Genomic Profiling of Pediatric Low-Grade Gliomas

Guillaume Bergthold

To cite this version:

Guillaume Bergthold. Genomic Profiling of Pediatric Low-Grade Gliomas. Genetics. Université Paris Sud - Paris XI; Harvard university. Medical school, 2015. English. NNT : 2015PA11T053 . tel-01374734

UNIVERSITE PARIS-SACLAY

ÉCOLE DOCTORALE 418 DE CANCEROLOGIE
Laboratoire de Vectorologie et thérapeutiques anticancéreuses
&
Cancer Biology Department Dana-Farber Cancer Institute, Boston, USA

THÈSE DE DOCTORAT

Soutenue le 30/09/2015

par

Guillaume BERGTHOLD

Genomic Profiling of Pediatric Low-Grade Gliomas

Etude des profils génétiques des gliomes de bas-grade pédiatriques

Directeur de thèse :	Jacques GRILL	Praticien des CLCC (Gustave Roussy, Villejuif, France)
Co-directeurs de thèse :	Mark W KIERAN	Assistant Professor (Harvard Medical School, Boston, USA)
	Rameen BEROUKHIM	Assistant Professor (Harvard Medical School, Boston, USA)
Composition du jury :		
Président du jury :	Pascale VARLET	PU-PH (Service de Pathologie, Hôpital Sainte-Anne, Paris)
Rapporteurs :	Natacha ENTZ-WERLE	PU-PH (Onco-Hématologie Pédiatrique, CHU Strasbourg)
	Daniel GAUTHERET	PU (Institut de Génétique&Microbiologie, Paris-Sud Orsay)
Examinateurs :	Pascale VARLET	PU-PH (Service de Pathologie, Hôpital Sainte-Anne, Paris)
	Marc SANSON	PU-PH (CRICM, Paris 13)

Genomic Profiling of Pediatric Low-Grade Gliomas

Etude des profils génétiques des gliomes de bas-grade pédiatriques

A tous les enfants et leurs familles affectés par le cancer,
Vous inspirez chaque jour mon travail.

A tous les professionnels qui travaillent au quotidien pour améliorer le confort
Et la survie des enfants atteints par le cancer.

A mon épouse Elsa et nos enfants Sarah, Nathan, Timothée et Victor
Vous représentez pour moi ce qu'il y a de plus précieux au monde.

Abstract

Low-grade gliomas represent the most frequent brain tumor arising during childhood. They are characterized by a broad spectrum of tumor types. For most of the patients, the overall survival is good, although the morbidity due to the treatments is affecting considerably their quality of life.

In the past, the definition of low-grade gliomas (LGG) has been mainly based on morphology. This histological classification of pediatric low-grade gliomas (PLGG), recently revised in 2007 suffers from the lack of reproducibility and precision, as a non-negligible number of PLGG are hard to classify in a specific category defined by the WHO classification. The recent progress in molecular biology and genetics has brought new insights in the biology of those tumors and allows better understanding of the biology of the tumors.

This work provides a comprehensive analysis of two different genetic approaches in PLGGs. The first part is based on the description of somatic genetic alterations of the DNA. Using a large PLGG cohort, we have been able to dissect the genome of those tumors and draw the landscape of the genetic alteration of those tumors. Although *BRAF* and *FGFR1* alterations are predominantly altered in those tumors, we have discovered a new translocation, *MYB-QKI*, that is almost exclusively present in a specific histological subgroup; angiocentric gliomas.

The second part of the thesis describes another angle of the biology of PLGGs by analyzing the transcriptome of bulk tumors. This work describes molecular differences between PLGGs from distinct histologies and arising from different locations in the brain as well as different *BRAF* mutation status.

As bulk tumor analysis is contaminated by several cell types, we were also able to test single cell expression analyses in three pilocytic astrocytomas (PAs) using RNA-sequencing. In this experimental work we have successfully tested the hypothesis that we can isolate single cells from fresh PLGG tumors in order to analyze the trasncriptome at a large scale. We observed that single cells expressing A2B5, a glial progenitor marker, isolated in pediatric pilocytic astrocytomas are characterized as a distinct biological population in comparison with A2B5 negative cells. These results underline the importance to improve the precision of the transcriptomic studies to capture the molecular signal of tumor cells and further understand the different pattern between normal cells and tumor cells.

Acknowledgments

First I want to thank my laboratory advisor Rameen Beroukhim. It has been such an honor for me that you accepted me as a research fellow in your laboratory. I appreciated so much working and learning in your team! Thank you for all your time, ideas, and funding to make my fellowship so productive and stimulating. You taught me with patience bioinformatic skills, statistics, maths as well as genomics. I was also able to improve considerably my English while learning to write structured grants and scientific papers as well as oral communications. The passion and knowledge you have in your research as well as your human qualities represent an example for me.

I would also like to express my deep gratitude to Mark W. Kieran and Jacques Grill, my two thesis advisors. Without you this project would not have been possible.

Thank you Mark for your mentoring and guidance during my fellowship. You were always available to discuss about our project, provide new insights for my work and enlarge my knowledge about low-grade gliomas. I also appreciated your human qualities as a doctor when I was attending the neuro-oncology conferences as well as shadowing clinical rounds.

Thank you Jacques for your friendship and your incredible trust in me. Since the first time we met in Paris, you were always willing to help me and provide me the best training as a young pediatric oncologist. Your enthusiasm and ideas in the neuro-oncology field were so contagious for me! Thank you for allowing me to perform this research fellowship in Boston.

Thank you Birgit for your continuous support during my Master in your laboratory. I have learned a lot with you and I thank you for your trust in me that encouraged me to pursue the path of research.

I would like to express deep gratitude to the Cancer Biology and the Molecular Biology Department members at Dana-Farber Cancer Institute. Thank you Chuck for your incredible team and staff! Thank you Keith for all your guidance and help during my stay at Dana-Farber. It has been such an asset for me to collaborate with you and your laboratory. Thank you Mimi for our collaboration during my stay in the Beroukhim lab. I have learned so much working with you. It has been such a great experience working together as pediatric oncologists in the lab. Thank you Peleg for your help and support when I started my project in the lab. Thank you Brent for being such a great desk neighbor in the lab! Thank you Steve, Barbara, Travis, Will, Linda, Jeremiah, Ruben, Rebecca and Laura, Lori, Cécile, Ben for all your help and support. Overall I appreciated working in such a stimulating environment. Thank you Suzan and John as well as your friendly staff at the Dana-Farber Flow Cytometry Core Facility.

I would like to acknowledge all the scientists I met at the Broad Institute. Your ideas and professionalism were so inspiring to me. I would like to give a special thanks to Alex that trained me for the single-cell whole transcriptome amplification and sequencing analyses. Thanks also to Tricia at the Flow Cytometry Core Facility in Harvard University. Your availability for afterhours sorting was so useful.

I thank also Natacha Entz-Werlé, Daniel Gautheret, Pascale Varlet and Marc Sanson for their availability as they accepted to be members of my thesis committee.

I gratefully acknowledge the funding sources that made my work possible. Thank you to the Nuovo-Soldati Foundation that trusted in my project and allowed me to start the work in Boston. Thank you to the Philippe Foundation, Inc, the Friends of DFCI as well as the PLGG Foundation that helped considerably my project. I also thank all the financial support I obtained from Rameen Beroukhim's numerous grants to allow working on such innovative and expensive projects.

Lastly, I would like to thank my amazing wife Elsa and our four wonderful children for their unconditional love, faithfulness and encouragement. Thank you for your patience and all the sacrifices you made to allow me working for my thesis. Thank you for my parents and all my family who supported me all those years in my work. Thank you for your prayers. Thank you.

Citations to Previously published Work

- Part of Chapter 1 was previously published in BBA Reviews. Cancer as:

  Bergthold G, Bandopadhayay P, Bi WL, Ramkissoon L, Stiles C, Segal RA, Beroukhim R, Ligon KL, Grill J, Kieran MW. (2014). Pediatric low-grade gliomas: how modern biology reshapes the clinical field. Biochim Biophys Acta. Apr;1845(2):294-307.

- Chapter 2 was submitted to Nature Genetics
- Chapter 3 was published in Neuro-Oncology as:

  Guillaume Bergthold, Pratiti Bandopadhayay, Yujin Hoshida, Shakti Ramkissoon, Lori Ramkissoon, Benjamin Rich, Cecile L Maire, Brenton R Paolella, Steven E Schumacher, Barbara Tabak, Ruben Ferrer-Luna, Memet Ozek, Aydin Sav, Sandro Santagata, Patrick Yung Wen, Liliana C Goumnerova, Azra H Ligon, Charles Stiles, Rosalind Segal, Todd Golub, Jacques Grill, Keith L. Ligon, Jennifer A Chan, Mark W. Kieran, Rameen Beroukhim. (2015). Expression profiles of 151 pediatric low-grade gliomas reveal molecular differences associated with location and histological subtype. Neuro-Oncology. Feb 26.

Table of Contents

Chapter 1: Introduction

1. Clinical and biological characteristics of Pediatric Low-Grade Gliomas (PLGGs)

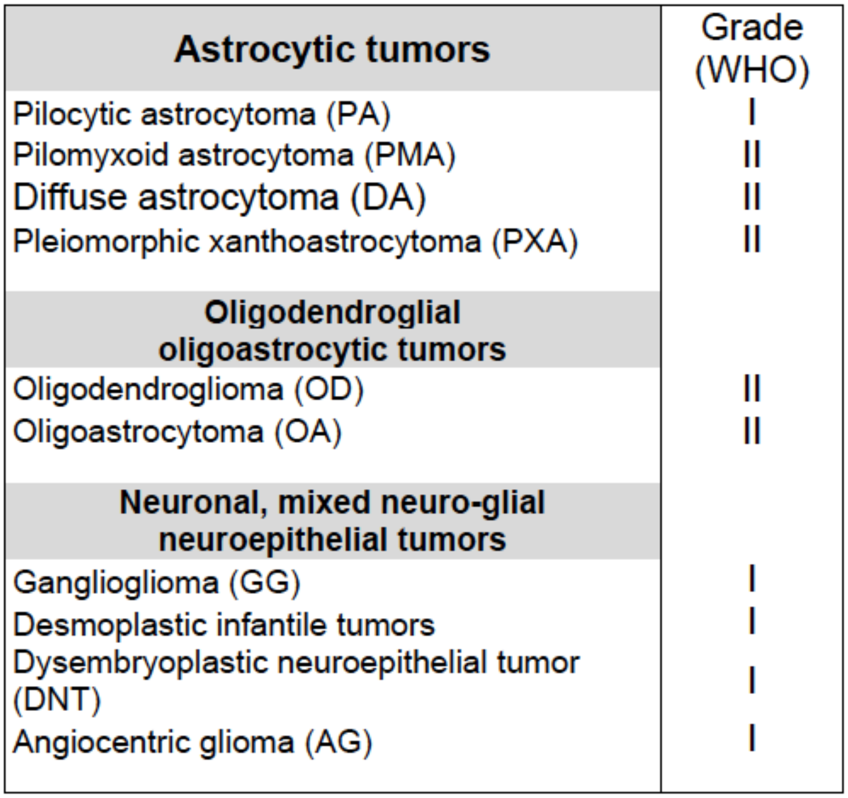
1.1. Definition

The World Health Organization (WHO) currently classifies PLGGs according to their morphological features[1]. Defined as grade I or II they represent a heterogeneous group of tumors, with more than 20 histological subtypes. PLGGs have been distinguished as either astrocytic tumors, oligodendroglial or oligoastrocytic tumors, neuronal, mixed neuro-glial or neuroepithelial tumors (Table 1). Tumors that do not meet the typical criteria of any single category are commonly labeled LGG-NOS for 'not-otherwise specified', which comprise more than a third of all PLGGs[2]. This sometimes results from small biopsy samples that lack sufficient material which to assign a WHO grade, and at other times, as a result of pathologic features that do not fit any one category.

Grade I pilocytic astrocytomas (PAs) are classically characterized by the presence of Rosenthal fibers, biphasic architecture, vascular proliferation, and eosinophilic granular bodies[1] (Figure 1a). Eosinophilic granular bodies are often located near cystic areas and may be implicated in cyst formation[3]. Less commonly, PAs contain regions of calcification[4]. Useful positive immunohistochemical markers include oligodendroglial markers OLIG2, myelin basic protein (MBP), platelet-derived growth factor (PDGF)[5-8] as well as the astrocytic marker Glial Fibrillary Acid Protein (GFAP), which is also considered a stem cell marker [9,10]. Gangliogliomas (GGs) are also grade I tumors and are characterized by perivascular chronic inflammation, granular bodies, binucleated neurons, calcification, and cystic degeneration[11] (Figure 1b). Dysembryoplastic neuroepithelial tumors (DNTs) and Angiocentric gliomas (AGs) are recently described subtypes also defined as grade I tumors. DNTs include a specific entity characterized by GFAP-negative oligodendroglia-like cells and floating neurons with a mucinous eosinophilic background [12,13] (Figure 1c). AGs, initially described by Lellouch-Tubiana et al., also named angiocentric neuroepithelial tumors (ANET), encompasses classically fusiform and bipolar astrocytic cells which stain positively for GFAP and S-100 arranged around blood vessels creating palisade-like structures [14,15] (Figure 1d). Microcalcifications are infrequently present. WHO grade II lesions include diffuse astrocytomas (DAs), pilomyxoid astrocytoma (PMAs), pleomorphic xanthroastrocyoma (PXAs), and oligodendroglial tumors. DAs are characterized by the presence of nuclear atypia, a low mitotic rate, and absence of vascular proliferation or palisading necrosis (Figure 1e). PMAs are characterized by astrocytic pleomorphism, significant cellular atypia, and multinucleated giant cells with intracellular lipid accumulation (Figure 1f). PXAs consist of pleomorphic and lipidized cells and tend to follow a more aggressive course with an increased frequency of leptomeningeal disease[16,17] (Figure 1g). Oligodendroglial tumors contain monomorphic cells with uniform round nuclei and perinuclear halos, microcalcifications and network of capillaries (Figure 1h and 1i). PLGGs typically have a low proliferative index, with MIB-1 scores between 0.1-10%[18-22]. This index is often higher in younger children where MIB-1 index higher than 10% can be seen in true PLGGs. However, correlation to either overall or progression-free survival in most studies has been variable and it remains unclear as to whether there is any prognostic significance[18,23-31].

Table 1:

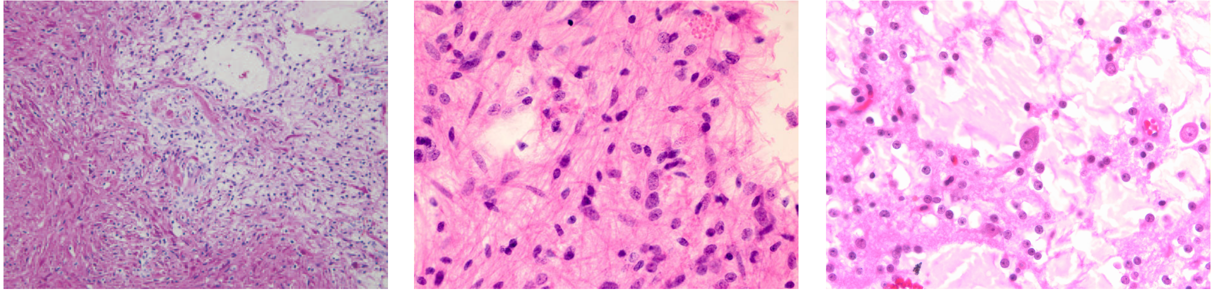
Major different subtypes of pediatric low-grade gliomas according to the latest WHO classification



Astrocytic tumors	Grade (WHO)
Pilocytic astrocytoma (PA)	I
Pilomyxoid astrocytoma (PMA)	II
Diffuse astrocytoma (DA)	II
Pleiomorphic xanthoastrocytoma (PXA)	II
Oligodendroglial oligoastrocytic tumors	
Oligodendroglioma (OD)	II
Oligoastrocytoma (OA)	II
Neuronal, mixed neuro-glial neuroepithelial tumors	
Ganglioglioma (GG)	I
Desmoplastic infantile tumors	I
Dysembryoplastic neuroepithelial tumor (DNT)	I
Angiocentric glioma (AG)	I

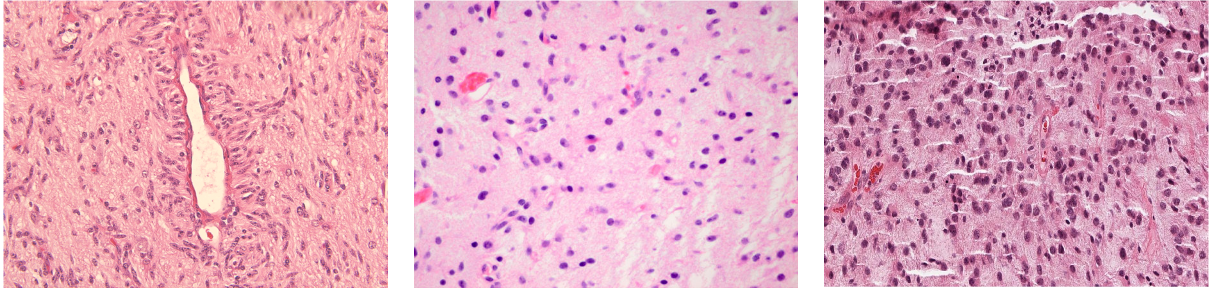
Figure 1:
Morphological aspect of the major subtypes of PLGGs defined by the WHO classification (H&E staining)

a. Pilocytic Astrocytoma b. Ganglioglioma c. Dysembryoplastic Neuro-epithelial tumor



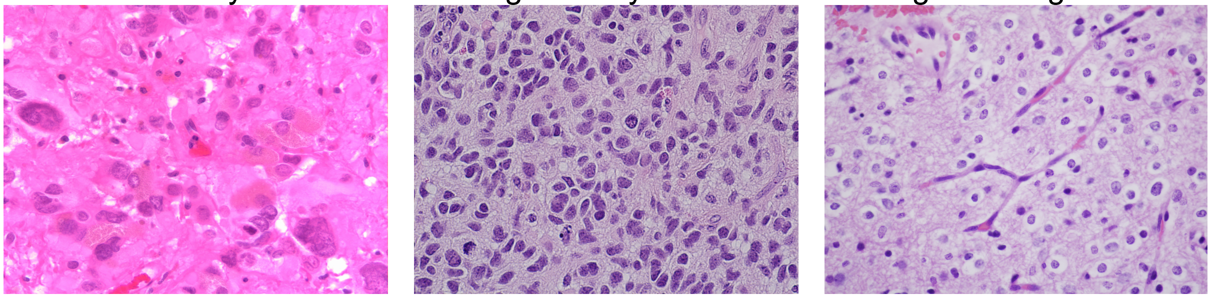
x200 x100 x100

d. Angiocentric Glioma e. Diffuse Astrocytoma f. Pilomyxoid Astrocytoma



x100 x100 x100

g. Pleiomorphic Xanthoastrocytoma h. Oligoastrocytoma i. Oligodendroglioma



x60 x60 x60

1.2. Epidemiology

Brain tumors represent the most common solid tumor of childhood, of which PLGGs are the most frequent[32]. The annual incidence of PLGGs is 2.1 per 100,000 persons in the United States[33,34], accounting for 1600 new diagnosis each year.

Table 2 summarizes the frequency of the major PLGG subtypes reported in recent epidemiologic studies[2,35-39]. The relative incidence of each LGG histological subtype varies with age, with clear differences in distribution between pediatric and adult LGGs (Figure 2). PAs most frequently develop during childhood and are extremely rare in adults. They represent the most common PLGG, accounting for 15% of all pediatric brain tumors[2,36,37,40,41] (Figure 3). DAs, oligodendrogliomas and oligoastrocytomas are more common in adults but extremely rare in children, representing less than 5% of PLGG [10,2,27-29]. Similarly, neuronal and mixed glial-neuronal tumors occur more commonly in the pediatric population. In addition to the defined groups of tumors, LGG-NOS tumors represent the second most prevalent diagnosis and have been reported to account for at least 17% of all PLGGs [2]. This is despite the fact that in most historical studies and governmental databases, this category is not included. This highlights the increasing need for integration of histology with molecular data to improve categorization of PLGG tumors. Although PLGG tumors can occur anywhere throughout the CNS, different subtypes demonstrate predilection for specific sites within the brain or the spine[42].

Pediatric DAs, AGs, PXAs and oligodendrogliomas are most frequently supratentorial [43-45,46], GGs occur most frequently within the temporal lobes[17,18,47] while PAs tend to localize to the cerebellum or the brainstem[48].

A fair fraction of PLGG can arise in the optical pathway as well as in the diencephalic/hypothalamic region; the incidence of those tumors is significantly higher in patients with neurofibromatosis type 1 (NF1).

Table 2:

Overview of six epidemiological studies including PLGGs performed in various countries around the world

	GRADE (WHO)	WHO International Classification of Diseases	USA	Germany	France	Denmark	Brazil
Number of total CNS tumors		5200	20,709	3268	1017	911	1058
Type of study		Retrospective	Retrospective CBTRUS Statistical Report	Retrospective German Childhood Cancer Registry	Prospective French Brain Tumor Database	Retrospective Multi-institution	Retrospective Single-institution Hospital das Clinicas of Sao Paulo
Age at diagnosis		0-19 years	0-19 years	1-15 years	0-19 years	0-15 years	0-21 years
Follow-up period		1980-1999	2005-2009	1990-1999	2004-2006	1960-1984	1974-2003
Frequency of tumors (%)							
Astrocytic tumors				16.4			
PA	I	14.8	15.5		23.1	16.5	18.2
DA	II	1.8	5.2	NA	1	13	6.2
PXA	II	0.4	NA		0.3		
Oligodendroglial tumors							
Oligodendroglioma	II	1.4	1.1	1.1	4	1.6	0.9
Oligoastrocytic tumors							
Oligoastrocytoma	II	0.6	0.7	NA	1.1	NA	NA
Neural and mixed neuro-glial tumors							
GG	I	2.5	7	3.2	4.6	2.2	3.6
Desmoplastic infantile astrocytoma	I	NA		NA	0.1	NA	0.3
DNT	I	NA		NA	3.1	NA	1.3
LGG-NOS tumors	I, II	0.1-6	11.3	NA	1.8	0.4	NA

Figure 2:

Comparison of the distribution of histological subtypes developing during childhood (0-19 years) and adulthood, according to the CBTRUS Statistical Report, 2012

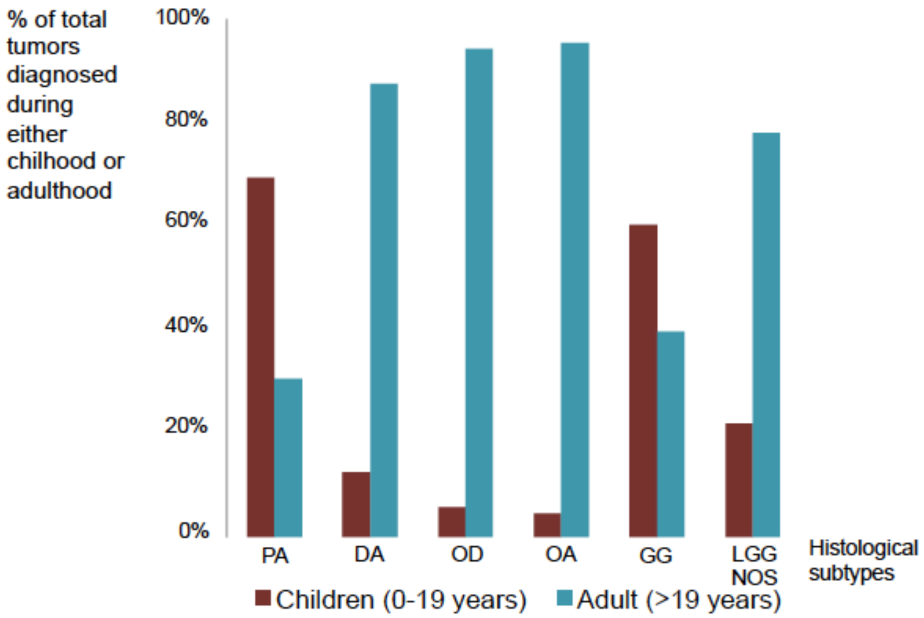
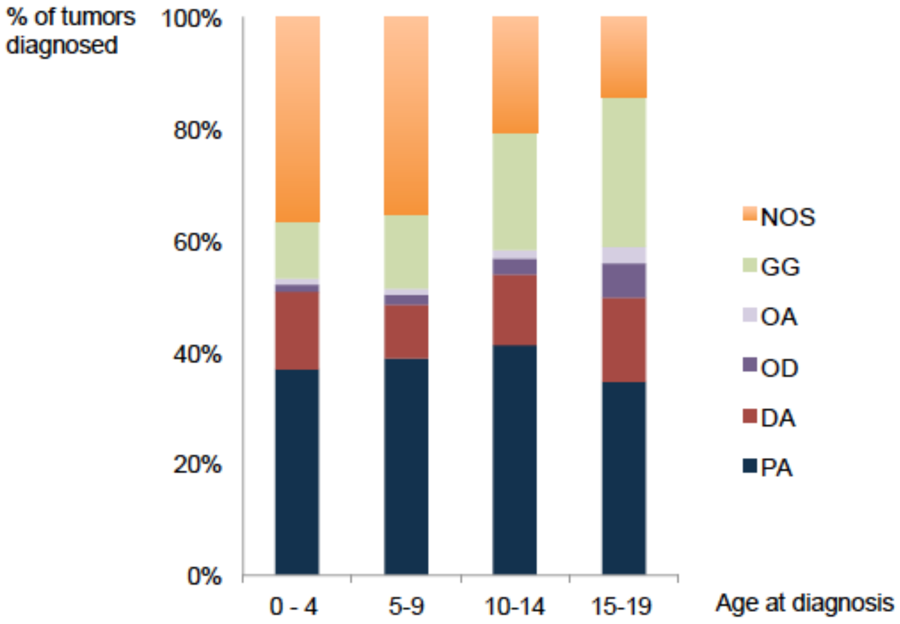


Figure 3:

Distribution of PLGGs histological subtypes during the 4 stages of development, according to the CBTRUS Statistical Report, 2012



Five percent of all PLGGs primary tumors are located in the spine and these are most frequently PAs[49]. PLGG can also develop in the cervicomedullary region[50] as well as in the tectum.

1.3. Genetic predisposition syndromes

Initial insight into the molecular characteristics of PLGGs was derived from the subset of non-sporadic tumors associated with genetic syndromes. Among these, the most frequent association is with NF1, also known as von Recklinghausen disease. PAs and DAs are the most common subtypes associated with NF1[51] and most commonly involve the optic pathway and hypothalamus[52-54]. NF1 is characterized by a germline mutation of *neurofibromin 1* (*NF1*), located on chromosome 17q, which results in activation of the RAS/MAPK signaling pathway. Importantly, only 30% of the tumors become symptomatic and require treatment, which suggests a unique biology underlying these tumors[55,56].

Tuberous Sclerosis (TS) is another neuro-cutaneous disorder with increased predilection for LGG, with brain tumors found in 5-14% of patients[57]. The most frequent brain tumor associated with TS is sub-ependymal giant cell astrocytoma (SEGA)[58]. TS is caused by mutations in two tumor suppressor genes, *TSC1* (hamartin, on chromosome 9q34) and *TSC2* (tuberin, on chromosome 16p13)[59]. These genes are part of the Rheb-mTOR pathways that function in regulation of cell proliferation.

These genetic syndromes contributed to our understanding of the importance of the ras/mTOR pathway in the oncogenesis of PLGGs. Additional findings from recent genomic studies have added further insights into the vital role of this pathway in the pathogenesis of PLGGs.

1.4. Clinical aspects

The clinical presentation of PLGGs is dictated by their location. Tumors in the posterior fossa typically present with acute signs and symptoms of elevated intracranial pressure secondary to obstructive hydrocephalus, as well as cerebellar signs[60], whereas LGGs of the optic pathway impair vision. PLGGs affecting the cerebral cortex typically present with focal neurological manifestations such as seizures or behavioral changes. Seizures are particularly associated with temporal, frontal, or parietal localization and oligodendrogliomas, GGs, DNTs or AGs subtypes[61-65]. Tumors involving the hypothalamus manifest with endocrinopathies or the diencephalic syndrome[66-68]. Tectal gliomas are often associated to hydrocephalus due to their expansion to the periaqueductal space. Compared to sporadic PLGGs, the clinical spectrum of NF1-related PLGGs diverges. NF1 patients more commonly present with multifocal tumors compared to sporadic cases[69].

PLGGs are most frequently localized at diagnosis, although they can present with disseminated disease. Leptomeningeal dissemination is reported in approximately 3-5% of children at presentation, especially in the setting of spinal cord or diencephalic/hypothalamic lesions[70-72], and may be associated with inferior overall survival compared to those who present with localized disease[73-75].

Radiological features of PLGGs are variable. These neoplasms are usually hypodense on CT compared to more malignant neoplasms. Grade I PLGGs are typically well-circumscribed tumors, with T1-hypointensity and T2-hyperintensity on MRI imaging. Following gadolinium administration, grade I astrocytomas usually demonstrate homogeneous enhancement. In contrast, grade II gliomas, especially DAs, are typically non-enhancing and may be less circumscribed[76-80]. PMAs express usually heterogeneous enhancement[80]. PLGGs are not usually associated with peri-tumoral edema or restricted diffusion on diffusion-weighted MRI sequences[81]. Magnetic resonance spectrometry (MRS), diffusion-weighted MRI (DWI) and diffusion tensor imaging (DTI) serve as useful adjuncts in further characterizing PLGGs. PET-scan and single-photon emission CT (SPECT) may also aid assessment of treatment efficiency and tumor recurrence. GGs typically exhibit contrast enhancement on CT scans and can have variable gadolinium enhancements on MRI – from absence of contrast enhancement to nodular or circumferential. Similar to astrocytomas and oligodendroglial tumors, they appear T1-hypointense and T2-hyperintense on MRIs. The contrast enhancement for oligodendroglial tumors is variable and is related to the infiltrative aspect of the tumors with a higher gadolinium contrast enhancement in solid and non-invasive tumors. DNTs do not displace brain structures but tend to infiltrate and usually have low or no contrast enhancement. They appear as bright T2-weighted and hypointense T1 tumors with typically neither mass effect nor peritumoral edema. Their slow growth may be associated with skull deformation when located in the cortex.

The natural history of pediatric LGGs is distinct from that of adult LGGs. On the whole, PLGGs exhibit slow rates of growth. Thus, the majority of children are diagnosed at least six months after symptom onset[82]. PLGG have been reported to spontaneously regress, especially in patients with NF1[83-87], who have been reported to have superior outcomes compared to sporadic cases[88-90]. Tumors that can be completely resected often require no further therapy highlighting the importance of location on outcome. In a recent prospective population-based study of a large cohort of 639 PLGGs, the 5-year PFS (progression free survival) was 69.4%[91], which is comparable to other studies[82,92-97]. Given the fact that two thirds of NF1 patients never progress, the recurrence rate of sporadic PLGGs is near 55%, as reported in the recent COG study[98]. The most significant risk factors for progression identified on multivariate analysis were young age at diagnosis (<1 year), subtotal resection, and DA histology[99]. Due in part to a better chance of complete resection, tumors involving the optic nerve or cerebellum have better progression-free survival (PFS) compared to those involving the chiasm and hypothalamus. Even if progression occurs, children diagnosed with PLGG have an excellent overall survival long-term, as described in a recent analysis of the SEER (Surveillance Epidemiology and End Results) database showing a 20 year overall survival of 87%[100]. In contrast to adults, PLGGs are characterized by a low incidence of malignant transformation[101-103]. Importantly, adult survivors of PLGG have low glioma related mortality, suggesting a very low propensity for malignant transformation of PLGG[100].

1.5. Therapeutic strategies

Given the excellent overall survival for the majority of PLGG patients, the treatment goal is to achieve tumor control while minimizing long-term tumor and treatment-related morbidity[104]. Most patients require only surveillance after surgery. If progression, recurrence and/or symptoms occur, then treatment modalities including surgery, chemotherapy (including biologic therapy), or less frequently, radiation therapy are indicated.

Surgical resection remains the cornerstone of PLGG management. Patients with gross total resection of tumor typically do not need further treatment. However, gross total resection is not always achievable without significant neurological impairment for some tumor locations, such as the optic pathway, hypothalamus, diencephalon, and brainstem. In these instances, the goal of surgery is to achieve maximal resection without risking severe neurologic deficits. Even in the event of a subtotal resection, the overall survival of patients remains excellent [100,105-110].

Chemotherapy is usually initiated for radiological and/or symptomatic progression. Over the last few decades, many protocols using either monotherapy or poly-chemotherapy have been tried for PLGG. Platinum-based chemotherapy such as carboplatin[111-113], cisplatin[114], oxaliplatin[115], ibroplatine[116] alone or in combination with vincristine[117-125] or etoposide (VP16)[126] have been widely utilized and evaluated. The combination of vincristine and carboplatin is commonly used as first-line therapy, with 5-year overall and progression-free survival rates of 86 to 97% and 39 to 61%. Carboplatin hypersensitivity is the most frequent adverse event[127-129], which can be effectively managed with pre-medication[32]. Ototoxicity is another issue that is important to monitor during treatment with platinum compounds. A combination of thioguanine, procarbazine, lomustine, and vincristine (TPCV) is another well-established chemotherapy regimen for progressive PLGG[125,130-132]. A prospective randomized clinical trial comparing outcomes of vincristine/carboplatin versus TPCV revealed that treatment with TPCV had a trend towards superior 5-year event-free survival (EFS) compared to vincristine/carboplatin (52% vs 39%, respectively), although this did not reach statistical significance[125]. However, the potential long-term morbidity associated with alkylating agents such as infertility and increased risk of secondary malignancy has led most oncologists to use vincristine/carboplatin as a first-line therapy over TPCV. Hematologic dyscrasias are other potential complications, especially of alkylating agents.

Alkylating agents have also been tested in combination with tamoxifen [133] or vinblastine [134] as well as in polychemotherapy regimens with other agents including procarbazine, cyclophosphamide, lomustine, vincristine, VP16 or 5-fluorouracile [135-138]. Monotherapy using temozolomide [139-142], vinblastine [143-145] or cyclophosphamide [146] have been used in progressive PLGGs with variable results in terms of outcome, depending on the ages of the children and the tumor locations enrolled in the studies.

Other protocols including vincristine/VP16 [147] or vincristine/carmustine [148] associated with intrathecal injection of methotrexate have shown 50-70% tumor control (defined as radiologic response or stable disease) in progressive PLGGs. Other chemotherapy regimens tested include vincristine alone [149], vincristine in combination with actinomycin [150], high dose ifosfamide [151], high dose cyclophosphamide [152], bleomycine [153], topotecan [154], idarubicin [155] or lenalinomide [156].

The anti-VEGF agent bevacizumab has recently been evaluated in combination with irinotecan for PLGG disease progression[157-159]. A recent phase II study which included 35 recurrent PLGGs reported at 2-year PFS of 47.8% using this treatment strategy [159]. Bevacizumab is generally well tolerated, however, patients need close monitoring for the development of hypertension or proteinuria and there are concerns for premature ovarian failure.

