×10^{-4}, LRT, respectively) all of which were INDELS and one of which caused a frameshift in *Itga6* (Fig 3.29).

In contrast to the relatively small number of recurrent mutations, *EGFR*vIII tumors had complex genomes by DNA copy number analysis (Fig 3.40). Significant focal amplifications and deletions, identified by GISTIC2[115], were evident in regions with known cancer genes, for example significant focal *Cdkn2a* deletions (GISTIC q-value = 1.39×10^{-5}) were evident and *EGFRvIII* (in *Col1a1* locus, GISTIC q-value = 0.017) was recurrently amplified. Significantly recurrent focal deletions were present in novel putative glioma drivers *Nlrp1b* and *Adgrl2* (GISTIC q-value = 2.92×10^{-14} and 2.19×10^{-6} respectively, Fig 3.31). Several of the most significantly mutated genes were also in regions with frequent deletions, including *Trp53*, *Tead2* and *Uimc1*, supporting putative tumor suppressive roles (Fig 3.31).

The potential significance and translational relevance of the most frequently mutated and / or focally deleted genes detected in the mouse gliomas were assessed by comparison with human low-grade glioma (LGG) data from 283 patients from The Cancer Genome Atlas (TCGA)[173], using the online tool Cbioportal (see Materials and Methods for further details). This revealed that *TEAD2* is recurrently deleted in 48% of human LGGs in a mutually exclusive manner with *TP53* (Bonferroni-adjusted $p < 0.001$, Fisher's exact test, Fig 3.32), identifying *TEAD2* as a putative cancer gene. Recurrent deletions in previously unknown glioma genes *NT5C2*, *ADGRL2* and *UIMC1* were observed whilst *SUB1, CES1*, *NLRP1* and *ITGA6* were frequently methylated in human LGGs (Supplementary Fig S8). These data help cross-validate the relevance of these novel putative cancer genes in human patients.

These mouse gliomas were all wild-type for *Idh1* (17/17 tumors sequenced), consistent with gliomas in humans in which *IDH1* and *EGFR* mutations tend to be mutually exclusive [174].

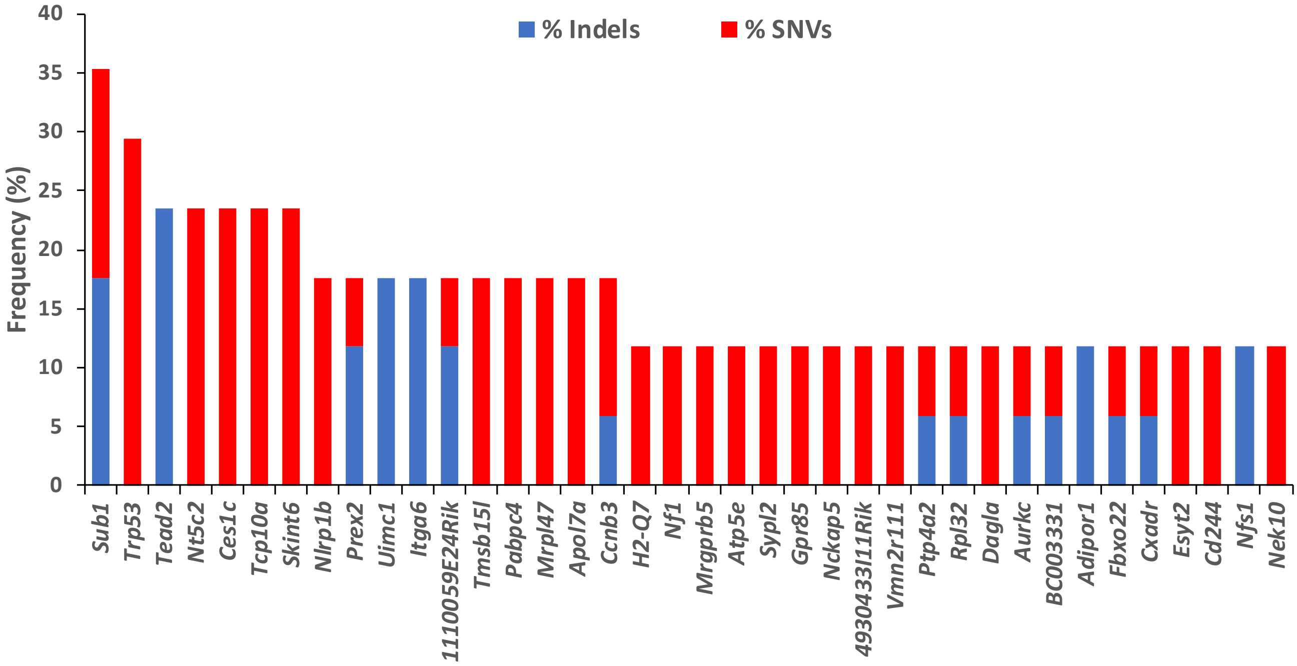


Fig 3.29. Mutational profile of 17 brain and spinal tumors. Genes are ranked according to the frequency of mutations (indels or SNVs). Known glioma drivers include *Trp53* and *Nf1*, and novel ones found mutated are *Sub1* and *Tead2*.

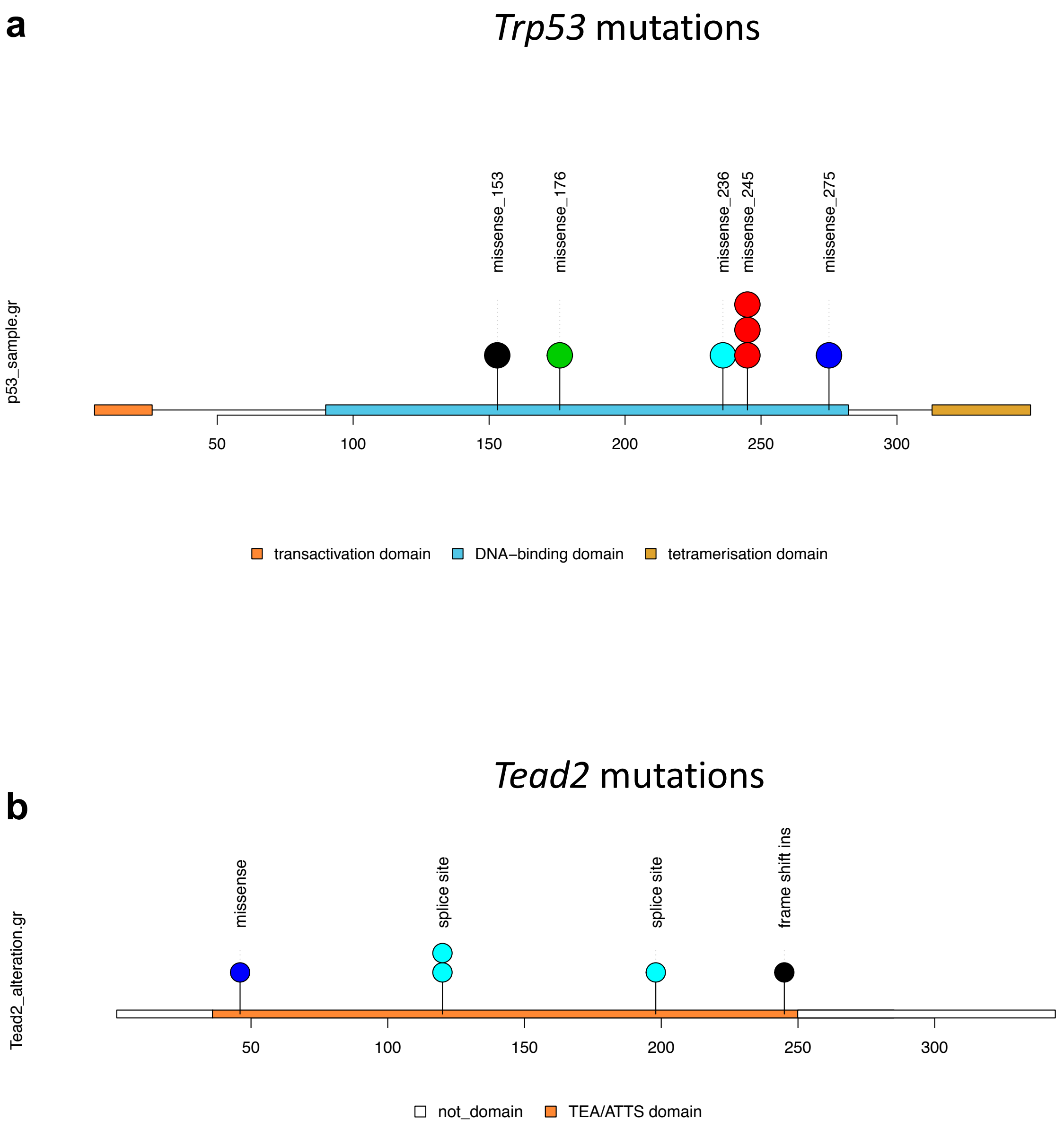


Fig 3.30. Mutations of *Trp53* and *Tead2* in *EGFRvIII*-only and *EGFRvIII*-PB gliomas are in DNA-binding domains. A. Plot outlining the location of *Trp53* mutations across all exome-sequenced mouse tumors. Five *EGFRvIII*-only and two *EGFRvIII*-PB tumors had in *Trp53*, all residing within its DNA-binding domain; 3 occurred in the same location. B. Plot outlining locations of *Tead2* mutations, all residing in the TEA/ATTS domain which is the DNA-binding domain of *Tead2*; 2 mutations were in same splice site location.