Radiation therapy was once standard-of-care for PLGG, however its use has decreased in PLGGs with increased awareness of its devastating long-term morbidities including cognitive deficits, increased risk of secondary high-grade malignancies, vasculopathy, endocrinopathy, and effects on growth [160-163]. Given the excellent overall survival of children with PLGG and the numerous available chemotherapy regimens, the use of radiation therapy for PLGG is generally avoided to minimize long-term and irreversible morbidity, and is used for those in whom disease control cannot be achieved with either surgery or chemotherapy (including targeted therapies). Several protocols using conformal external beam radiotherapy at doses between 50-59 Gy have been reported in the treatment of non-operable or progressive PLGGs with 5-year PFS ranging from 74%-88% [120,164-166]. Over last decade, through the advances in radiotherapy techniques, significant progress has been made in minimizing scatter doses to normal brain. These techniques include stereotactic conformational external radiotherapy [167-169], gamma-knife stereotactic radiotherapy techniques [170-172] and proton beam radiotherapy [173,174].

2. Interrogating tumor genomes

One fundamental aspect of research in oncology is to characterize the structural and sequences alterations of the genome in tumor cells and to identify how those genetic disruptions might influence tumor growth or deregulate the differentiation state of cells.

2.1. Distinction of copy-number variations (CNVs)

In the late 19th century, Sutton was one of the first scientists that documented the presence of chromosomes in grasshoppers. Later in the 1970's, early studies of tumor cells originating from biopsies provided the first insight into tumor cells that were essentially arrested in metaphase or prometaphase, where the DNA is the most condensed. Despite a resolution of a few megabases, karyotypes from these cells allow identification of certain chromosomal abnormalities such as aneuploidy, trisomy and the more subtle disruptions such as deletions or insertions. By providing an overview of the tumor genome, this technique is still used in clinical routine and has been extremely useful to stratify treatment decisions as specific chromosomal disruptions have been related to distinct patient outcome, especially in leukemia[175,176].

Later techniques using chip-based methods such as CGH (comparative genomic hybridization), array-CGH and SNP array, have dramatically improved the resolution of the detection of chromosomal imbalances[177]. These techniques are based on a competitive fluorescence in situ hybridization method. Two different sources of DNA, labeled with distinct colors of fluorophores are matched on each chromosome for identification of chromosomal differences between the two sources. Comparing to traditional techniques such as giemsa banding or fluorescence in situ hybridization (FISH), which are limited by the resolution of the microscope, those techniques provide a resolution from 5-10 megabases (CGH) to a locus-by-locus resolution of 100 kilobases (array CGH). Although conventional CGH and array CGH have been used abundantly to detect chromosomal imbalances, they are not able to detect structural alterations such as reciprocal translocations or inversions. SNP arrays, that also genotype common single nucleotide variants, have the advantage of being able to detect copy-neutral LOH (loss of heterozygosity) that can also be involved in oncogenesis. All those DNA chip-based techniques have been applied to a variety of adult and pediatric cancer subtypes providing a comprehensive description of copy-number profiles in those tumors [178,179] [180]. One major obstacle of those methods is the limitation of the resolution due to the number of probes or markers designed for the platform. Small alterations in the sequence (such as indels or short rearrangements) are not identified, as the size and the number of markers used does not span all nucleotides.

2.2. The development of DNA sequencing techniques and its challenges

The early DNA sequencing techniques were based on the chain termination method, developed by Frederic Sanger in the late 1970's. This method is based on the selective incorporation of dideoxynucleotides radioactively or fluorescently labeled for the detection. Although this method has been widely used in the last decades, it often provides poor quality of the first 15-40 bases due to the primer binding process. Additionally, as cloning of the DNA fragments is often used upfront the sequencing, one can find parts of the cloning vector in the resulting sequence. Pyrosequencing represents an alternative technique, based on sequencing by synthesis; it allows the sequencing of 100-400 megabases in a few hours.

The recent development of next-generation sequencing (NGS) represents a new avenue in the exploration of the structure of the cancer genome[181]. Through digital sequencing of both normal (mostly blood) and tumor genomes (issued from biopsies or surgical tumor resection), those techniques offer the possibility to detect, at a high-throughput level, somatic genetic alterations at a single base resolution. Whole genome sequencing (WGS) represents the most comprehensive genomic exploration[182]. Alternatively, target sequencing (of which whole-exome sequencing is a special case[183]) has recently been developed in order to reduce the costs of sequencing[184]. All those techniques generate millions of digital sequences called reads. In order to rebuild the tumor genome from the millions of reads generated several analytical methods have been developed. Multiple analytical steps are necessary to transform and normalize the raw reads (stored in a .fq.gz format) into a collection of mapped reads and their mapping location (stored in a .bam file format[185]). Digital sequencing results in multiple technical biases and artifacts that generate noise (biases due to flow cell, lane, dinucleotide context, machine cycles, generation of duplicate reads). The challenge of NGS interpretation is to reduce to noise in order to provide the best signal. Reads are sorted, normalized, cleaned and marked for duplicates. Most of those steps can be performed using SAMtools and Picard tools (URLs : samtools.sourceforge.net, picard.sourceforge.net). These preprocessing steps provide the best quality sequencing data that is crucial for further calling applications.

Genome Analysis Toolkit (GATK)[186] is a powerful computational kit to call somatic mutations.

2.3. Detection of somatic point mutations

Two major approaches have been developed to detect structural variants in an individual genome from NGS data; *de novo* assembly and classical alignment strategy[187]. The algorithms used in *de novo* assembly are trying to reconstruct the entire genome of the individual from overlaps between reads generated[188]. The assembled genome is then matched to the reference genome in order to call somatic variants. This computational intensive method requires a long reads in order to improve the assembly and provide high quality data. In the second strategy, currently the most commonly used, structural differences are found by aligning reads from the individual genome to the reference genome. Several bioinformatic tools such as BWA, BFAST, Maq have been developed to align the billions of reads generated to a reference genome[189-191]. Those approaches are more straightforward for the user and less labor intensive computationally. However, sensitive and specific detection of structural variations remains still a challenge due to both technological limitations and biological features of single variations. Errors in the production of reads, limited read lengths and insert sizes, sampling biases (such as CG- rich regions) are the most frequent limitations that are causing technical artifacts in DNA sequencing[192]. Biologically, the presence of numerous repetitive sequences is limiting the confidence of the mapping prediction of the algorithms.

After the assembly step, further algorithms are used to call point mutations. Noisy reads (with a low quality score and multiple mismatches) are ignored. Through statistical analyses, sites that are likely to carry somatic mutations with high confidence are identified[193]. These detection methods account for the fraction of discordant nucleotides between the sequence and its reference, the coverage (average number of reads representing a given nucleotide in the reconstructed sequence) at this site as well as the quality of the reads.

2.4. Detection of somatic indels (short insertions and deletions)

By comparing the tumor and the reference sequences, genomic regions that may harbor small insertions or deletions (above a specified threshold and having sufficient coverage) are called through statistical tools (GATK). In order to decrease the number of false positive single nucleotide variations caused by misaligned reads, particularly at the 3' end, a joint local-realignment of all samples from a same individual ("co- cleaning") is highly recommended (GATK). All sites potentially harboring small insertions or deletions in either the tumor or the matched normal are realigned in all samples.

2.5. Detection of chromosomal rearrangements

Tracking the exact breakpoints of small rearrangements or indels has become particularly feasible with whole genome sequencing (WGS). When performing paired-end sequencing of the tumors (both ends of a fragment are sequenced), depending on the coverage and the length of the reads it is possible to identify rearrangements by detecting the presence of discordant reads (the expected distance between two pairs of reads is higher or smaller than predicted from the fragment size, or the reads are disoriented) that support the presence of structural variations in the genome. Rearrangements are predicted from read pairs mapping to different chromosomes or to unexpected positions or orientations on the same chromosome. Further tools fish for reads spanning a fusion and maps the exact breakpoint using modified alignment algorithms. Each breakpoint is called and assigned with a score reflecting the number of tumor reads supporting a breakpoint, the fraction of nearby reads with MAPQ0, the prevalence of other nearby discordant pairs and the standard deviation of breakpoint starting positions[194].

2.6. Distinction of passenger and driver mutations

Since the development of NGS techniques, the number of tumor samples sequenced has exponentially increased[195]. This represents a huge resource in order to provide a comprehensive description of the genome of different tumor types. Indeed, one crucial aspect in oncology is to distinguish driver mutations, likely to cause directly the cells to transform and generate a tumor, from passenger mutations that occur randomly across the genome during the replication. Being able to identify driver mutations for each cancer type represents an important challenge in oncology as these alterations represent potential therapeutic targets. As passenger mutations tend to occur randomly across the genome, it becomes easier to reveal potential driver mutations when a same mutation is recurrent across multiple tumors. Several analytical tools have been recently developed in order to identify significant recurrent mutations across a wide panel of tumors that have been sequenced[196,197]. Briefly, the concept is to detect the mutations that occur in a set of individuals with a higher prevalence, above a mutation rate. One example of a tool developed by the Broad Institute is called Mutsig. The algorithm determines first the number of mutations and the number of covered bases for each gene. The counts are then broken down by mutation context category: four context categories discovered by MutSig, and one for indel and 'null' mutations, which include indels, nonsense mutations, splice-site mutations, and non-stop (read-through) mutations. For each gene, Mutsig calculates the probability of seeing the observed distribution of mutations or a more extreme one, given the background mutation rates calculated across the dataset[198]. These algorithms are continuously improved in order to increase the specificity to detect potential driver mutations. Several parameters such as the gene footprint length, the replication time, the level of expression of genes and the GC content have been identified as influencing the mutation rate and are therefore considered as confounders.

2.7. Detection of CNVs using NGS data

In sequencing files, once reads are normalized it is possible to detect structural variations using the copy-number ratio obtained through the coverage. Computational tools are combining a local change-point analysis and subsequent merging of adjacent chromosomal segments with similar copy number.

2.8. Analyses of the transcriptome

For RNA analysis, various techniques have been developed in the past to explore the transcriptome of tumors. The most frequent one was the microarray technology that measures the relative expression of target genes based on the hybridization of mRNA to probes designed on the chip. This cost-effective methodology has been applied widely at a large scale across multiple tumor types. However, an important limitation is related to the fact that this technique represents a semi-quantitative detection of the gene expression. Indeed, the signal is quantified by the amount of chemiluminescence or fluorophore emission activated by the amount of nucleid acid that binds to the target. The signal can thus be quickly saturated. Alternatively, an analytical technology, Serial analysis of gene expression (SAGE) and all the following derivative versions (LongSAGE, RL-SAGE, SuperSAGE) provide more a quantitative measure of the concentration of mRNA in a cell. This tag-based method allows also the potential discovery of new transcripts, as it is not based on a known target. These more costly methods are progressively replaced by RNA-sequencing that provides the most comprehensive information of the transcriptome of a cell at a particular time. Based on the abundance of cDNA reads sequenced, this method provides not only the quantitative gene expression level in an organism but is also able to detect alternative splice variants, somatic point mutations and rearrangements.

2.8.1. Distinguishing known classes

One of the important aspects of transcriptomic analysis is to identify specific patterns across the dataset. Supervised learning is one task that is commonly used in gene expression analyses. The task consists of analyzing a training set that produces an inferred function, which serves as a model to test new examples and test whether the new input vector is correctly predicted by the model and therefore belongs to the class. Several supervised learning algorithms are available such as support vector machines[199], linear regression[200], logistic regression and neuronal networks[201].

2.8.2. Detecting subclasses

A complementary approach used in gene expression analysis is unsupervised classification or clustering. Cluster analysis groups objects in clusters in such a way that objects in a cluster are more similar to each other than to those in other groups. Different approaches exist. Hierarchical clustering (also called connectivity models) are grouping samples based on their distance connectivity[202]. Centroid models (for example k-mean algorithms) defined each group by a single mean vector[203]. In the distribution models, clusters are defined by their statistical distribution[204]. Density models or subspace models are other types of algorithms used in unsupervised clustering[205]. By providing the number of subclasses in a specific dataset, those tools give an insight of the degree of heterogeneity in the dataset and encourages to explore further the molecular specificity of each subgroups.

2.8.3. Detecting pathway activation

Several tools are available to identify genes that are significantly overexpressed or underexpressed between specific groups, related to different phenotypes or defined by the unsupervised clustering methods. It is also possible to identify gene-sets related to a specific biological task that are enriched in a specific group. One example of such a method is called gene-set enrichment analyses (GSEA)[206]. This statistical tool identifies groups of genes (directly uploaded from the NCBI databases such as Biocarta and Gene Ontology) that are regulated together in a predefined group of samples.

2.9. Detecting mutations and rearrangements in RNA-seq data

Several analytical tools are currently developed to explore the architecture of the RNA[207-209]. RNA-seq analyses allow the detection of fusion genes as well as different splice variant in an usupervised approach.

Identification of all splice variants represents an important asset. All RNA-Seq reads identified as clusters supporting each splice variant are identified. Reads are then aligned to a reference genome containing annotated spliced transcript sequences and unspliced gene sequences. It is therefore possible to identify alternative splicing that occurs in the RNA sequence.

RNA-sequencing is also powerful to detect rearrangements as long as transcripts are expressed. Numerous bioinformatic packages have recently been developed in order to detect chimeric reads supporting rearrangements (FusionSeq, TopHat-Fusion, Fusion-Finder, FusionHunter,…). Those tools are based on the discovery of reads mapping two different genes and the presence of chimeric reads spanning the fusion between two genes.

Similar to DNA-sequencing, using a mapping algorithm, it is possible to identify candidate mutations by comparing the reads from the tumor to the reference sequence.

The detection of somatic mutations in RNA-seq data is possible if the gene is enough expressed so that it is characterized by a reasonable coverage.

As for DNA sequencing analyses, multiple technical artifacts (transcript size, GC content, and sequencing depth) interplay in the analyses of RNA-seq data and need to be optimized in order to provide better quality data especially to call translocations or somatic mutations with a high sensibility, despite the low quality of samples [210].

3. Genomic features in PLGGs

Very few reports in the literature describe karyotypes in PLGGs and those that have been characterized have been commonly reported to be normal[211]. A striking finding of PLGGs is the low number of genetic alterations present in the tumors. Early cytogenetic studies revealed almost normal diploid karyotypes across multiple subtypes of PLGGs[211-213]. The most frequent recurrent chromosomal alteration identified was a gain of chromosome 7, especially in PAs [212,214-216]. Other chromosomal structural abnormalities included gains of chromosome 4, 5, 6, 8, and 11 and deletion of 17p in a subset of PAs, inversion in chromosome 8, and loss of chromosome 1q[212,214-222].

In the late 2000's, the first copy-number profiles in PLGGs have been performed.

Several studies profiling essentially PAs using array-CGH have been performed and described a recurrent gain of chromosome region 7q34, representing the only detectable genomic alterations in these tumors[223-226]. Additional genomic profiling of diffuse PLGG subtypes, such as DAs, AGs and GGs revealed recurrent gain of chromosome MYBL1 and loss of MYB and CDKN2A/B[227].

Recent advances in high-throughput genetic sequencing and gene expression profiling have shed important insights into the genomic alterations of PLGGs[181,228,229].

One important limitation to these studies is the lack of sufficient tumor tissue from rarer subtypes of PLGGs such as tectal gliomas, thalamic and optic pathway tumors.

3.1. Neurofibromatosis 1 (NF1)

The increased risk of LGGs in children with NF1 germline mutation was one of the first clues that dysregulation of the mitogen-activated protein kinases (MAPK) pathway may be important in the pathogenesis of PLGGs. *NF1* encodes neurofibromin, which is ubiquitously expressed at variable levels in different tissue types during development. Structurally, neurofibromin contains a central domain homologous to Ras-GTPase-activating (Ras-GAP) proteins and acts as a negative regulator of the Ras-Raf-MEK-ERK pathway[230]. In neurofibromatosis, *NF1* mutations produce a loss of function of neurofibromin that leads to the constitutive activation of the Ras pathway and results in proliferation of astrocytes[32], among other phenotypes. Thus, MAPK pathway activation has long been known to contribute to the pathogenesis of LGGs in NF1 patients[231]. In addition, constitutive expression of MEK1 causes an increase in astrocytic proliferation.

3.2. BRAF duplication-fusions

Genetic rearrangements of the oncogene *BRAF* are the most common genomic alterations found in sporadic PLGGs. Early studies utilizing comparative genomic hybridization (CGH) identified a gain of the specific chromosomal region 7q34 containing the *BRAF* locus as the most frequent copy number alteration in PLGGs[32], involving 50-100% of pediatric PAs[225,232,233]. The *BRAF* duplication is found more frequently in cerebellar and hypothalamic-chiasmatic tumors[225].

The 7q34 gain has been characterized to represent a duplication of *BRAF* with a tandem insertion in the *KIAA1549* gene[32]. This BRAF duplication results in the activation of the downstream effectors of the MAPK pathway, MEK and ERK[223,226]. Subsequently, variants of the fusion transcript involving *BRAF* gene have been described, involving not only *KIAA1549* but also other fusion partners, *SRGAP3, FAM131B, MACF1*, *RNF130*, *CLCN6*, *MKRN1* and *GNAI1*[234-237] [228,229]. *RAF1*, which encodes a protein that leads to the stabilization and activation of *BRAF*, has also been described to harbor gene fusions with *SRGAP3* and *QK1,* leading to the constitutive activation of MAPK pathway[229,234,236]. These *BRAF* rearrangements tend to occur frequently in cerebellar lesions. Strikingly, all of the fusion protein variants are characterized by loss of the N-terminal inhibitory domains of BRAF, resulting in constitutive activation of the BRAF kinase and downstream activation of MAPK and its effectors, MEK and ERK.

Although the BRAF fusion protein has been shown to result in a tandem duplication of the *BRAF* locus, further studies are necessary to explain the precise mechanism by which the fusions contribute to the formation of tumor and the specific role of KIAA1549 and SRGAP3 segments within the BRAF fusion transcripts. One recent study reported that regions flanking the breakpoints of the *RAF* gene fusion are enriched with microhomologous sequences. This has led to the hypothesis that tandem duplications of the *RAF* gene might be generated by microhomology-mediated break-induced replication[238]. *In vitro* evaluation of the effect of the BRAF fusion protein has suggested that this protein has oncogenic properties and is able to activate the MAPK pathway. The short form of *KIAA1549-BRAF* fusion induces anchorage-independent growth in multiple cell lines[236,239]. Furthermore, pharmacologic inhibition of MEK1/2 in short-term cultured PLGG cell lines significantly diminishes cell proliferation[223], supporting a role of the MAPK pathway in promoting proliferation. Taken together, BRAF and RAF1 fusion transcripts, leading to constitutive activation of MAPK pathway, may play a crucial role in the pathogenesis of sporadic PAs and may also present potential therapeutic targets for PLGGs.

3.3. BRAF V600E and other less frequent mutations

Another frequent genomic alteration in PLGGs is the *BRAF* V600E mutation[169], which also results in deregulation of the MAPK pathway[32]. This mutation has been described in other cancer subtypes, including melanoma[240], colorectal cancer[241], leukemia[242], and high-grade gliomas[243]. *BRAF* is one of the most mutated genes in cancer[195]. The *BRAF* V600E point mutation occurs most commonly in PXAs, GGs, DAs, and PMAs[228,229,234,244-248] and is only rarely detected in PAs[249]. Thus *BRAF* duplications and V600E point mutation are almost always mutually exclusive. The *BRAF* V600E alteration confers constitutive BRAF kinase activation, and transforms NIH3T3 fibroblasts *in vitro*[236]. Other rare forms of small amino-acid insertions in *BRAF* have been identified in PAs[228]. The *BRAF* V600E mutation has been shown to promote transformation of human neural stem cells, followed by senescence[250]. However, it remains unclear whether this recurrent alteration is sufficient to drive the development of PAs.

3.4. Other mutations and rearrangements involving the MAPK pathway

Recent landmark sequencing projects including large cohorts of PLGGs identified recurrent genomic alterations in fibroblast growth factor receptor type 1 (FGFR1) [228,229]. *FGFR1* genomic alterations have also been described in breast cancer, lung cancer, and glioblastomas. *FGFR1* point mutations (N546K and K656E) were found in 5% of supratentorial PAs. Both mutations have been described to transform cells *in vitro*. In 2% of cases, *FGFR1* mutations were associated with the presence of a *PTPN11* mutation, another downstream effector of FGFR1[228]. In the same study, one PA possessed a tandem duplication of *FGFR1*. Importantly, gene expression analysis revealed that FGF2, a ligand of FGFR1, was significantly over-expressed in PAs compared to other astrocytic tumors, suggesting that the FGF/FGFR pathway alteration plays an important role in tumorigenesis of PLGGs. Additionally, *FGFR1* mutations and duplication of its tyrosine kinase domain have also been described in PAs, DAs, and DNTs[229].

Alterations of other MAPK members have also been described in PLGG. These include genomic alterations affecting the kinase domain of neurotrophic tyrosine kinase type 2 (NTRK2), which have been described in pediatric PAs[228]. Finally, *KRAS* activating mutations have also been described in 3-5% of sporadic PAs[228,229,234,251,252].

3.5. PI3K and RTK Signaling

After the MAPK pathway, the other most frequently altered pathways in PLGGs include the phosphatidylinositol 3-kinase (PI3K)/AKT/mammalian target of rapamycin (mTOR) pathway, the epidermal growth factor receptor (EGFR) pathway, sonic hedgehog (SHH) signaling, and the vascular endothelial growth factor (VEGF) signaling pathway.

PI3K is an intracellular protein that is recruited to the cell membrane after stimulation of a transmembrane growth receptor such as EGFR or platelet derived growth factor receptor A (PDGFRA - which also signals along the Ras-Raf-MEK-MAPK pathway), resulting in activation of downstream effectors, such as AKT and mTOR, to induce cell proliferation and inhibition of apoptosis. As initially suggested by early studies of tuberous sclerosis, activation of mTOR through mutations of its upstream inhibitor result in increased predisposition for PLGGs, in particular the SEGA subtype. In a series of PLGG, 44% of tumors were demonstrated to have evidence of PI3K/Akt/mTOR pathway activation[32]. Over-expression of the BRAF-fusion transcript in neural stem cells results in activation of mTOR pathway, leading to the formation of glioma-like lesions and further supports the cross communication between these two pathways[253]. Additionally, the deregulation of Rheb and further mTOR activity in TS patients is another important insight for the role of MAPK pathway in PLGGs as mTOR pathway is connected to the MAPK pathway. In contrast, MEK1/2 knockdown in mice results in the absence of glial cell differentiation and proliferation[254].

Activation of the EGFR pathway has been shown in a small series of PLGGs. Comparative genomic hybridization and fluorescent in situ hybridization (FISH) studies of six disseminated PLGGs demonstrated EGFR amplification, while none was observed in a cohort of localized tumors. This led to speculation that deregulation of the EGFR pathway may play a role in the pathogenesis of disseminated PLGGs[255]. Additionally, rare mutations of PDGFRA have been reported in PAs, GGs, and LGG-NOS tumors[248].

Although the sonic hedgehog pathway is most commonly associated with tumorigenesis of medulloblastoma and high-grade gliomas[256,257], a recent study suggests that this pathway could play a role in a subset of pediatric PAs via the over-activity of *PTCH*[258]. In this series of 20 pediatric PAs, 45% of tumors demonstrated over-expression of *PTCH* mRNA. Interestingly, a significant inverse correlation between *PTCH* expression level and patient age suggests that the SHH pathway is more frequently activated in young patients.

Finally, the potential role of angiogenesis is highlighted through studies involving the VEGF pathway, one of the major signaling pathways in cancer biology, contributing to neovascularization which is essential for tumor growth[32]. Comparative analysis of vessel architecture in 59 pediatric PAs and adult high-grade gliomas showed that vessel immaturity and instability are present in both tumor types[259]. Another study of 17 pediatric PAs demonstrated immunohistochemical reactivity for activated VEGF receptors. However, further validation studies are necessary to confirm altered VEGF signaling in pediatric PAs.

3.6. Transcription factors

Genomic alterations affecting key transcription factors have been described in PLGGs. These include *MYB* amplification in DAs and focal deletions of *MYB* in AGs[260]. *MYB* is an oncogene that is mutated or altered in T-ALL[261,262], breast cancer, pancreatic cancer, and CNS tumors, including primitive neuroectodermal tumors and medulloblastoma[263,264]. In PLGGs, *MYB* expression has been shown to be upregulated in a proportion of diffuse LGGs (60%) and PAs (41%). Its role in the normal development of the CNS and tumorigenesis remains unknown.

More recently, a novel recurrent genetic rearrangement involving another member of the MYB transcription factor family, *MYBL1*, was identified in a cohort of grade II DAs and AGs[229,227]. Importantly, this specific duplication-truncation of *MYBL1* has demonstrated tumorigenic properties *in vitro*.

3.7. Epigenetic alterations

Aberrant epigenetic regulation has been increasingly described in human cancers and has become a major focus in a number of pediatric cancers [265]. Epigenetic regulation of the genome can be defined as heritable modifications in gene expression that do not directly affect the DNA sequence[266]. Epigenetic modifications include multiple mechanisms affecting the chemical properties of DNA, histones, or other proteins involved in DNA packaging[267]. The frequency of alterations in epigenetic modifiers in cancer has been shown in multiple cancer types including hematologic tumors[268,269], Wilms' tumors[270], retinoblastoma[271], neuroblastoma, thyroid carcinoma, hepatocellular carcinoma, sarcoma[272], and brain tumors such as medulloblastoma[273] and Atypical teratoid rhabdoid tumors (ATRTs) with *SMARCB1* mutations [274,275].

The evidence that epigenetics is a major factor in pediatric glioma biology is extremely strong. Direct mutations in the chromatin modifier H3F3A have been described in pediatric GBMs[276] as well as DAs and PAs[228,229]. This suggests that dysregulation of chromatin remodeling effectors are also acting with genomic alterations in the tumorigenesis of a small subset of PLGGs. Other genomic alterations include HIPK2 genomic gains and increased mRNA expression level in a subset of sporadic PAs arising from the cerebellum[232,277] and BCR gene rearrangement in one PMA[278].

3.8. Prognostic implications

Recently attempts have been made to correlate specific genomic alterations to clinical outcome with controversial results. A multivariate analysis of 146 patients reported that the presence of the KIAA1549-BRAF fusion protein was the most significant favorable prognostic factor in pediatric PAs following subtotal resection[279]. Another study including 106 PLGGs, most of which were sporadic PAs, showed no statistical superior progression-free survival rates among tumors with the *BRAF* duplication compared to the wild-type tumors[280]. The observation that *BRAF* duplicated tumors behave differently than the others remains an open question, especially with the recent discovery of new BRAF fusion types that might have biased the previous studies. Further larger and controlled or prospective analyses are needed to address this question. It has been hypothesized that improved outcome in PAs conferred by the BRAF duplication may be due to oncogene-induced senescence (OIS), which occurs through the activation of p16Ink4a pathway[281]. OIS is a mechanism of tumor suppression that has been implicated in other cancer subtypes[282]. In contrast, p16 deletion has been identified as a negative prognosticator in 198 PLGG[283]. This remains to be further validated. Similarly, a recent study performed on GGs has showed that the presence of the V600E point mutation was associated with significant lower recurrence-free survival[284]. The recent discovery of other genomic alterations such as *FGFR1* mutations will also enlarge the field of exploration between clinical outcome and biology.

3.9. PLGG alterations differ from the adults

Genetic alterations in pediatric LGGs differ from adult LGGs. Concomitant deletion of chromosome 1p and 19q is one of the most frequent recurrent genetic alterations in adult oligodendrogliomas, aiding in diagnosis as well as serving as a favorable prognostic marker[285,286]. In contrast, concomitant deletion of chromosome 1p and 19q is rare in children with oligodendrogliomas[287,288], and does not confer similar chemosensitivity when present[289]. Similarly, mutations in TP53, a tumor suppressor gene that codes for a nuclear phosphoprotein and regulates cycle cell arrest, apoptosis, and genetic stability, are frequently found in adult but rarely in pediatric LGGs[290,248,291-294]. *IDH1* and *IDH2* mutations are also rarely observed in PLGGs while they are frequent in adults. In a recent study examining *IDH1* and *IDH2* in 445 CNS tumors and 494 non-CNS tumors, *IDH1/2* mutations were described to occur with a frequency of more than 70% in adult patients across a variety of glial tumors including low-grade astrocytomas, anaplastic astrocytomas, oligodendrogliomas and oligoastrocytomas and secondary glioblastomas derived from the lower-grade gliomas[295]. In contrast, *IDH1/2* mutations are rare in children, although when found in adolescent patients they may be a harbinger of the adult form of the disease, meriting concordant treatment recommendations[296,297].

4. Goal of the study

Our study was articulated around two major goals. In the first study we performed a wide DNA sequencing analysis in order to reveal potential new driver mutations, especially in more rare histological subtypes. We wanted to extend the recent effort of DNA sequencing in PLGG tumors to increase our power to detect new potential driver mutations. Based on the recent sequencing studies, we can reasonably hypothesize that *BRAF* and *FGFR1* alterations likely drive a large fraction of PLGGs. However, there is still a fraction of tumors that exhibit neither *BRAF* nor *FGFR1* mutations (approximately 10% of supratententorial tumors). By increasing the number of samples sequenced, we increase the probability to detect new potential driver mutation in other PLGG subtype[195]. We also applied algorithms developed at the Broad Institute in order to validate independently the previous findings published and extend the genomic landscape by adding our local cohort of tumors.

The other study was a large RNA expression profiling study. By a transcriptomic approach, we wanted to understand how the clinical heterogeneity of PLGGs was reflected at a molecular level and whether we could identify potential transcriptional patterns that might inform us about the potential lineage of origin of those tumors.

Chapter 2: The Genetic Landscape of PLGGs

1. Introduction

In Chapter 1 we described the various genomic alterations in PLGGs as well as the new opportunities brought by the next-generation sequencing technology. One of the most important parameters in the detection of driver events from genomic sequencing analyses is to obtain enough power to detect potentially new recurrent mutations[298]. In this section we describe a meta-analysis of the genetic alterations in PLGGs which we performed by integrating next generation sequencing data from the two major PLGG sequencing projects that had been recently published[229,299], combined with our local cohort which includes samples obtained from the Dana-Farber Cancer Center/Boston Children's Hospital Cancer and Blood Disorder Center (DFCI) and others obtained through a collaboration with Children's Hospital of Philadelphia (CHOP) and Necker Hospital in Paris (NEM). By increasing the number of samples, especially the non-PA tumors, we have been able to discover new recurrent mutations that might be involved in oncogenesis of PLGGs.

2. Methods

2.1. Data downloading and sequencing

Our local cohort of tumors (DFCI and CHOP) was sequenced at the BGI platform at CHOP. For WGS, DNA was randomly fragmented and libraries were prepared for paired-end sequencing on an Illumina HiSeq 2000. Exome hybrid capture was performed with NimbleGen SeqCap EZ Exome Library SR v2.2, Agilent. The other sequencing files were obtained through recently published datasets (one from the German Cancer Center – DKFZ in Germany[299] and the other one from the Pediatric Cancer Genome Project –PCGP, led by St-Jude children's research Hospital and Washington, University of St Louis[229]).

2.2. WGS/WES analyses pipeline

Raw data from recently published datasets obtained from normalized bam files [229,299] were generated using the Picard SamToFastq command. The preprocessing steps are summarized in Figure 4. Read pairs were aligned to the reference genome hg19 (Build 37) using the Burrows-Wheeler Aligner (bwa) with options −q 5 −l 32 −k 2 −o 1[190].

Figure 5 summarizes the different downstream calling analyses we performed. Copy-number alterations were evaluated using SegSeq[300]. We then identified significantly recurrent copy-number alterations by adding to segmented copy-number data pseudo-markers every 10,000 bases and analyzing these by GISTIC 2.0[301]. Somatic point mutations and short indels were called using Mutect and IndelLocator[193]. In order to identify genes that are significantly recurrently mutated above a background mutation rate, we used the Mutsig algorithm developed at the Broad Institute[197]. Rearrangements and their exact breakpoints were identified using a combination of dRanger and BreakPointer algorithms[194]. These analyses were performed within the Firehose environment[302].

Figure 4:

Summary of the WGS/WES preprocessing steps performed for all samples

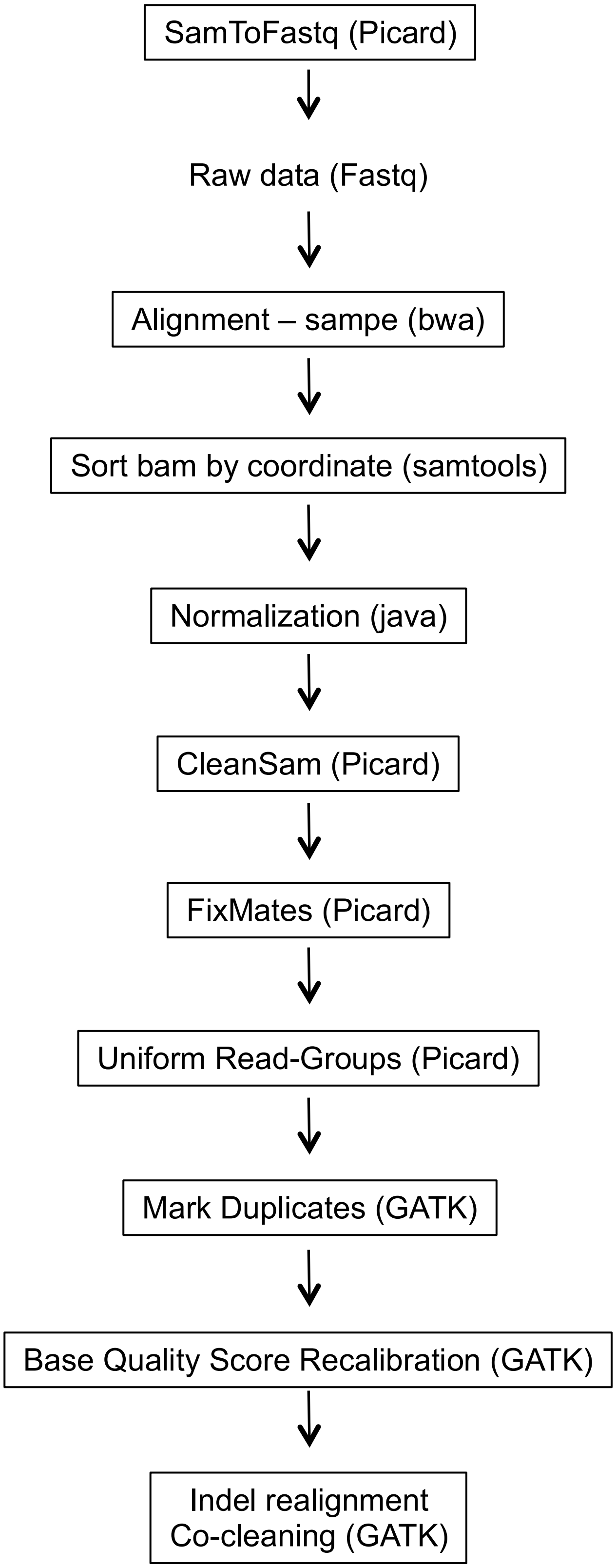
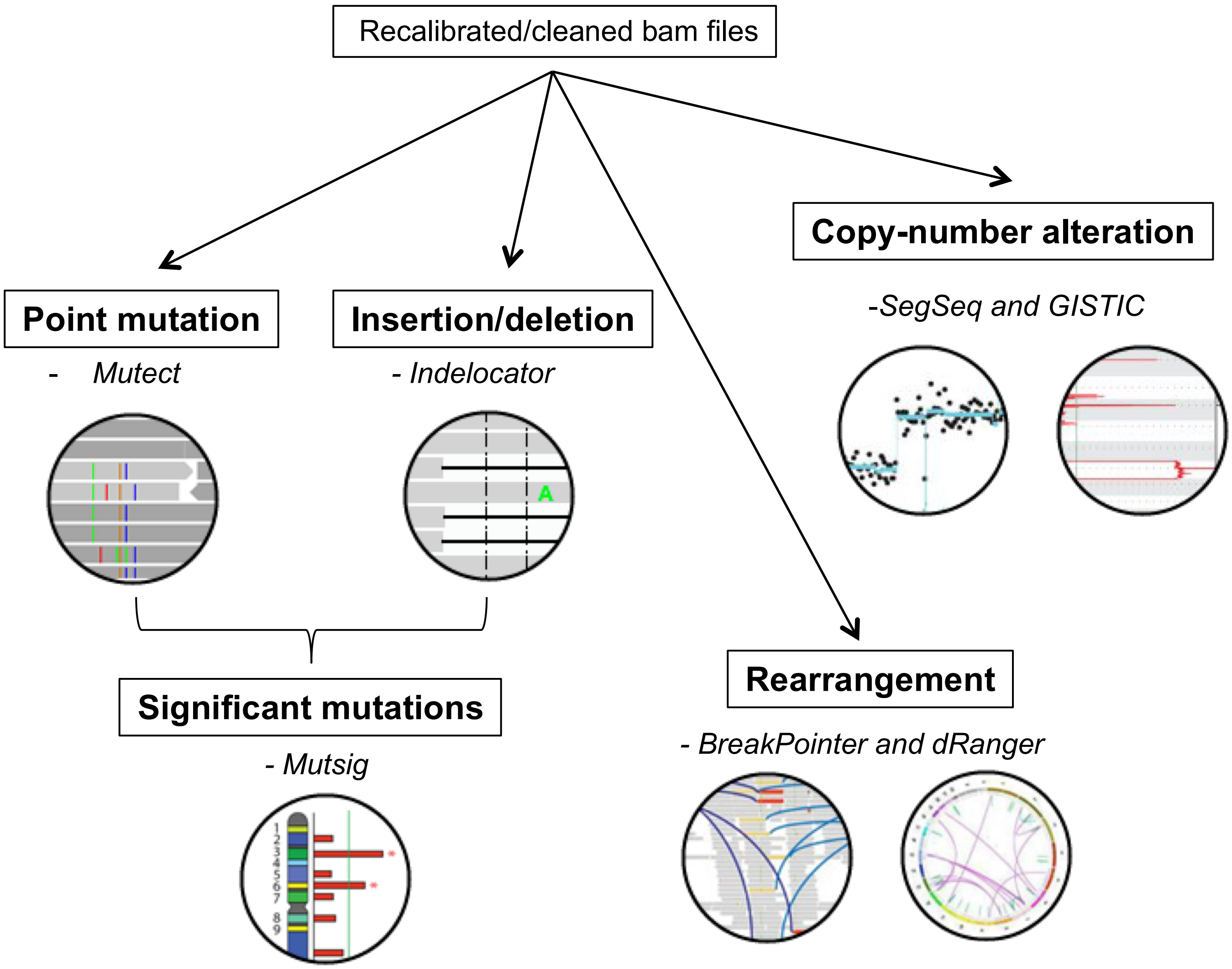


Figure 5:

Summary of all the downstream applications performed on the WGS files with the algorithms names used within the Firehose environment

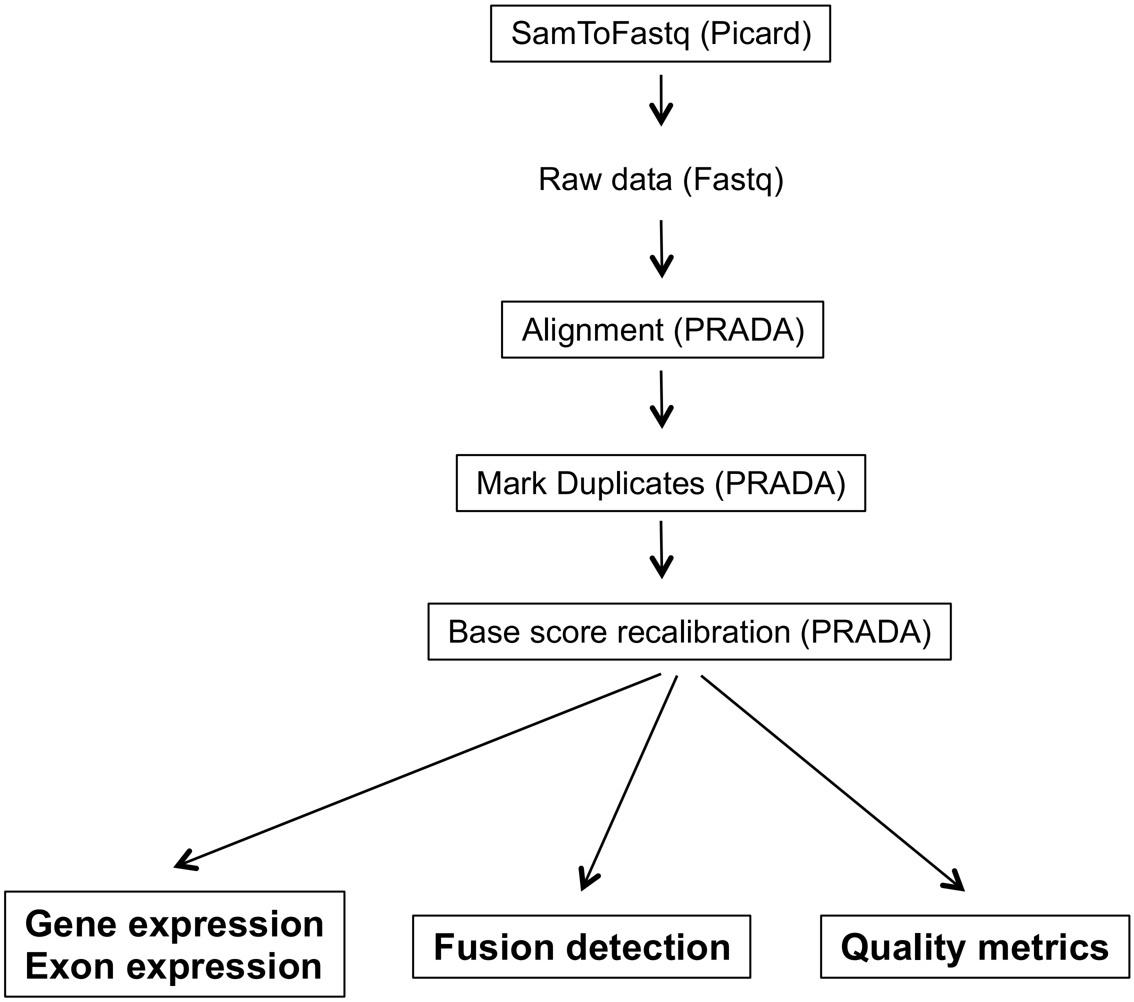


2.3. RNA-sequencing analysis pipeline

For RNA-sequencing, total RNA was quantified using the Quant-iT™ RiboGreen® RNA Assay Kit and normalized to 5ng/ul. An aliquot of 200ng for each sample was transferred into library preparation. This method uses oligo dT beads to select mRNA from the total RNA sample followed by heat fragmentation and cDNA synthesis from the RNA template. The resultant cDNA then goes through library preparation using the Broad designed indexed adapters substituted in for multiplexing. After enrichment the libraries were quantified with qPCR using the KAPA Library Quantification Kit for Illumina Sequencing Platforms and then pooled equimolarly. Pooled libraries were normalized to 2nM and denatured using 0.1 N NaOH prior to sequencing. Flowcell cluster amplification and sequencing were performed according to the manufacturer's protocols using either the HiSeq 2000 or HiSeq 2500. Each run was a 76bp paired-end with an eight-base index barcode read. Data was analyzed using the Broad Picard pipeline, which includes de-multiplexing and data aggregation. The preprocessing steps are summarized in Figure 6. The other RNA-sequencing files were downloaded from recently published datasets. Downloaded RNA-seq bam files were transformed in fastq files using the Picard SamToFastq algorithm. Raw paired-end reads were aligned to the reference genome hg19 and preprocessed using the PRADA (Pipeline for RNA-sequencing Data Analysis)[208]. Gene-expression levels, exon expression levels, quality metrics as well as detection of fusion transcripts were used from the PRADA software platform within the Firehose environment.

Figure 6:

Summary of the preprocessing steps and downstream applications performed on the RNA-seq files



3. Results

3.1. Description of the cohort

We analyzed the genomes of 169 samples, including 138 that had undergone WGS (one of which had also undergone WES) and 53 that had undergone RNA-seq (of which 22 samples were also analyzed by WGS). These included data from 88 tumors previously published in a survey of PAs (88 WGS only)[299], 57 tumors (25 PAs/57) previously published in a survey of various PLGG subtypes (15 by WGS only, 21 by RNA-seq only and 21 both by WGS and RNA-seq), [229], and 24 tumors (0 PAs/24) whose data have not been published (13 by WGS only, 10 by RNA-seq only and 1 by both).

Clinical and pathological information of the cohort are summarized in Table 3. The largest group of tumors was PAs (112 tumors, or 67%). The other tumor subtypes were DAs (24 tumors; 14%), angiocentric gliomas (AGs) (6 tumors; 3.5%), oligoastrocytomas (OAs) (5 tumors; 3%), oligodendrogliomas (ODs) (4 tumors; 2.4%), pilomyxoid astrocytomas (PMAs) (4 tumors; 2.4%), and gangliogliomas (GGs), dysembryoplastic neuroepithelial tumors (DNTs), and pleomorphic xanthroastrocytomas (PXAs) (one of each; 0.4% each). Eight tumors (4.9%) were classified as PLGG not otherwise specified (NOS) and three tumors (1.6%) were not documented in terms of histology. Histological central review was not possible for the samples originated from DKFZ or St-Jude.

The median age in the whole cohort was 6 (range, 0.4-21).

Overall, the coverage in the three cohorts was comparable with a mean of total fraction of callable bases of 89% (range, 75.6-90.6), although the sequencing depth used by the centers was different (Figure 7).

Table 3:

Clinical information of the cohort analyzed related to the sequencing type

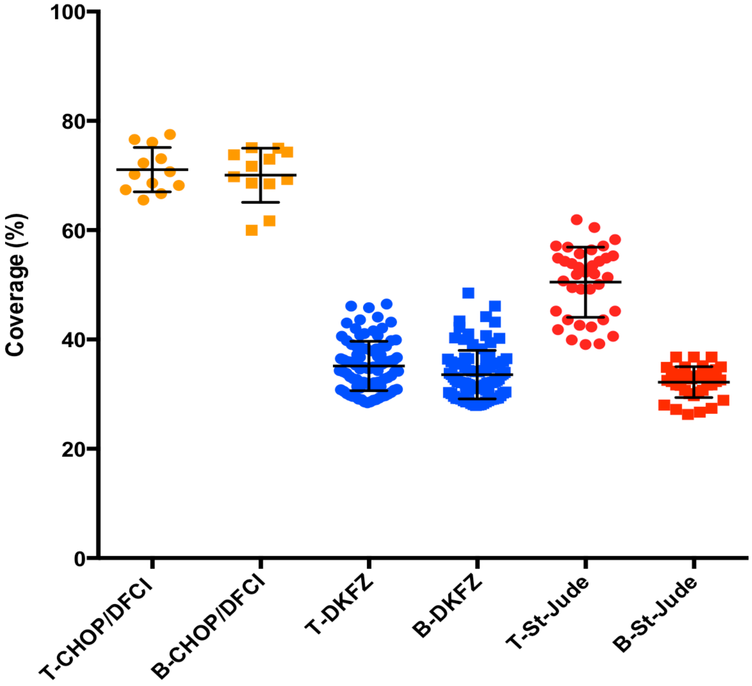
Sample ID	Age	Location	Histology	Sequencing
PLGG DFCI/CHOP 1	16.1	IT	DA	DNA
PLGG DFCI/CHOP 2	2.1	IT	DA	DNA
PLGG DFCI/CHOP 3	5.6	Pons	DA, grade II/III	DNA
PLGG DFCI/CHOP 4	10.5	Brainstem	DA	DNA
PLGG DFCI/CHOP 5	11.8	CBL	DA pilo features	DNA
PLGG DFCI/CHOP 6	14.8	CBL	DA pilo features	DNA
PLGG DFCI/CHOP 7	3.3	Temporal Lobe	DA	DNA
PLGG DFCI/CHOP 8	1.8	CBL	DA	DNA
PLGG DFCI/CHOP 9	21,8	ST	DA	DNA
PLGG DFCI/CHOP 10	13	ST	NOS	DNA
PLGG DFCI/CHOP 11	ND	ST	DA	DNA
PLGG DFCI/CHOP 12	ND	ST	DA	DNA
PLGG DFCI/CHOP 13	ND	Parietal Lobe	AG	DNA
PLGG DFCI/CHOP 14	3	ST	AG	DNA/RNA
PLGG DFCI/CHOP 15	ND	CBL	DA	RNA
PLGG DFCI/CHOP 16	ND	Optic chiasm	ND	RNA
PLGG DFCI/CHOP 17	ND	Optic chiasm	AG	RNA
PLGG DFCI/CHOP 18	ND	CBL	AG	RNA
PLGG DFCI/CHOP 19	ND	4th ventricle	NOS	RNA
PLGG DFCI/CHOP 20	ND	CBL	NOS	RNA
PLGG DFCI/CHOP 21	ND	CBL	PMA	RNA
PLGG DFCI/NEM 22	5	4th ventricle	NOS	RNA
PLGG DFCI/NEM 23	ND	3rd ventricle	ND	RNA
PLGG DFCI/NEM 24	2	CBL	NOS	RNA
PLGG DKFZ 1	3	Optic chiasm	PA	DNA
PLGG DKFZ 2	1	4th ventricle	PA	DNA
PLGG DKFZ 3	4	Optic nerve	PA	DNA
PLGG DKFZ 4	7	CBL	PA	DNA
PLGG DKFZ 5	15	Hypothalamus	PA	DNA
PLGG DKFZ 6	3	CBL	PA	DNA
PLGG DKFZ 7	11	CBL	PA	DNA
PLGG DKFZ 8	14	CBL	PA	DNA
PLGG DKFZ 9	5	CBL	PA	DNA
PLGG DKFZ 10	4	CBL	PA	DNA
PLGG DKFZ 11	4	Brainstem	PA	DNA
PLGG DKFZ 12	2	CBL	PA	DNA
PLGG DKFZ 13	3	Brainstem	PA	DNA
PLGG DKFZ 14	3	Thalamus	PA	DNA
PLGG DKFZ 15	8	3rd ventricle	PA	DNA
PLGG DKFZ 16	12	Brainstem	PA	DNA
PLGG DKFZ 17	14	CBL	PA	DNA
PLGG DKFZ 18	4	CBL	PA	DNA
PLGG DKFZ 19	11	CBL	PA	DNA
PLGG DKFZ 20	9	Brainstem	PA	DNA
PLGG DKFZ 21	9	CBL	PA	DNA
PLGG DKFZ 22	15	CBL	PA	DNA
PLGG DKFZ 23	3	Brainstem	PA	DNA
PLGG DKFZ 24	6	Brainstem	PA	DNA
PLGG DKFZ 25	9	Thalamus	PA	DNA
PLGG DKFZ 26	3	Thalamus	PA	DNA
PLGG DKFZ 27	6	CBL	PA	DNA
PLGG DKFZ 28	10	Optic chiasm	PA	DNA
PLGG DKFZ 29	5	CBL	PA	DNA
PLGG DKFZ 30	9	Thalamus	PA	DNA

PLGG DKFZ 31	13	CBL	PA	DNA
PLGG DKFZ 32	5	CBL	PA	DNA
PLGG DKFZ 33	4	3rd ventricle	PA	DNA
PLGG DKFZ 34	17	Brainstem	PA	DNA
PLGG DKFZ 35	4	3rd ventricle	PA	DNA
PLGG DKFZ 36	16	CBL	PA	DNA
PLGG DKFZ 37	6	CBL	PA	DNA
PLGG DKFZ 38	6	CBL	PA	DNA
PLGG DKFZ 39	2	CBL	PA	DNA
PLGG DKFZ 40	9	Brainstem	PA	DNA
PLGG DKFZ 41	14	Brainstem	PA	DNA
PLGG DKFZ 42	6	4th ventricle	PA	DNA
PLGG DKFZ 43	14	CBL	PA	DNA
PLGG DKFZ 44	12	CBL	PA	DNA
PLGG DKFZ 45	8	CBL	PA	DNA
PLGG DKFZ 46	9	Brainstem	PA	DNA
PLGG DKFZ 47	4	CBL	PA	DNA
PLGG DKFZ 48	11	CBL	PA	DNA
PLGG DKFZ 49	4	CBL	PA	DNA
PLGG DKFZ 50	6	CBL	PA	DNA
PLGG DKFZ 51	8	3rd ventricle	PA	DNA
PLGG DKFZ 52	3	CBL	PA	DNA
PLGG DKFZ 53	12	3rd ventricle	PA	DNA
PLGG DKFZ 54	21	CBL	PA	DNA
PLGG DKFZ 55	4	CBL	PA	DNA
PLGG DKFZ 56	7	CBL	PA	DNA
PLGG DKFZ 57	15	Parietal Lobe	PA	DNA
PLGG DKFZ 58	8	CBL	PA	DNA
PLGG DKFZ 59	9	4th ventricle	PA	DNA
PLGG DKFZ 60	1	4th ventricle	PA	DNA
PLGG DKFZ 61	15	CBL	PA	DNA
PLGG DKFZ 62	7	CBL	PA	DNA
PLGG DKFZ 63	13	Brainstem	PA	DNA
PLGG DKFZ 64	4	CBL	PA	DNA
PLGG DKFZ 65	5	CBL	PA	DNA
PLGG DKFZ 66	5	CBL	PA	DNA
PLGG DKFZ 67	12	3rd ventricle	PA	DNA
PLGG DKFZ 68	14	CBL	PA	DNA
PLGG DKFZ 69	9	CBL	PA	DNA
PLGG DKFZ 70	20	CBL	PA	DNA
PLGG DKFZ 71	7	CBL	PA	DNA
PLGG DKFZ 72	3	CBL	PA	DNA
PLGG DKFZ 73	14	Brainstem	PA	DNA
PLGG DKFZ 74	11	Optic chiasm	PA	DNA
PLGG DKFZ 75	10	CBL	PA	DNA
PLGG DKFZ 76	17	Brainstem	PA	DNA
PLGG DKFZ 77	2	Brainstem	PA	DNA
PLGG DKFZ 78	7	Hypothalamus	PA	DNA
PLGG DKFZ 79	12	IT	PA	DNA
PLGG DKFZ 80	1	IT	PA	DNA
PLGG DKFZ 81	2	DE	PA	DNA
PLGG DKFZ 82	11	DE	PA	DNA
PLGG DKFZ 83	5	ST	PA	DNA
PLGG DKFZ 84	5	ST	PA	DNA
PLGG DKFZ 85	1	DE	PA	DNA
PLGG DKFZ 86	7	ST	PA	DNA
PLGG DKFZ 87	5	ST	PA	DNA
PLGG DKFZ 88	8	IT	PA	DNA
PLGG SJUDE 1	4	IT	PA	DNA/RNA
PLGG SJUDE 2	6	IT	PA	DNA/RNA
PLGG SJUDE 3	1	DE	PA	DNA/RNA
PLGG SJUDE 4	0.5	DE	PA	DNA/RNA

PLGG SJUDE 5	5	ST	DA	DNA/RNA
PLGG SJUDE 6	1	ST	OA	DNA/RNA
PLGG SJUDE 7	1	DE	PMA	DNA/RNA
PLGG SJUDE 8	17	IT	PA	DNA/RNA
PLGG SJUDE 9	6	DE	PA	DNA/RNA
PLGG SJUDE 10	3	IT	PA	DNA/RNA
PLGG SJUDE 11	12	IT	PA	DNA/RNA
PLGG SJUDE 12	13	ST	PA	DNA/RNA
PLGG SJUDE 13	7	ST	PA	DNA
PLGG SJUDE 14	NA	ST	ND	DNA
PLGG SJUDE 15	10	IT	PA	DNA/RNA
PLGG SJUDE 16	14	ST	PA	DNA
PLGG SJUDE 17	15	ST	PA/PMA	DNA/RNA
PLGG SJUDE 18	5	DE	PA	DNA
PLGG SJUDE 19	5	IT	PA	DNA
PLGG SJUDE 20	4	DE	PA	DNA
PLGG SJUDE 21	6	ST	PA	DNA
PLGG SJUDE 22	3	Brainstem	PA	DNA
PLGG SJUDE 23	5	DE	PA	DNA/RNA
PLGG SJUDE 24	0.4	ST	NOS	DNA
PLGG SJUDE 25	10	DE	PA	DNA
PLGG SJUDE 26	1	Spinal cord	PA	DNA
PLGG SJUDE 27	5	ST	PA	DNA
PLGG SJUDE 28	3	ST	DA	DNA/RNA
PLGG SJUDE 29	11	ST	GG	DNA
PLGG SJUDE 30	9	CBL	PA	DNA
PLGG SJUDE 31	9	ST	DA	DNA
PLGG SJUDE 32	15	ST	OD	DNA/RNA
PLGG SJUDE 33	5	ST	DA	RNA
PLGG SJUDE 34	3	ST	DA	DNA/RNA
PLGG SJUDE 35	2	ST	DA	DNA/RNA
PLGG SJUDE 36	13	ST	PXA	DNA/RNA
PLGG SJUDE 37	10	ST	PA	DNA/RNA
PLGG SJUDE 38	8	ST	DA	RNA
PLGG SJUDE 39	9	ST	PA/PMA	RNA
PLGG SJUDE 40	6	ST	OA	RNA
PLGG SJUDE 41	11	ST	DA	RNA
PLGG SJUDE 42	4	ST	DA	RNA
PLGG SJUDE 43	11	ST	AG	RNA
PLGG SJUDE 44	4	ST	OD	RNA
PLGG SJUDE 45	4	ST	NOS	RNA
PLGG SJUDE 46	13	ST	DNT	RNA
PLGG SJUDE 47	6	ST	OD	RNA
PLGG SJUDE 48	5	ST	OD	RNA
PLGG SJUDE 49	5	ST	NOS	RNA
PLGG SJUDE 50	3	ST	AG	RNA
PLGG SJUDE 51	4	ST	DA	RNA
PLGG SJUDE 52	4	ST	OA	RNA
PLGG SJUDE 53	6	ST	DA	RNA
PLGG SJUDE 54	3	ST	PA	RNA
PLGG SJUDE 55	13	ST	DA	RNA
PLGG SJUDE 56	14	ST	OA	RNA
PLGG SJUDE 57	2	ST	OA	RNA

Figure 7:

Summary of the coverage for the WGS samples in the three cohorts



(T=tumor, B=blood).

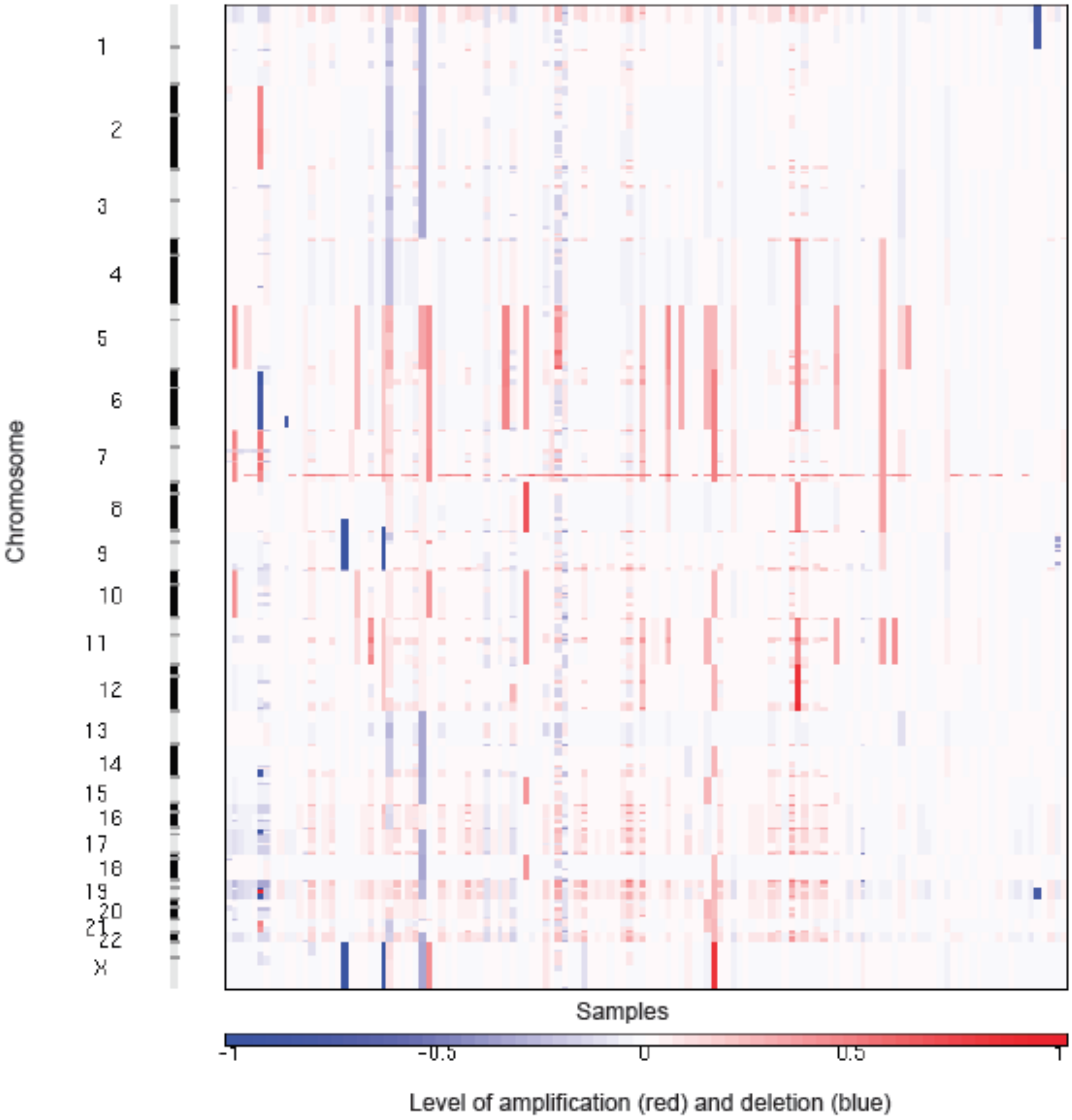
3.2. Structural alterations in PLGGs

High depth sequencing of PLGGs revealed that these tumors harbor very few somatic genetic events. Using the copy-number ratio provided by WGS sequencing we were able to explore the somatic copy-number alterations (SCNA) of the tumors. As previously reported, we observed that PLGGs are very silent in terms of CN alterations. Indeed, we observed in our cohort that only 2.4% of the genome was affected by broad copy-number alterations (spanning more than half of a chromosome arm) and that 0.15% of the genome was involved in focal copy-number alterations.

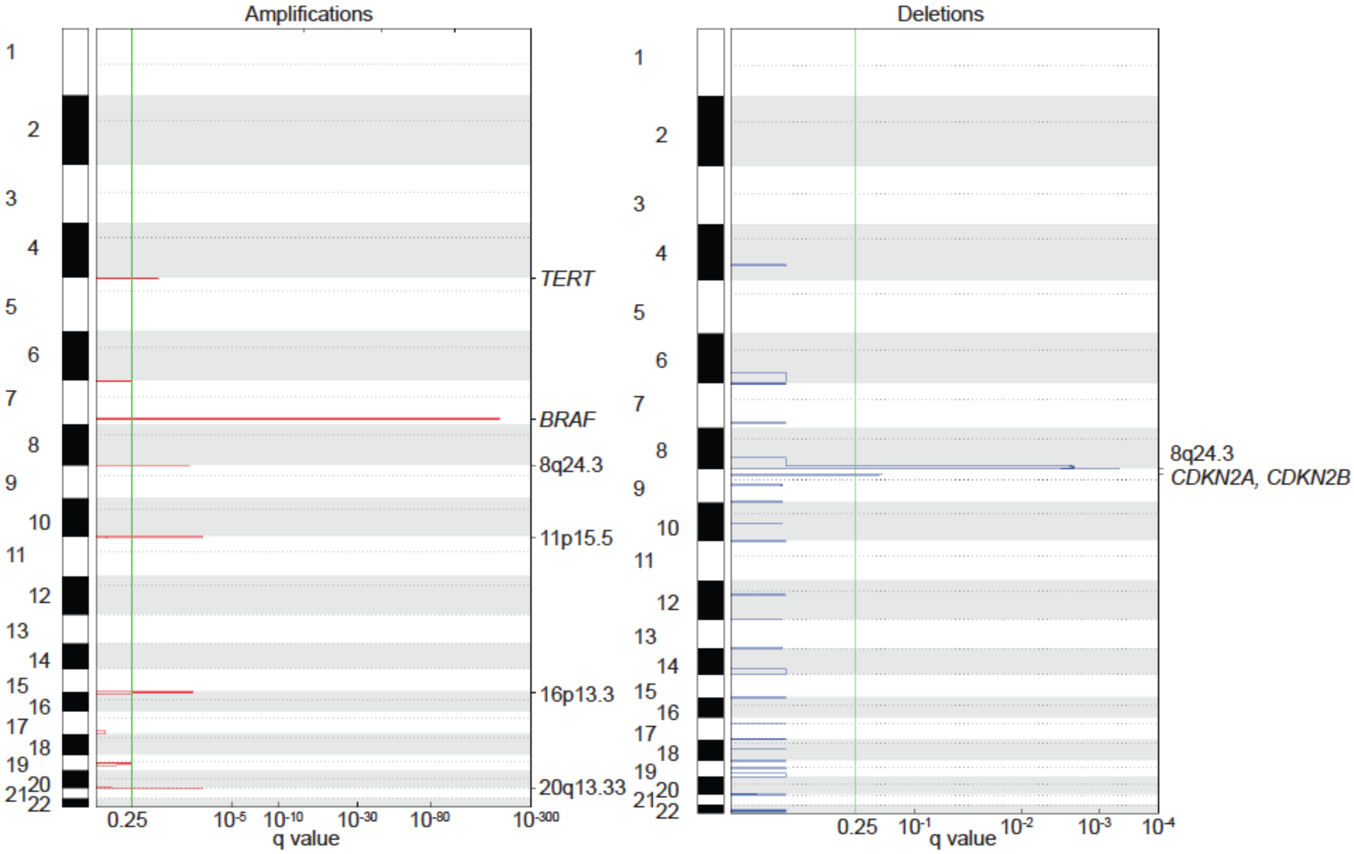
GISTIC analysis revealed that the most frequent copy-number alterations were amplifications in chromosome 5, 6, 7 and 11 and deletions in chromosome 6, 9 and 1p (Figure 8A). We observed 6 significant focal copy-number gains and 2 focal deletions (Figure 8B). The most significant focal amplification peak was observed in chromosome 7q34, which corresponds to the BRAF duplications. This event is so recurrent that it can be distinguished in the raw copy-number profile. The second most significant focal amplification peak encompassed 34 genes on chromosome 5p15 including *TERT*. In the focal deletions peak, we observed at the 9p21.3 region that was significantly deleted, *CDKN2A* and *CDKN2B* were the 2 genes among only 5 genes reported.

Figure 8: Copy-number profiles of the WGS cohort

A. Raw copy-number heat-map of the WGS cohort



B. Focal amplifications (left panel) and deletions (right panel) represented by GISTIC



3.3. Exploration of the somatic mutations in the cohort

PLGGs exhibit low frequency of somatic mutations. Among the 138 samples analyzed by WGS, the median of the mutation rate was 1.02×10^{-5} mutations/base (range 4.16×10^{-6} to 2.91×10^{-5} mutations/base) (Table 4). In the exonic regions, the most frequent mutation category observed was the transitions of C to G or A (Table 5). A median of one rearrangement was detected per sample (range 0 – 59).

The overall spectrum of mutations and rearrangements in our cohort is presented in Figure 9.

We observed genetic alterations in 152 of the 169 tumors (90%). 127 tumors described rearrangements (75%). We were unable to detect recurrent or known driver alterations in 17 of the 169 tumors, all of which were profiled with RNA-sequencing alone, most likely resulting from lower sensitivity of detection.

Mutsig analysis performed in our WGS cohort revealed 5 significant genes mutated in at least two samples (Table 6).