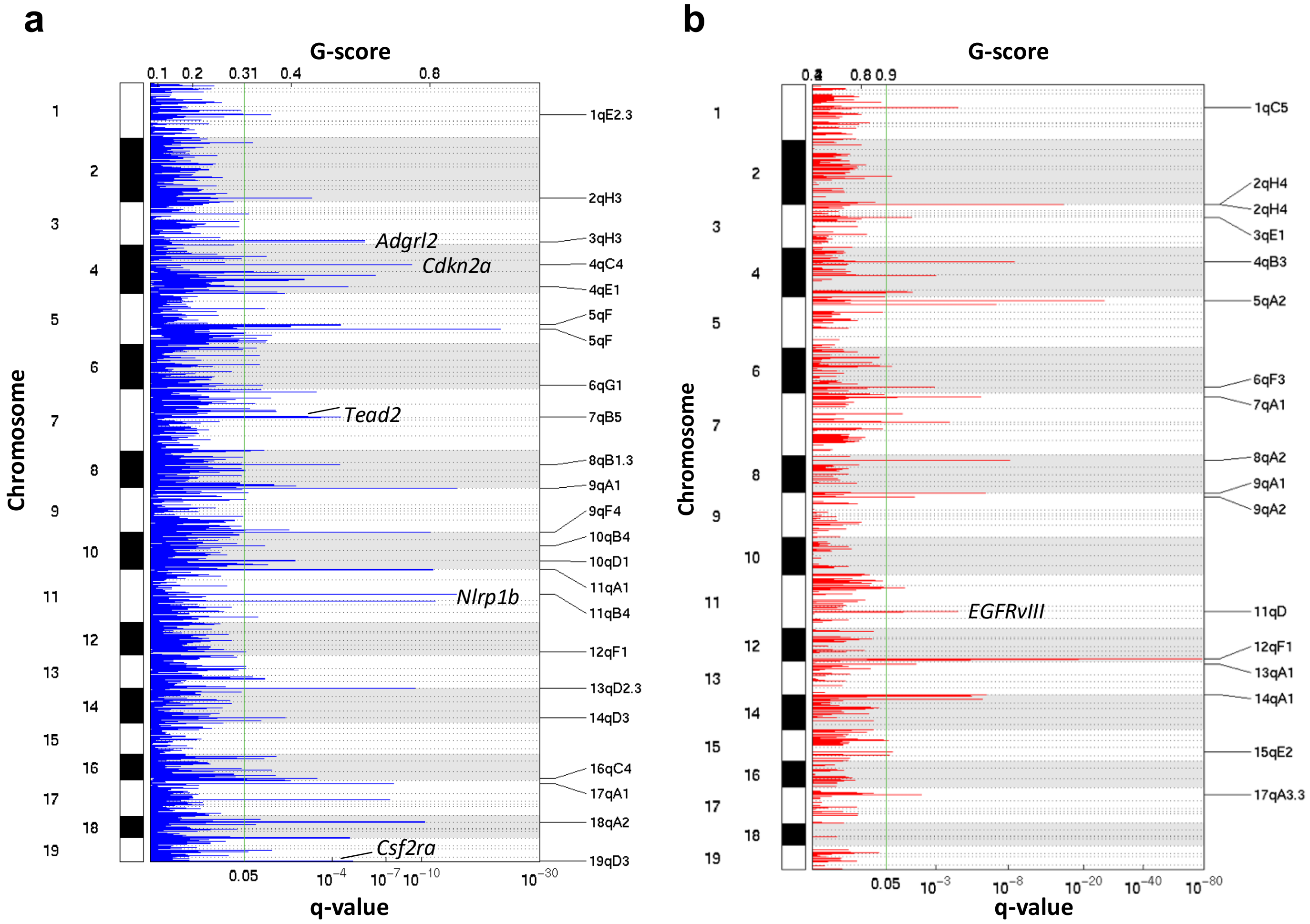


Figure 3.31. Plots showing focal copy number variations across *EGFRvIII*-only and *EGFRvIII*-PB mice. Significant focal deletions as determined by GISTIC2 are displayed in A, and significant focal amplications are displayed in B. Lower x-axis represents q-value (significance at < 0.05) and top x-axis represents the G-score. G-score and p-values are calculated as described in [115].

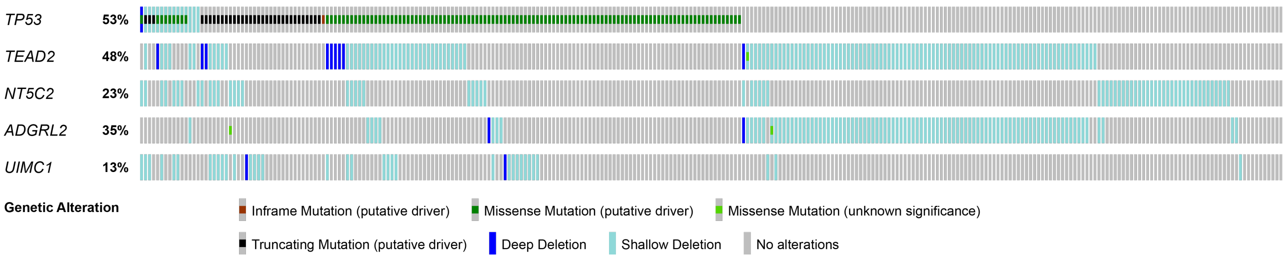


Figure 3.32. Analysis of TCGA data on human LGGs for genetic alterations in the genes most significantly mutated in mouse *EGFRvIII* gliomas. These genes are recurrently mutated in *EGFRvIII*-gliomas in mice, and are seen here to be altered with high frequency in patients with LGG; these alterations in *TP53* and *TEAD2* are mutually exclusive in this patient cohort (Bonferroni-adjusted $p < 0.001$, Fisher's exact test), as are alterations in *NT5C2* and *TP53* (Bonferroni-adjusted $p = 0.003$, Fisher's exact test). These data were analysed using the publicly available software Cbioportal, http://www.cbioportal.org/. Each column represents one tumor.

Transposon Mutagenesis Replaces Genomic Instability in Glioma Progression

Transposons have been successfully used for identifying cancer driver genes[54, 55, 57, 58, 60, 61, 67, 70, 175]. Mobilized *PiggyBac* transposons randomly integrate in the genome and activate and/or disrupt gene expression[69] (see Chapter Four for further discussion of the background to transposon mutagenesis). We performed a conditional *PiggyBac* transposon mutagenesis screen *in vivo* to further identify genes that cooperate with mutant-*EGFR* in gliomagenesis.

As was the case for *EGFRvIII*, To limit transposition to the central nervous system a conditional *PiggyBac* transposase allele was activated by *nes-cre*. An experimental cohort of quadruple transgenic mice carrying conditional *EGFRvIII*, 20 copies of a *PiggyBac* transposon (ATP1S2)[69], a conditional *PiggyBac* transposase and *nes-cre* were generated (*EGFRvIII*-PB, n=47; Fig 3b, see Methods). As controls, we established transgenic mice expressing *EGFRvIII* but lacking transposition (*EGFRvIII*; *nes-cre* = EGFRvIII-only, n=31; and *EGFRvIII*; *nes-cre*; ATP1S2, n=10) and a set with transposition but lacking *EGFRvIII* (transposase; ATP1S2; *nes-cre* = PB-only, n=20). Mean survival times between *EGFRvIII*-PB and *EGFRvIII*-only cohorts were similar (36.2 vs 38.1 weeks, *p*=0.95, log-rank test, Fig 3.33) and both groups had similar incidences and pathological grades of brain and spinal gliomas (Fig 3.34 and 3.35, Table 3.6). After one year, 10 PB-only mice without any neurological signs were culled to determine if transposon mobilisation in the CNS was sufficient to induce tumors or tumor precursor lesions - histological analysis did not reveal any tumors in the brain or spine of these 10 mice.

Survival Times of Mice

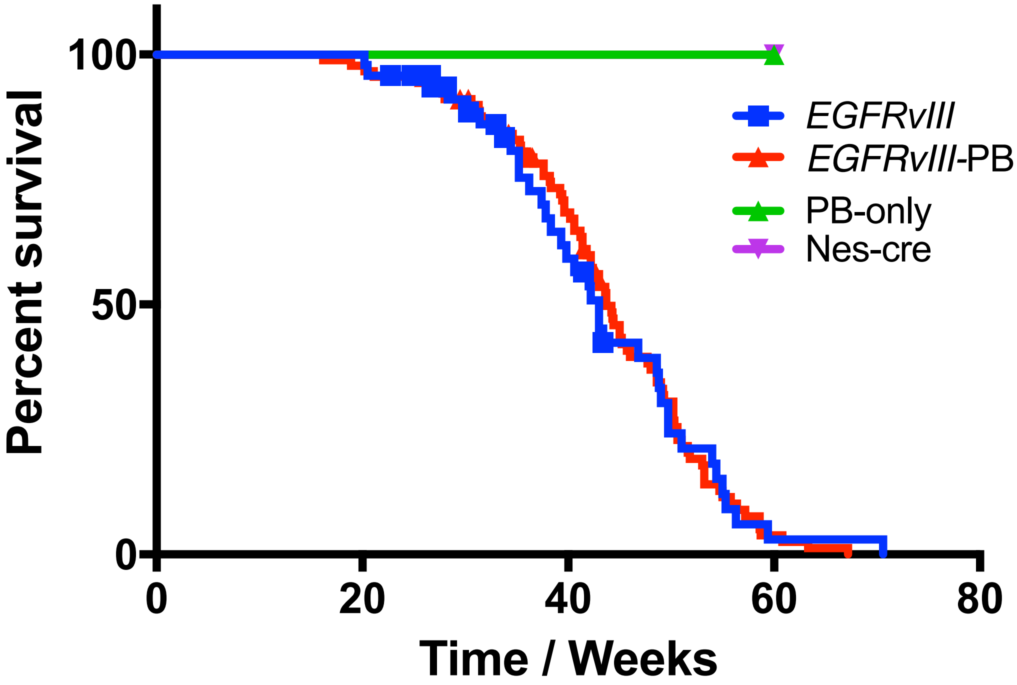


Fig 3.33. Kaplan-Meier curves for survival times of *EGFRvIII*-mice with and without *piggyBac* transposition. Kaplan-Meier survival curves of EGFRvIII-only (n=31) and EGFR-PB (n=47) mice, with no significant difference between them (p = 0.95, log-rank test). No differences in survival or 129 pathology were observed between EGFRvIII; nes-cre and EGFRvIII; nes-cre; ATP1S2 mice. Tumors were not observed in PB-only (TSPB; ATP1S2; nes-cre, n=20) or nes-cre (n=10) mice after 60 weeks.

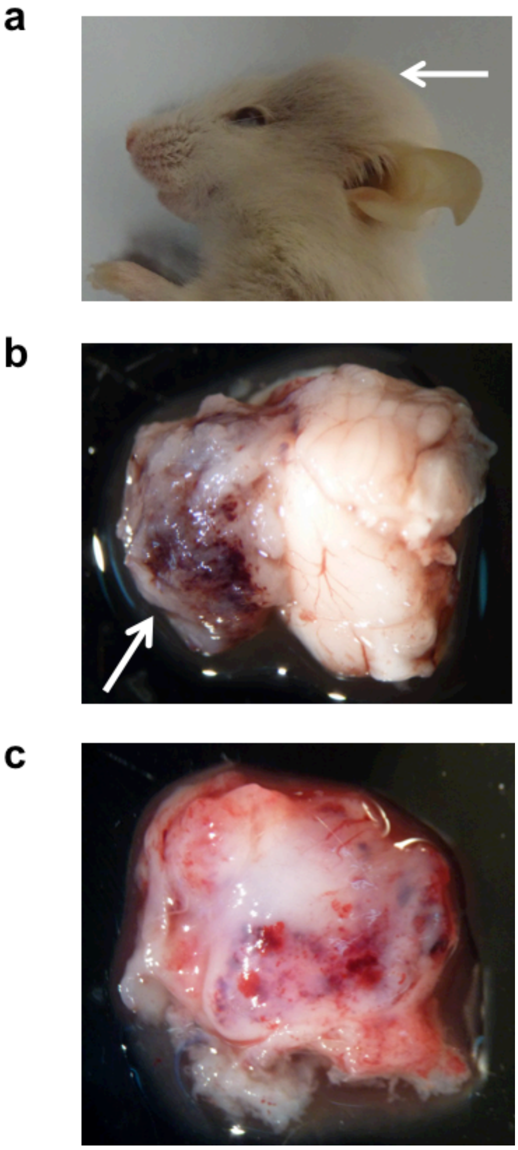


Fig 3.34. Clinical Phenotype of *EGFRvIII*-PB mice. A. Photograph of an *EGFRvIII*-PB mouse with enlarged head due to an underlying brain glioma. B, C. Macroscopic photographs of areas of the brain from the same mouse showing the presence of a tumor on the brain surface.

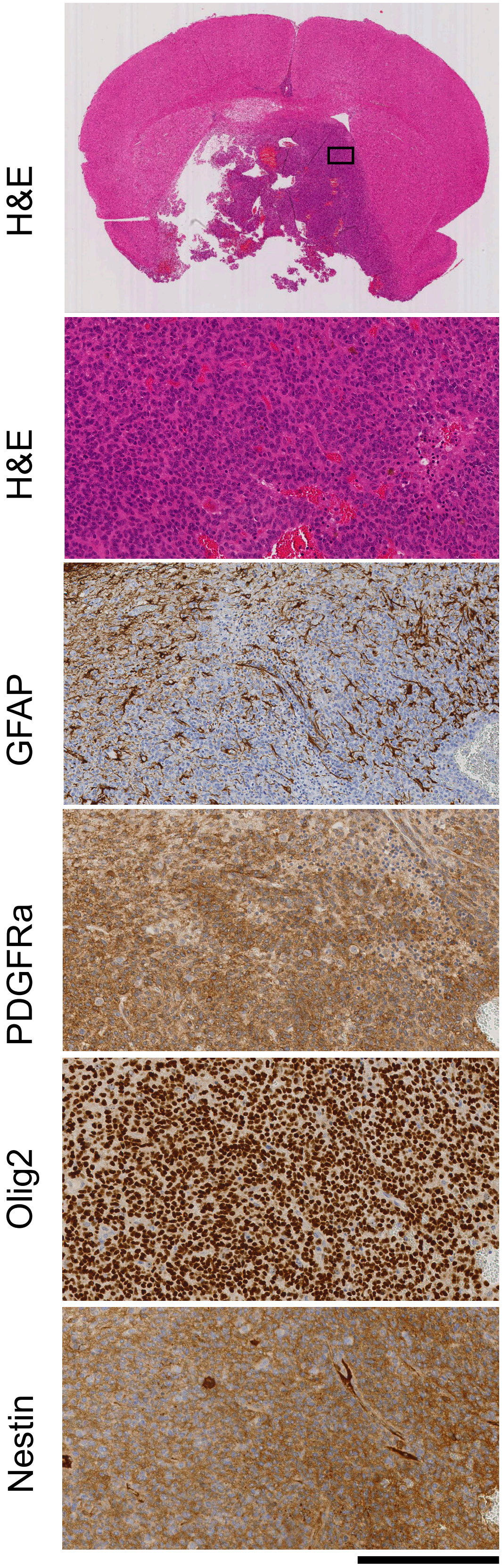

H&E
H&E
GFAP
PDGFRa
Olig2
Nestin

Fig 3.35. H&E and immunostaining profile of a typical grade III brain glioma from an *EGFRvIII*-PB mouse, which demonstrates strong expression of neural stem and transit-amplifying cell markers. The scale bar at the bottom corresponds to 2.8mm for the top panel, and 200μm for all panels below.

Mouse ID	Age (weeks)	Clinical Phenotype	Brain Pathology	Spine Pathology
26.2b	29.5	head greatly enlarged; difficulty breathing; minimal mobilisation	SVZ expansion, widespread leptomeningeal tumor, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
5.2e	30	moderate -severe macrocephaly, pilorection, laboured breathing, hunched, lethargic	subarachnoid spread, large basal tumor, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
36.1b	30.3	moderate macrocephaly, severe ataxia, weak hindlimbs	SVZ growth, possible subarachnoid spread, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
21.2a	30.6	moderate macrocephaly, very lethargic, hyperventilating	widespread subarachnoid tumor growth (grade II glioma), thickened SVZ	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
24.1a	35.5	moderate - severe macrocephaly, lethargy, difficulty breathing; tilted	tumor cell clusters in ventricle, subarachnoid spread, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
22.1b	38.3	acute hindlimb paralysis; mild macrocephaly, head tilt.	small tumor nests in subarachnoid space, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
8.4b	36.5	moderate macrocephaly, rapid breathing (for several weeks), very lethargic / hunched.	expansion of svz, small subarachnoid tumor clusters/ spread, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
5.3c	41.8	moderate macrocephaly, rapid / deep breathing acutely, hunched / immobile.	no SVZ changes, no tumor seen	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
5.5b	36	mild macrocephaly, hyperventilating, very weak hindlimbs (dragging); very big bladder (likely spinal)	intraventricular, subarachnoid tumor growth grade 2/3	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
25.2b	34.2	severe macrocephaly, rapid breathing, slow	intraventricular tumor cell clusters, subarachnoid spread, grade 2/3	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
25.2d	34.2	severe macrocephaly, rapid breathing, all limbs weak (?spinal parenchyma invasion)	tumor spread grade 2/3	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
35.1e	35.3	severe macrocephaly, rapid breathing, severe tilt; hunched	large tumor spread grade 4	widespread subarachnoid/ leptomeningeal spread, minimal

				infiltration into spinal cord but strong infiltration into nerve roots; grade II
38.1e	34.6	moderate body tilt, culled early for tail injury	intraventricular, subarachnoid spread, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
25.3a	31.6	severe macrocephaly, hyperventilating, lethargy	grade 4; basis of brain, frontal, subarachnoidal	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
21.1j	42.2	acute paralysis, urinary retention	grade 2 glioma, no surface attachment	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
5.3b	45.7	moderate macrocephaly, weight loss, lethargy	grade 2 glioma, Ventricular growth, base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
8.5c	37.6	moderate / severe macrocephaly, lethargy	grade 2 glioma, Ventricular growth, base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
49.1b	31.3	moderate macrocephaly, lethargy, hyperventilating	grade 2 glioma, Tumor in lateral ventrical, base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
21.1h	45	acute paralysis (complete) of hindlimbs, prior abnormal gait	grade 2 glioma, small tumor nest in base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
53.1e	29.6	hunched, hyperventilation, severe macrocephaly	grade 2 glioma, small tumor nest in base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
21.2f	40.2	severe ataxia (head bob), walking low	grade 2 glioma, Ventricular growth, base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
22.2b	30	moderate macrocephaly, walking low / weak hindlimbs, big bladder/retention	grade 2 glioma, Ventricular growth, base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
48.1a	37.6	severe macrocephaly, hyperventilating, lethargy	grade 2 glioma, Ventricular growth, base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
35.1D	42.4	Mild hydro, very weak hindlimbs, urine retention	grade 2 glioma, intraventricular tumor and small tumor at the base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II

30.1a	42.2	moderate macrocephaly (hydro), lethargic, weak limbs	grade 2 glioma, intraventricular, small tumor nest at the base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
30.3a	36	severe macrocephaly, slow, hyperventilation	grade 3 possibly 4, large circumscribed extra cerebral tumor	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
48.1b	39.6	moderate body tilt, lethargy	grade 2 glioma, intraventricular tumor, small tumor islands at the base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
49.2e	33.3	moderate body tilt, lethargy	grade 2 glioma, small intraventricular tumor	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
48.1i	39.6	moderate body tilt, hydrocephalus, lethargy, low walk, head bob	grade 2 glioma, very small tumor nests at the base of the brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
21.1b	49.3	moderate body tilt, uncoordinated; no hydrocephalus	grade 4 glioma, large circumscribed extra cerebral tumor	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
32.1d	43.7	severe macrocephaly, lethargy, body tilt, hyperventilation	grade 4, large extra cerebral tumor, glioma at the base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
13.2c	40	hydrocephalus, limb paralysis	grade 2 glioma, small intraventricular tumor, small tumor islands at the base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
21.1a	50.6	hydrocephalus, paralysis, hyperventilation	grade 4 glioma, large extracerebral tumor, tumor islands at the base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
21.2g	43.7	hydrocephalus, inactive / strong body tilt	grade 4 glioma, large tumor at the base of brain with MVP (microvascular proliferation)	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
53.2b	31.4	hydrocephalus, hunched, piloerection, uncoordinated movements	no tumor	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
22.2j	44.3	moderate/severe hydrocephalus, hyperventilation, lethargy	grade 2 glioma, multiple intarventricular tumor islands, widespread tumor cell subseeding	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
13.3a	49.2	moderate hydrocehapalus, uncoordinated / slow	grade 2 glioma, small intraventricular tumor,	widespread subarachnoid/ leptomeningeal spread, minimal

			small tumor islands at the base of brain	infiltration into spinal cord but strong infiltration into nerve roots; grade II
35.1b	45	moderate hydrocephalus, very weak hindlimbs (walking low)	grade 2 glioma, small intraventricular tumor, small tumor islands at the base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
20.1j	30	head tilt, hydrocephalus, lethargy	hydrocephalus, no tumor	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
25.3b	40.6	head tilt, hydrocephalus, lethargy	ventricular tumor and tumor islands at the base of brain, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
20.3d	41.4	moderate hydrocephalus, head tilt, lethargy	grade 2 glioma, small tumor islands at the base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
48.1g	23.9	head tilt, hydrocephalus, lethargy	grade 2 glioma, ventricular tumor and tumor islands at the base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
21.1k	41	hydrocephalus, abnormal gait, lethargy	grade 2 glioma, small tumor islands at the base of brain	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
22.2l	46	hydrocephalus, paralysis of hindlimbs	ventricular tumor and tumor islands at the base of brain, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
25.3f	42.7	severe hydrocephalus, body tilt, lethargy	ventricular tumor and tumor islands at the base of brain, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
51.2h	35.4	moderate hydrocephalus, hyperventilation, slow	ventricular tumor and tumor islands at the base of brain, grade II glioma	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II
35.1c	43	walking low, urogenital staining	no tumor	widespread subarachnoid/ leptomeningeal spread, minimal infiltration into spinal cord but strong infiltration into nerve roots; grade II

Table 3.6. Clinical and pathological details of all *EGFRvIII*-PB mice. As with *EGFRvIII*-only mice, brain tumors showed heterogeneity in pathology, but spinal tumors were homogenous.