Table 4:

Summary of the breakdown of exonic mutations by type in the WGS cohort

Type of mutation	**n**
Frame Shift Del	32
Frame Shift Ins	10
In Frame Del	12
In Frame Ins	17
Missense Mutation	808
Nonsense Mutation	44
Nonstop Mutation	2
Silent	402
Splice Site	40
Translation Start Site	12
TOTAL	***1379***

Table 5:

Summary of the breakdown of exonic mutations by category in the WGS cohort

Category	n	Rate per Mb
*CpG->T	200	8.2 e-07
*Cp(A/C/T)->T	146	7.4 e-08
C->(G/A)	275	1.2 e-07
A->mut	199	9.5 e-08
indel+null	157	3.6 e-08
Double null	0	0
Total	**977**	***2.2 e-07***

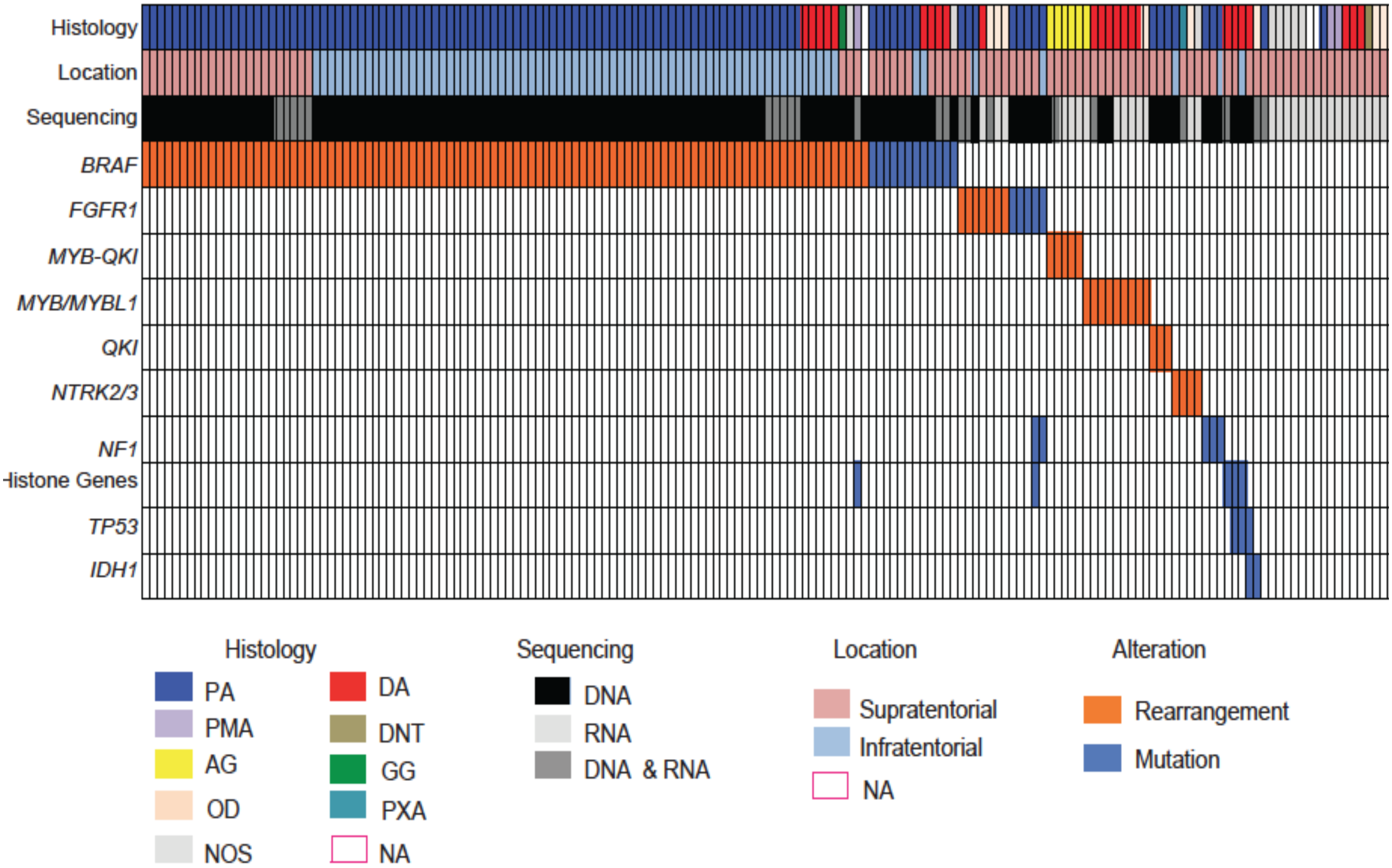
Table 6:

Summary of the significant recurrent mutated genes identified by Mustig in the WGS cohort

GENE	DESCRIPTION	n	p	q
BRAF	v-raf murine sarcoma viral oncogene homolog B1	10	0	0
FGFR1	fibroblast growth factor receptor 1	6	0	0
H3F3A	H3 histone, family 3A	4	0	0
NF1	neurofibromin 1	5	0	0,003
TP53	tumor protein p53	3	0	0,007

Figure 9: Landscape of somatic mutations in 169 PLGGs

Co-mut plot representing the driver alterations identified in our cohort by whole genome sequencing (DNA) and/or RNA-sequencing (RNA). Each column is representing a sample. Histological subtypes include Pilocytic Astrocytoma (PA). Pilomyxoid Astrocytoma (PMA), Angiocentric Glioma (AG), Oligodendroglioma (OD), Diffuse Astrocytoma (DA), Dysembryoplastic Neuroepithelial Tumor (DNT), Ganglioglioma (GG), Pleomorphic Xanthoastrocytoma (PXA), PLGG not otherwise specified (NOS). Tumors for which histology is unavailable are designated NA.



3.3.1. BRAF is the most frequently altered gene in PLGGs

The gene most frequently altered was *BRAF*. Duplications were present in 98 samples (58%), including 89 PAs (representing 80% of the PAs overall: 66/71 or 93% of the IT PAs and 23/41 or 56% of the ST PAs), as well as in 5/24 DAs and in 1/4 PMAs (this sample is originated from the St-Jude cohort and the histology was not centrally reviewed by our group). Table 7 summarizes the 7 different types of *BRAF* duplications related to their different histological subtypes. We confirmed that the most frequent *BRAF* duplication was *BRAF-KIAA1549* (92%), involving different exons of *KIAA1549* and *BRAF*. All the BRAF fusions, however, were characterized by loss of the 3' end of *BRAF* containing the regulatory domain, suggesting a gain of function mutation.

Table 7: Summary table of the 7 different BRAF duplication variants

Details	Number	Histology
KIAA1549-BRAF	90	84 PA, 4 DA, 1 PMA, 1 ND
RNF130-BRAF	2	PA
FAM131B-BRAF	2	1 PA, 1 DA
MKRN1-BRAF	1	PA
GNAI1-BRAF	1	PA
BRAF-MACF1	1	GG
FXR1-BRAF	1	NOS

12 tumors showed 4 different mutations in *BRAF*. The hotspot V600E point mutation[245,246] was observed in 7 tumors (4 PAs and 3 DAs, 6/7 located in supratentorial regions). Three tumors exhibited a previously observed p.599 599T>TT insertion[303] (one DA, one NOS tumor and 1 PA), one ST PA described an insertion of the three amino acids VLR at position 506 as described[299] and one tumor (ST PA) was characterized by a new mutation, p.E451D (Figure 10).

FGFR1 was the second most frequently altered gene. We observed 12 mutated samples in our cohort: 6 missense mutations affecting 5 PAs (Figure 9 and 10), 4 duplications within the exonic region of the gene, and 3 duplication-inversions between *FGFR1* and *TACC1* (one identified by WGS and the two other one by RNA-sequencing). This spectrum of mutations is affecting the Tyrosine Kinase domains (TK1 and TK2), leading to a potential hyper activated *FGFR1* protein. We also found one rearrangement involving *FGFR3* with *TACC4* in one ST DA.

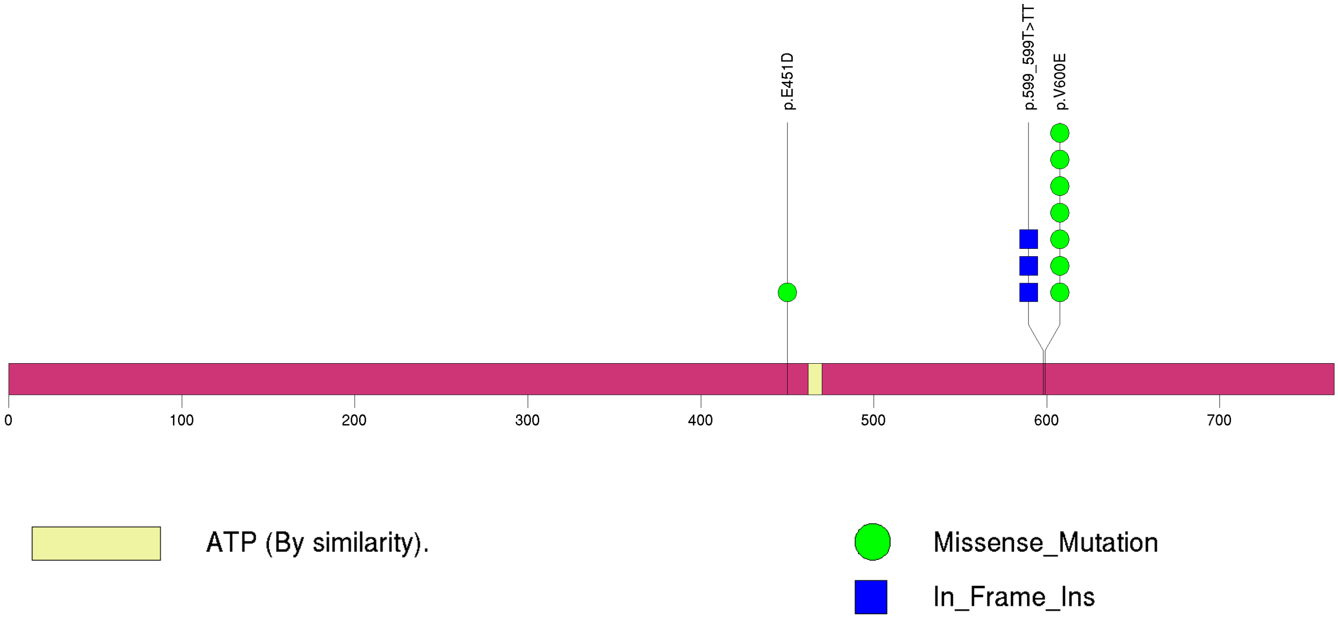
NTRK2 rearrangements were present in 4 tumors (3 PAs and one OA). In all 4 tumors, the alterations were mutually exclusive to *BRAF, FGFR1* or other potential oncogenic mutations (Figure 9). Two tumors carried the same rearrangement between the exon 15 of *NTRK2* and the exon 6 of *QKI*. In the other tumors *NTRK2* was fused to the 3' end of *NAV1* and *NACC2*. We also identified a rearrangement involving *NTRK3* and *NAV6* in one PXA tumor analyzed by RNA-sequencing. None of the *NTRK2* rearrangements were similar to those recently described in pediatric high-grade gliomas. Conversely, *ETV6-NTRK3* mutation has been described in one pediatric gliobastoma although the breakpoint sites among *ETV6* and *NTRK3* were different.

NF1 was mutated in 5 tumors (one tumor carried two mutations in the same gene) (Figure 10).

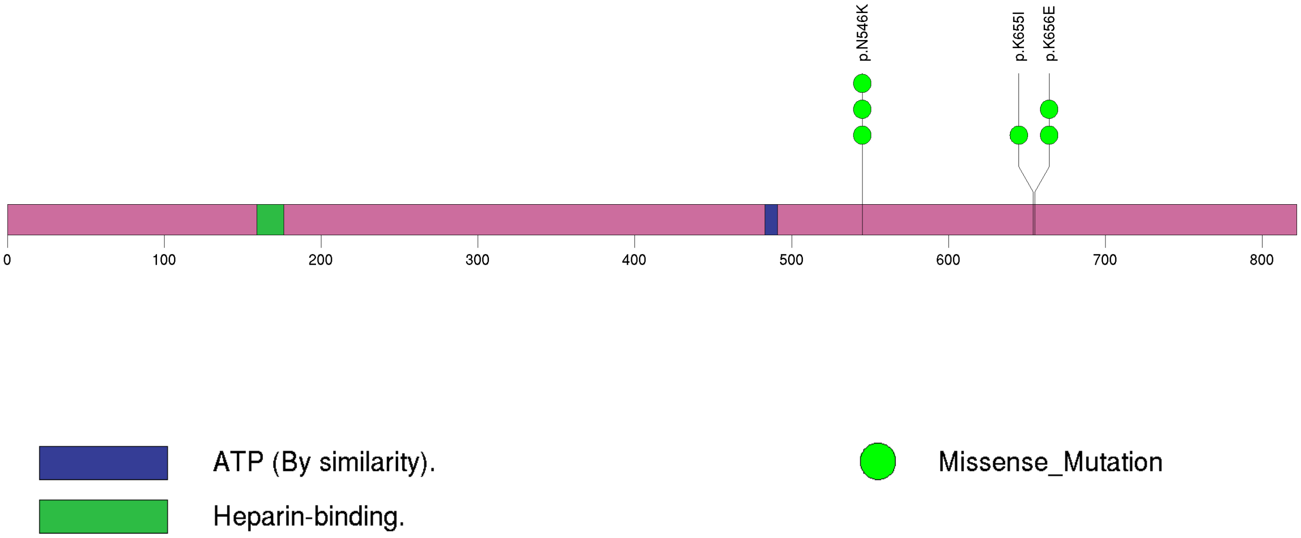
We observed mutations in genes encoding histones in five PLGGs, 2 PAs/Pilomyxoid Astrocytomas (PMAs) and 3 DAs (Figure 9). 4/5 were *H3F3A* mutation, also described in pediatric high-grade gliomas and one *HIST1H3B*. 1/5 mutation was A115G.

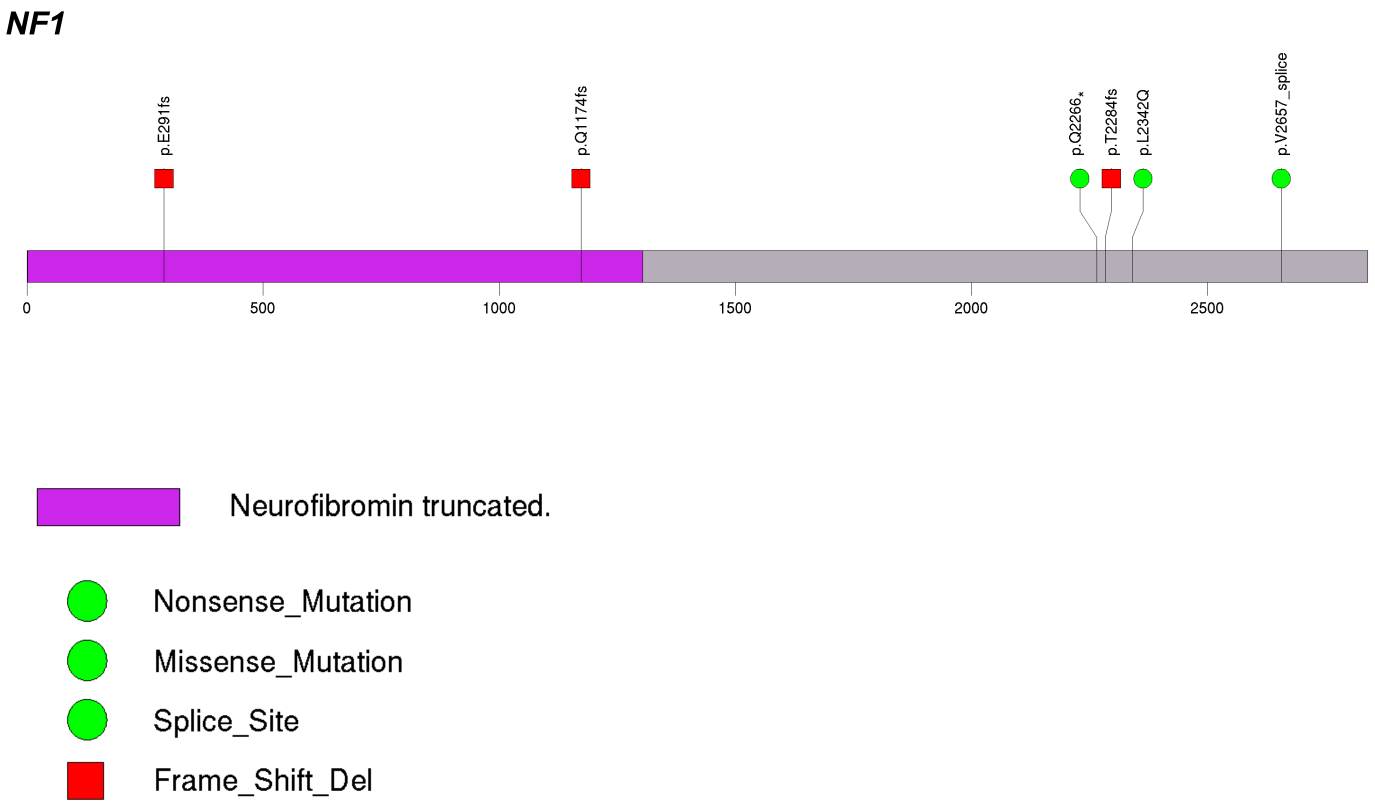
Figure 10: Mutation sites in the most frequently mutated genes of the WGS cohort
For each gene represented (with the number of amino-acids), the different type of mutations are summarized.

BRAF



FGFR1

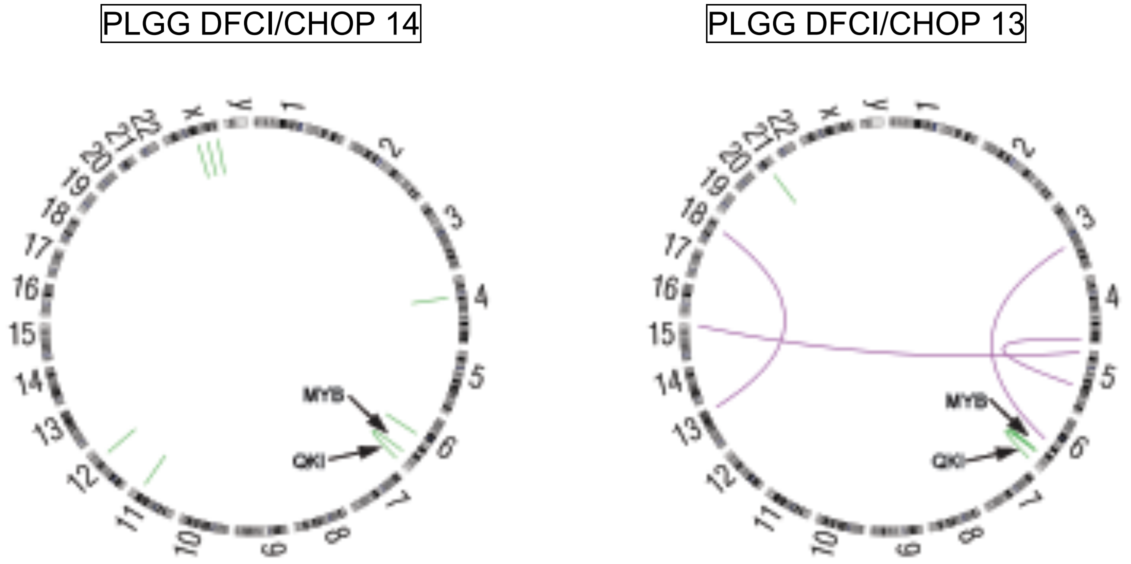



NF1
p.E291fs
p.Q1174fs
p.Q2266*
p.T2284fs
p.L2342Q
p.V2657_splice
0
500
1000
1500
2000
2500
Neurofibromin truncated.
Nonsense_Mutation
Missense_Mutation
Splice_Site
Frame_Shift_Del

3.3.2. MYB-QKI rearrangement is the unique recurrent mutation in Angiocentric Gliomas (AGs)

PAs as well as other known histologic subtypes have been characterized so far to harbor multiple genomic events. Here we show that among the 4 DFCI/CHOP AG tumors analyzed, we identified a recurrent translocation between *MYB* and *QKI* genes (Figure 9). Additionally, one AG from the PCGP cohort showed the similar *MYB-QKI* translocation. In our 2 tumors analyzed by WGS, we were able to identify discordant reads using dRanger (Figure 11A). Using the copy-number ratio we were then able to visualize the location of the breakpoints in MYB and *QKI* (Figure 11B).

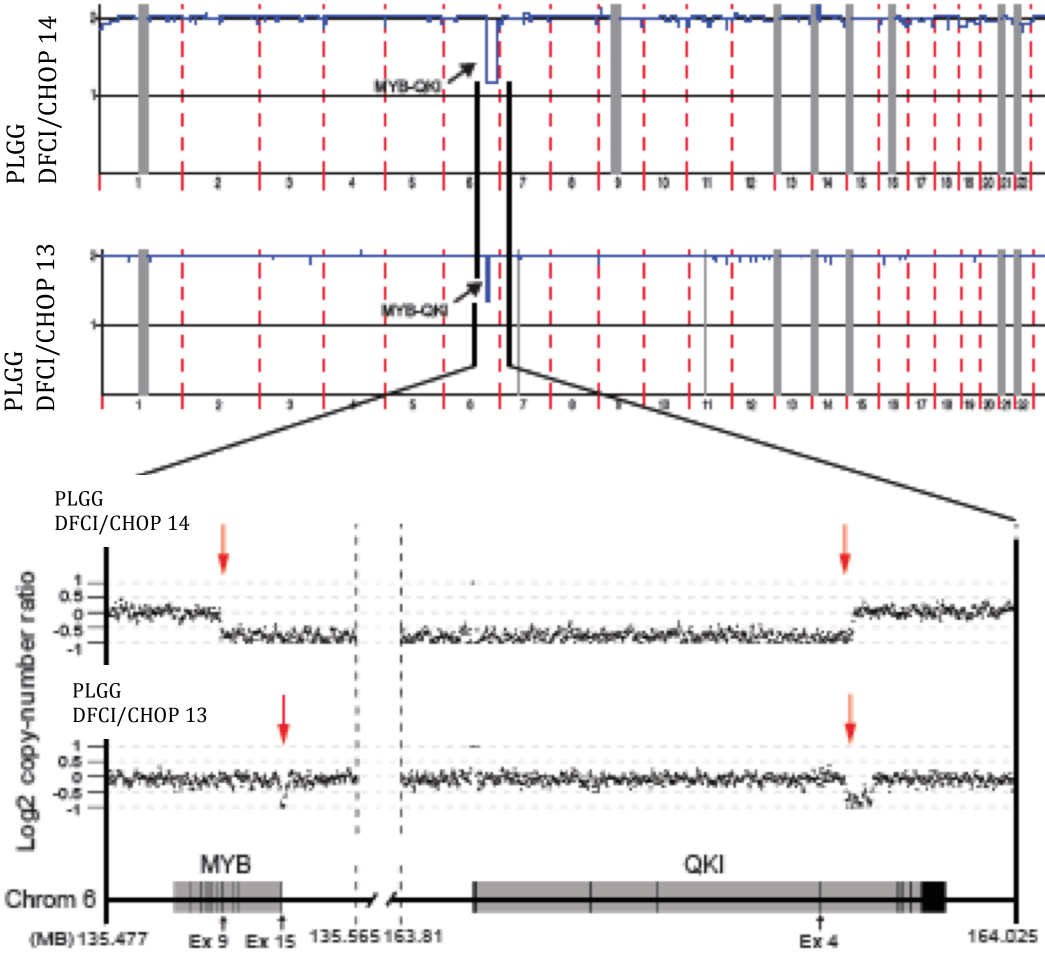
Figure 11: MYB-QKI rearrangements identified in two AG tumors by WGS analysis A. Rearrangements in PLGG DFCI/CHOP 14 (left panel) and PLGG DFCI/CHOP 13 (right panel). Circos plots are showing intrachromosomal (green) and interchromosomal (purple) rearrangements.


PLGG DFCI/CHOP 14

PLGG DFCI/CHOP 13

B. Copy-number profiles in PLGG DFCI/CHOP 14 and PLGG DFCI/CHOP 13

Top panels are representing the whole genome for both samples. The two bottom panels are focusing the Log-2 copy number ratio (obtained from the WGS data) for the samples for *MYB* and *QKI* genes.



For the first tumor (PLGG DFCI/CHOP14), the breakpoints were located between exon 9 and 10 of *MYB* and between exon 4 and 5 of *QKI*, resulting in a in-frame protein fusion between exon 9 of *MYB* and exon 5 of *QKI* (Figure 12). For the second tumor (PLGG DFCI/CHOP 13), we observed a more complex rearrangement involving *EYA4* with *MYB* and *QKI* (Figure 12), resulting in the loss of the last exon of *MYB* fused to the exon 5-7 of *QKI*. The four first exons of *QKI* were fused to the 11 first exons of *EYA4* (Figure 12). We also confirmed the *MYB-QKI* rearrangement in PLGG DFCI/CHOP 13 using RNA-sequencing by identifying reads spanning the junction between the 3’ end of MYB to the 5’ extremity of QKI (Figure 12).

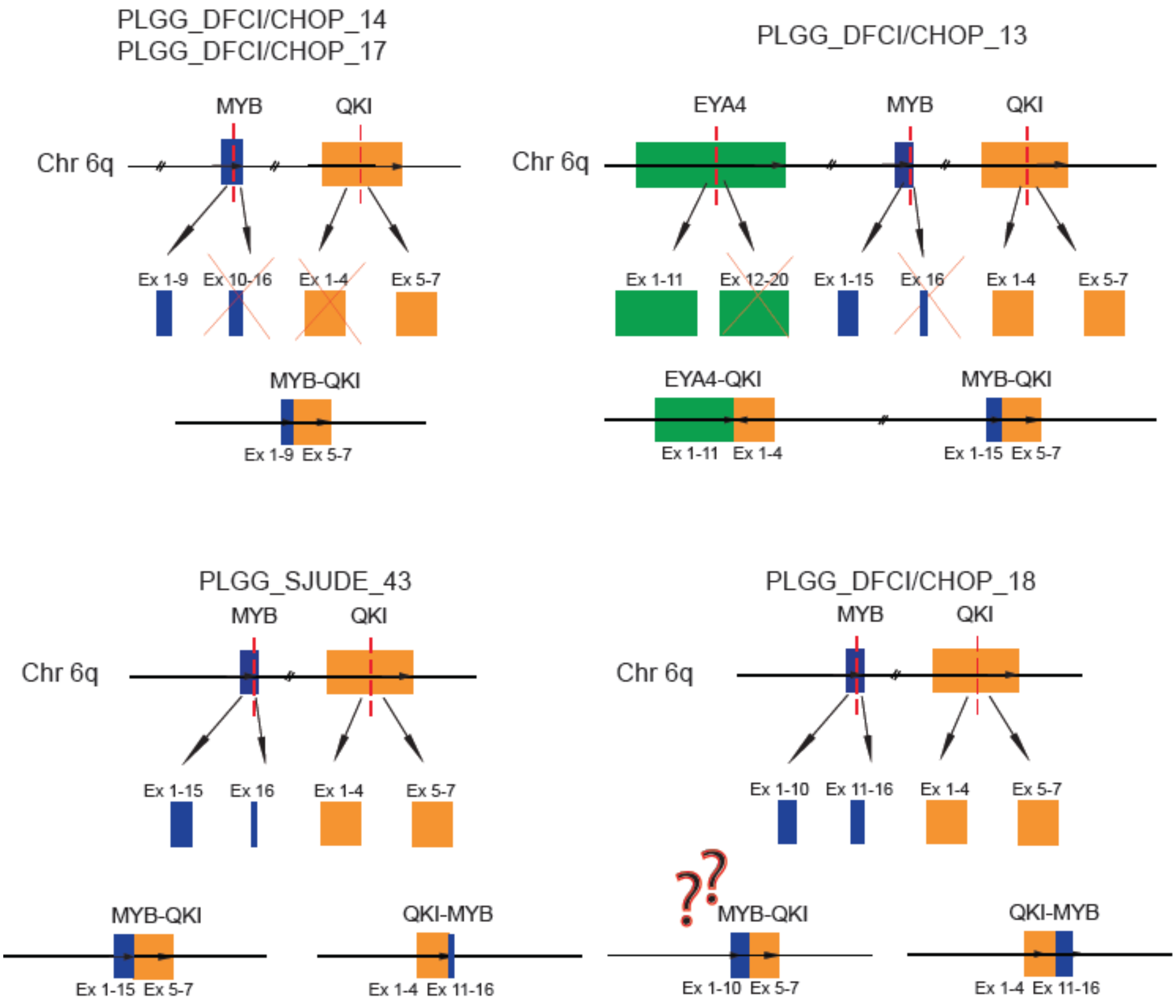
We also performed RNA-sequencing in an additional set of 2 typical AGs (PLGG DFCI/CHOP 17 and PLGG DFCI/CHOP 18) and observed the same rearrangement occurring between the 5’ end of MYB and the 3’ end of QKI in PLGG DFCI/CHOP 17 with the exact same breakpoints in the exons of MYB and QKI as in PLGG DFCI/CHOP 13. In PLGG DFCI/CHOP 18, we were able to identify the recurrent fusion reads spanning the QKI-MYB junction (between the 5’ end of exon 4 of QKI fused to the 3’ end of MYB) (Figure 13A).

In total, using WGS and RNA-seq, among the 6 AGs, five tumors showed a MYB-QKI rearrangement. We identified 4 different types of MYB-QKI rearrangements, all resulting into a loss of the 3’ end of MYB and the 5’ part of QKI except for one tumor where we were not able to identify the reciprocal fusion transcript.

Figure 12:

Schematic representation of the MYB-QKI alterations identified in our cohort

For each of the four panels, *MYB* and *QKI* breakpoints are represented by the red dashed lines. For PLGG DFCI/CHOP 13, *EYA4* breakpoint is represented with the red dashed line. Exons that were not expressed in the RNA-sequencing data are represented by a red cross. For PLGG DFCI/CHOP 18, no MYB-QKI forward reads were identified in the RNA-seq data.



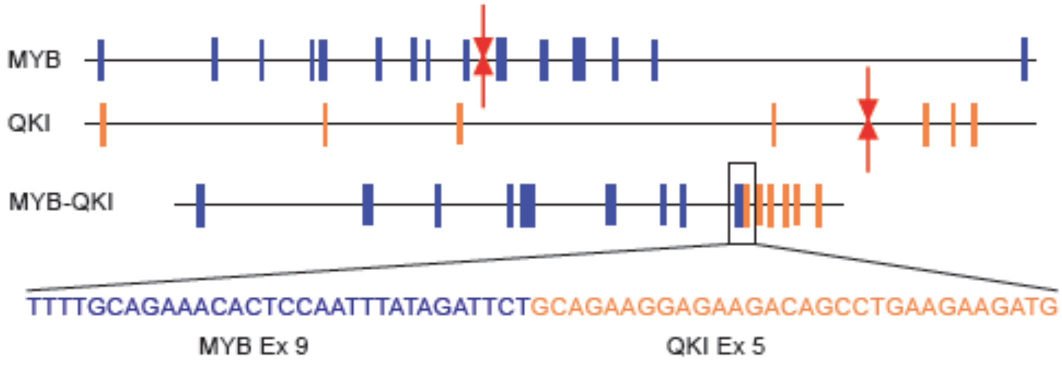
We further analyzed using RNA-seq the expression of the different exons of MYB and QKI, relative to matched normal brain expression (Brainspan dataset). For one AG tumor (PLGG DFCI/CHOP 13), we observed a dramatic decrease of expression after exon 9 of MYB, relative to normal brain, corresponding to the exact site of the breakpoint in that tumor (Figure 13B). For the 7 exons in QKI, we did not notice a significant change in the expression of the different exons, relative to normal brain. For the 2 other MYB-QKI tumors that we analyzed by RNA-seq, we did not observe the same pattern of decrease of the expression in the exon of MYB after the breakpoint. The expression of the QKI exons was always stable, relative to normal brain. This difference might be due to the fact that QKI is expressed at a low level in normal brain tissues, whereas MYB might only be expressed in the embryonic stage with no significant level of expression during childhood.

Besides *MYB-QKI* rearrangement, we also identified nine additional tumors with a *MYB* family alteration (Figure 9). One ST DA was characterized by a *MYBL1-MMP16* rearrangement as previously described[227] and eight ST tumors (6 DAs, 1 DNT and 1 AG) showed different rearrangements between *MYB* and several other partners as previously described in the PCGP cohort[229]. We also observed rearrangements involving *QKI* but not *MYB* in three PAs.

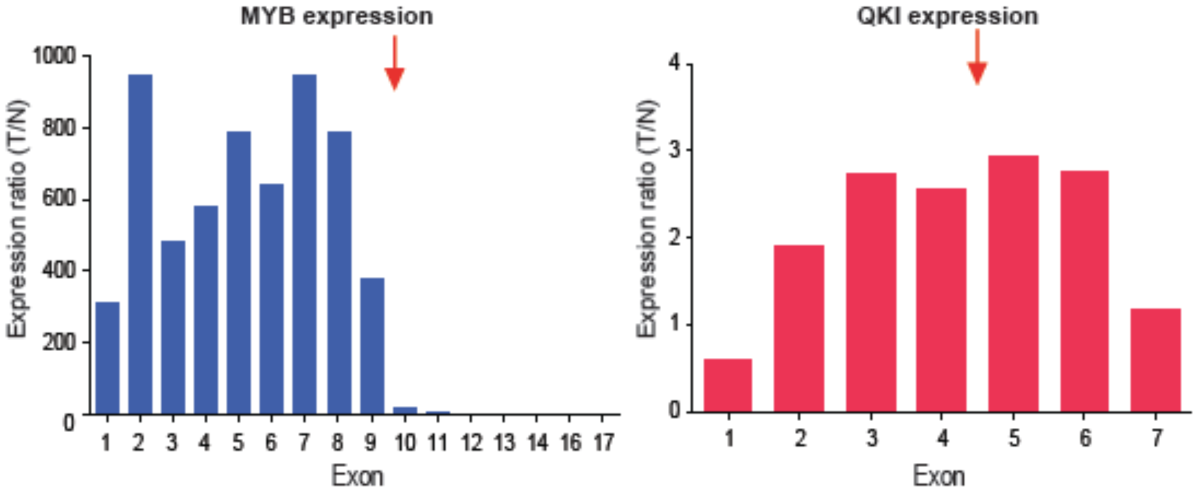
Figure 13:

A. Validation of the exact breakpoint using RNA-sequencing data in PLGG DFCI/CHOP 18

Red arrows represent the site of the breakpoint within the *MYB* and *QKI* genes. Example of a fusion read (blue for MYB and orange for QKI) is represented.



B. MYB (left panel) and QKI (right panel) expression in PLGG DFCI/CHOP 13 relative to normal brain expression obtained with the RNA-seq data



3.4. MYB-QKI rearrangement is oncogenic through multiple mechanisms

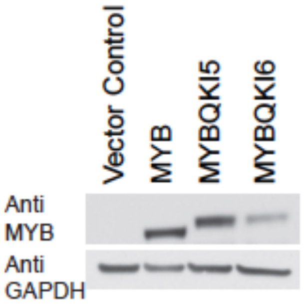
(Work preformed exclusively by Pratiti Bandopadhayay, Lori Ramkissoon and Payal Jain)

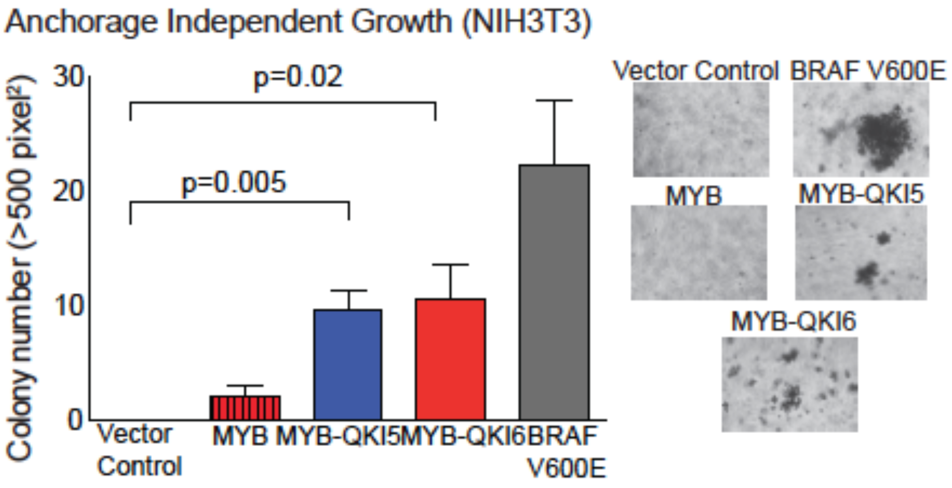
MYB-QKI is an oncogenic fusion protein

To test whether the MYB-QKI fusions were oncogenic, we stably expressed MYB-QKI5, and MYB-QKI6 in two heterologous cell model systems: NIH3T3 cells and in murine neural stem cells. Overexpression of both MYB-QKI5 fusion and MYB-QKI6, but not full-length MYB, was sufficient to induce anchorage independent growth in NIH 3T3 cells (Figure 14).