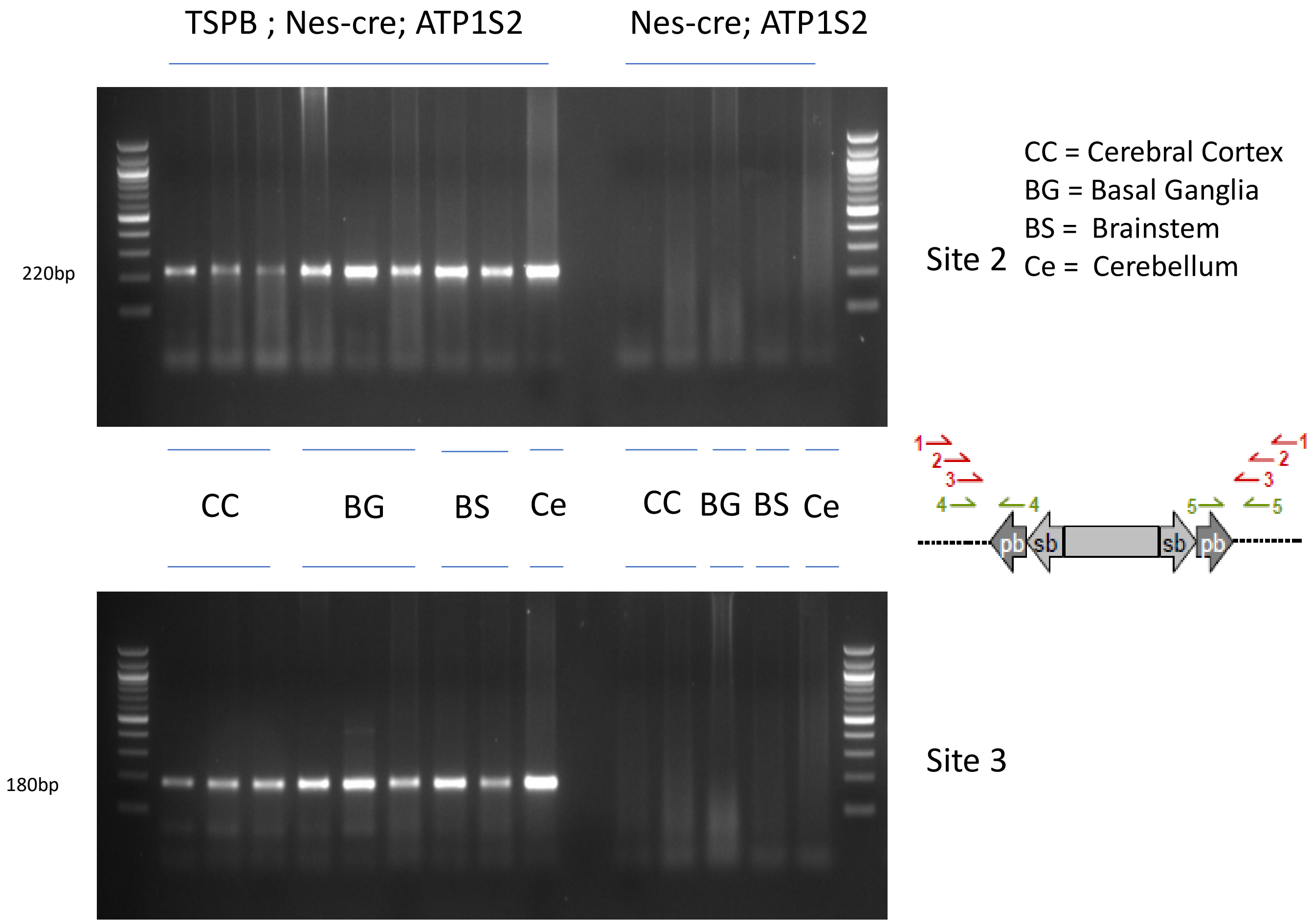

TSPB ; Nes-cre; ATP1S2
Nes-cre; ATP1S2
220bp
Site 2
CC = Cerebral Cortex
BG = Basal Ganglia
BS = Brainstem
Ce = Cerebellum
CC
BG
BS
Ce
CC
BG
BS
Ce
1
2
3
4
4
5
5
3
2
1
pb
sb
sb
pb
180bp
Site 3

Fig 3.36. PCR to identify mobilisation of the *piggyBac* transposon in the brain. Two independent PCR experiments were performed (sites 2 and 3 as shown here), and only mice with the transposase allele showed jumping of the transposon.

To identify mobilisation of the *piggyBac* transposon in the brain of the relevant mice, PCRs were performed as described in the Methods section. Fig 3.36 demonstrates that in mice carrying both the transposon and transposase alleles (in addition to nestin-cre), specific bands were produced that signify transposon mobilisation: a 220bp band for PCR site 2 and 180bp for PCR site 3. These bands were present for all sites of the brain tested in these mice, including the cerebral cortex, basal ganglia, brainstem, and cerebellum. In contrast, mice lacking the transposase allele did not produce these bands, indicating that the transposon was not able to mobilise in the brain tissue of these mice.

DNA from brain tissue of the same mice in the previous experiment was also tested for non-mobilisation of the *piggyBac* transposon through a PCR as described in the Methods section. The results shown in Fig 3.37 are for PCR non-mobilisation. For mice carrying the transposon (and nestin-cre); the specific 423bp band was always present, indicating transposon non-mobilisation. This also occurred even for mice that also carry the transposase allele, indicating that mobilisation of the transposon is never complete and there is always an element of non-mobilisation. Multiple brain sites were tested (cerebral cortex, basal ganglia, brainstem and cerebellum), and the non-mobilisation band was present for all sites.

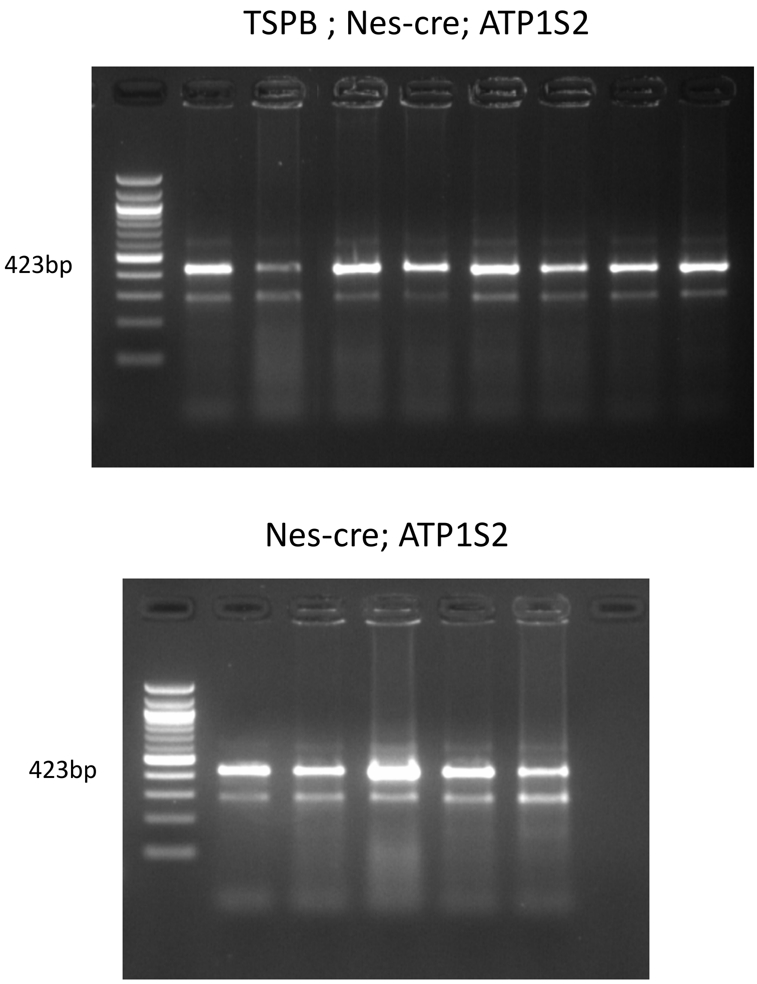

TSPB ; Nes-cre; ATP1S2
423bp
Nes-cre; ATP1S2
423bp

Fig 3.37. PCR for detecting non-mobilisation of the *piggyBac* transposon. This experiment was performed on mice with and without the transposase (mouse genotypes are displayed above) – both sets of mice had an element of non-mobilisation, implying that not all transposons jump even in the presence of the transposase.

Most previous successful transposon mutagenesis forward genetic screens have found that mice with transposition tend to survive for significantly less time than those lacking transposition. We hypothesized that the absence of reduction in survival times of the *EGFRvIII*-PB mice may reflect genetic instability in the *EGFRvIII*-only mice, perhaps through oncogene-induced replicative stress[176], that is similar in consequence to the transposon mutagenesis in *EGFRvIII*-PB mice. To test this, FISH cytogenetic analysis was conducted; primary cultures were established from these tumors and cells were used from these for FISH (SKY). Chromosomal aberrations, typically appearing in subsets of metaphases rather than all metaphases, were counted as described in the Methods for 3 *EGFRvIII*-only tumors and 6 *EGFRvIII*-PB tumors. This analysis revealed a significantly higher frequency of chromosomal aberrations in *EGFRvIII*-only tumors compared to *EGFRvIII*-PB tumors (19 vs 6.4 mean number of chromosomal aberrations, $p = 0.013$, unpaired two-tailed t-test; Fig 3.38, 3.39).

Whole-exome sequencing of 20 brain and spinal gliomas from *EGFRvIII*-PB mice confirmed these had substantially less complex tumor-genomes with fewer copy number changes than *EGFRvIII*-only tumors (Fig. 3.40). Nevertheless, whole chromosome 11 amplification was still common as well as focal amplifications of *EGFRvIII* (*Col1a1* locus) and localized deletions in *Cdkn2a, Nlrp1b* and *Adgrl2* in tumors arising from both cohorts. GISTIC2 analysis shows these alterations occur significantly more frequently than expected by chance (q-value < 0.05; Fig 3.31), suggesting they provide a selective advantage for tumor progression. The wild-type *EGFR* gene was not amplified however.

Whole-exome sequencing analysis revealed that while the median number of mutations was similar between the cohorts, their mutational profiles differed substantially. The top 5 mutated genes identified in the *EGFRvIII*-PB tumors were *Obscn*, *Hspg2*, *Rrbp1*, *Rpgrip1* and *Atp5o* which have unknown functions in cancer (Fig 3.41). Although, the frequency of mutations in these genes was high (70-40%), *Obscn* and *Hspg2* are particularly large genes (so more likely to harbor mutations) and contained many synonymous changes, suggesting they were passengers. Nevertheless, in the *EGFRvIII*-PB cohort there were low-frequency mutations in a subset of the putative drivers we previously identified in *EGFRvIII*-only tumors, including frequent splice site mutations in *Sub1* and *Nt5c2*, and mutations in *Nlrp1b, Trp53,* *Tead2, Uimc1* and *Itga6* (Fig 3.42). Of note, three tumors across both cohorts had a mutation and focal deletion in *Nlrp1b,* suggesting selection for complete loss of *Nlrp1b* which may occur via multiple mechanisms.

Together, these results suggest that *PiggyBac* mutagenesis substitutes for genetic instability thus reducing copy number variation in *EGFRvIII*-PB tumors and highlights the relevance of the transposon-mediated genetic alterations for glioma progression. Mapping and analysis of *piggyBac* insertion sites in these tumors are detailed in the next Chapter.

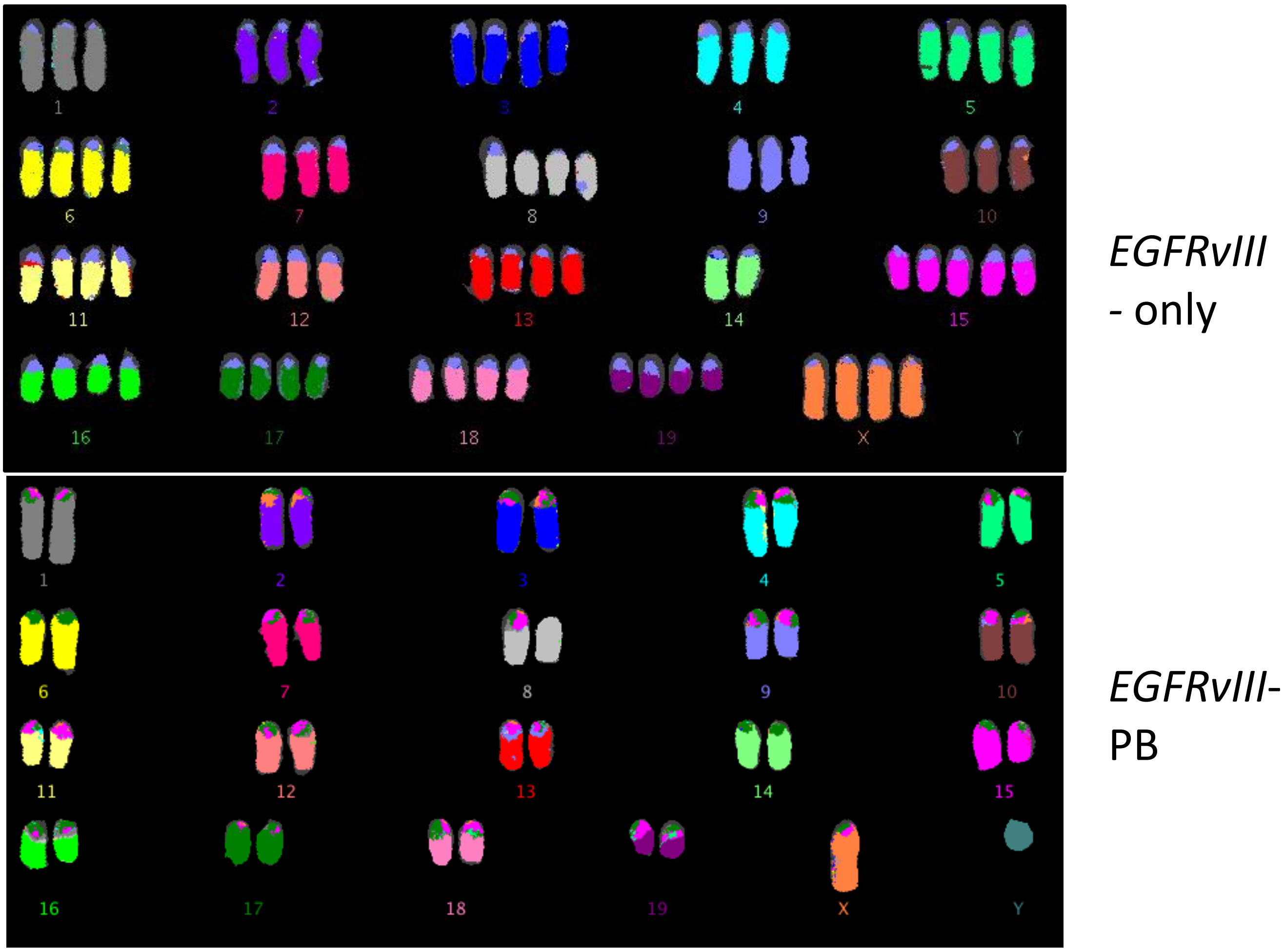

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
X
Y
EGFRvIII - only
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
X
Y
EGFRvIII-PB

Fig 3.38. Representative karyotype of *EGFRvIII*-only and *EGFRvIII*-PB brain tumors, showing polyploidy in the non-PB tumor.

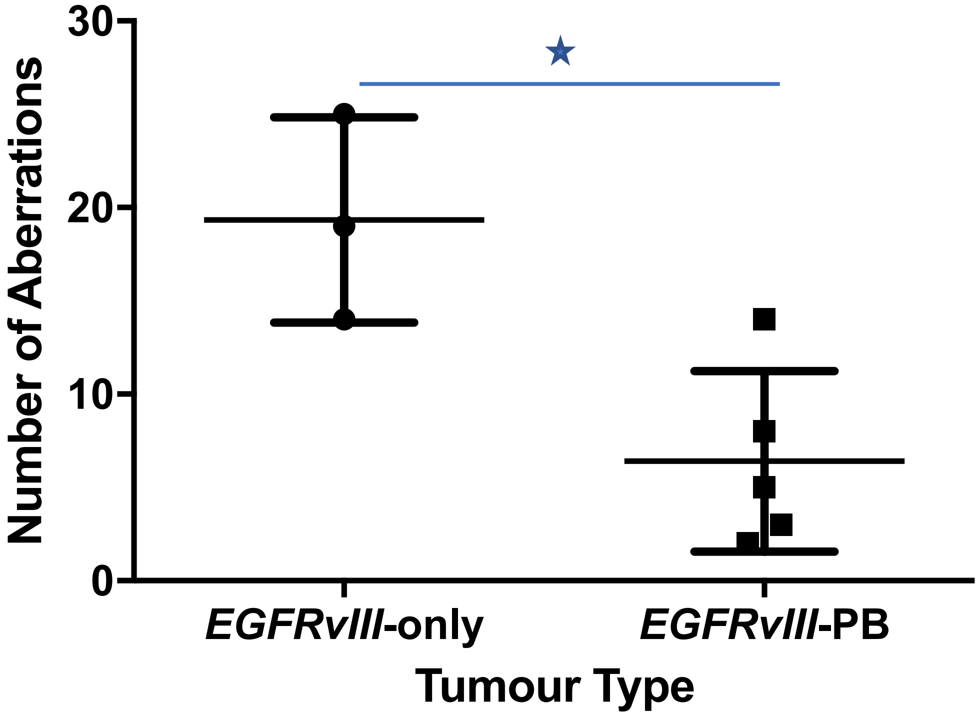
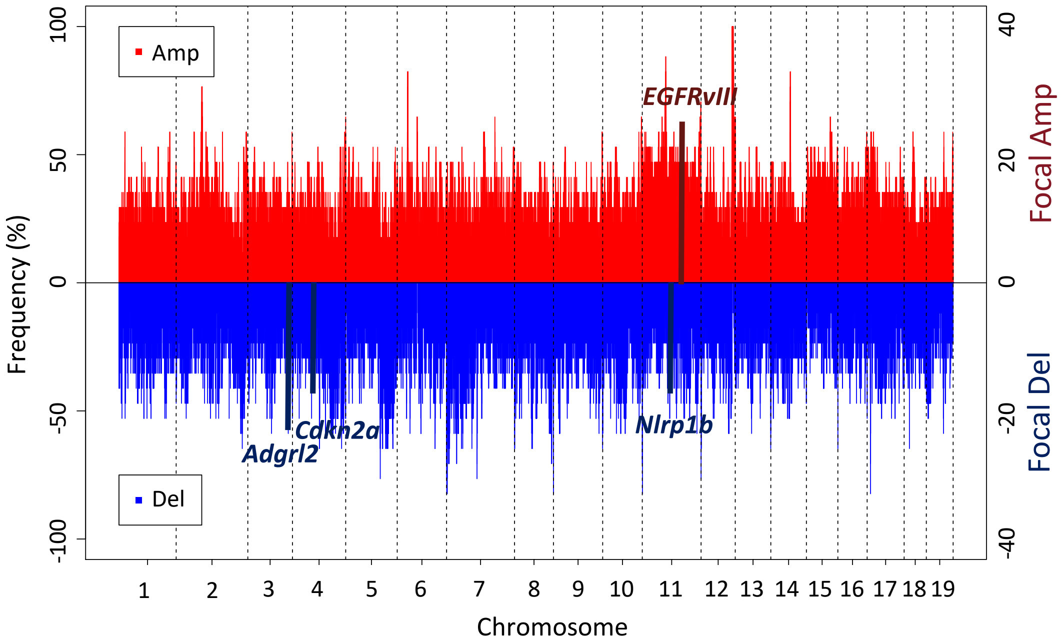
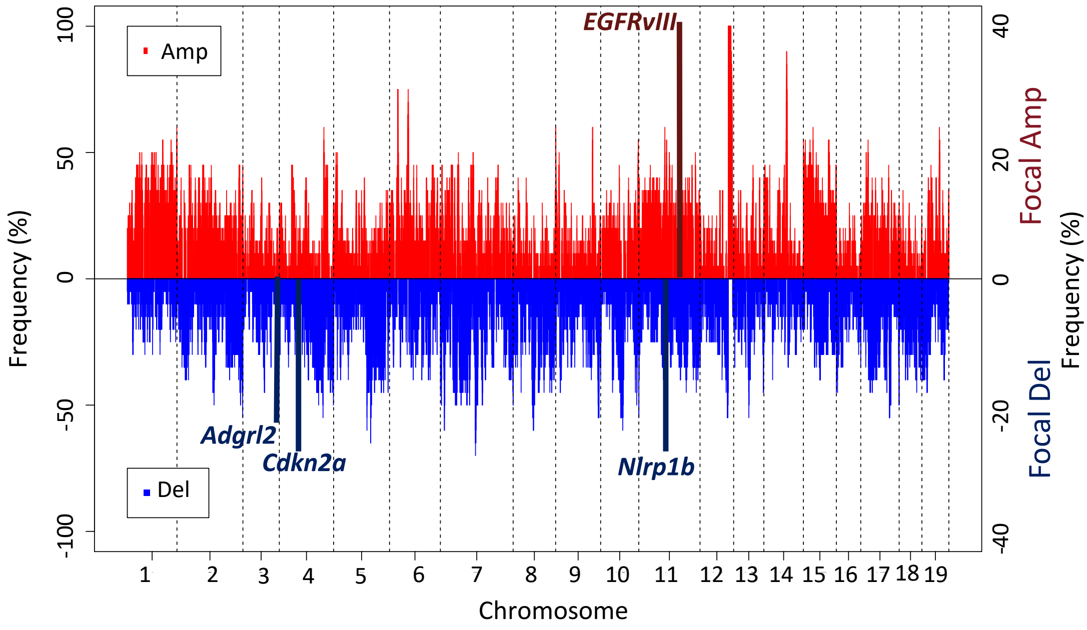


Fig 3.39. Chromosomal aberrations in *EGFRvIII*-only and *EGFRvIII*-PB tumors (n=3 and n=5 tumors respectively; mean chromosomal aberrations 19 vs 6.4, p = 0.013, unpaired two-tailed t-test; plots show mean +/- standard deviation). * denotes significance at level p < 0.05.