Figure 14: MYB-QKI protein fusion leads to anchorage independent growth in vitro

Expression of MYB, MYB-QKI5 and MYB-QKI6 in NIH3T3 cells



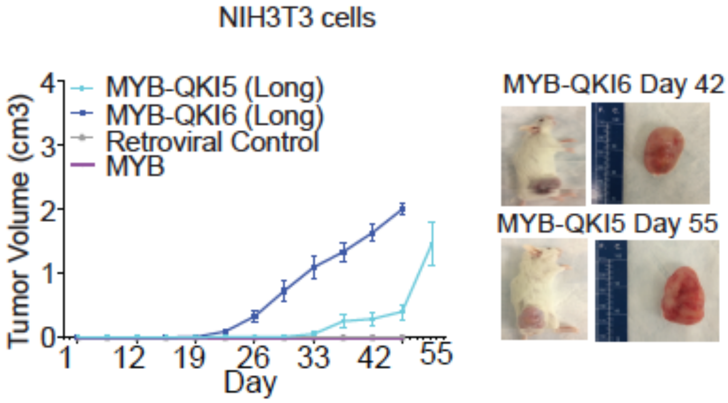


Number of colonies of NIH3T3 cells expressing MYB, MYB-QKI5, MYB-QKI6 or a vector control in soft agar (Left) and representative images (right). NIH3T3 cells over-expressing BRAFV600E are shown as a positive control. Mean of three replicate measurements are shown. Error bars represent SEM.

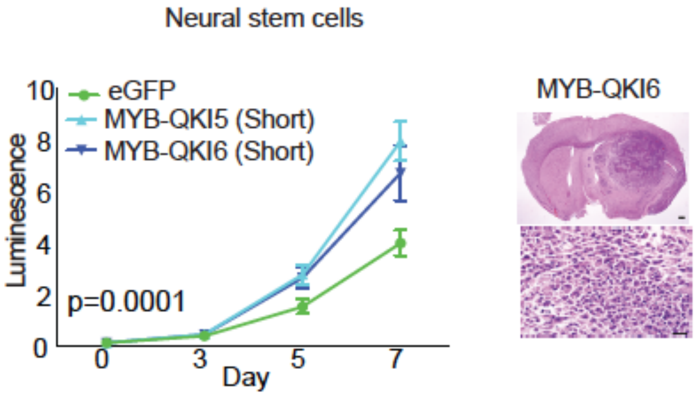
Furthermore, upon injection as flank xenograft in mice, MYB-QKI fusion expressing NIH 3T3 cells lead to robust tumor formation as compared to controls (Figure 15A). Similarly, in murine neural stem cells (mNSCs), overexpression of either MYB-QKI5 or MYB-QKI6 led to significant increases in cell proliferation compared to eGFP across five replicate pools (Figure 15B) and was sufficient to induce gliomagenesis on average 98 days post-injection in an intracranial allograft model. Tumors had histologic features of infiltrating gliomas with some evidence of enhanced growth around vessels and a clustered growth pattern distinct from models of adult glioblastoma. However the tumors differed from human AGs in that they had high-grade features with frequent mitoses and marked cytologic atypia. Immunohistochemistry of the induced tumors showed diffuse expression of GFAP while OLIG2 was present in a subset of cells (Figure 15C), a pattern that is similar to that seen in human AGs.

Figure 15: MYB-QKI fusion protein is oncogenic

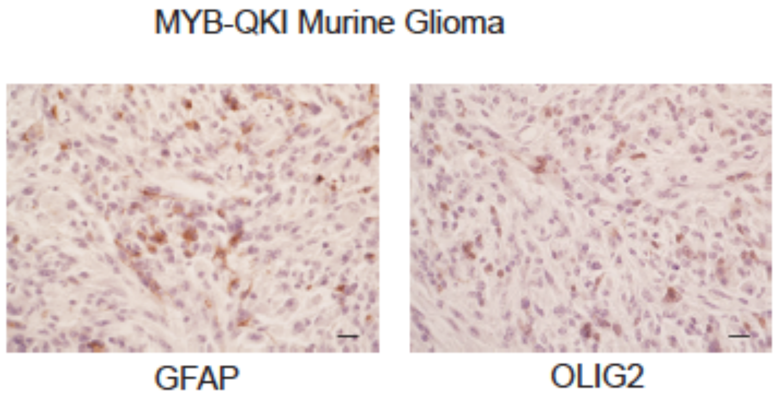
A. Tumor growth following flank injections of NIH3T3 cells overexpressing MYB, MYB-QKI5 (Long), MYB-QKI6 (Long) or a vector control. Mean of five measurements are depicted. Error bars represent SEM.



B. In vitro cell proliferation rates of mNSCs over-expressing eGFP, MYB-QKI5 (Short) or MYB-QKI6 (Short). Mean of five independent pools are depicted. Error bars represent SEM. Representative images intracranial mNSC-MYB-QKI6 tumors.



C. Expression of GFAP and Olig2 in murine MYB-QKI gliomas.

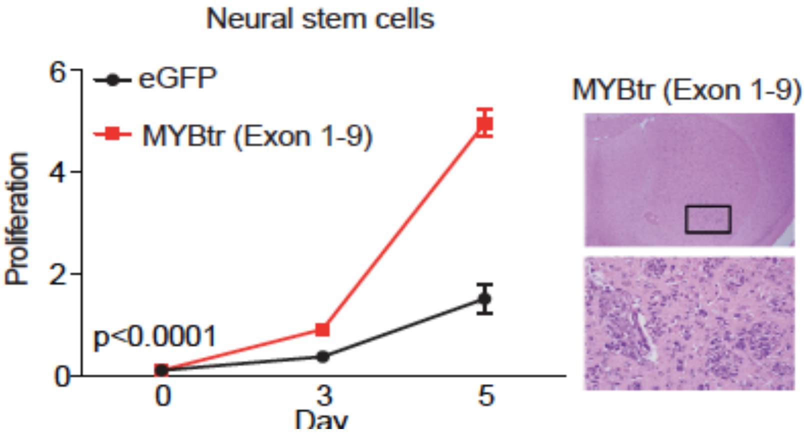


We performed genome-wide gene expression analyses of three independently-generated pools of mNSCs overexpressing MYB-QKI5, MYB-QKI6, MYBtr$^{Exon1\text{-}9}$, QKItr, or eGFP. Relative to eGFP-expressing cells, cells expressing MYB-QKI5 and MYB-QKI6 exhibited significantly different expression of 1621 and 1947 genes, respectively, with 1029 genes overlapping. We defined a MYB-QKI gene expression signature comprising the 50 genes whose differential expression was most correlated with its expression. These genes include KIT and CDK6, which have previously been reported to be associated with MYB activation[45]. Expression of either MYBtr$^{Exon1\text{-}9}$ or MYB-QKI was associated with enrichment of signatures of MYB pathway activation.

Role of aberrant expression of MYB

*In neural stem cells, overexpression of MYB exons 1-9 increased cell proliferation rates compared to eGFP controls and induced tumor formation less than 100 days following intracranial implantation (Figure 16)***.**

Figure 16: In vitro cell proliferation of mouse neural stem cells *which over-express MYBtr or eGFP controls. Mean of five independent pools is depicted. Error bars represent SEM. Representative images of intracranial mNSC-MYBtr tumors.*

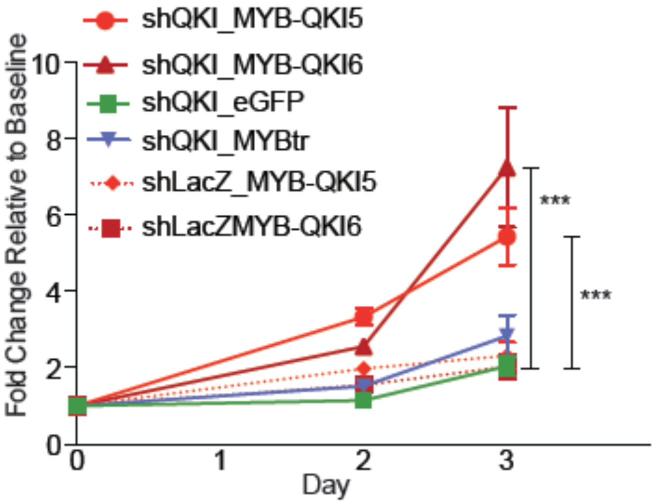


Role of loss of expression of QKI

Reduced Qk expression also resulted in increased proliferation of mNSCs, with the greatest increase observed in the context of pre-existing MYB-QKI expression (p=0.0007 for MYB-QKI5, p=0.0030 for MYB-QKI6 and p=0.03 for MYBtr; Figure 17). These data suggest overexpression of MYB-QKI and suppression of QKI has cooperative functional effects in AGs.

Figure 17: MYB-QKI disrupts QKI that appears as a tumor suppressor gene

Cell proliferation of mouse neural stem cells expressing MYBtr, MYB-QKI5, MYB-QKI6 or eGFP control with suppression of wild-type qk. Values represent mean of four independent experiments ± SEM.

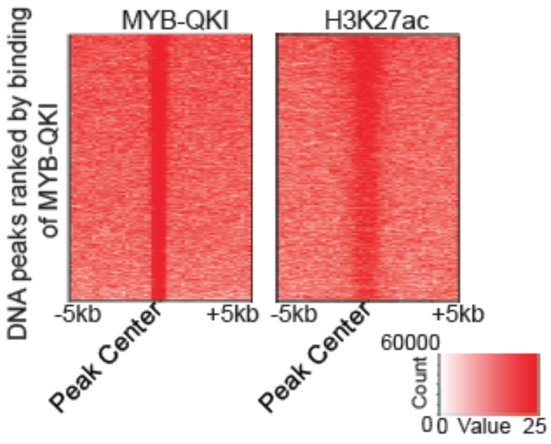


MYB-QKI rearrangement translate QKI enhancer elements within the MYB promoters

To assess the DNA binding pattern and potential transcriptional targets of MYB-QKI, we performed chromatin immunoprecipitation with parallel sequencing (ChIP-seq) in mNSCs expressing MYB-QKI and compared these results with H3K27ac ChIP-Seq to define the location of enhancer regions in the same cells. MYB-QKI5 bound 3672 sites and H3K27ac bound 9122 sites across the genome (p threshold 1e-6), with overlap at 1907 sites (52% of MYB binding sites, $p<0.0001$) (Figure 18A). These findings are consistent with recent reports in the context of T-cell ALL, where MYB binding was also correlated with regions of H3K27ac defined enhancers (89% of super-enhancer regions in Jurkat cells). The MYB-QKI5 binding sites were located within a 100kb distance of 88% (22/25) of the upregulated genes in the MYB-QKI signature but only 40% (10/25) of the downregulated genes. Each of the MYB-QKI binding sites associated with an upregulated gene was associated with an H3K27ac enhancer peak, while only 70% of MYB-QKI binding sites at downregulated genes overlap with enhancers ($p=0.003$). Two of the H3K27ac-binding enhancer regions were within 100kb of the 5' and 3' ends of Qk, the mouse QKI gene (Figure 18B), and each of these was associated with a MYB-QKI binding peak. This suggests a possible auto-regulatory feed-back loop between MYB-QKI and Qk in mouse neural stem cells. MYB was also associated with a MYB-QKI binding peak, but there was no associated enhancer.

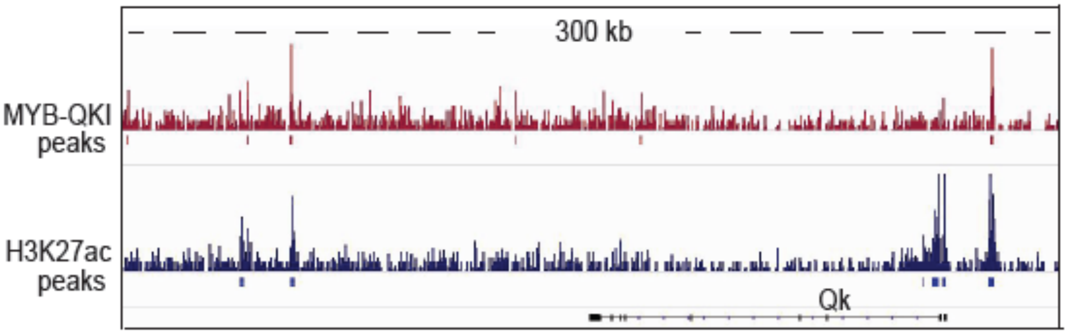
In normal human cortical brain samples, MYB is not associated with H3K27ac peaks, consistent with the finding that MYB is not expressed in normal human brain cortex. In contrast, QKI, which is expressed in human cortical brain, is associated with several H3K27ac peaks, including sequences at the 3' end of QKI. The MYB-QKI rearrangement is predicted to bring these enhancer elements within 15kb of the MYB promoter (Figure 18C), raising the possibility that these may contribute to the observed aberrant expression of MYB.

Figure 18: MYB-QKI co-localizes with H3K27ac binding across the genome

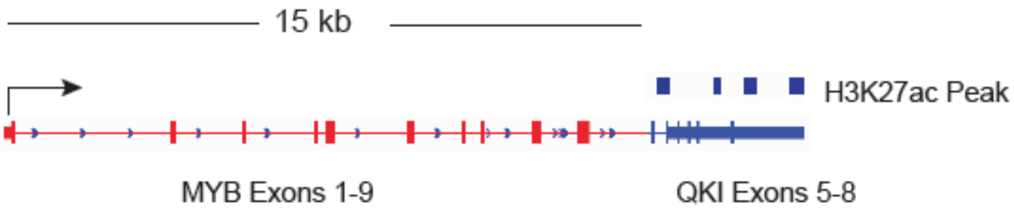


A. Heatmap of H3K27ac and MYB-QKI levels at MYB-QKI regions. Each row shows +/-5 kb centered on MYB-QKI peaks. These regions are rank-ordered by MYB-QKI signal. Scaled intensities are in units of rpm/bp

B. Binding of MYB-QKI and H3K27ac to Qk in mNSC. Binding of MYB-QKI to genes included in the MYB-QKI gene expression signature.



C. Predicted H3K27ac enhancer elements in MYB-QKI, with translocation of genomic enhancers on 3'QKI within 15kb of 5'MYB.



4. Discussion

In this section we have described the results of a meta-analysis pooling the largest genomic dataset of PLGGs.

We confirmed that the structure of the genome of PLGGs is poorly affected by copy-number alterations and mutations. The strong amplification peak identified by GISTIC is related to the *BRAF* duplication present in 57% of the tumors analyzed. The fact that the genome of those tumors is simple and poorly affected by genetic disruptions can be considered on one hand as an advantage to identify potential oncogene, as every recurrent event is highly likely to become a driver mutation[298].

As previously described we confirmed that *BRAF* duplications were the most frequent genetic alteration in those tumors[299]. We showed 7 different rearrangement types all involving BRAF with different partners, although the mechanism of BRAF rearrangement was identical with a loss of the 3' terminal domain, which result in the activation of the BRAF protein and further MAPK signaling activation.

We also confirmed that *FGFR1* mutations (point mutation, duplication and rearrangements), *NF1*, *NTRK2*/*NTRK3* as well as histone genes alterations are affecting significantly a fraction of PLGGs. *FGFR1* point mutations were exclusively found in PAs (mostly ST, 4/5 PAs), whereas *FGFR1* rearrangements were not restricted to PAs (2 PAs) with alterations described in one DA and 2 OAs. *NTRK2* rearrangements were found in 3/4 of PAs (mostly ST, 2/3) and the *NTRK3* rearrangement was found in one PXA. Histone family genes (*H3F3A* and *HIST1H3B*) have already been described to be mutated in higher grade gliomas in the pediatric population[276]. Further functional studies are needed to explore the exact role of those mutations in PLGG tumorigenesis.

This study underlines the importance of increasing the number of samples in order to gain in power and enhance the chances to identify rare oncogenic events that are occurring as a lower frequency[195]. By increasing the number of samples, especially the non-PAs, we were able to identify a recurrent rearrangement between MYB and QKI strongly correlated to AGs. This pattern is unique in brain tumors, as a specific mutation appears to be restricted to a specific histological subtype. This pattern has been also described in other tumor subtypes such as *BCR-ABL* translocation in chronic myelogenous leukemia [304] or *MITF* (Microphthalmia-associated transcription factor) amplification in melanoma[305]. AGs are rare PLGGs, recently identified as a distinct entity[14,306,307]. Previous copy-number studies performed on AGs have already identified CN loss in 6q corresponding to the MYB-QKI region[227,308]. This rearrangement between MYB and QKI was recently described in the PCGP genomic analyses in one AG sample[229]. *MYB* (v-myb avian myeloblastosis viral oncogene homolog) codes for a transcription factor regulating essentially hematopoiesis[309]. Recently *MYB* alterations have been described in hematopoietic cancer[287] but also several other cancer subtypes, such as adenoid cystic carcinomas[310] and breast[298]. Additionally, a recent study on grade II PLGGs performed by our group has identified that recurrent loss of the 3’ terminal end of *MYBL1*, a member of the superfamily MYB is present in a subset of grade II PLGGs[227]. In our cohort of AGs we were able to describe that the rearrangement results in a loss of the 3’ terminal end of *MYB* that contains the regulatory domains and fused to the 3’ extremity of *QKI.* QKI is an RNA-binding protein that regulates multiple important key processes in the cell such as pre-mRNA splicing, export of mRNAs from the nucleus, protein translation, and mRNA stability. The encoded protein is involved in myelinization and oligodendrocyte differentiation[311,312]. QKI is known to harbor different splice variants that are characterized by different downstream applications in the cell[313,314]. QKI has been identified to be significantly lost in a recent pan-cancer analysis[195] and has recently been described as a tumor suppressor candidate in various cancer subtypes such as gastric cancer[315], prostate cancer[313] as well as glioblastomas[316].

Further functional work performed by Pratiti Bandopadhayay and Lori Ramkissoon, collaborators in the lab, allowed to explore further the role of the MYB-QKI rearrangement in various *in vitro* and *in vivo* models. They described that MYB-QKI disrupts two cancer-associated genes, and likely contributes to the neoplastic process by three distinct mechanisms: 1) creation of an oncogenic fusion protein; 2) aberrant, elevated expression of MYB, possibly by bringing *QKI*-related enhancers into proximity with MYB; and 3) hemizygous loss of the tumor suppressor *QKI*, which has specific CNS-contextual roles in glial maturation (Work submitted to Nature Genetics).

In our study we observed a non-negligible number of tumors do not show any potential driver mutation or rearrangement (18/169, 10%). The fact that all those samples have only been assessed by RNA-sequencing raises the concern that restricted RNA-sequencing analyses is not sufficient to detect potential genomic alterations due to low quality of the tissue or the poor expression level of a fraction of genes that might play an important role in oncogenesis mechanisms. Alternatively, this observation suggests that genetic alterations might be located outside the coding regions, as it has been identified recently described in melanomas[317]. Additionally, alterations affecting epigenetic regulators such as the methylation, acetylation of histones might contribute to the particular biology of those tumors.

Finally, this work raises the concern about the new challenges emerging from the computational tools used to explore the genomic sequencing of tumors. Depending on the aligner and the various calling methods used for indels, rearrangements and mutations, one might find different results that can affect the downstream functional studies. It might be interesting to compare and match different algorithm to understand how the differences observed are emerging and whether there could be a uniform way to approach the genomic analysis in oncology in order to increase the accuracy of the results.

Chapter 3: Bulk expression profiling of PLGGs

1. Introduction

High-throughput mRNA expression profiling is a useful approach to characterize molecular alterations in tumors. However, such profiling has not been done in large numbers for non-PAs in the pediatric population[318-320]. The small numbers of patients with these diseases has resulted in limited amounts of tissue available for comprehensive genomic characterization, and in particular very little fresh frozen tissue, which is required by many genomic assays.

In this chapter we describe the results of gene expression profiles of 151 paraffin-embedded PLGGs, from a large variety of histological subtypes, using an innovative platform that allows expression profiling in paraffin-embedded samples of 6100 genes that are commonly dysregulated in cancer. Using these data, we were able to classify the tumors according to phenotypic features and identified differences in molecular patterns between different types of tumors, including supratentorial (ST) and infratentorial (IT) PAs and *BRAF*-duplicated, V600E-mutated and *BRAF* wild-type tumors. We also compared expression profiles as function of age.
We related these differences to publicly available expression profiles from normal developing brains to determine whether the differences we observed were reflected in normal brains.

2. Methods

2.1. Tumor sample collection

All tissues used were paraffin-embedded and obtained from Boston Children's Hospital (Boston, USA) through IRB-approved protocols. Diagnoses were made according to the WHO classification scheme by consensus of three neuropathologists (Jennifer Chan, Keith Ligon and Sandro Santagata). For 137 samples (91%), we performed FISH to assess the presence of *BRAF*-KIAA fusion transcripts[321]. We genotyped for *BRAF-V600E* mutations in 94 samples (62%) using targeted sequencing [246]. All tumors were clinically annotated with outcomes.

2.2. RNA isolation and microarray analysis

RNA was extracted using TRIzol (Life Technologies) followed by purification using the RNeasy MinElute Cleanup Kit (Qiagen). Expression profiles were generated using the Illumina DASL platform (San Diego, CA)[322] and normalized using cubic spline interpolation[323,324]. This platform used 6100 genes that harbor the largest variation across a large sample panel across tissue types, as described [323]. Samples were run in three batches; we applied batch correction using ComBat[325]. Z-scores were generated for each gene across samples. Self Organizing Maps[326], hierarchical clustering[202], Comparative Marker Selection Analysis[327,328], and Gene-set Enrichment Analysis (GSEA)[206,329] were performed using GenePattern[330]. Principal Component Analyses were generated in Matlab. For clustering, PCA plots, and distance measurements, we used the 250 genes with highest variance. Use of 100, 500, or 1000 genes gave similar results but poorer average silhouette values. For Comparative Marker Selection, we only report genes with differences in mean z-scores of greater than 0.5. The data have been deposited in the Gene Expression Omnibus[331] (accession number GSE60898).

2.3. Normal brains

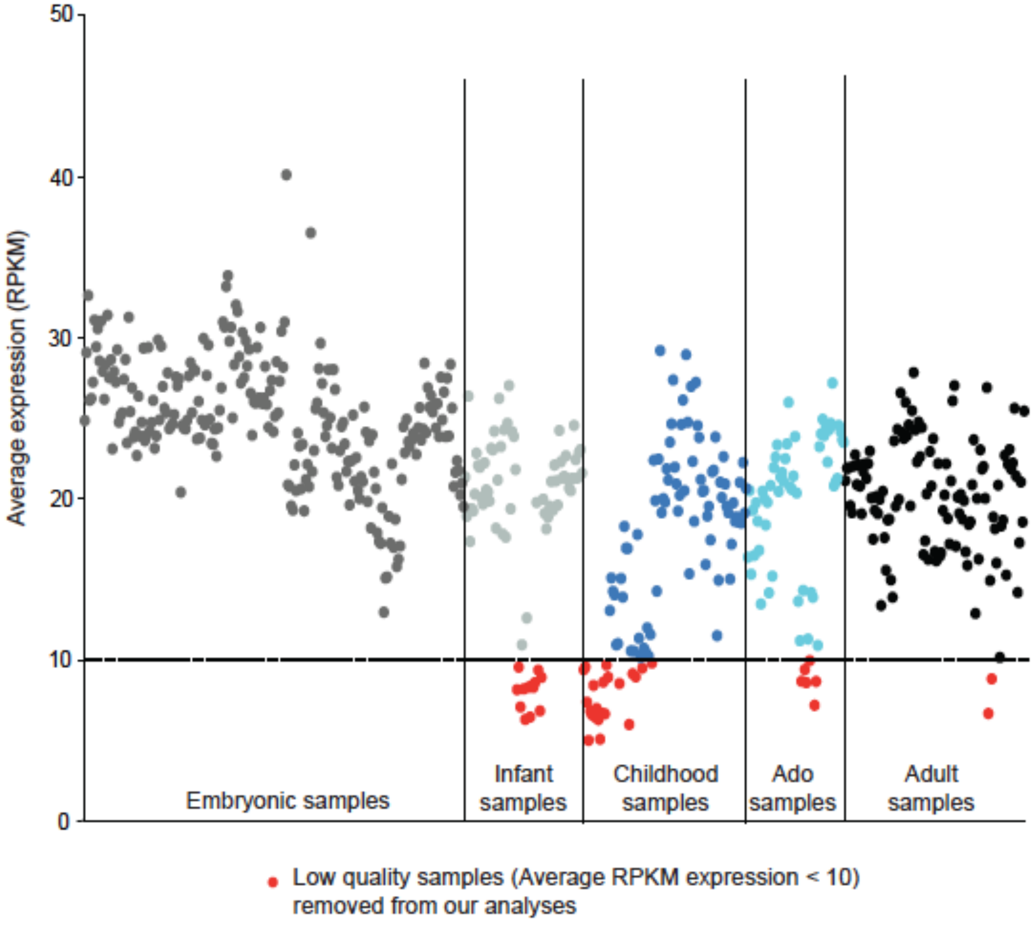
Normalized RNA-sequencing data across brain regions and ages (infancy, childhood and adolescence) were uploaded from BRAINSPAN[332]. To allow comparison between BRAINSPAN and DASL data, we selected the 5483 genes assessed by both. Forty-three BRAINSPAN samples (0/234 embryos, 13/73 infants, 22/101 childhood samples, 6/59 adolescents, and 2/112 adults; Figure 19) exhibited low expression levels across most genes and average levels <10 RPKM, two SDs below the mean, suggesting tissue degradation. These were removed.

Figure 19:

Average expression among the 5483 genes we selected across all samples in the BRAINSPAN cohort, arranged by age.

Samples whose average expression levels are below 2 standard deviations below the mean (10 RPKM) are indicated with red dots.

We considered these samples to be low quality and excluded them from further analyses.



2.4. Immunohistochemistry

HLA-DRA immunohistochemistry (Abcam, 1:250) was performed in an additional cohort of 18 PA tumors using the Dual Link Envision+ (Dako) detection system. Citrate was used for antigen retrieval and diaminobenzidine (DAB) as the chromogen. We assessed DAB surface and intensity staining using Cell Profiler (Broad Institute)[333] in 15 representative tumor sections for each sample.

2.5. Statistical analysis

Fisher's exact and Mann-Whitney tests were used to generate p-values as appropriate. False discovery rates (FDRs) were determined to account for multiple hypotheses[334] and FDR<0.25 was considered significant. Log-rank (Mantel-Cox) tests were used to examine differences in event-free survival.

3. Results

We analyzed expression of 6100 genes in 151 paraffin-embedded PLGG samples spanning six histological subtypes: pilocytic astrocytomas (PAs), gangliogliomas (GGs), diffuse astrocytomas (DAs), dysembryoplastic neuroepithelial tumors (DNTs), oligodendroglial tumors (ODs) and PLGGs Not Otherwise Specified (NOS) (Figure 20A).

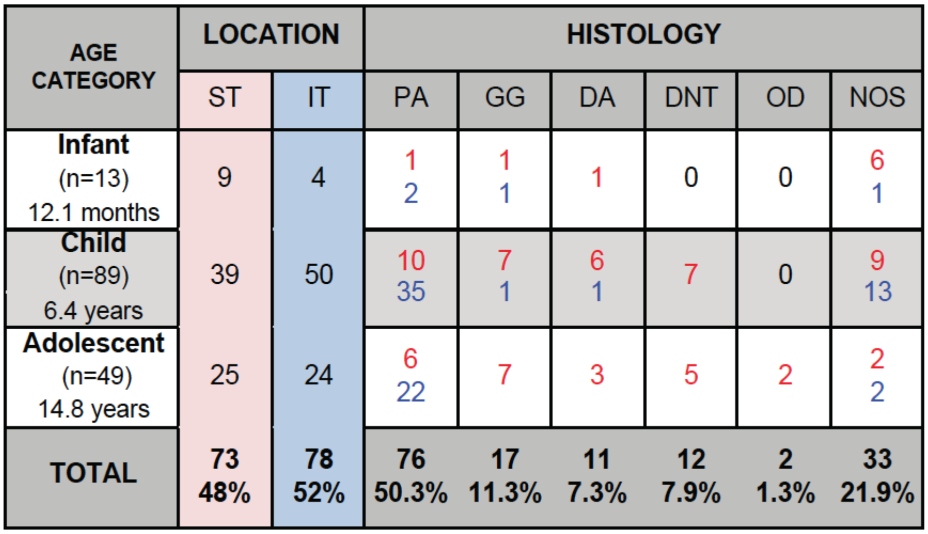
The cohort included tumors from 13 infants (birth to 18 months), 89 children (19 months to 11 years) and 49 adolescents (12 to 19 years). Primary tumors constituted 135 samples; 16 were collected after second surgery for relapse or progression, including nine for which we had paired primary and recurrent samples.

NOS tumors were the only histological subtype that occurred with equal frequency in both the supratentorium and infratentorium. DAs, GGs, DNTs and ODs occurred most commonly in the supratentorium ($p<0.0001$); PAs were primarily infratentorial (and primarily cerebellar; $p<0.0001$) (Figure 20A).

Consistent with prior reports[225,228,229,248], we observed *BRAF*-KIAA duplications most frequently in PAs ($p<0.0001$); *BRAF*-V600E mutations were significantly associated with GGs ($p<0.0002$) (Figure 20B). Similar fractions of NOS tumors harbored *BRAF*-duplications (38%) and V600E mutations (26%) (Figure 20B). Tumors with *BRAF*-duplications were predominantly infratentorial (81%, $p<0.0001$); those with *BRAF-V600E* mutations were predominantly supratentorial (90%, $p<0.0001$).

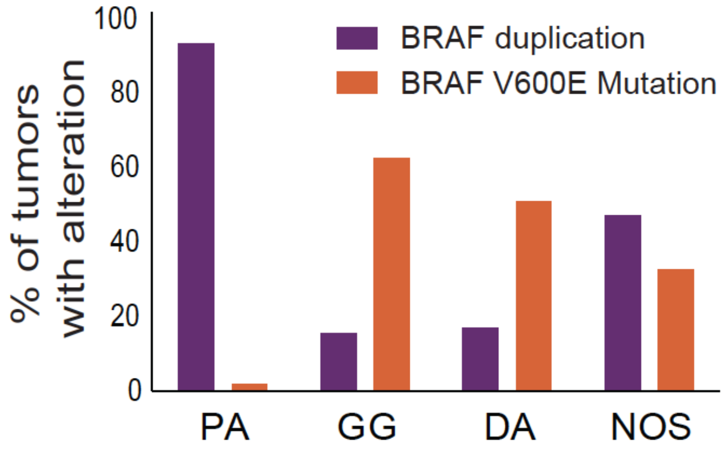
Figure 20: Description of the sample set

A. Summary of tumor locations and histologies by age in the pediatric population. Red and blue numbers indicate ST and IT tumors, respectively.



AGE CATEGORY	LOCATION		HISTOLOGY					
	ST	IT	PA	GG	DA	DNT	OD	NOS
Infant (n=13) 12.1 months	9	4	1 2	1 1	1	0	0	6 1
Child (n=89) 6.4 years	39	50	10 35	7 1	6 1	7	0	9 13
Adolescent (n=49) 14.8 years	25	24	6 22	7	3	5	2	2 2
TOTAL	**73** **48%**	**78** **52%**	**76** **50.3%**	**17** **11.3%**	**11** **7.3%**	**12** **7.9%**	**2** **1.3%**	**33** **21.9%**

B. *BRAF* genomic status by histology. DNT and OD tumors did not described *BRAF* alterations.



3.1. PLGGs cluster in three molecular groups according to location, histological subtype, and *BRAF* genomic status

The 135 newly diagnosed PLGGs segregated into three clusters ("Supratentorial", "BRAF-duplicated", and "Mixed"; Figure 21A), associated with different histologies (Figure 21B) by Euclidean SOM[326] (Figure 22A and 22B).

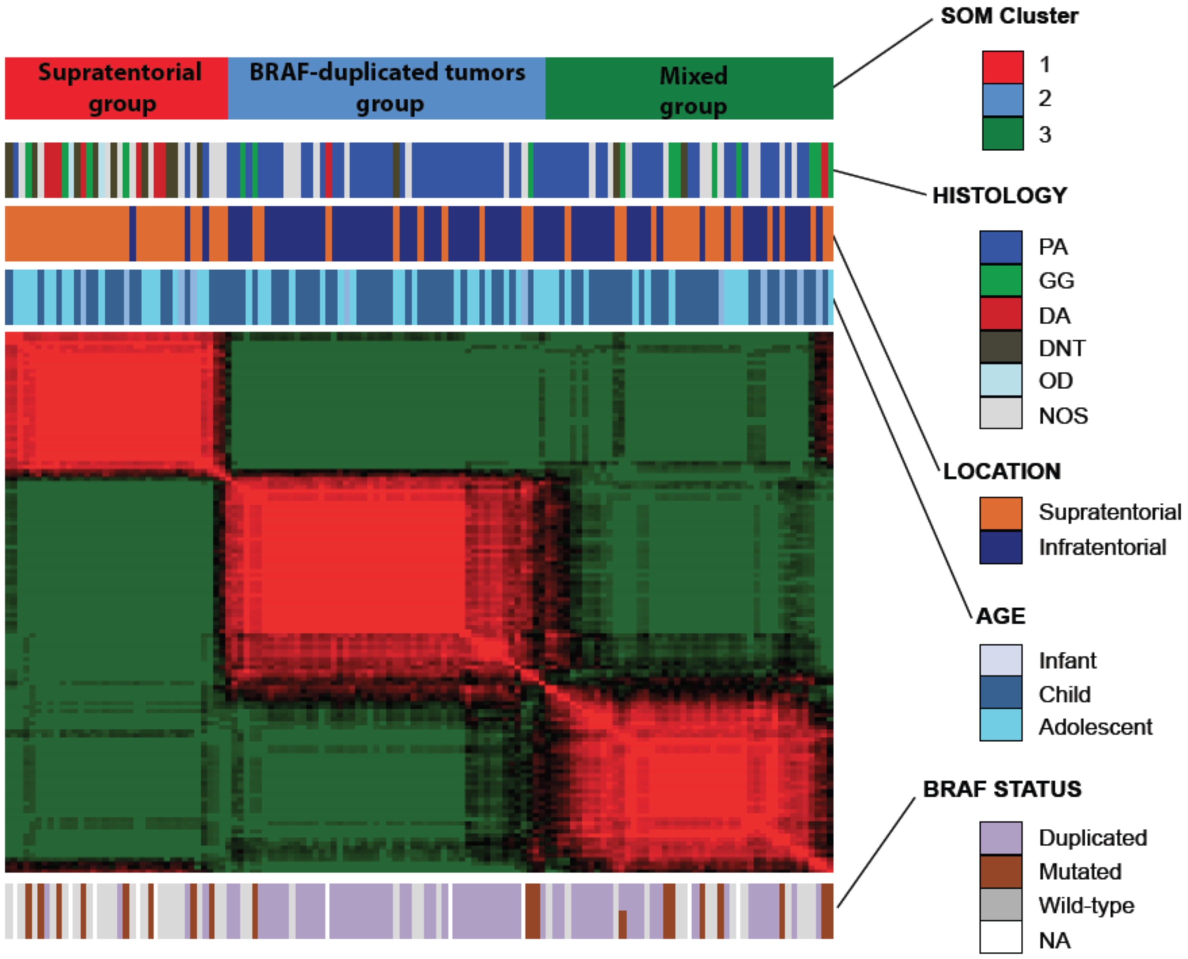
DAs, were enriched in the Supratentorial cluster (p=0.0014), as were ODs and DNTs (p values < 0.001). PAs were enriched in the BRAF-duplicated cluster (p<0.0001), with most of the remaining in the Mixed cluster. GGs were primarily distributed between clusters 1 and 3 (p=0.160). NOS tumors were distributed almost equally across all three clusters, indicating that they are molecularly heterogeneous. NOS tumors with *BRAF* duplications preferentially segregated into the BRAF-duplicated cluster (p=0.05), which was otherwise dominated by *BRAF*-duplicated IT PAs. NOS tumors with V600E point mutations mostly segregated in clusters 1 and 3 (p=0.2) with the ST GGs and other ST tumors.