***EGFRvIII*-only**



***EGFRvIII*-PB**

Fig 3.40. Copy number analysis from whole-exome sequencing data of mouse gliomas. Top panel if from *EGFRvIII*-only mice and demonstrates there are more copy number alterations, particularly gains, in this cohort compared with *EGFRvIII*-PB mice in the bottom panel. Each row represents an independent tumor and each column is a chromosome. Red lines are copy number gains (amplifications) and blue lines are copy deletions.

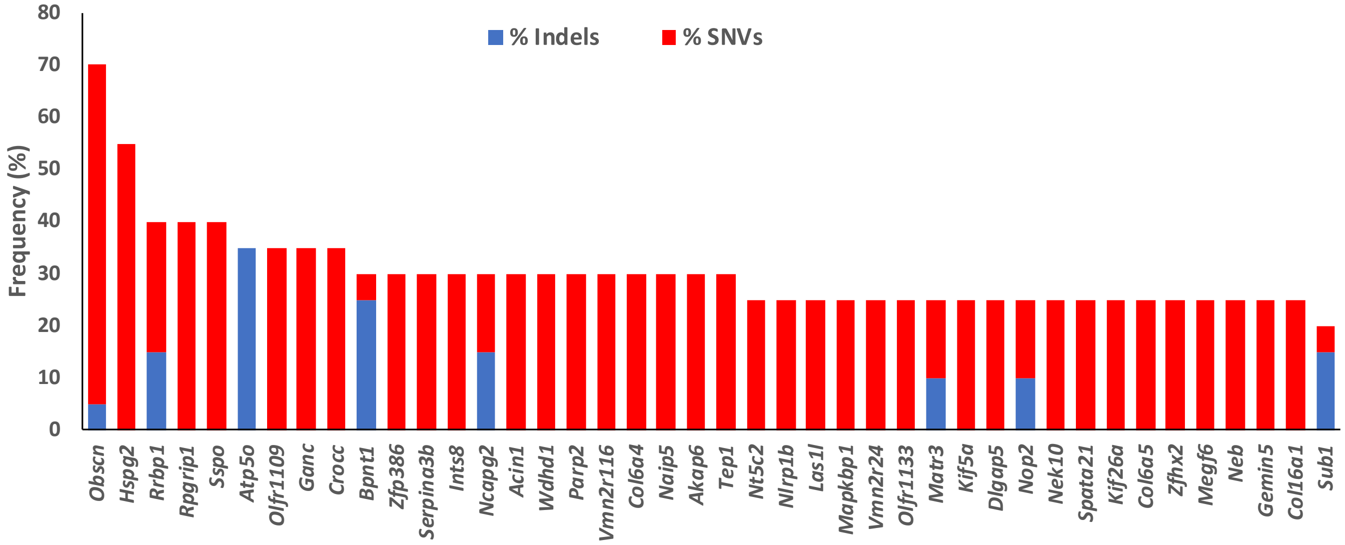


Fig 3.41. Mutational profile of 20 *EGFRvIII*-PB brain and spinal tumors from whole-exome sequencing.

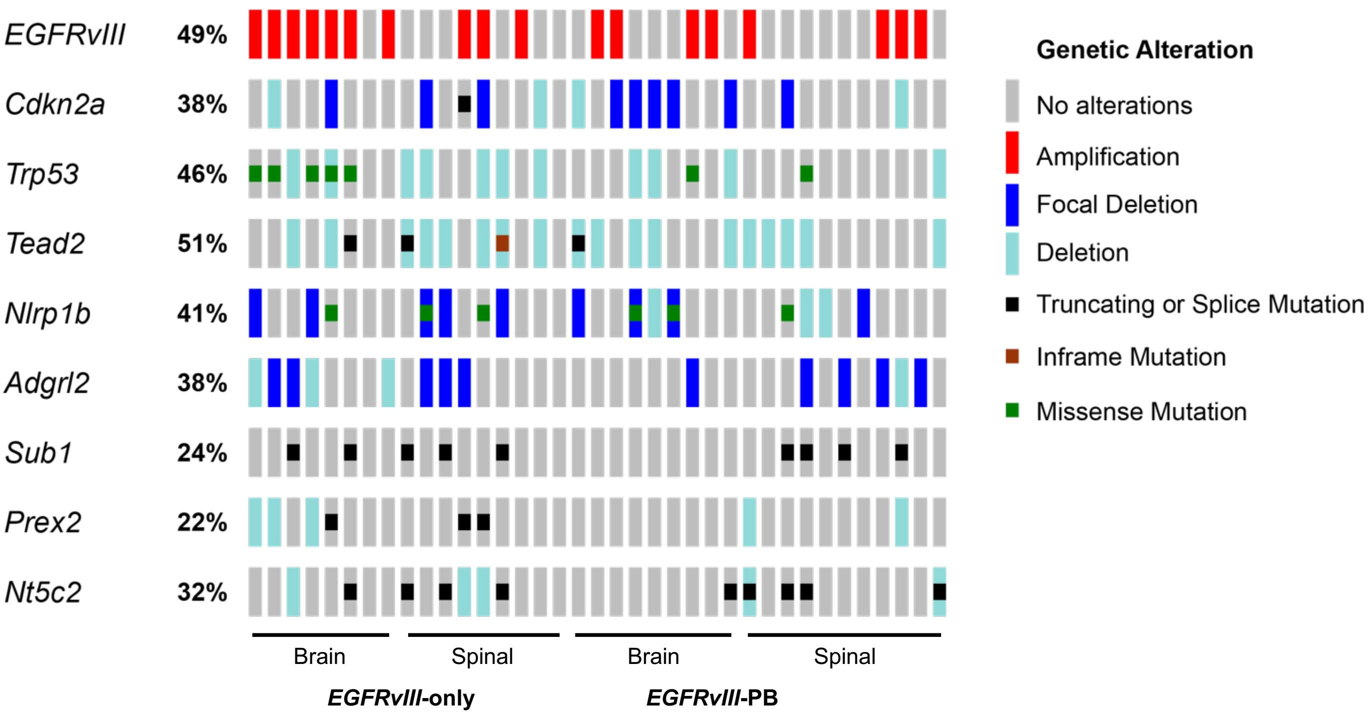


Fig 3.42. Key cancer genes identified, either as significantly mutated from MuSiC or copy number altered from GISTIC2, across all mouse brain and spinal tumors in both cohorts; each column represent one tumor.

Discussion

EGFRvIII can initiate gliomagenesis *in vivo*

In this study, we have demonstrated that *EGFRvIII* when expressed in the central nervous system using the *nestin*-cre driver, initiates glioma formation in the brain and spinal cord. In our work, the earliest glioma-like lesions were observed in the subventricular zone and at the brain surface, and later gliomas were observed both in the brain and spinal cord with 100% penetrance.. Other studies in mice have concluded that *Egfr* activating mutations alone are insufficient to generate gliomas *in vivo*, and instead they require cooperation with other mutations such as loss of tumor suppressors *Pten* and *Cdkn2a* [92, 96, 128]. In our study, microneoplasias were observed from 12 weeks of age and gliomas were observed with median latency of 36 weeks, suggesting *EGFRvIII* is sufficient to initiate gliomagenesis *in vivo*. Although here the mice only carried a single mutation (in *EGFR*), the latency for formation of gliomas was rather long. Similarly, there was a low incidence (~10%) of glioblastomas. These findings suggest that *EGFRvIII* can be the first tumor-driving genetic event in gliomagenesis but there is a requirement for activation or suppression of additional cooperative pathways for tumor progression. RNA-sequencing analysis of the *EGFRvIII*-induced tumors showed differential expression between these and normal brain and spinal cord samples, with significant enrichment for neural development pathways and oncogenic pathways such as P53 signalling and the MAPK pathway. Although *EGFRvIII* starts the tumorigenic process by stimulating cellular proliferation, activation in these other pathways must later become important in tumor formation and can at least partially explain the delay between *EGFRvIII* is first expressed and when a tumor is fully developed. Additionally, during glioma formation there is somatic acquisition by tumor cells of genetic alterations in other oncogenes and tumor suppressor genes, as shown in our whole-exome sequencing data from *EGFRvIII*-driven gliomas. This likely occurs because *EGFRvIII* is a potent stimulator of cell proliferation, and mitosis is associated with a low but non-zero rate of errors in DNA replication; if these errors occur in cancer genes, then this will set up a Darwinian natural selection process whereby mutations providing a survival advantage to the cells will be selected for, enabling evolution of the tumor. Indeed, *EGFRvIII* has been previously associated with genetic instability *in vitro [177]*.

Microneoplasias as precursors to gliomas have been described in mice previously, for example the work by Jacques and colleagues demonstrated such lesions in mice SVZ in the presence of *Trp53* and *Pten* double homozygous loss [97]. In their study, the investigators micro-injected adenovirus expressing cre under control by the GFAP-promoter into the ventricles of *Trp53*$^{-/-}$; *Pten*$^{-/-}$ mice, which then developed microneoplasias in the SVZ followed later by high-grade gliomas, leading the authors to conclude that SVZ type-B neural stem cells (which express GFAP) were the origin of gliomas in their model.

It is worth considering why previous studies in mice did not find that *EGFRvIII* alone induced gliomas, unlike in our study. There are likely multiple reasons for this discrepancy:

1. The relatively long latency for tumor formation in our model demanded long observation times, whereas some previous studies had shorter observation times of their mice expressing *EGFRvIII*. For example, Holland and colleagues histologically assessed their mouse brains carrying this mutation after 10 weeks [92], in contrast the median latency of tumor formation in our model is 36 weeks. Similarly, Zhu and colleagues observed their mice for around only 30 weeks after induction of EGFRvIII expression [96].
2. Several studies targeted expression of *EGFRvIII* from a particular brain region. For example, Zhu et al injected cre (carried by adenovirus) into the striatum of their mice [96], and Holland et al used the replication competent ALV splice acceptor (RCAS) viral system for *EGFRvIII* gene transfer into the frontal lobes of mice. In contrast the tumors seen in our mouse model were typically in the ventricles and / or surface of the brain / spinal cord, rather than deep parenchymal lesions; the *nes-cre* driver used here targets neurogenic niches.
3. Previous studies on *EGFRvIII* in mice did not report assessment of the spinal cord, whereas mice in this study developed spinal gliomas with 100% penetrance.