The Supratentorial cluster was significantly enriched with ST PLGGs (p<0.0001), whereas IT tumors were significantly enriched in the BRAF-duplicated cluster (p<0.0001; Figure 21C).

The Mixed cluster was not significantly enriched with tumors from either location (p=0.591).

Figure 21: PLGGs divide into three molecular clusters

A. Self-organizing map (SOM) clustering heat map including the 135 primary PLGGs.



B-D. Distributions among the three clusters of histology (B), ST and IT tumors (C) and *BRAF* genomic status (D) (* p< 0.001, Fisher's exact test between members of the starred cluster vs both other clusters).

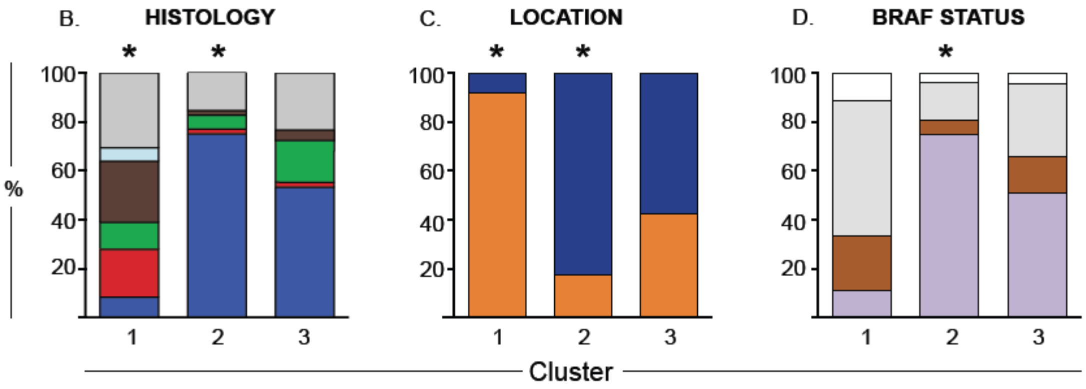
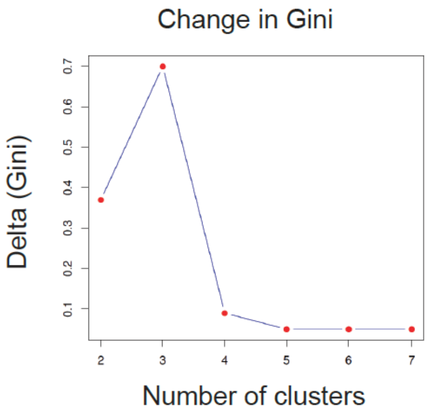


Figure 22:

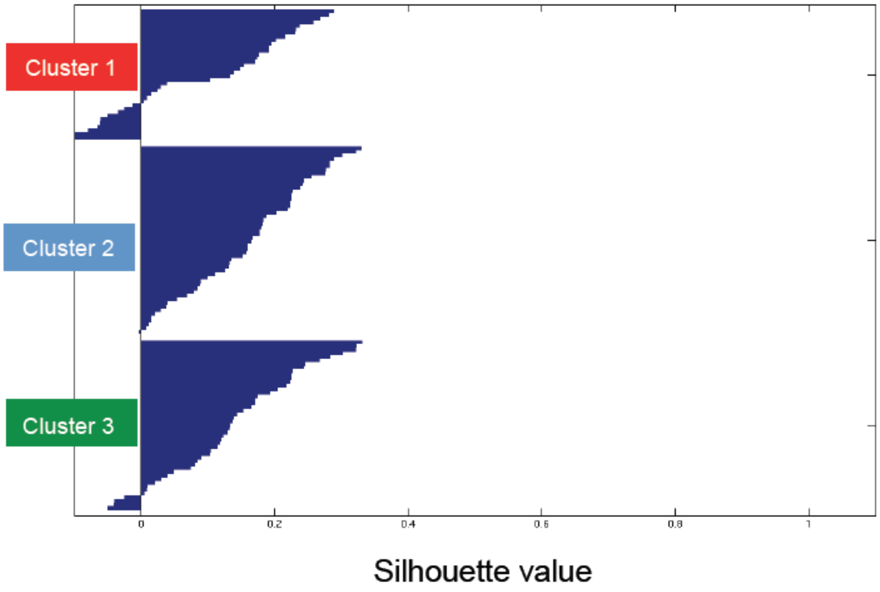
A. Delta Gini Index obtained in the SOM clustering analysis among the 135 newly diagnosed PLGGs.

Increased numbers suggest improved dispersion between clusters.



B. Silhouette plot analysis using the SOM clustering results of the 135 newly diagnosed PLGGs.

Each horizontal bar represents one sample. Positive silhouette values indicate samples that lie well within the cluster. A separate K-nearest neighbors (KNN) analysis classifying tumors according to their three nearest neighbors generated identical assignments to those here.



Tumors from different age groups segregated almost evenly across the three clusters. To determine how adult low-grade gliomas (ALGGs) relate to these clusters, we added 15 primary ALGGs, all supratentorial (7 ODs, 5 DAs, 2 PAs, 1 NOS tumor), to the pediatric cohort and performed SOM clustering on the combined cohort of 150 tumors. We again obtained three clusters that substantially aligned with the three clusters in the original analysis (Figure 23). Of the 15 ALGGs, 12 (80%) segregated with the supratentorial PLGGs in cluster 1.

BRAF-duplicated tumors were significantly enriched in cluster 2 (58%, p<0.0001), whereas *BRAF-V600E* mutated tumors were enriched in clusters 1 and 3 (p=0.066; Figure 20D). This observation reflects the finding that *BRAF*-duplicated tumors are more frequent among infratentorial PAs, whereas ST PLGGs more commonly harbor the *BRAF*V600E point mutation. The *BRAF* status of the NOS tumors followed the same distribution, with a significant enrichment of NOS *BRAF*-duplicated tumors in cluster 2 and *BRAF-V600E* mutated tumors in cluster 1 and 3 (p=0.045).

The three clusters exhibited no differences in event-free survival (Figure 24).

Figure 23 :

Self-organizing map (SOM) clustering heat map including 135 primary PLGGs and 15 primary ALGGs.

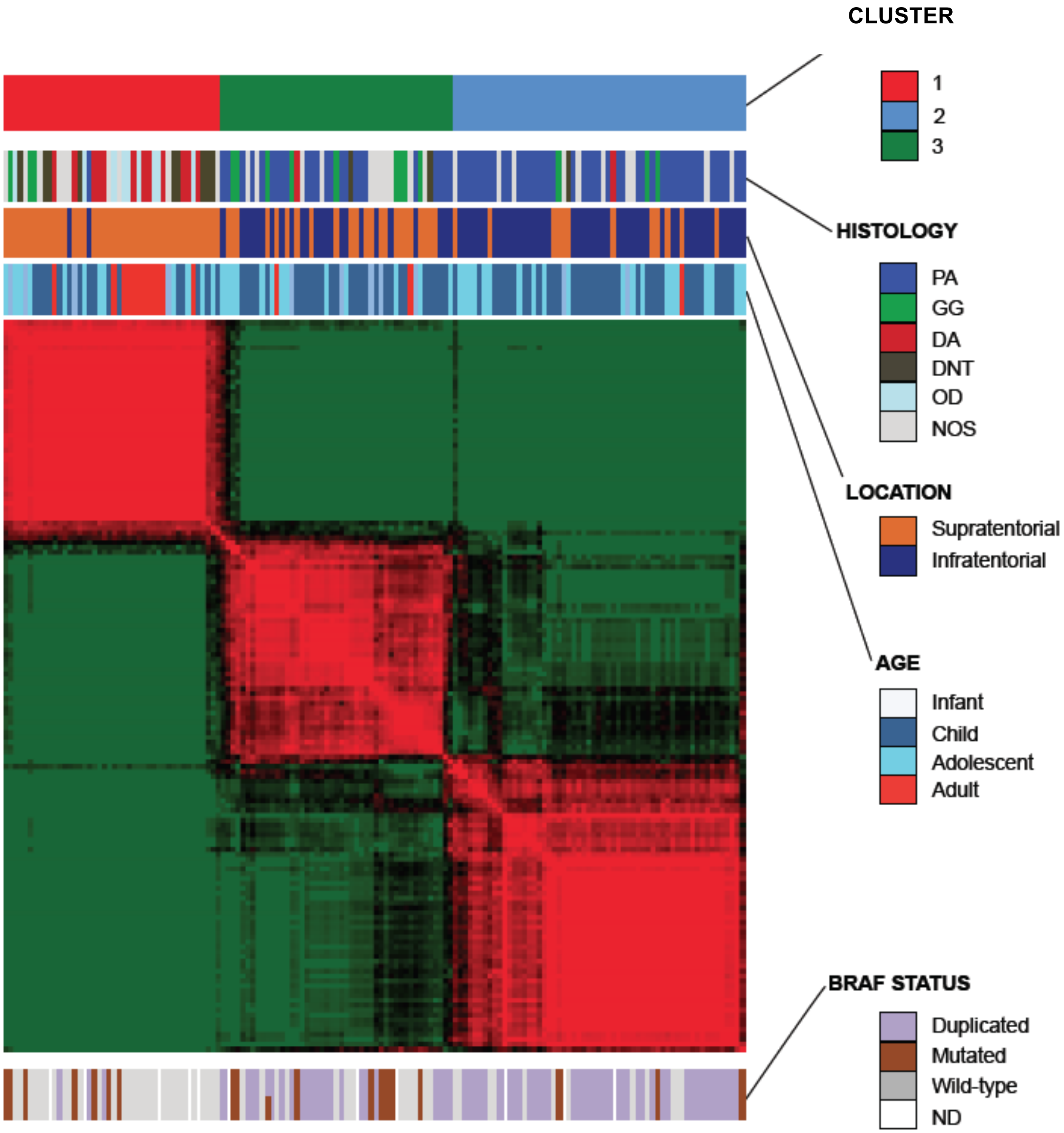
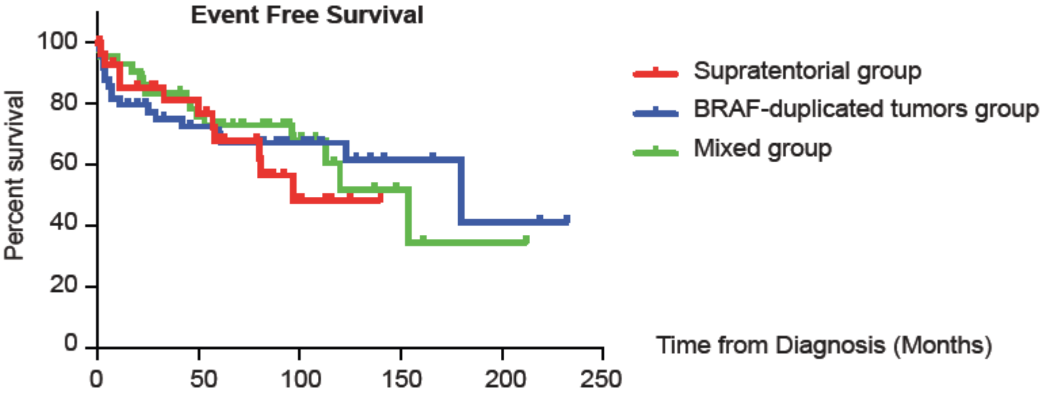


Figure 24: Event-free survival curves for the three clusters.



We compared molecular signatures between the different clusters using GSEA and the 1320 gene-sets in the C2 canonical pathway (CP) set.

The Supratentorial cluster was significantly enriched for 34 gene-sets (q-value<0.25; Table 8A). Of these, seven were associated with neurotransmission mechanisms (of 18 total; p<0.0001), suggesting a neuronal signature. No gene-sets were significantly enriched in the BRAF-duplicated cluster. The Mixed cluster was significantly enriched for 33 gene-sets (Table 8B), including three of 21 gene-sets associated with cell cycle control (p=0.0139) and five of 67 gene-sets associated with immune response (p=0.0227).

Table 8:

A. Gene-sets significantly enriched in cluster 1.

NAME	p value	FDR
KEGG NITROGEN METABOLISM	0.00	0.00
REACTOME AMINO ACID AND OLIGOPEPTIDE SLC TRANSPORTERS	0.00	0.03
REACTOME TRANSPORT OF INORGANIC CATIONS ANIONS AND AMINO ACIDS OLIGOPEPTIDES	0.00	0.09
REACTOME NEUROTRANSMITTER RELEASE CYCLE	0.00	0.10
REACTOME ION CHANNEL TRANSPORT	0.00	0.11
REACTOME NEURONAL SYSTEM	0.00	0.12
REACTOME TRANSMISSION ACROSS CHEMICAL SYNAPSES	0.00	0.12
REACTOME GABA RECEPTOR ACTIVATION	0.00	0.13
KEGG PROXIMAL TUBULE BICARBONATE RECLAMATION	0.00	0.14
KEGG CALCIUM SIGNALING PATHWAY	0.00	0.16
KEGG CITRATE CYCLE TCA CYCLE	0.01	0.17
REACTOME TRANSMEMBRANE TRANSPORT OF SMALL MOLECULES	0.00	0.18
REACTOME PHOSPHOLIPASE C MEDIATED CASCADE	0.01	0.18
REACTOME FGFR LIGAND BINDING AND ACTIVATION	0.02	0.19
REACTOME NEUROTRANSMITTER RECEPTOR BINDING AND DOWNSTREAM TRANSMISSION IN THE POSTSYNAPTIC CELL	0.01	0.20
KEGG PARKINSONS DISEASE	0.02	0.20
REACTOME ION TRANSPORT BY P TYPE ATPASES	0.02	0.20
REACTOME NEGATIVE REGULATION OF FGFR SIGNALING	0.02	0.20
REACTOME INWARDLY RECTIFYING K CHANNELS	0.02	0.20
REACTOME SIGNALING BY FGFR MUTANTS	0.00	0.20
REACTOME RESPIRATORY ELECTRON TRANSPORT ATP SYNTHESIS BY CHEMIOSMOTIC COUPLING AND HEAT PRODUCTION BY UNCOUPLING PROTEINS	0.03	0.20
REACTOME POTASSIUM CHANNELS	0.01	0.21
KEGG OXIDATIVE PHOSPHORYLATION	0.02	0.21
REACTOME INHIBITION OF VOLTAGE GATED CA2 CHANNELS VIA GBETA GAMMA SUBUNITS	0.03	0.21
BIOCARTA PGC1A PATHWAY	0.03	0.21
KEGG CARDIAC MUSCLE CONTRACTION	0.02	0.21
REACTOME MITOTIC PROMETAPHASE	0.02	0.21
REACTOME RESPIRATORY ELECTRON TRANSPORT	0.04	0.21
REACTOME SLC MEDIATED TRANSMEMBRANE TRANSPORT	0.00	0.21
REACTOME NITRIC OXIDE STIMULATES GUANYLATE CYCLASE	0.04	0.21
REACTOME DARPP 32 EVENTS	0.02	0.24
BIOCARTA BIOPEPTIDES PATHWAY	0.03	0.24
REACTOME GABA B RECEPTOR ACTIVATION	0.00	0.00
ST WNT CA2 CYCLIC GMP PATHWAY	0.00	0.03

🟨 Gene-sets involved in neurotransmission mechanisms

B. Gene-sets significantly enriched in cluster 3.

NAME	p value	FDR
REACTOME CDK MEDIATED PHOSPHORYLATION AND REMOVAL OF CDC6	0.00	0.18
REACTOME DESTABILIZATION OF MRNA BY AUF1 HNRNP D0	0.01	0.18
REACTOME CDT1 ASSOCIATION WITH THE CDC6 ORC ORIGIN COMPLEX	0.00	0.18
REACTOME INNATE IMMUNE SYSTEM	0.00	0.19
REACTOME ER PHAGOSOME PATHWAY	0.00	0.19
REACTOME VIF MEDIATED DEGRADATION OF APOBEC3G	0.01	0.19
KEGG VIRAL MYOCARDITIS	0.02	0.19
REACTOME P53 INDEPENDENT G1 S DNA DAMAGE CHECKPOINT	0.00	0.20
REACTOME SCF BETA TRCP MEDIATED DEGRADATION OF EMI1	0.01	0.20
REACTOME SIGNAL TRANSDUCTION BY L1	0.01	0.20
REACTOME HIV INFECTION	0.00	0.20
REACTOME METABOLISM OF RNA	0.01	0.21
BIOCARTA PROTEASOME PATHWAY	0.02	0.21
REACTOME SCFSKP2 MEDIATED DEGRADATION OF P27 P21	0.01	0.21
BIOCARTA NKCELLS PATHWAY	0.03	0.21
REACTOME SIGNALING BY WNT	0.01	0.21
REACTOME METABOLISM OF MRNA	0.01	0.22
REACTOME REGULATION OF APOPTOSIS	0.01	0.22
REACTOME AUTODEGRADATION OF THE E3 UBIQUITIN LIGASE COP1	0.00	0.22
REACTOME ACTIVATION OF NF KAPPAB IN B CELLS	0.01	0.22
KEGG SYSTEMIC LUPUS ERYTHEMATOSUS	0.04	0.22
REACTOME REGULATION OF MRNA STABILITY BY PROTEINS THAT BIND AU RICH ELEMENTS	0.00	0.23
REACTOME TGF BETA RECEPTOR SIGNALING ACTIVATES SMADS	0.01	0.23
REACTOME NUCLEOTIDE BINDING DOMAIN LEUCINE RICH REPEAT CONTAINING RECEPTOR NLR SIGNALING PATHWAYS	0.02	0.23
REACTOME THE ROLE OF NEF IN HIV1 REPLICATION AND DISEASE PATHOGENESIS	0.02	0.23
PID TGFBRPATHWAY	0.01	0.23
PID NFKAPPABCANONICALPATHWAY	0.02	0.24
REACTOME REGULATION OF ORNITHINE DECARBOXYLASE ODC	0.02	0.24
REACTOME CLASS I MHC MEDIATED ANTIGEN PROCESSING PRESENTATION	0.00	0.24
REACTOME CROSS PRESENTATION OF SOLUBLE EXOGENOUS ANTIGENS ENDOSOMES	0.00	0.24
REACTOME P53 DEPENDENT G1 DNA DAMAGE RESPONSE	0.00	0.24
KEGG PROTEASOME	0.01	0.24
REACTOME HOST INTERACTIONS OF HIV FACTORS	0.00	0.24

🟪 Gene-sets involved in cell cycle control

🟧 Gene-sets involved in inflammatory pathways

3.2. Variations in expression patterns across histological subtypes

We analyzed how expression profiles vary across the four major histologies (PAs, DAs, DNTs, and GGs) using Comparative Marker Selection Analysis and GSEA. We controlled for location by comparing only the ST tumors.

We found that 205 genes and 90 gene-sets were significantly differentially expressed between the ST PAs and the ST DAs (Figure 25A, Table 9A and 9B) and 230 genes and 120 gene-sets were significantly differentially expressed between the ST PAs and the 12 ST DNTs (Figure 25B, Table 9C and 9D). In both comparisons, 57 gene-sets were enriched among PAs, including three gene-sets from each of two families: inflammatory pathways (KEGG complement disease, KEGG antigen processing and presentation, and reactome interferon gamma signaling) and extracellular matrix organization (pid integrin 3 pathway, reactome extracellular matrix organization, and reactome collagen formation). Similarly, 99 genes were significantly differentially expressed in ST PAs relative to both ST DAs and ST DNTs; two of the five genes with the largest differences in expression, two (*SERPING1* and *CD74*) are in inflammatory pathways. Twenty-eight gene-sets were enriched in ST DNTs relative to PAs. Thirteen of these related to a neurons or neuronal transmission, including four of the five most significant gene-sets (reactome post NMDA receptor activation events, reactome neurotransmitter release cycle, and reactome neurotransmitter receptor binding, reactome neurotransmission across chemical synapses). The four genes with the greatest expression in ST DNTs relative to ST PAs are all related to signal transduction (*SCN2A2, STMN2, SNAP25*, and *VSNL1*). No gene-sets were significantly upregulated in ST DAs relative to ST PAs;

among the five genes that had the greatest expression ST DAs relative to ST PAs, two (*SL17A7* and *ATP1A2*) are involved in ion transport. We found no genes or gene-sets that were significantly differentially expressed between the ST PAs and the 15 ST GGs and either the ST PAs or DAs.

Using the first three components of a PCA analysis, we observed differences between the ST PAs and DAs (Figure 25C) and between the ST PAs and DNTs (Figure 25D); the ST PAs and GGs appear to be mixed (Figure 25E). However, these global differences only trend towards statistical significance (p=0.09, 0.08, 0.13, respectively). We compared the Euclidean distances within and between all PAs and DAs, PAs and DNTs, and PAs and GGs and determined the significance of these differences by permuting class labels. We controlled for the location by permuting separately ST and IT tumors within each group of tumors.

Figure 25: **Variation in gene expression profiles across different histological subtypes**

A-B. Heat maps representing genes found to be differentially expressed between ST PAs and DAs (A) and between ST PAs and DNTs (B). C-E. PCA analysis comparing the expression pattern of ST PAs and ST DAs (C), ST PAs and ST DNTs (D) and ST PAs and ST GGs (E).

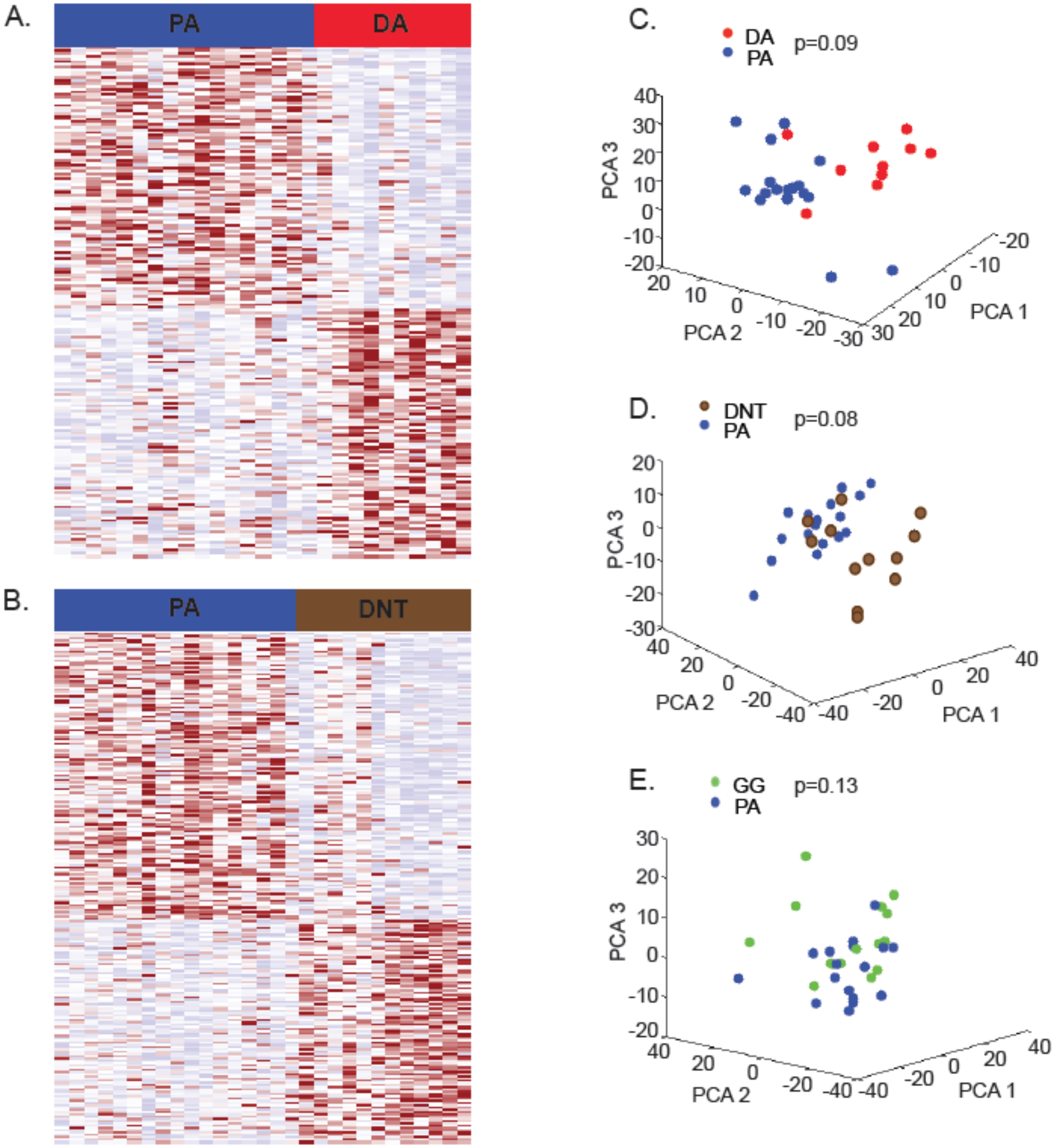


Table 9:

A. List of the genes significantly differentially expressed between ST PAs and DAs

upregulated	Gene	p value	FDR	Difference mean z-score PA-DA
PA	RARRES1	0.00	0.06	2.42
PA	SERPING1	0.00	0.08	1.93
PA	CHI3L1	0.01	0.17	1.92
PA	MGP	0.00	0.14	1.82
PA	PLA2G2A	0.00	0.09	1.81
PA	TM4SF1	0.00	0.08	1.76
PA	SERPINA3	0.00	0.13	1.67
PA	PHLDA1	0.00	0.09	1.56
PA	GPX3	0.00	0.10	1.53
PA	CTGF	0.01	0.18	1.46
PA	CAV1	0.00	0.12	1.45
PA	APOC2	0.01	0.19	1.42
PA	SAT	0.01	0.21	1.39
PA	S100A4	0.00	0.13	1.36
PA	LMOD1	0.00	0.11	1.30
PA	CD74	0.00	0.07	1.29
PA	LYZ	0.00	0.08	1.27
PA	S100A11	0.00	0.09	1.22
PA	ACTN1	0.00	0.07	1.21
PA	AHR	0.00	0.08	1.20
PA	DUSP1	0.01	0.14	1.20
PA	PBEF1	0.00	0.13	1.18
PA	COL1A2	0.01	0.15	1.17
PA	CAV2	0.00	0.13	1.16
PA	MSR1	0.00	0.07	1.14
PA	C1R	0.00	0.12	1.14
PA	FCGR2A	0.00	0.06	1.14
PA	CNN2	0.00	0.06	1.11
PA	PODXL	0.01	0.16	1.10
PA	APOD	0.00	0.13	1.08
PA	BGN	0.00	0.09	1.06
PA	RASSF2	0.01	0.20	1.05
PA	MHC2TA	0.01	0.15	1.04
PA	HLA-DPA1	0.00	0.08	1.04
PA	PRSS23	0.00	0.06	1.04
PA	VGF	0.01	0.19	1.02
PA	A2M	0.00	0.06	0.98
PA	DOC1	0.00	0.09	0.97
PA	IFI16	0.00	0.12	0.97
PA	ACTA2	0.01	0.22	0.96
PA	MAP1B	0.01	0.14	0.96
PA	ABHD2	0.00	0.07	0.93
PA	LAMB1	0.00	0.09	0.93
PA	CALB2	0.01	0.14	0.92
PA	TRPM8	0.00	0.06	0.91
PA	RUNX1	0.01	0.14	0.90
PA	GBP2	0.00	0.09	0.87
PA	COL3A1	0.00	0.14	0.87
PA	CSPG4	0.00	0.13	0.87
PA	GADD45A	0.02	0.22	0.83
PA	PLAU	0.00	0.06	0.83
PA	ALDH1A3	0.01	0.16	0.80
PA	COL4A1	0.00	0.09	0.80
PA	TAGLN	0.01	0.18	0.80
PA	HLA-DPB1	0.00	0.06	0.77
PA	SIX1	0.00	0.06	0.75
PA	HLA-DMA	0.00	0.10	0.74
PA	HLA-DRA	0.00	0.10	0.74
PA	IFI30	0.00	0.08	0.73
PA	FRZB	0.01	0.14	0.73
PA	IF	0.00	0.08	0.72
PA	EMP2	0.00	0.08	0.72
PA	VWF	0.00	0.12	0.71
PA	GAD1	0.00	0.10	0.71
PA	IGJ	0.01	0.18	0.70
PA	ARL7	0.00	0.09	0.70
PA	CTSC	0.00	0.07	0.69
PA	IL1RAP	0.00	0.11	0.69
PA	LMAN2	0.00	0.08	0.69
PA	BTG1	0.00	0.12	0.68
PA	C1S	0.00	0.08	0.68
PA	IQGAP1	0.02	0.23	0.68
PA	COL1A1	0.00	0.08	0.68
PA	EVI2B	0.01	0.18	0.68
PA	SLC26A2	0.00	0.09	0.67
PA	WWTR1	0.00	0.08	0.67
PA	DPEP2	0.01	0.17	0.67
PA	TMSB10	0.01	0.17	0.66
PA	PAM	0.00	0.08	0.65
PA	SARA1	0.00	0.11	0.65
PA	CASP4	0.00	0.08	0.65
PA	TRIM14	0.01	0.16	0.64
PA	TUSC3	0.00	0.12	0.64
PA	HSPA6	0.01	0.17	0.64
PA	DSCR1	0.00	0.13	0.64

PA	NPC2	0.02	0.23	0.63
PA	CDH11	0.00	0.11	0.63
PA	FYB	0.00	0.07	0.62
PA	S100A10	0.00	0.12	0.62
PA	HMGA1	0.02	0.24	0.62
PA	FOSL2	0.00	0.11	0.62
PA	HCLS1	0.00	0.13	0.62
PA	OSMR	0.00	0.08	0.61
PA	MEST	0.01	0.20	0.61
PA	BTG3	0.01	0.20	0.61
PA	PKM2	0.00	0.12	0.60
PA	GYPC	0.02	0.23	0.59
PA	CTHRC1	0.00	0.06	0.59
PA	AKAP13	0.01	0.22	0.57
PA	TEAD4	0.00	0.13	0.57
PA	HLA-DQA1	0.00	0.06	0.57
PA	MMP17	0.01	0.20	0.56
PA	ECM1	0.00	0.08	0.55
PA	SERPINF1	0.01	0.15	0.55
PA	CD59	0.01	0.15	0.55
PA	ARHGDIB	0.02	0.24	0.55
PA	GNB2L1	0.01	0.17	0.55
PA	QSCN6	0.00	0.09	0.55
PA	BCAT1	0.01	0.15	0.54
PA	CPVL	0.00	0.07	0.54
PA	LIF	0.00	0.13	0.54
PA	LBR	0.00	0.06	0.53
PA	VAMP5	0.00	0.07	0.53
PA	STXBP2	0.01	0.19	0.52
PA	ITGB3	0.02	0.23	0.51
PA	ALOX5AP	0.00	0.12	0.50
DA	ACAA1	0.01	0.21	-0.50
DA	CTH	0.00	0.09	-0.50
DA	HRPT2	0.00	0.13	-0.50
DA	CAMK2G	0.00	0.09	-0.51
DA	CROP	0.01	0.16	-0.51
DA	ADIPOR2	0.01	0.19	-0.52
DA	THRA	0.00	0.12	-0.52
DA	PPP1R12B	0.00	0.08	-0.52
DA	APC	0.02	0.23	-0.52
DA	FDFT1	0.00	0.06	-0.52
DA	BTBD3	0.01	0.14	-0.52
DA	AKT3	0.00	0.13	-0.55
DA	PSIP1	0.00	0.07	-0.55
DA	DAPK1	0.00	0.08	-0.55
DA	COG2	0.00	0.08	-0.56
DA	SEPW1	0.01	0.18	-0.57
DA	PHGDH	0.00	0.08	-0.57
DA	KCNJ3	0.01	0.16	-0.58
DA	TRIM2	0.01	0.18	-0.58
DA	KCNJ10	0.01	0.17	-0.58
DA	DST	0.01	0.17	-0.59
DA	SLC14A1	0.00	0.07	-0.59
DA	TLOC1	0.00	0.06	-0.59
DA	TTYH1	0.00	0.08	-0.60
DA	MAPK8IP2	0.01	0.18	-0.60
DA	RASGRF1	0.01	0.15	-0.60
DA	PLCL1	0.01	0.14	-0.61
DA	SLC7A11	0.00	0.08	-0.61
DA	NPY	0.01	0.16	-0.62
DA	PECI	0.01	0.16	-0.62
DA	ARHGAP5	0.00	0.10	-0.63
DA	WASF3	0.00	0.07	-0.63
DA	GAB1	0.00	0.07	-0.63
DA	SLC1A3	0.01	0.14	-0.63
DA	KIAA0103	0.00	0.07	-0.63
DA	EFHD1	0.01	0.19	-0.64
DA	PTPNS1	0.00	0.09	-0.64
DA	RASA4	0.00	0.08	-0.65
DA	ADRA1A	0.00	0.09	-0.66
DA	WSB1	0.02	0.22	-0.66
DA	DCTN1	0.01	0.19	-0.67
DA	MOBP	0.02	0.23	-0.67
DA	PBP	0.01	0.18	-0.67
DA	CYP2J2	0.00	0.13	-0.68
DA	ITPKB	0.01	0.19	-0.68
DA	DAAM2	0.01	0.15	-0.68
DA	TMEM47	0.00	0.08	-0.70
DA	USP1	0.02	0.23	-0.71
DA	GJA4	0.00	0.07	-0.71
DA	RAB30	0.00	0.12	-0.72
DA	PON2	0.02	0.25	-0.73
DA	GPM6B	0.00	0.07	-0.73
DA	SCHIP1	0.01	0.17	-0.73
DA	RAGE	0.00	0.10	-0.73
DA	PCDH9	0.00	0.10	-0.74
DA	PCDH17	0.01	0.18	-0.74
DA	KCNN2	0.00	0.12	-0.76
DA	MACF1	0.01	0.14	-0.77
DA	ITGA6	0.00	0.11	-0.78
DA	TUBB4	0.02	0.25	-0.78
DA	ACTN2	0.01	0.18	-0.80
DA	CAMK4	0.00	0.07	-0.80
DA	PRODH	0.00	0.10	-0.82