4. Some studies induced expression of *EGFRvIII* in ‘adult’ mice, for example Klingler and colleagues employed a tetracycline-inducible *EGFRvIII* with the mutant allele induced from mice aged 4 weeks[128], and Zhu et al induced expression of the allele with cre from mice aged over 3 months[178]. Other studies have suggested that the age at which the genetic lesions occur affect tumor latency – Llaguno and colleagues from Luis Parada’s laboratory showed that the latency for glioma formation from neural progenitors in mice with *Nf1*, *Trp53* and *Pten* was substantially reduced when these genetic alterations occurred from embryonic stages compared with 4 week old mice[179]. The *nes*-cre driver employed in our model induces recombination from embryonic day 13.5, and this cre line was similarly employed in the study by Llaguno et al.
5. The mice in this study are of mixed background, which may influence cancer phenotypes. However, the complete penetrance of gliomas in our mice provides strong evidence for *EGFRvIII* being the key initiating drive for these tumors. Interestingly, a recent systematic review of many published studies found no evidence for inbred mice (lacking genetic heterogeneity) having greater stability in phenotypic traits compared with outbred mice [180].

The cell of origin of gliomas is not the focus of our work, but it is worth considering this issue and how future studies can help elucidate this question for our model. Tumors in our mice may originate from a single source such as the SVZ, or the origins may be multi-focal. The latter explanation is more likely given that there are some cases where there are tumors on the brain surface without well-developed neoplastic lesions in the SVZ. The origin of these tumors is unclear. It is possible that the brain surface tumors therefore arise for neural stem cell or neural precursors much like the SVZ gliomas; indeed, a recent study has demonstrated meningeal neural stem cells in the developing central nervous system in mice, and these cells have a matching transcriptomic profile to neural stem cells from the SVZ [181]. Further work will be needed to shed light on whether these meningeal NSCs are the origin of a subset of gliomas in these mice and in humans. Such studies are likely to use cellular lineage-tracing approaches to directly interrogate the fate of the meningeal neural stem cells that carry oncogenic mutations such as *EGFRvIII*. Knowledge of the origin of gliomas, even if dependent on genetic context, may be important clinically for treatment. For example, understanding that *EGFRvIII*-driven gliomas arise from the subventricular zone and brain surface with subarachnoid seeding would suggest that treatments that fail to tackle these regions would not be sufficient to treat the cancer: removal of the primary tumor will still lead to recurrence because tumor cells would be left behind in the SVZ and / or subarachnoid space. Therefore, it may be sensible to add adjunctive treatments targeted to these regions if they are thought to be the origin of tumors in individual cases, such as using radiotherapy to the SVZ and intrathecal or intraventricular chemotherapy to eradicate tumor cells in the ventricular system that would otherwise lead to distal tumor recurrence.

Nevertheless, our data demonstrates that *EGFRvIII* can be the initiating driver mutation for gliomas *in vivo*. At first glance, these observations appear to be at odds with the observation in many human glioblastomas with same *EGFR* mutation that show heterogeneous expression of the mutant EGFR protein across the tumor (suggesting this mutation was a late acquisition in clonal evolution) [182]. However, these observations may be reconciled by the notion that although not all cells in the tumor express the protein, all cells may carry the mutation and some cells may simply downregulate expression of the EGFRvIII protein. This is particularly likely to be the case late in tumor evolution when other somatic mutations can act as tumor drivers making the initiating *EGFRvIII* less important for tumor progression. Several publications have described internalisation of the EGFRvIII protein in gliomas [183, 184], supporting this possibility. Our work is consistent with a study suggesting that EGFRvIII cells from GBMs have cancer stem-cell properties such as self-renewal and tumor-initiating capacity [185], although serial transplantation of tumor cells into mice is required to demonstrate this from our tumors. Our findings are also in keeping with the observations in humans that the presence of EGFR alterations in low grade gliomas is poor prognostic indicator for survival and that this mutation is much more common in glioblastomas compared with low grade tumors, suggesting EGFR alterations are driving glioma progression in humans as well.

As mentioned previously, our understanding of the timing and prevalence of *EGFRvIII* is confounded by significant intratumor heterogeneity in gliomas. Tumor heterogeneity may even be partially maintained by EGFR itself which has been reported to actively drive genetic heterogeneity through a cytokine circuit involving IL-6 [186]. A recent study has demonstrated that most patients with who had *EGFRvIII* in their primary GBM maintain this mutation in their recurrent GBM, suggesting this mutation may be still be, at least partially, driving gliomagenesis at recurrence [187]. Studies have been conflicting on the timing of acquisition of *EGFR* mutations and amplifications in gliomagenesis. A recent report found that *EGFRvIII* was present throughout different sites of GBMs in a small patient cohort, despite not being expressed throughout the same tumors suggesting there are mechanisms for downregulating expression of the mutant receptor; moreover, it was found that demethylation of *EGFRvIII* led to re-expression of the protein implying that epigenetic mechanism underpin control of its expression [188]. These results would be consistent with our work here showing *EGFRvIII* can be an initiating event in a subset of gliomas. In contrast however, another study concluded that the *EGFR* amplification occurs early in GBM patients but that *EGFRvIII* occurred later as there was substantial intratumoral heterogeneity in its expression; there were only three patients sampled in this study, limiting the strength of these findings however [189]. Previous work also suggested that glioma cell subpopulations develop an *EGFR* amplification first and later develop a mutation in the gene [190] ; it is more difficult however to draw conclusions about the timing of these genetic alterations from studying fully formed ‘end-stage’ GBMs in human patients as opposed to transforming normal neural cells with these events as we have done here.

Interesting differences between the brain gliomas, including the high-grade ones, discovered in these *EGFRvIII*-mice and human gliomas is that the mouse tumors tend to have little invasion into the brain parenchyma beyond the tumor margins; in contrast human gliomas tend to widely infiltrate brain parenchyma and can be situated at virtually any location. These differences may reflect unique features of tumors that are initiated by *EGFRvIII* as opposed to tumors triggered by different initiating mutations; it could be that distinct additional mutations need to be acquired at later stages for successful invasion into brain parenchyma. An alternative explanation for this difference is that mouse gliomas may be generally less invasive than their human counterparts.

Relationship to human brain gliomas

In human gliomas the initiating event in *IDH1*-wild-type tumors is poorly understood, yet activating mutations in *EGFR* (including *EGFRvIII*) are particularly common in these gliomas. Our intention therefore was not to model a histologically low-grade glioma versus a glioblastoma (GBM), but rather to determine the precise effect of introducing an activating *EGFR* mutation into the CNS and study the resulting phenotype and genetics. It is true that the majority of the gliomas generated in this model appear histologically low-grade, whereas in humans the majority with *EGFR* mutations are histologically GBM. Likely explanations for this include:

1. In humans very recent work has demonstrated histologically low-grade appearing, *IDH1*-wildtype astrocytomas are in fact representing *IDH1*-wildtype, early forms of GBM with their corresponding molecular features (particularly *EGFR* amplification) and poor prognosis [159, 191]. Von Deimling's group showed In a large cohort of human patients with histologically low-grade gliomas (LGGs) (n=544) that were reclassified as glioblastomas (GBMs) based on methylation profiling, *EGFR* amplification was found in 36% (n=196)[191]. They further demonstrated *EGFR* amplification as the single parameter of those tested with the highest specificity (>99%) for upgrading histologically *IDH1*-wt LGGs to *IDH1*-wt GBMs. Although this study analysed *EGFR* amplification than *EGFRvIII* specifically (which requires further study), this is the most likely subtype of glioma we have modelled with *EGFRvIII* in

mice. Consistent with this, all of our mice succumbed to their tumors by approximately 1 year of age, all tumors were *Idh1*-wildtype and transcriptomic analysis of mouse tumors showed significant enrichment for the human mesenchymal GBM gene set.

2. Some tumors in humans may first acquire tumor suppressor losses (such as *CDKN2A*, *PTEN* and other genes identified in this study) and subsequently acquire an *EGFR* mutation leading to transformation to a high-grade histological phenotype.

Eye Lesions

The *EGFRvIII* / nestin-cre mice developed eye abnormalities suggestive of underlying vascular pathology: there was a high incidence of bleeding (typically minor, but occasionally more major haemorrhages occurred) and histology showed neovascularisation of the retina and a reactive proliferation in the ciliary body. There was also apoptosis of the lens. This constellation of pathologies has not been previously reported to be induced by *EGFRvIII* expression in the eye. As discussed above, it is thought that expression of a strong oncogene such as *EGFR* may trigger a cell to activate cell death programs as a protective mechanism against tumor formation. Apoptosis of the lens therefore may be explained in this way. *EGFR* has been linked to angiogenesis, and this may be an additional reason as to why it is oncogenic given that cancers require development of their own blood supply in order to progress; the neovascularisation of the eye in these mice may be a reflection of the pro-angiogenic characteristic of *EGFRvIII*. Further characterisation of these eye lesions is outside the scope of this study.

Spinal Gliomas

It is certainly of interest that mice conditionally expressing *EGFRvIII* under *nestin*-cre control develop spinal gliomas. Very few models of spinal tumors have been published in the literature, and the genetic basis of these tumors is poorly understood not only because these are rare tumors but also because surgical options are often constrained by the tumor invading the spinal cord limiting the extent of resection and therefore availability of material for DNA sequencing. This is the first time to our knowledge that it has been demonstrated that *EGFR* activating mutations can drive spinal gliomagenesis *in vivo*. Although histologically classified as benign, these tumors despite being rare are highly clinically relevant because they often lead to severe neurological impairments such as paralysis due to their eloquent location, and these clinical signs were observed in the mice of this study. Human spinal astrocytomas tend to strongly invade the spinal cord parenchyma; the tumors in the mice here tend to weakly invade the parenchyma but more strongly invade the nerve roots, based on histopathological analysis. This difference in pathology may be due to the species differences. The minimal invasion into CNS parenchyma seems to resemble optic pathway gliomas in humans, which are benign tumors but fervently invade the optic nerve leading to significant neurological deficits (blindness). These spinal gliomas are most likely primary tumors in the mice because in many cases there are only small and early tumor lesions in the brain but larger and more pervasive spinal tumors, making these tumors less likely to be a result of metastasis from the brain tumors. Conversely, in cases where the mice had aggressive high-grade gliomas in the brain, the spinal tumors were often low grade, in keeping with these being independent primary spinal tumors. RNA-sequencing also identified different transcriptomic profiles for brain compared with spinal gliomas in these mice, reflecting the locations in which these tumors arose. The unique location of these tumors (leptomeningeal) and benign histology share do not reflect all spinal gliomas in humans, but they are similar to a subset of tumors – paediatric leptomeningeal low-grade gliomas.