DA	TJP1	0.00	0.06	-0.82
DA	BMPR1B	0.01	0.15	-0.82
DA	KCNJ4	0.00	0.11	-0.84
DA	CPE	0.01	0.19	-0.85
DA	SC5DL	0.00	0.07	-0.86
DA	PTPRS	0.00	0.08	-0.87
DA	INSIG1	0.00	0.08	-0.88
DA	GJA1	0.01	0.16	-0.91
DA	PSAT1	0.00	0.08	-0.92
DA	VIL2	0.01	0.18	-0.95
DA	ELOVL2	0.01	0.17	-0.96
DA	AQP4	0.01	0.16	-0.98
DA	SOX9	0.00	0.13	-0.99
DA	CTNND2	0.00	0.08	-1.01
DA	DIO2	0.01	0.14	-1.05
DA	GPR37	0.00	0.09	-1.10
DA	NDRG2	0.00	0.13	-1.12
DA	FXYD1	0.01	0.15	-1.16
DA	ATP1A2	0.00	0.11	-1.17
DA	CLEC3B	0.01	0.16	-1.18
DA	SLC6A11	0.01	0.16	-1.23
DA	ADD3	0.00	0.08	-1.29
DA	SLC17A7	0.00	0.09	-1.29
DA	RORB	0.00	0.06	-1.40
DA	AHCYL1	0.00	0.07	-1.40
DA	MLC1	0.00	0.09	-1.44

B. List of the gene-sets significantly differentially expressed between ST PAs and DAs

NAME	p value	FDR
KEGG AUTOIMMUNE THYROID DISEASE	0.00	0.02
KEGG COMPLEMENT AND COAGULATION CASCADES	0.00	0.01
KEGG TYPE I DIABETES MELLITUS	0.00	0.02
KEGG SYSTEMIC LUPUS ERYTHEMATOSUS	0.00	0.02
KEGG VIRAL MYOCARDITIS	0.00	0.02
KEGG ALLOGRAFT REJECTION	0.00	0.01
KEGG ASTHMA	0.00	0.01
KEGG LEISHMANIA INFECTION	0.00	0.03
KEGG GRAFT VERSUS HOST DISEASE	0.01	0.03
KEGG ANTIGEN PROCESSING AND PRESENTATION	0.00	0.03
PID SYNDECAN 1 PATHWAY	0.01	0.05
KEGG INTESTINAL IMMUNE NETWORK FOR IGA PRODUCTION	0.00	0.06
REACTOME MHC CLASS II ANTIGEN PRESENTATION	0.00	0.06
REACTOME GROWTH HORMONE RECEPTOR SIGNALING	0.00	0.08
BIOCARTA PML PATHWAY	0.01	0.08
PID AMB2 NEUTROPHILS PATHWAY	0.00	0.07
PID INTEGRIN3 PATHWAY	0.00	0.07
KEGG CYTOKINE CYTOKINE RECEPTOR INTERACTION	0.00	0.07
REACTOME TCR SIGNALING	0.00	0.07
REACTOME GENERATION OF SECOND MESSENGER MOLECULES	0.00	0.08
PID INTEGRIN2 PATHWAY	0.02	0.07
BIOCARTA ETS PATHWAY	0.01	0.08
REACTOME IMMUNOREGULATORY INTERACTIONS BETWEEN A LYMPHOID AND A NON LYMPHOID CELL	0.01	0.08
REACTOME INTERFERON GAMMA SIGNALING	0.01	0.08
KEGG JAK STAT SIGNALING PATHWAY	0.00	0.08
REACTOME EXTRACELLULAR MATRIX ORGANIZATION	0.01	0.08
BIOCARTA INTRINSIC PATHWAY	0.02	0.08
REACTOME RESPONSE TO ELEVATED PLATELET CYTOSOLIC CA2	0.01	0.08
REACTOME GLYCOSAMINOGLYCAN METABOLISM	0.01	0.08
PID IL4 2PATHWAY	0.01	0.08
REACTOME AMYLOIDS	0.00	0.08
REACTOME COMPLEMENT CASCADE	0.03	0.09
PID TCR PATHWAY	0.01	0.09
REACTOME ASPARAGINE N LINKED GLYCOSYLATION	0.00	0.08
REACTOME INNATE IMMUNE SYSTEM	0.00	0.08
REACTOME INTRINSIC PATHWAY	0.01	0.09
REACTOME DOWNSTREAM TCR SIGNALING	0.02	0.09
BIOCARTA TH1TH2 PATHWAY	0.01	0.09
KEGG PATHOGENIC ESCHERICHIA COLI INFECTION	0.03	0.09
REACTOME FORMATION OF FIBRIN CLOT CLOTTING CASCADE	0.03	0.09
REACTOME CYTOKINE SIGNALING IN IMMUNE SYSTEM	0.00	0.10
REACTOME COLLAGEN FORMATION	0.01	0.10
PID PTP1BPATHWAY	0.01	0.10
KEGG HEMATOPOIETIC CELL LINEAGE	0.01	0.12
PID AJDISS 2PATHWAY	0.01	0.11
PID CXCR4 PATHWAY	0.01	0.14
REACTOME CHONDROITIN SULFATE DERMATAN SULFATE METABOLISM	0.04	0.14
PID SYNDECAN 4 PATHWAY	0.02	0.14
REACTOME INTEGRIN CELL SURFACE INTERACTIONS	0.02	0.14
KEGG CELL ADHESION MOLECULES CAMS	0.01	0.14
REACTOME REGULATION OF IFNA SIGNALING	0.04	0.14
REACTOME SIGNALING BY ILS	0.01	0.14
PID ATF2 PATHWAY	0.02	0.14
REACTOME CELL SURFACE INTERACTIONS AT THE VASCULAR WALL	0.02	0.14
PID INTEGRIN1 PATHWAY	0.04	0.14
REACTOME PLATELET ACTIVATION SIGNALING AND AGGREGATION	0.00	0.15
BIOCARTA NFKB PATHWAY	0.03	0.15
PID AVB3 INTEGRIN PATHWAY	0.03	0.15
PID CD8TCRPATHWAY	0.01	0.16

REACTOME NCAM1 INTERACTIONS	0.04	0.16
REACTOME SMOOTH MUSCLE CONTRACTION	0.04	0.16
BIOCARTA LAIR PATHWAY	0.05	0.17
PID SMAD2 3NUCLEARPATHWAY	0.02	0.17
PID INTEGRIN A9B1 PATHWAY	0.03	0.17
REACTOME DEGRADATION OF THE EXTRACELLULAR MATRIX	0.07	0.17
KEGG LEUKOCYTE TRANSENDOTHELIAL MIGRATION	0.02	0.17
PID IL12 2PATHWAY	0.02	0.17
KEGG FOCAL ADHESION	0.02	0.17
PID IL12 STAT4PATHWAY	0.03	0.17
REACTOME COSTIMULATION BY THE CD28 FAMILY	0.01	0.17
REACTOME IL1 SIGNALING	0.04	0.17
REACTOME INTERFERON SIGNALING	0.02	0.18
REACTOME HEPARAN SULFATE HEPARIN HS GAG METABOLISM	0.04	0.18
PID UPA UPAR PATHWAY	0.05	0.17
REACTOME TOLL RECEPTOR CASCADES	0.01	0.21
PID FOXM1PATHWAY	0.05	0.21
PID P53DOWNSTREAMPATHWAY	0.02	0.22
BIOCARTA CTLA4 PATHWAY	0.06	0.22
KEGG P53 SIGNALING PATHWAY	0.04	0.22
PID TNFPATHWAY	0.05	0.23
KEGG ECM RECEPTOR INTERACTION	0.07	0.23
BIOCARTA SPPA PATHWAY	0.04	0.23
REACTOME SEMAPHORIN INTERACTIONS	0.04	0.23
PID GLYPICAN 1PATHWAY	0.04	0.23
REACTOME LIPID DIGESTION MOBILIZATION AND TRANSPORT	0.04	0.20
PID DELTANP63PATHWAY	0.05	0.24
REACTOME INTERFERON ALPHA BETA SIGNALING	0.08	0.24
PID IL27PATHWAY	0.08	0.24
REACTOME PRE NOTCH EXPRESSION AND PROCESSING	0.08	0.24
BIOCARTA CSK PATHWAY	0.08	0.24

C. List of the genes significantly differentially expressed between ST PAs and DNTs

upregulated	Gene	p value	FDR	Difference mean z-score PA-DNT
PA	SERPING1	0.00	0.08	1.97
PA	SERPINA3	0.00	0.05	1.79
PA	CD74	0.00	0.03	1.76
PA	RARRES1	0.00	0.04	1.69
PA	HLA-DQB1	0.00	0.09	1.57
PA	GFPT2	0.00	0.10	1.41
PA	CTGF	0.01	0.16	1.40
PA	MHC2TA	0.00	0.04	1.38
PA	LYZ	0.00	0.03	1.34
PA	LMOD1	0.00	0.05	1.34
PA	POSTN	0.00	0.04	1.33
PA	CD44	0.00	0.08	1.31
PA	ANXA1	0.01	0.14	1.26
PA	C1R	0.00	0.07	1.23
PA	DOC1	0.00	0.05	1.21
PA	F13A1	0.00	0.07	1.20
PA	GPX3	0.00	0.07	1.19
PA	PFKFB3	0.00	0.10	1.16
PA	S100A4	0.01	0.16	1.15
PA	MS4A6A	0.00	0.03	1.14
PA	AZGP1	0.01	0.17	1.11
PA	MSR1	0.00	0.08	1.10
PA	KIAA0247	0.01	0.13	1.08
PA	C3	0.00	0.03	1.08
PA	HLA-DRA	0.00	0.03	1.08
PA	ZFP36L2	0.00	0.07	1.08
PA	SRPX	0.01	0.11	1.08
PA	AEBP1	0.01	0.13	1.05
PA	GBP2	0.00	0.03	1.04
PA	LGALS1	0.01	0.17	1.04
PA	PHLDA1	0.01	0.12	1.03
PA	HLA-DMA	0.00	0.03	1.01
PA	IGFBP3	0.00	0.05	1.01
PA	CXCR4	0.01	0.11	0.98
PA	CNN2	0.00	0.09	0.95
PA	HLA-DPA1	0.01	0.12	0.95
PA	TM4SF1	0.03	0.25	0.95
PA	TRIM22	0.01	0.11	0.95
PA	IFI16	0.02	0.18	0.95
PA	ACTN1	0.00	0.08	0.94
PA	CHPF	0.00	0.04	0.94
PA	AHR	0.01	0.11	0.93
PA	EVI2B	0.00	0.10	0.90
PA	C1S	0.00	0.03	0.90
PA	SRPX2	0.00	0.03	0.88
PA	ACTA2	0.02	0.23	0.88
PA	WWTR1	0.00	0.07	0.87
PA	PBEF1	0.03	0.25	0.87
PA	RUNX1	0.01	0.12	0.87
PA	SPARC	0.00	0.03	0.86
PA	FCGR2A	0.01	0.12	0.85
PA	PLTP	0.00	0.04	0.83
PA	AKAP13	0.00	0.07	0.83
PA	TNFRSF12A	0.03	0.24	0.81

PA	NPC2	0.01	0.11	0.81
PA	LAMB1	0.00	0.08	0.81
PA	ITGB4	0.00	0.07	0.81
PA	TYROBP	0.03	0.25	0.80
PA	TAGLN	0.01	0.13	0.80
PA	LMAN2	0.00	0.03	0.80
PA	PODXL	0.02	0.21	0.79
PA	HEXA	0.00	0.03	0.79
PA	LAMP1	0.00	0.08	0.77
PA	LGALS3BP	0.02	0.22	0.77
PA	LY96	0.02	0.19	0.76
PA	TRPM8	0.00	0.08	0.74
PA	ACTG2	0.02	0.21	0.74
PA	BGN	0.02	0.20	0.74
PA	HLA-DPB1	0.00	0.04	0.74
PA	ECM2	0.02	0.19	0.72
PA	OSMR	0.01	0.11	0.72
PA	HSPA6	0.00	0.07	0.71
PA	SLC4A7	0.01	0.13	0.70
PA	ARHGDIB	0.02	0.19	0.69
PA	HCLS1	0.00	0.11	0.69
PA	IF	0.00	0.07	0.69
PA	PRSS23	0.00	0.07	0.68
PA	IGJ	0.00	0.10	0.68
PA	TMSB10	0.01	0.15	0.67
PA	SP100	0.01	0.15	0.67
PA	RCN1	0.01	0.14	0.67
PA	CHMP2A	0.00	0.07	0.67
PA	ENG	0.01	0.17	0.67
PA	COL1A1	0.01	0.13	0.67
PA	STK17A	0.01	0.17	0.66
PA	CD14	0.02	0.21	0.65
PA	NME4	0.02	0.21	0.64
PA	MAP3K8	0.00	0.09	0.63
PA	P4HB	0.01	0.15	0.63
PA	DPEP2	0.01	0.13	0.63
PA	HMGA1	0.00	0.08	0.62
PA	AIF1	0.01	0.16	0.62
PA	TXNIP	0.01	0.16	0.62
PA	COL6A1	0.01	0.16	0.61
PA	C12orf8	0.01	0.18	0.60
PA	FER1L3	0.00	0.04	0.60
PA	GNB2L1	0.00	0.07	0.60
PA	CTNNB1	0.00	0.05	0.59
PA	COL18A1	0.02	0.20	0.59
PA	RGS2	0.01	0.18	0.58
PA	ITGB1	0.00	0.03	0.57
PA	MMP14	0.00	0.09	0.57
PA	PLAU	0.00	0.09	0.57
PA	COL4A1	0.01	0.17	0.57
PA	CALR	0.01	0.14	0.56
PA	CSDA	0.00	0.04	0.56
PA	CALU	0.02	0.18	0.56
PA	PFKL	0.02	0.19	0.56
PA	TRAM1	0.00	0.10	0.56
PA	TFE3	0.00	0.07	0.56
PA	ITGB2	0.00	0.05	0.56
PA	SIX1	0.01	0.12	0.55
PA	SLC20A1	0.01	0.14	0.55
PA	STXBP2	0.02	0.20	0.55
PA	MFGE8	0.00	0.07	0.55
PA	AKAP12	0.01	0.17	0.54
PA	MCFD2	0.01	0.11	0.54
PA	PCNA	0.01	0.14	0.54
PA	MYH9	0.00	0.10	0.53
PA	FGL2	0.00	0.10	0.53
PA	SERPINB6	0.00	0.07	0.53
PA	RNASE6	0.01	0.12	0.52
PA	VAMP5	0.00	0.08	0.51
PA	ITGA3	0.02	0.19	0.51
PA	S100A10	0.02	0.24	0.50
PA	TGFBI	0.00	0.03	0.50
PA	TAGLN2	0.00	0.10	0.50
PA	FYB	0.01	0.18	0.50
PA	FRZB	0.02	0.20	0.50
DNT	NR1D2	0.02	0.23	-0.50
DNT	KLHDC3	0.01	0.12	-0.51
DNT	EDIL3	0.01	0.11	-0.51
DNT	KIAA0103	0.00	0.07	-0.51
DNT	SLC11A2	0.01	0.16	-0.51
DNT	BRP44	0.03	0.25	-0.53
DNT	OMG	0.00	0.05	-0.53
DNT	RASGRF1	0.00	0.03	-0.54
DNT	GLRB	0.00	0.09	-0.54
DNT	UGT8	0.01	0.14	-0.54
DNT	RAB33A	0.00	0.05	-0.54
DNT	ATP8A2	0.00	0.08	-0.55
DNT	MAPRE2	0.00	0.10	-0.55
DNT	SNCB	0.01	0.13	-0.55
DNT	CTNND2	0.00	0.10	-0.55
DNT	OPCML	0.03	0.25	-0.56
DNT	PPP2CA	0.02	0.22	-0.57
DNT	NFYB	0.02	0.18	-0.57
DNT	ARHGAP5	0.01	0.14	-0.57

DNT	RAN	0.01	0.12	-0.57
DNT	PIK3R1	0.01	0.12	-0.57
DNT	MYO5A	0.01	0.11	-0.57
DNT	ATP6V0A1	0.00	0.03	-0.58
DNT	KLRC3	0.00	0.07	-0.58
DNT	GSTA4	0.01	0.17	-0.58
DNT	STMN1	0.00	0.04	-0.59
DNT	PLCL1	0.00	0.03	-0.59
DNT	GNG4	0.01	0.13	-0.60
DNT	PDE9A	0.00	0.05	-0.60
DNT	STAU2	0.00	0.04	-0.60
DNT	CDK5R1	0.00	0.07	-0.60
DNT	HES1	0.01	0.13	-0.61
DNT	PRKAR1B	0.01	0.11	-0.61
DNT	COX6C	0.01	0.12	-0.61
DNT	SHC3	0.02	0.22	-0.61
DNT	RPL13	0.00	0.05	-0.61
DNT	MOBP	0.01	0.15	-0.62
DNT	NET1	0.00	0.09	-0.62
DNT	CHGB	0.01	0.15	-0.63
DNT	GPSM2	0.00	0.06	-0.63
DNT	RAB6A	0.01	0.15	-0.64
DNT	RORB	0.02	0.20	-0.64
DNT	HRMT1L2	0.01	0.11	-0.64
DNT	PTPRD	0.00	0.05	-0.64
DNT	IDI1	0.00	0.09	-0.64
DNT	SYN2	0.00	0.03	-0.65
DNT	TJP1	0.00	0.05	-0.66
DNT	GPR17	0.03	0.24	-0.66
DNT	TRIM2	0.00	0.03	-0.68
DNT	GRP	0.00	0.07	-0.68
DNT	PTPRN2	0.01	0.18	-0.69
DNT	RGS5	0.00	0.06	-0.69
DNT	BTBD3	0.00	0.05	-0.70
DNT	AMPH	0.00	0.04	-0.71
DNT	CPSF6	0.00	0.05	-0.71
DNT	C11orf8	0.00	0.05	-0.71
DNT	ABCA5	0.00	0.09	-0.73
DNT	KCNN2	0.00	0.07	-0.73
DNT	NPC1	0.01	0.12	-0.74
DNT	CAMK2N1	0.02	0.23	-0.75
DNT	NPY	0.00	0.03	-0.76
DNT	RNF144	0.00	0.05	-0.76
DNT	CLDN10	0.00	0.06	-0.76
DNT	SEZ6L	0.00	0.07	-0.76
DNT	MAP2K4	0.00	0.04	-0.77
DNT	PRKACB	0.00	0.07	-0.77
DNT	PIP5K2A	0.01	0.13	-0.78
DNT	NRGN	0.01	0.14	-0.81
DNT	SPTBN2	0.03	0.25	-0.82
DNT	PCDH17	0.01	0.11	-0.83
DNT	EDG2	0.02	0.22	-0.83
DNT	R3HDM	0.00	0.10	-0.84
DNT	GRIA2	0.01	0.13	-0.85
DNT	RGS7	0.00	0.05	-0.87
DNT	INA	0.00	0.10	-0.88
DNT	CHGA	0.01	0.18	-0.89
DNT	GPR37	0.00	0.10	-0.91
DNT	ELAVL3	0.01	0.16	-0.91
DNT	CAMK4	0.00	0.04	-0.92
DNT	EIF4A2	0.00	0.08	-0.92
DNT	MAPK8IP2	0.00	0.05	-0.93
DNT	KCNK1	0.00	0.03	-0.93
DNT	CSMD1	0.00	0.03	-0.95
DNT	ACVR2A	0.00	0.06	-0.99
DNT	KCNJ4	0.00	0.08	-1.00
DNT	PRKCB1	0.00	0.05	-1.03
DNT	DLGAP1	0.00	0.08	-1.03
DNT	FGF12	0.00	0.06	-1.04
DNT	ATP1B1	0.00	0.04	-1.04
DNT	CNTN1	0.00	0.04	-1.12
DNT	PCP4	0.01	0.11	-1.20
DNT	CHD5	0.01	0.12	-1.21
DNT	ADD3	0.00	0.08	-1.26
DNT	MT1G	0.00	0.07	-1.36
DNT	SLC17A7	0.00	0.03	-1.44
DNT	PLP1	0.01	0.14	-1.50
DNT	CLEC3B	0.00	0.08	-1.51
DNT	VSNL1	0.00	0.06	-1.56
DNT	SNAP25	0.00	0.03	-1.58
DNT	STMN2	0.00	0.06	-1.63
DNT	SCN2A2	0.00	0.06	-1.67

D. List of the gene-sets significantly enriched between ST PAs and ST DNTs

GENE-SETS ENRICHED IN THE ST PAs	p value	FDR
KEGG COMPLEMENT AND COAGULATION CASCADES	0.00	0.00
PID INTEGRIN2 PATHWAY	0.00	0.00
REACTOME INTEGRIN CELL SURFACE INTERACTIONS	0.00	0.02
PID UPA UPAR PATHWAY	0.00	0.02
KEGG VIRAL MYOCARDITIS	0.00	0.03
KEGG INTESTINAL IMMUNE NETWORK FOR IGA PRODUCTION	0.00	0.03
KEGG ASTHMA	0.00	0.03
KEGG GRAFT VERSUS HOST DISEASE	0.01	0.03
KEGG ALLOGRAFT REJECTION	0.01	0.03
KEGG AUTOIMMUNE THYROID DISEASE	0.00	0.03
KEGG LEISHMANIA INFECTION	0.01	0.03
REACTOME COMPLEMENT CASCADE	0.00	0.03
PID INTEGRIN CS PATHWAY	0.00	0.03
KEGG HEMATOPOIETIC CELL LINEAGE	0.00	0.03
KEGG ANTIGEN PROCESSING AND PRESENTATION	0.01	0.03
BIOCARTA INTRINSIC PATHWAY	0.01	0.03
REACTOME IMMUNOREGULATORY INTERACTIONS BETWEEN A LYMPHOID AND A NON LYMPHOID CELL	0.01	0.04
KEGG SYSTEMIC LUPUS ERYTHEMATOSUS	0.01	0.04
REACTOME RESPONSE TO ELEVATED PLATELET CYTOSOLIC CA2	0.00	0.04
PID INTEGRIN1 PATHWAY	0.01	0.04
BIOCARTA TH1TH2 PATHWAY	0.01	0.04
PID INTEGRIN3 PATHWAY	0.01	0.06
KEGG ECM RECEPTOR INTERACTION	0.01	0.07
REACTOME GLYCOSAMINOGLYCAN METABOLISM	0.01	0.07
PID SYNDECAN 4 PATHWAY	0.00	0.10
REACTOME INTERFERON GAMMA SIGNALING	0.02	0.10
KEGG TYPE I DIABETES MELLITUS	0.03	0.10
PID AVB3 INTEGRIN PATHWAY	0.01	0.11
PID PTP1BPATHWAY	0.02	0.11
PID ARF6 TRAFFICKINGPATHWAY	0.00	0.12
PID INTEGRIN A9B1 PATHWAY	0.02	0.12
BIOCARTA AMI PATHWAY	0.02	0.12
REACTOME FORMATION OF FIBRIN CLOT CLOTTING CASCADE	0.03	0.12
REACTOME POST TRANSLATIONAL PROTEIN MODIFICATION	0.00	0.12
PID IL4 2PATHWAY	0.04	0.12
PID AMB2 NEUTROPHILS PATHWAY	0.02	0.13
REACTOME GENERATION OF SECOND MESSENGER MOLECULES	0.01	0.13
KEGG CYTOKINE CYTOKINE RECEPTOR INTERACTION	0.01	0.13
REACTOME EXTENSION OF TELOMERES	0.02	0.13
KEGG JAK STAT SIGNALING PATHWAY	0.00	0.13
REACTOME HEPARAN SULFATE HEPARIN HS GAG METABOLISM	0.02	0.13
KEGG NUCLEOTIDE EXCISION REPAIR	0.01	0.13
REACTOME LIPOPROTEIN METABOLISM	0.02	0.13
PID CXCR4 PATHWAY	0.01	0.13
REACTOME CELL SURFACE INTERACTIONS AT THE VASCULAR WALL	0.02	0.13
REACTOME CHONDROITIN SULFATE DERMATAN SULFATE METABOLISM	0.02	0.13
PID HIVNEFPATHWAY	0.01	0.13
PID AJDISS 2PATHWAY	0.01	0.13
KEGG CELL ADHESION MOLECULES CAMS	0.02	0.14
REACTOME SPHINGOLIPID METABOLISM	0.02	0.14
REACTOME PLATELET ACTIVATION SIGNALING AND AGGREGATION	0.00	0.14
BIOCARTA LAIR PATHWAY	0.03	0.14
BIOCARTA SPPA PATHWAY	0.01	0.15
REACTOME REGULATION OF INSULIN LIKE GROWTH FACTOR IGF ACTIVITY BY INSULIN LIKE GROWTH FACTOR BINDING PROTEINS IGFBPS	0.03	0.15
KEGG FOCAL ADHESION	0.01	0.15
BIOCARTA INFLAM PATHWAY	0.05	0.15
REACTOME GLOBAL GENOMIC NER GG NER	0.01	0.16
REACTOME EXTRACELLULAR MATRIX ORGANIZATION	0.05	0.16
KEGG LEUKOCYTE TRANSENDOTHELIAL MIGRATION	0.00	0.18
PID TCR PATHWAY	0.02	0.18
BIOCARTA HIVNEF PATHWAY	0.01	0.18
PID SYNDECAN 1 PATHWAY	0.06	0.19
PID A6B1 A6B4 INTEGRIN PATHWAY	0.02	0.19
KEGG REGULATION OF ACTIN CYTOSKELETON	0.01	0.19
REACTOME GLYCOSPHINGOLIPID METABOLISM	0.06	0.19
REACTOME LAGGING STRAND SYNTHESIS	0.03	0.19
BIOCARTA DEATH PATHWAY	0.04	0.19
KEGG LYSOSOME	0.03	0.19
REACTOME ASPARAGINE N LINKED GLYCOSYLATION	0.03	0.19
BIOCARTA NFKB PATHWAY	0.05	0.19
PID TAP63PATHWAY	0.05	0.19
PID DELTANP63PATHWAY	0.04	0.19
REACTOME HS GAG DEGRADATION	0.06	0.19
PID ILK PATHWAY	0.02	0.20
KEGG PRIMARY IMMUNODEFICIENCY	0.07	0.20
REACTOME COLLAGEN FORMATION	0.08	0.20
PID MYC REPRESSPATHWAY	0.03	0.20
KEGG ARACHIDONIC ACID METABOLISM	0.04	0.21
PID FRA PATHWAY	0.09	0.22
KEGG P53 SIGNALING PATHWAY	0.04	0.22
PID AR PATHWAY	0.04	0.22
PID P53DOWNSTREAMPATHWAY	0.04	0.22
REACTOME COSTIMULATION BY THE CD28 FAMILY	0.02	0.22
PID AVB3 OPN PATHWAY	0.03	0.22
BIOCARTA DC PATHWAY	0.09	0.22
BIOCARTA RHO PATHWAY	0.03	0.23
REACTOME DNA STRAND ELONGATION	0.07	0.23

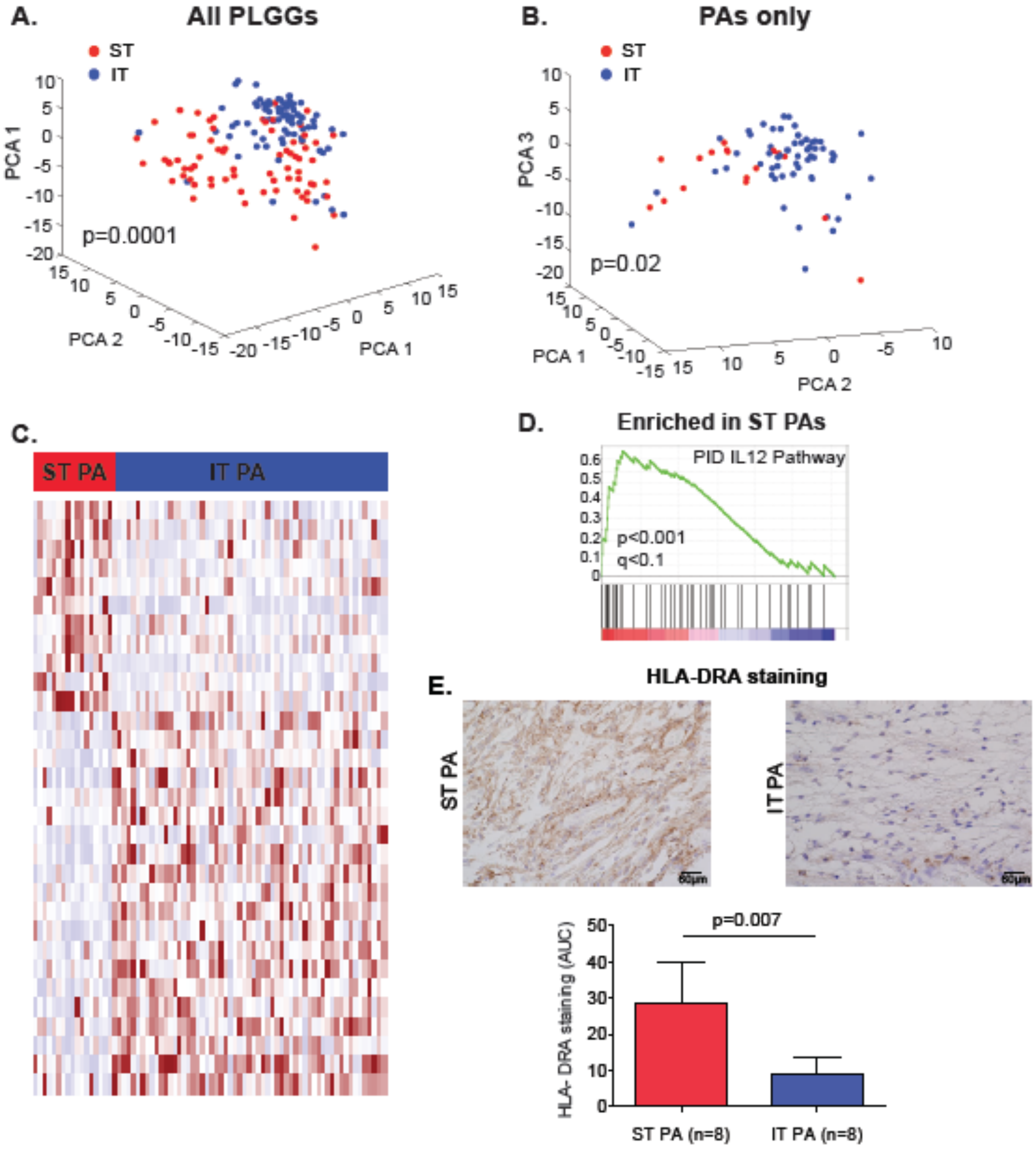
REACTOME NCAM1 INTERACTIONS	0.08	0.23
REACTOME NUCLEOTIDE EXCISION REPAIR	0.02	0.23
BIOCARTA UCALPAIN PATHWAY	0.05	0.23
KEGG SMALL CELL LUNG CANCER	0.04	0.23
KEGG DNA REPLICATION	0.07	0.24
GENE-SETS ENRICHED IN THE ST DNTs		
REACTOME POST NMDA RECEPTOR ACTIVATION EVENTS	0.00	0.08
REACTOME NEUROTRANSMITTER RELEASE CYCLE	0.01	0.08
BIOCARTA NO1 PATHWAY	0.01	0.08
REACTOME NEUROTRANSMITTER RECEPTOR BINDING AND DOWNSTREAM TRANSMISSION IN THE POSTSYNAPTIC CELL	0.00	0.08
REACTOME TRANSMISSION ACROSS CHEMICAL SYNAPSES	0.00	0.09
BIOCARTA NOS1 PATHWAY	0.00	0.09
REACTOME NITRIC OXIDE STIMULATES GUANYLATE CYCLASE	0.00	0.11
REACTOME PLC BETA MEDIATED EVENTS	0.00	0.13
REACTOME OPIOID SIGNALLING	0.00	0.13
REACTOME CA DEPENDENT EVENTS	0.01	0.14
REACTOME DAG AND IP3 SIGNALING	0.01	0.14
REACTOME ACTIVATION OF NMDA RECEPTOR UPON GLUTAMATE BINDING AND POSTSYNAPTIC EVENTS	0.00	0.14
KEGG LONG TERM POTENTIATION	0.00	0.14
KEGG OLFACTORY TRANSDUCTION	0.00	0.15
KEGG CARDIAC MUSCLE CONTRACTION	0.01	0.15
PID RAS PATHWAY	0.00	0.16
REACTOME PHOSPHOLIPASE C MEDIATED CASCADE	0.02	0.16
BIOCARTA GPCR PATHWAY	0.00	0.16
REACTOME DARPP 32 EVENTS	0.01	0.16
REACTOME ION CHANNEL TRANSPORT	0.01	0.16
REACTOME CREB PHOSPHORYLATION THROUGH THE ACTIVATION OF RAS	0.02	0.16
BIOCARTA PGC1A PATHWAY	0.01	0.17
BIOCARTA VIP PATHWAY	0.00	0.17
KEGG GAP JUNCTION	0.01	0.17
REACTOME NEURONAL SYSTEM	0.00	0.20
REACTOME POTASSIUM CHANNELS	0.02	0.22
REACTOME SIGNALING BY FGFR IN DISEASE	0.01	0.23
BIOCARTA AT1R PATHWAY	0.03	0.24

3.3. Variations in expression patterns by tumor location

We also observe significant differences in expression between ST and IT tumors on both a global scale (p=0.0001; Figure 26A) and at the individual gene level, after controlling for histology. We compared Euclidean distances within and between the two groups, controlling for histology by maintaining histology class labels. We also found significant differences after restricting the analysis to PAs (p=0.02). These differences can be visualized using the first three principle components of a Principal Components Analysis (PCA) across the whole cohort (Figure 26A) and across the PAs (Figure 26B). Among PAs, 31 genes and 21 gene-sets were significantly differentially expressed between ST and IT tumors (Figure 26C, Table 10A-B). All 21 differentially expressed gene-sets were enriched among the ST tumors, including three that were associated with IL12 pathway activation (out of four total; p<0.0001; Figure 26D).

Figure 26: ST and IT PAs exhibit different molecular patterns

A-B. PCA analyses comparing ST and IT PLGGs (A) and ST and IT PAs (B). C. Heat-map of genes found to be differentially expressed between ST and IT PAs. D. Enrichment plot of an IL12 gene-set found to be significantly enriched in ST PAs relative to IT PAs. E. HLA-DRA immunohistochemical staining of individual ST and IT PAs (top) and summary results across 8 ST and 8 IT tumors (bottom). The y-axis represents the area under the curve (AUC) metric from Cell Profiler. * p<0.05.



Among our ST tumors, 58 were cortical and 15 were midline. After controlling for histology as a confounder, we observed 73 genes and 3 gene-sets significantly differentially expressed between the two ST regions (Table 10D and 10E).