Although the frequency and nature of *EGFR* alterations in human spinal tumors remains to be determined in larger genomic studies than the few smaller scale studies thus far conducted, *EGFR* amplification and expression has been detected in a small cohort of disseminated paediatric spinal LGGs[160]. Although our mice have *EGFRvIII* as the driver, these tumors may possibly be generated by other mechanisms for increased EGFR signalling including alternative *EGFR* mutations, amplification and / or overexpression. Further work overexpressing the wild-type *EGFR* in mice will help to address this question. In human patients, germline *NF1*-loss predisposes to spinal glioma[192] and a study of spinal gliomas detected *CDKN2A* deletion and loss of heterozygosity at 10q23 (containing *PTEN*)[193].

We hope our novel mouse models of spinal gliomas will provide further opportunities for insights into the pathogenesis of this disease and development of therapeutics *in vivo*.

Molecular alterations

Solid tumors typically contain a number of chromosomal copy number changes reflective of chromosomal instability due to loss of mitotic fidelity in DNA replication. Chromosomal instability and aneuploidy correlates with a worse prognosis and resistance to treatment [194]. Chromosomal analysis of the *EGFRvIII*-driven brain gliomas using FISH revealed a number of chromosomal aberrations. FISH is a useful technique in enabling detection of large translocations, duplications and deletions. These tumors substantial inter-tumor heterogeneity in their chromosomal changes. Certain alterations occurred in multiple tumors however, including amplification of chromosomes 11 and 15 (either through whole chromosome gains or through Robertsonian translocations), indicating that there must be a selective advantage conferred to cells carrying these chromosomal changes. One possible explanation for chromosome 11 gains is that the *EGFRvIII* transgene is on the *Col1a1* locus, which is situated on this chromosome; moreover, the mouse wild-type *EGFR* gene is itself located on this same chromosome. Whole-exome sequencing also revealed there was amplification of the *Col1a1* locus (where *EGFRvIII* transgene is situated). Gain of chromosome 11 therefore leads to copy number gains of the *EGFRvIII* transgene, amplifying the oncogenic EGFR signalling, suggesting the ‘dose’ of oncogenic EGFR is important for tumorigenesis. It is unclear what the reason is behind selection for chromosome 15 amplification in the gliomas from this study; although *C-myc* is a prominent proto-oncogene located in this chromosome, *C-myc* itself is not amplified according to the copy number analysis of these tumors.

We examined tumor-genomic evolution through whole-exome-sequencing and RNA-seq. Given the number of exonic mutations in each tumor was modest, genetic drivers could be discerned based on recurrent mutations as well as the impact of these on gene function. The exome-sequencing data also demonstrated that additional driver mutations are needed for glioma expansion after *EGFRvIII* mutation and amplification initiate tumorigenesis. The most frequently mutated genes *Sub1, Trp53* and *Tead2* had loss-of-function mutations in more than 20% of samples. A number of recurrently mutated genes were observed in frequently deleted regions and/or they had focal deletions, including *Tead2*, *Uimc1* and *Nlrp1b*. *Cdkn2a* and *Adgrl2* had recurrent focal deletions too. Correlation of the mouse with human glioma genetic data suggested many of the mutated and deleted genes are also altered in patients, such as recurring deletions of *TEAD2* and methylation of *NLRP1*.

To further characterise these spinal gliomas at a molecular level and to determine if they are related to the brain tumors, we compared their transcriptomic profiles from the RNA-sequencing data with those of brain gliomas from *EGFRvIII* mice. This comparison revealed that genes representing brain development were significantly upregulated and enriched in the brain tumor cohort, whereas genes involved in spinal cord processes such as sensory perception and regulation of motor activity were upregulated and enriched in the spinal tumors. These data point towards different and independent origins of these two types of tumor: the brain gliomas have an origin from cells in the brain, and the spinal gliomas originate from the spinal cord. This result is certainly in keeping with the spinal tumors being primary gliomas rather than metastases from the brain, despite that these two types of tumor often co-occurred in the same mice.

Genomic copy number analysis revealed that these tumors had very complex genomes with significant copy number changes throughout. The extent of these alterations make it hard to discern causal alterations by this method, aside from the focal changes described. Given their genomic complexity, the transcriptomes of mouse tumors exhibited many changes from normal tissue. Recurrent amplification of *EGFRvIII* was observed, suggesting strong selection for increased expression. RNA-seq analysis confirmed the endogenous mouse *Egfr* was upregulated in all tumors, implying that *EGFRvIII* signalling involves collaboration with the endogenous gene.

Genomic instability includes both structural and numerical chromosomal abnormalities, and is key driving force and also a hallmark of cancer [195-197]. By using conditional *PiggyBac* insertional mutagenesis the expected reduction in mouse survival[70] was not observed - one explanation for this is that the chromosomal instability observed in the absence of transposition is able to provide an adequate reservoir of additional mutations to facilitate oncogenesis. Supporting this is the obvious difference in ploidy observed in tumors from the two cohorts, with reduced ploidy in tumors with PB transposition compared to those without. A previous study from David Largaespada's laboratory reported similar findings with *Sleeping Beauty* transposition in osteosarcomas [58]. A difference in the spectrum of mutations was also apparent from exome sequencing data: in the absence of the transposon the most frequently mutated genes are more plausible cancer genes, such as *Trp53*. Together, these data imply that *piggyBac* mutagenesis replaces the need for genomic instability in providing secondary molecular alterations needed to drive gliomagenesis. The reduced copy number variation in tumors with *piggyBac* mutagenesis can potentially greatly simplify interpretation of the cancer genomes: as the transposon integrations in cancer genes will be clonally selected for, sequencing for and identifying common integration sites should provide comprehensive information on the functional cancer genes at play in these tumors. This will be the focus of my next Chapter.

Deeper understanding of the molecular basis of spinal gliomas is needed in order to advance treatment options, which are currently limited. Demonstration that *EGFR* is a driver of spinal gliomagenesis suggests these tumors may be amenable to therapeutic targeting with EGFR inhibitors that have had some success in lung and colorectal cancers for example. It remains to be seen however what the frequency of *EGFR* mutations is in human spinal gliomas, and this can only occur if concerted efforts are made to collect sufficient material for DNA sequencing in a large cohort of patients likely from multiple clinical centres.

Study Limitations

One potential limitation of this work is the control used for our RNA-seq experiment; I used SVZ tissue dissected from a *nes*-cre mouse lacking *EGFRvIII.* A more accurate control would be the cell of origin of the gliomas we produced in our model, such as the neural stem cell. However, as mentioned earlier, our aim was not to study the glioma cell of origin in this investigation and thus the cell of origin is currently unknown for our model; using cerebral cortex, although admittedly will contain a mixture of cell types, avoids making an assumption about the cell of origin here. Future work is warranted to answer this question; one possibility is to compare the RNA-seq data from the brain gliomas with gene expression profiles of different 'normal' brain cell types , including neural stem cells, oligodendrocyte precursor cells, astrocytes and neurons , in order to determine the most similar normal cell type as this may point towards the cell of origin. Alternatively, lineage tracing studies such as mosaic analysis with double markers as utilized by Liu and colleagues (see Introduction, [101]) can greatly help in identifying the cell of origin in the model we have generated here.

A potential limitation is that one of the EGFR antibody we used for immunohistochemical staining of tumors was not specific for the *EGFRvIII* mutation, so may not accurately define the proportion of tumor cells carrying this mutation. The monoclonal antibody we used for immunohistochemistry was manufactured using the human EGFR protein (purified from A431 cells, see https://www.abcam.com/egfr-antibody-31g7-ab218383.html and[96]) as an immunogen, with specificity for human over mouse EGFR, but further work would be needed to assess any cross-reactivity with mouse EGFR protein. The antibody cannot be staining human EGFR-wild type in these tumors - we only introduced the human *EGFRvIII* transgene in these mice (not the human EGFR-wild type gene). We confirmed this by PCR genotyping with primers spanning the junction between exons 1 and 8 that is created upon deletion of exons 2 -7, yielding the expected 670 bp fragment for *EGFRvIII.* However, to further confirm protein expression of EGFRvIII in tumor cells, we also sourced an EGFRvIII monoclonal antibody and this confirmed expression of the recombinant protein in glioma cells.

Another limitation of this work is that, although there are recurrent mutations in certain genes in the tumors generated from our model, further functional studies are required in order to definitively prove that some of these are driving tumorigenesis. Such work may include, for example, CRISPR knockout or siRNA knockdown of individual genes in cell lines or directly in mice to demonstrate these accelerate glioma development. However, the presence of recurrent genetic alterations, including many in known glioma tumor suppressor genes (*Trp53* and *Cdkn2a*), provides strong statistical support for the notion that, at least some , of these mutations are contributing to gliomagenesis.

One challenge in interpreting pathology of the brain of mice is distinguishing between the normal rostral migratory stream (RMS) and small tumors in the same location (such as SVZ), as these can look similar. The RMS is a specialised migratory route by which neural precursors migrate from the SVZ to the olfactory bulb in mice and certain other mammals. However, we overcame this challenge by taking advantage of an expert neuropathologist with substantial experience in interpreting normal and pathological mouse pathologies, examining mouse brains in exactly the same way (see Materials and Methods) with four coronal slices per brain, and blinding the pathologist to mouse genotype. Given that all *EGFR*-mutant mouse brains displayed microneoplasias or full tumors, many of which were also in locations the RMS does not reside (such as third ventricle and multiple locations on the brain surface), many mice had multiple such lesions and control mice had none, and the fact that older mice had full tumors in the same location (suggesting these had progressed from the earlier precursors), make compelling arguments that the phenotypes we are observing are truly tumors rather than normal variants of the RMS.

Another limitation of the study is that, although RNA-sequencing was performed and gene set enrichment analysis demonstrated enrichment for key oncogenic signatures, downstream signalling pathways in our tumors were not studied in detail. In particular, activation of specific signalling proteins, such as Akt, Erk, Mek, and Stat3, was not tested, largely because the focus of our work here is the role of genetic alterations in tumorigenesis. However, investigating the downstream signalling proteins that are activated in these tumors will be important in future work to further understand the molecular driving forces of gliomagenesis.

Conclusions

In this Chapter, I have demonstrated *EGFRvIII* can initiate gliomagenesis in the brain and spinal cord, with the long latency for tumor formation reflecting the need for additional secondary molecular alterations to be present as well. Through whole-exome sequencing, we identified significantly recurrent mutations present in these tumors, suggesting a landscape of genes contribute to tumorigenesis. Chromosomal instability was observed in these tumors; however, in the presence of *piggyBac* transposition this instability was reduced, implying *piggyBac* provides secondary mutations needed for gliomagenesis instead. Such mutations will be the focus of the next Chapter.