These differences between ST and IT tumors appear to be tumor-specific and not reproduced in normal brain tissue. We compared ST and IT expression profiles from 12 children's brains in the publicly available BRAINSPAN dataset[332]. Among the 31 genes that were differentially expressed between ST and IT PAs, 30 were also included in the BRAINSPAN dataset. Among these, 14 genes were differentially expressed in the same direction and 16 in the opposite direction between tumors and normal brains (p = 0.29). We also performed GSEA on 12 ST and 12 IT normal samples from children in the BRAINSPAN database (Table 10C). We did not observe significant enrichment in either direction of any inflammatory pathways.

However, we did find indications of increased IL12 pathway activation in ST tumors in expression data from an independent set of 16 ST and 30 IT PAs[229]. We evaluated all four IL12 gene-sets that are included in the C2 CP set. One of these was enriched in the ST PAs (p=0.03); the other three exhibited trends in that direction (p=0.0984, 0.1262, and 0.2350, respectively).

We further evaluated this finding by performing immunohistochemistry for HLA-DRA in an independent set of 8 ST and 8 IT PAs. HLA-DRA is a major effector in immune response and was one of the genes overexpressed in ST relative to IT PAs. We observed significantly higher expression of HLA-DRA in the ST (Figure 26E, p = 0.007). The cells that stained positive for HLA-DRA included cells within vessels (possibly lymphocytes) and also cells with glial morphology within the tumor, suggesting that both populations may contribute to the increase in inflammatory signatures in ST PAs.

Table 10:

A. List of the genes significantly differentially expressed between ST and IT PAs.

Upregulated in	Name	p value	FDR	Difference mean z-score ST – IT PAs
ST PA	GFPT2	0.00	0.14	1.23
ST PA	RARRES1	0.00	0.20	0.91
ST PA	LYZ	0.00	0.13	0.80
ST PA	F13A1	0.00	0.09	0.75
ST PA	CD74	0.00	0.24	0.67
ST PA	SIX1	0.00	0.09	0.62
ST PA	LAMB1	0.00	0.14	0.60
ST PA	TRIM22	0.00	0.19	0.59
ST PA	IGJ	0.00	0.09	0.58
ST PA	SLIT1	0.00	0.19	0.58
ST PA	HLA-DRA	0.00	0.11	0.54
IT PA	NOS1	0.00	0.16	-0.51
IT PA	KLHDC3	0.00	0.09	-0.53
IT PA	GRIA4	0.00	0.09	-0.53
IT PA	CHAD	0.00	0.14	-0.53
IT PA	ABAT	0.00	0.09	-0.55
IT PA	NTRK2	0.00	0.19	-0.61
IT PA	TRH	0.00	0.19	-0.65
IT PA	CSMD1	0.00	0.09	-0.69
IT PA	EDIL3	0.00	0.09	-0.70
IT PA	RHOB	0.00	0.20	-0.71
IT PA	PCDH17	0.00	0.09	-0.76
IT PA	AGT	0.00	0.19	-0.76
IT PA	FGF12	0.00	0.09	-0.79
IT PA	CPE	0.00	0.09	-0.81
IT PA	CNTN1	0.00	0.09	-0.81
IT PA	RIT2	0.00	0.16	-0.99
IT PA	COL9A1	0.00	0.19	-1.03
IT PA	CDH19	0.00	0.09	-1.03
IT PA	AGTRL1	0.00	0.09	-1.30
IT PA	S100A1	0.00	0.21	-1.54

C. Gene sets significantly enriched in ST PAs in comparison with IT PAs

NAME	P value	FDR
PID IL12 STAT4PATHWAY	0.00	0.01
KEGG ALLOGRAFT REJECTION	0.01	0.05
BIOCARTA CTLA4 PATHWAY	0.00	0.05
KEGG SYSTEMIC LUPUS ERYTHEMATOSUS	0.00	0.06
KEGG AUTOIMMUNE THYROID DISEASE	0.01	0.06
KEGG VIRAL MYOCARDITIS	0.00	0.06
PID IL12 2PATHWAY	0.00	0.07
KEGG INTESTINAL IMMUNE NETWORK FOR IGA PRODUCTION	0.00	0.07
KEGG ASTHMA	0.01	0.08
BIOCARTA SPRY PATHWAY	0.01	0.08
BIOCARTA IL12 PATHWAY	0.00	0.09
KEGG GRAFT VERSUS HOST DISEASE	0.01	0.11
BIOCARTA TH1TH2 PATHWAY	0.01	0.12
KEGG CYTOKINE CYTOKINE RECEPTOR INTERACTION	0.01	0.16
KEGG LEISHMANIA INFECTION	0.00	0.16
REACTOME IMMUNOREGULATORY INTERACTIONS BETWEEN A LYMPHOID AND A NON LYMPHOID CELL	0.02	0.17
KEGG TYPE I DIABETES MELLITUS	0.02	0.19
PID TCR PATHWAY	0.01	0.20
PID INTEGRIN3 PATHWAY	0.02	0.24
BIOCARTA CSK PATHWAY	0.01	0.24
REACTOME DOWNSTREAM TCR SIGNALING	0.03	0.24

D. List of the 30 genes significantly differentially expressed between ST and IT PAs related to the expression in the Brainspan dataset

Gene	Expression in normal brain (BRAINSPAN)
CDH19	OPPOSITE
CHAD	SAME
FGF12	OPPOSITE
HLA-DRA	OPPOSITE
KLHDC3	OPPOSITE
PCDH17	OPPOSITE
RHOB	SAME
RIT2	SAME
AGT	SAME
COL9A1	OPPOSITE
F13A1	SAME
RARRES1	SAME
TRH	OPPOSITE
CNTN1	SAME
EDIL3	OPPOSITE
GRIA4	SAME
LAMB1	OPPOSITE
NOS1	OPPOSITE
NTRK2	OPPOSITE
S100A1	OPPOSITE
SLIT1	SAME
ABAT	OPPOSITE
CD74	OPPOSITE
CPE	SAME
CSMD1	OPPOSITE
GFPT2	OPPOSITE
IGJ	SAME
LYZ	SAME
SIX1	SAME
TRIM22	SAME

D. List of the 73 genes significantly differentially expressed between midline and cortical tumors

Upregulated in	Name	p value	FDR	Difference mean Zscore midline - cortex
midline	LMOD1	0,01	0,21	1,33
midline	FOS	0,00	0,10	1,31
midline	CD44	0,00	0,14	1,18
midline	AEBP1	0,00	0,10	1,13
midline	CAV1	0,01	0,22	1,12
midline	AGTRL1	0,01	0,24	1,10
midline	C1R	0,01	0,21	1,09
midline	ID4	0,00	0,10	0,98
midline	RARRES3	0,00	0,10	0,95
midline	ZFP36L2	0,00	0,14	0,94
midline	ECM2	0,00	0,11	0,82
midline	GYPC	0,01	0,22	0,81
midline	LGALS3BP	0,00	0,14	0,80
midline	SIX1	0,00	0,10	0,77
midline	PRRX1	0,00	0,20	0,76
midline	NFIL3	0,01	0,22	0,74
midline	CHPF	0,01	0,21	0,73
midline	S100A10	0,00	0,10	0,72
midline	SLC39A14	0,00	0,15	0,71
midline	FCGBP	0,00	0,15	0,71
midline	GLUL	0,01	0,21	0,70
midline	PRSS23	0,00	0,10	0,69
midline	AHNAK	0,01	0,23	0,68
midline	ARHGDIA	0,00	0,16	0,67
midline	COL6A1	0,00	0,13	0,66
midline	FMOD	0,00	0,20	0,66
midline	TRPM8	0,01	0,22	0,65
midline	CP	0,00	0,20	0,64
midline	HCLS1	0,00	0,20	0,63
midline	MVP	0,00	0,10	0,61
midline	PLTP	0,01	0,22	0,61
midline	PLAU	0,00	0,10	0,59
midline	DDR1	0,00	0,16	0,59
midline	CTNNB1	0,00	0,14	0,57
midline	EMP2	0,00	0,17	0,57
midline	RFX1	0,00	0,10	0,56
midline	PKM2	0,01	0,24	0,56
midline	SPARC	0,00	0,11	0,55
midline	NFE2L1	0,01	0,23	0,51
midline	SERPINA5	0,00	0,11	0,51
cortex	NFYB	0,00	0,17	-0,51
cortex	RGS5	0,01	0,22	-0,52
cortex	CPSF6	0,00	0,14	-0,54
cortex	SLC1A3	0,00	0,18	-0,54
cortex	C11orf8	0,00	0,16	-0,54
cortex	PPP2CA	0,00	0,19	-0,55
cortex	COX7B	0,00	0,20	-0,58
cortex	NRGN	0,00	0,20	-0,58
cortex	R3HDM	0,01	0,22	-0,58
cortex	SKP1A	0,00	0,14	-0,59
cortex	DBI	0,00	0,14	-0,60
cortex	RGS7	0,00	0,10	-0,60
cortex	CTNND2	0,01	0,22	-0,61
cortex	FOXG1B	0,00	0,10	-0,62
cortex	ACSL3	0,00	0,19	-0,63
cortex	GPM6A	0,00	0,10	-0,65
cortex	NPY	0,01	0,22	-0,66
cortex	NGFRAP1	0,01	0,22	-0,68
cortex	AES	0,00	0,20	-0,68
cortex	NM 001069	0,01	0,24	-0,70
cortex	GP1BB	0,01	0,24	-0,73
cortex	PC4	0,00	0,20	-0,76
cortex	PIP5K1B	0,00	0,14	-0,78
cortex	PBP	0,00	0,14	-0,80
cortex	PLS3	0,00	0,19	-0,81
cortex	CAMK4	0,00	0,10	-0,83
cortex	ATP5G3	0,01	0,24	-0,85
cortex	DLGAP1	0,00	0,14	-0,91
cortex	SLC17A7	0,00	0,14	-0,92
cortex	ST13	0,01	0,22	-0,95
cortex	GABRB1	0,00	0,13	-0,98
cortex	SNAP25	0,00	0,10	-1,12
cortex	VSNL1	0,00	0,10	-1,26

E. List of the gene-sets significantly differentially expressed between midline and cortical tumors

Enriched in	Name	p value	FDR
midline	REACTOME MEMBRANE BINDING AND TARGETTING OF GAG PROTEINS	0,00	0,18
cortex	KEGG PARKINSONS DISEASE	0,00	0,18
cortex	REACTOME POST NMDA RECEPTOR ACTIVATION EVENTS	0,00	0,18

3.4. Molecular differences associated with *BRAF* alterations

We observed significant differences in levels of individual genes and gene-sets between *BRAF*-duplicated and *BRAF-V600E* mutated tumors. We performed Comparative Marker Selection between the *BRAF*-duplicated or mutated tumors and WT tumors, controlling for both location and histology as possible confounders by maintaining these class labels during all permutations. We evaluated both individual gene scores and sums of scores across all genes within the gene-sets listed in the MSigDBC2P database[335]. No genes or gene-sets were significantly differentially expressed between the *BRAF*-duplicated tumors and WT tumors or between *BRAF-V600E* mutated and WT tumors. We did find 15 genes that were differentially expressed between the *BRAF*-duplicated and -V600E mutated tumors (Figure 27A, Table 11). Eleven genes were overexpressed in *BRAF*-duplicated tumors, including the calcium-binding protein S100A1, involved in the regulation of cell cycle progression and differentiation, and the transcription factor ID4. Four genes were overexpressed in *BRAF-V600E*-mutant tumors, including *ANK1*, an ankyrin involved in cell motility and proliferation, and the tyrosine kinase *PTK2B*, which regulates neuronal activity.

To evaluate global differences, we compared the Euclidean distances between *BRAF*-duplicated, *BRAF-V600E* mutated, and *BRAF-WT* tumors to Euclidean distances between tumors within each of these classes, after controlling for location and histology. We found no significant differences between any of these groups (p=0.15, p=0.40, p=0.26, respectively; Figure 27B).

Figure 27: Variation in gene expression between *BRAF*-KIAA duplicated and *BRAF-V600E* mutated tumors

A. Heat-map representing genes found to be differentially expressed between *BRAF*-duplicated and V600E-mutated tumors.

B. Histogram representing the Euclidean distances between and within BRAF-duplicated and WT tumors, between and within *BRAF-V600E* mutated and WT tumors, and between and within *BRAF*-duplicated tumors and *BRAF-V600E* mutated tumors.

P-values represent the probability that the Euclidean distances between tumor groups are significantly higher than distances within tumor groups.

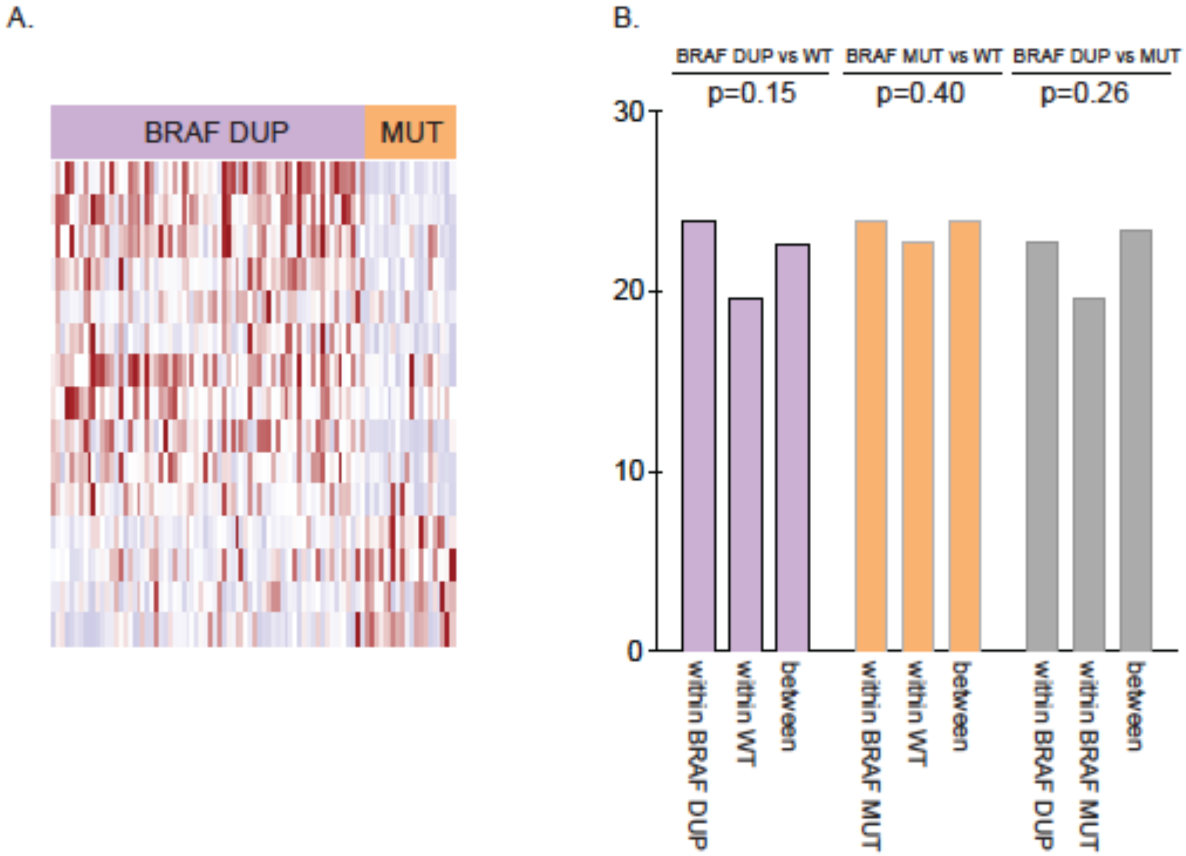


Table 11: **Genes exhibiting significantly different expression between *BRAF*-duplicated and V600E-mutated tumors.**

Upregulated in	gene	p value	FDR	Difference mean z-score BRAF DUP - BRAF MUT
DUP	S100A1	0.00	0.23	2.57
DUP	ID4	0.00	0.17	1.24
DUP	ABAT	0.00	0.24	1.00
DUP	NEBL	0.00	0.17	0.84
DUP	PCP4	0.00	0.17	0.80
DUP	AGT	0.00	0.21	0.77
DUP	LSAMP	0.00	0.23	0.74
DUP	NTRK2	0.00	0.17	0.73
DUP	NOS1	0.00	0.20	0.64
DUP	CNTN1	0.00	0.17	0.64
DUP	ADCYAP1R1	0.00	0.17	0.54
MUT	PTK2B	0.00	0.20	-0.53
MUT	ANK1	0.00	0.23	-0.90
MUT	POSTN	0.00	0.17	-0.91
MUT	PHLDA2	0.00	0.21	-0.98

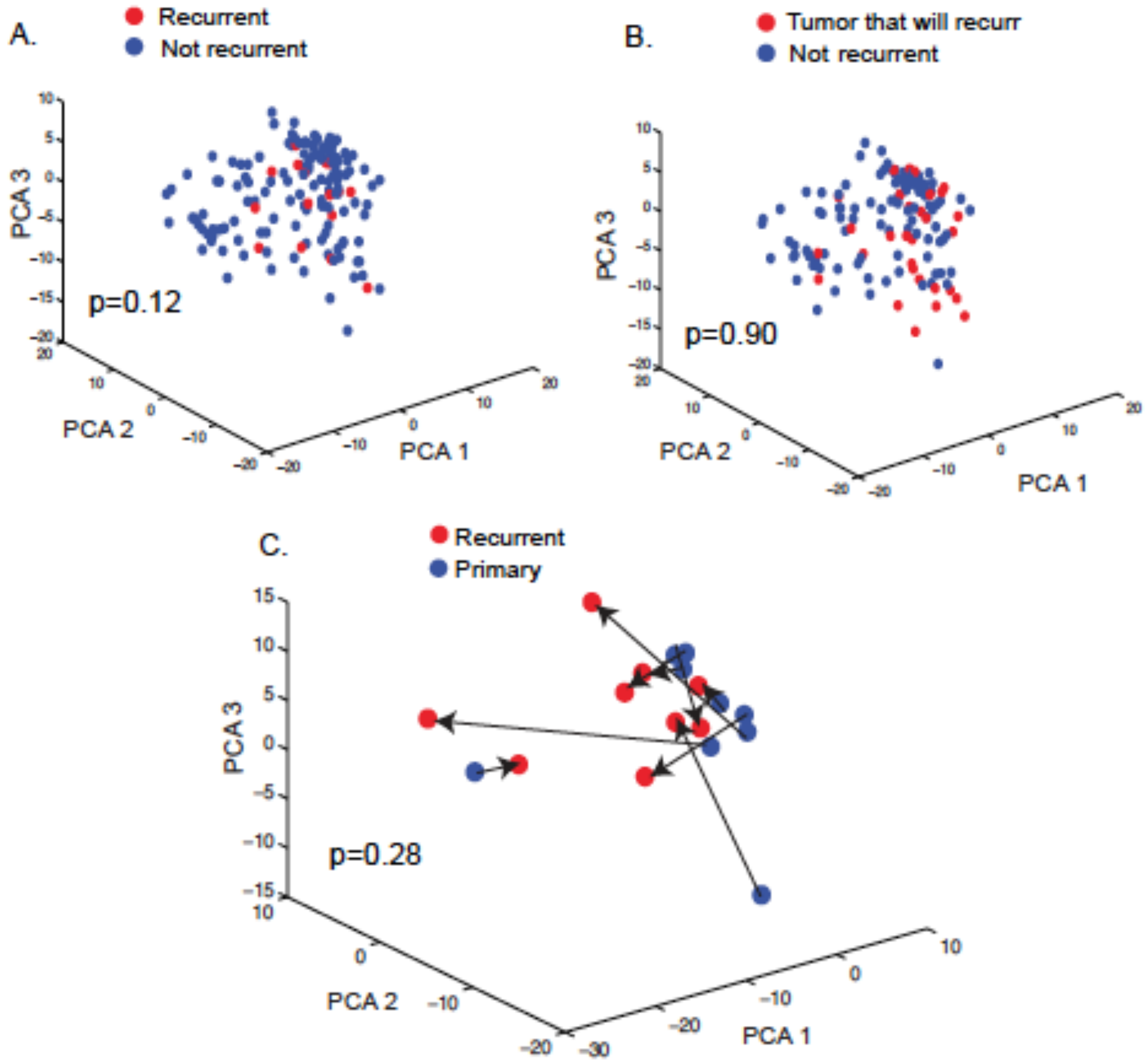
3.5. Differences between primary and recurrent tumors

Our cohort included 31 primary tumors from children who went on to have recurrences and 16 recurrent tumors, with overlap in nine cases. We observed no significant differences in the overall Euclidean distances between primary and recurrent tumor samples or between the 104 primary tumors with no record of recurrence and the 31 tumors that went on to progress (p=0.12, p=0.9, respectively; Figures 28A-B). We also found no genes or gene-sets whose expression differed significantly between primary and recurrent PLGGs after controlling for location and histology.

Compared to primary tumors from the same patients, recurrent tumors exhibited enrichment of two genes: the zinc finger *ZNF652* and the kinase anchor protein *AKAP9*, which interacts with signaling proteins from various signal transduction pathways. However, there was no significant difference in overall Euclidean distances between the paired primary and recurrent tumors (p=0.28), and the changes between them did not favor any direction or magnitude (Figure 28C).

Figure 28: Recurrent and primary PLGGs do not exhibit clearly distinct molecular patterns.

A-C. PCA analysis comparing the expression pattern of primary and recurrent PLGGs (A), the 104 primary tumors that have no record of recurrence and the 31 primary tumors that have relapsed or progressed (B), and the nine paired primary and recurrent samples (C). Arrows match each primary to its recurrent pairs.



3.6. Similar expression patterns in childhood and adolescent PAs

Tumors that occurred in young children also exhibited no significant differences compared to adolescents approaching adulthood. We evaluated both individual genes and gene-sets, and controlled for location and histology by selecting only IT PAs for comparison.

4. Discussion

PLGGs are characterized by a spectrum of histological subtypes, different ages and locations of development within the brain, and varied genetic alterations. We examined how expression profiles of PLGGs are shaped by their clinical, histopathological, and *BRAF* genomic characteristics, and whether differences associated with age and location were reflected in expression profiles from normal brain.

We identified substantial differences between ST and IT tumors after controlling for histology, indicating that tumor location significantly influences controlling for histology, indicating that tumor location significantly influences PLGG expression profiles. Only tumors with purities greater than 70% were included in this analysis, however. The differences were also not significantly similarly reflected in ST and IT normal brain expression, suggesting that the differences were specific to the cancer cells or to their interactions with normal cells within the tumor. ST and IT PAs have previously been found to harbor distinct expression [319,320] and methylation profiles[336]. However, the genes found to be differentially expressed in the prior study[319,320] did not overlap with our gene set. Among the 15 genes reported to be differentially expressed in that study that were represented in our DASL platform, two-thirds were similarly differentially expressed between normal brain ST and IT regions (p=0.3).

ST PAs are enriched for IL12 pathway activity compared to IT PAs. IL12 activates T-cells and regulates the production of IFN-γ via JAK2 and STAT4[337,338]. HLA-DRA, which drives various inflammatory pathways, was specifically enriched among ST PAs[339]. These findings suggest either an inflammatory signal in ST PA tumor cells or inflammatory activation in the tumor microenvironment. HLA-DRA expression in colorectal cancer has been reported to reflect an immunologic response in the tumor bed that is associated with a better outcome[340,341].

We identified molecular differences between PAs and DAs and between PAs and DNTs, even after controlling for location. Although DAs and PAs are both classified as astrocytic tumors by the WHO consensus, the differences we detect suggest either that they arise from distinct cells of origin or the distinct genomic abnormalities that drive these tumor subtypes are reflected in their expression profiles. Conversely, other PLGG subtypes such as PAs and GGs tend to cluster together in the same SOM group and did not exhibit statistically significant differences using a distance permutation test. This was surprising, as their morphologies and *BRAF* alterations differ. It is possible that differences will be revealed by larger tumor cohorts with increased statistical power.

NOS tumors distributed equally across the 3 clusters in our unsupervised analyses, suggesting that they do not constitute a unified molecular class. The finding that their distribution was dependent on their *BRAF* alteration status suggests the NOS class includes tumors with strong molecular similarities to existing histologic classes (eg PAs and GGs). It may be appropriate to include such molecular features in the diagnostic criteria for these tumors so as to reclassify the NOS tumors among the well-defined histologic groups with which they are most similar.

BRAF-duplicated and *BRAF-V600E* mutated tumors exhibited differences after controlling for location and histological subtype, including differences in expression levels of genes involved in cell development and differentiation. However, these differences were not extensive. These tumors also tended to cluster separately in our unsupervised analysis, but not to a statistically significant degree. It is possible that we would detect greater differences in a larger cohort or using an assay that evaluates larger numbers of genes. However, the minimal differences between these tumor subsets, as compared to the differences between ST and IT tumors, may indicate that these alterations lead to similar downstream effects. Similarly, our inability to detect substantial differences in gene expression profiles of *BRAF*-altered tumors compared to wild-type tumors might be related to a lack of power to detect small differences in our dataset, to a minimal impact of *BRAF* alterations on gene expression profiles or to the existence of genomic alterations in the wild-type tumors that lead to similar downstream effects. These alterations might include mutations or copy-number alterations of *FGFR1* or *NTRK2* rearrangements, all of which have recently been described in PLGGs and shown to lead to lead to MAPK activation[228,229].

We find that PLGG expression patterns do not substantially evolve with progression. This is consistent with clinical observations that although PLGGs do often progress or recur, they rarely undergo malignant transformation[100] and often respond to the same chemotherapy regimen[117].

We also found no significant differences in outcome between the three different clusters of primary tumors, possibly due to lack of power as PLGGs grow slowly.

A significant complication in the analysis of bulk gene expression profiles, as obtained here, is that they reflect the sum of profiles across the many subclones and cell states of both tumor and normal cells within the sample[342,343]. For example, the finding that ST PAs exhibit enhanced inflammatory pathways may reflect changes within the cancer cells or an increased inflammatory response. Single-cell expression profiling[344,345] may enable us to disentangle these different contributions to the bulk profile, thereby providing important information about the molecular programs that are altered specifically in the tumor cells and in their microenvironment.

Chapter 4: Single cell expression profiling

1. Introduction

One major limitation of bulk expression profiling is the impurity of the signal obtained. Indeed, in a standard punch biopsy, several cell types are mixed and contribute to the overall expression pattern. Using high purity samples decreases the fraction of contaminating normal cells in the analyses. However, beside the cancer cells that are analyzed, there is a mixture of interstitial cells, vascular cells, and epithelial cells that might influence the overall expression signal[343,346,347]. Additionally, a same cell type might present a different expression signal due to the stochastic process of those cells[343]. The cancer cells themselves may also reside in a number of different differentiation states.

Recent advances in cell isolation and sequencing have enabled the possibility to isolate and extract RNA and sequence the transcriptome of single cells[348]. This progress has enabled analyses of the transcriptome with higher resolution to understand deeper the biology of distinct cell types in cells and tumors. However these techniques have never been applied to a pediatric tumor type. We have previously performed FFPE bulk expression profiling on more than 150 PLGGs that revealed substantial tumor heterogeneity on the level of gene expression. We therefore wanted to test this cutting edge procedure to PLGG tumors in order to explore the gene expression profiles of individual cells from dissociated tumors obtained by flow cytometry.

This project was conducted with different steps. We first tested the hypothesis that we can dissociate and sort pediatric glial cells from a LGG tumor with good viability. We then tested different antibodies to explore the expression of those markers in several tumors. We then optimized the extraction, transcriptome amplification and library preparation protocols to our PLGG cells. Here we describe the results of single cell RNA-sequencing originated from 3 fresh PLGG tumors.

2. Methods

The main steps of the protocol are summarized in Figure 29.

2.1. Tumor dissociation, antibody staining and single-cell lysate preparation

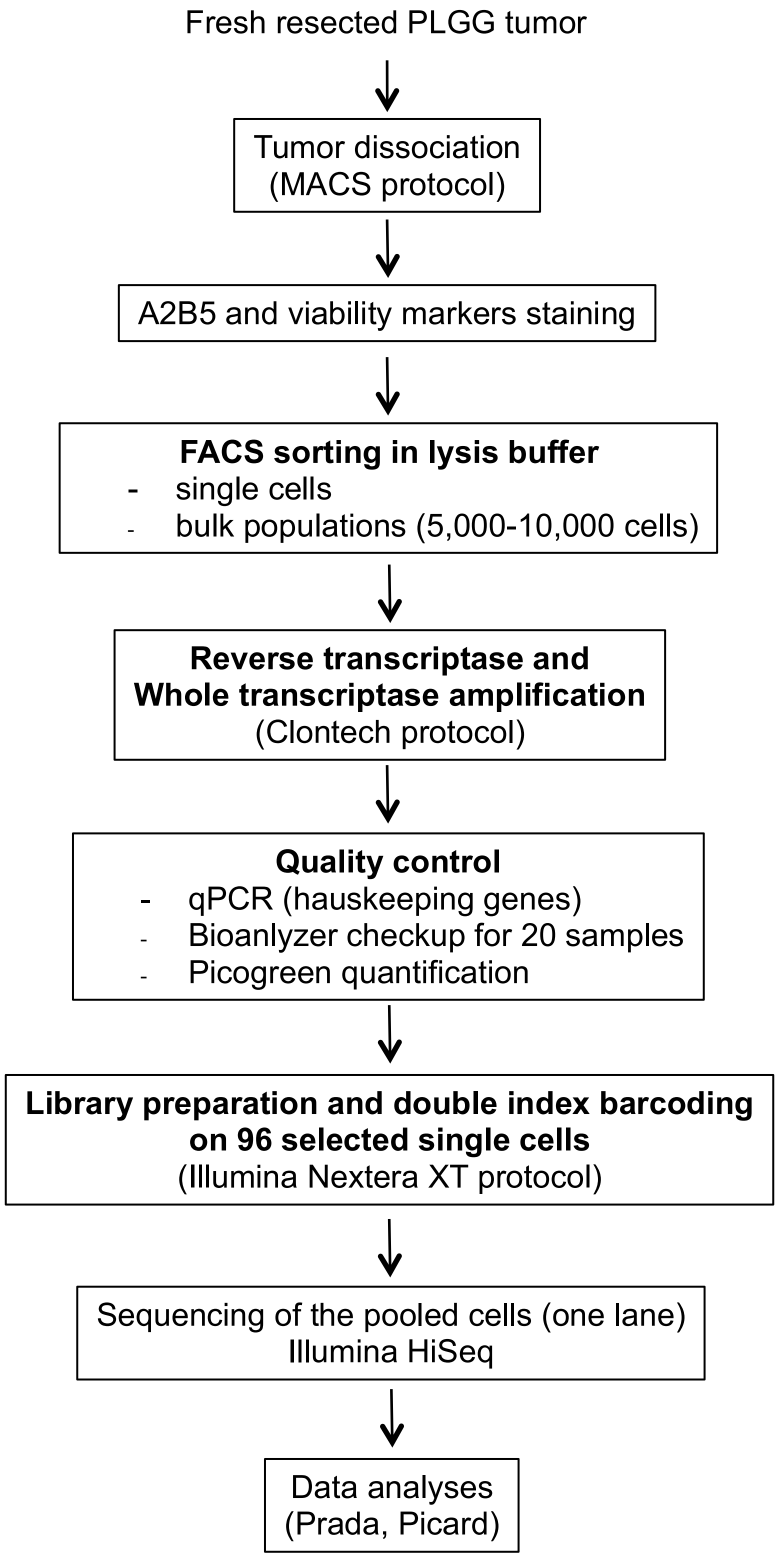
Fresh PLGG tumor was collected from the operating room with a maximum delay of 4 hours after surgical resection. Tumor dissociation was then performed using the Miltelyi Brain Tumor Dissociation Kit protocol (MACS, Miltenyi Biotec). After two filtering-washing steps (using a 70 µm filter), we prepared the cell suspension at a concentration of 1x10^6 cells/ml. We then stained the fresh dissociated cells using various antibodies (Abs) reflecting different developmental stages in the glioneuronal differentiation process. Anti-A2B5 Ab (staining glial precursor cells) conjugated with APC, anti-GLAST Ab (also know as ACSA-1 for astrocyte cell surface antigene 1, expressed predominantly expressed at the surface of differentiated astrocytes) conjugated with PE, anti-O4 Ab (expressed predominantly at the surface of differentiated oligodendrocytes) conjugated with biotin and anti-PSA-NCAM Ab (linked to the extra-cellular domain of neural adhesion molecule, staining preferentially neuronal progenitor as well as more differentiated neuronal tissues) were used at 1/10 concentration (MACS, Miltenyi Biotec). Anti-FGFR1 (fibroblast growth factor receptor 1) antibody (D8E4) was used at 1:250 concentration (Cell Signaling Technology).

For the PLGG tumors that we evaluated without further single cell sorting, we used the LSRFortressa Analyzer (BD Biosciences, Dana-Farber Cancer Institute Flow Cytomtry Core, Boston, USA).

Single cells were sorted using an Astrios FACS sorter (Beckman Coulter, Bauer Flow Cytometry, Harvard University, Cambridge, USA) in 96-cell plates containing 5µl of TCL buffer with 1% 2-mercaptoethanol. To improve the yield of the viable cells we used two cells death-viability markers; we gated the cells based on the positive viability marker Calcein AM (Life Technologies) and the absence of the cell death marker Near-IR (Life Technologies). After a short spin (800g, 1 min), plates were immediately freezed at - 80°C. Additional samples containing 200-5000 cells (depending on the yield) were sorted in 1,5ml lysis buffer as bulk controls.

Figure 29:

Summary of the workflow from the tumor to single cell RNA-sequencing



2.2. Reverse transcriptase and whole transcriptome amplification

We first performed a lysate purification using beads technology (Agencourt RNAClean XP SPRI beads). We added 2,2 volumes of RNA-SPRI beads to each well of lysate. After 10 min incubation we placed the 96-well plate on an adapted magnet (Life Technologies DynaMagTM-96 side skirted magnet) and incubated for 5 min. After three steps of washing with ethanol 80% we allowed beads to dry at room temperature for approximately 10 minutes and eluted the pellet with the 3' SMART CDS Primer II A (according to the protocol, SMARTer ultra Low Input RNA Kit, Clontech®). We then followed the reverse transcriptase and the whole transcriptome amplification (WTA) protocols (Advantage 2 Polymerase Kit) according to the manufacture recommendations. We then purified the amplified cDNA using the DNA SPRI beads (Agencourt AMPure XP). We added 0.8 volumes of the DNA SPRI beads buffer to the WTA solution. After three steps of washing with ethanol 80% we allowed beads to dry at room temperature for approximately 10 minutes and eluted the pellet with 20µl of TE buffer and transferred the eluent in a new 96-well plate.

For the bulk controls, we purified the RNA using the RNEasy Plus Micro Kit (Qiagen) according to the manufacturers recommendations. We then substitute the 1 μ L purified RNA in place of H2O at the first step of the reverse transcriptase buffer preparationl. We then followed the same procedure as for the single cells.

2.3. Quality control steps

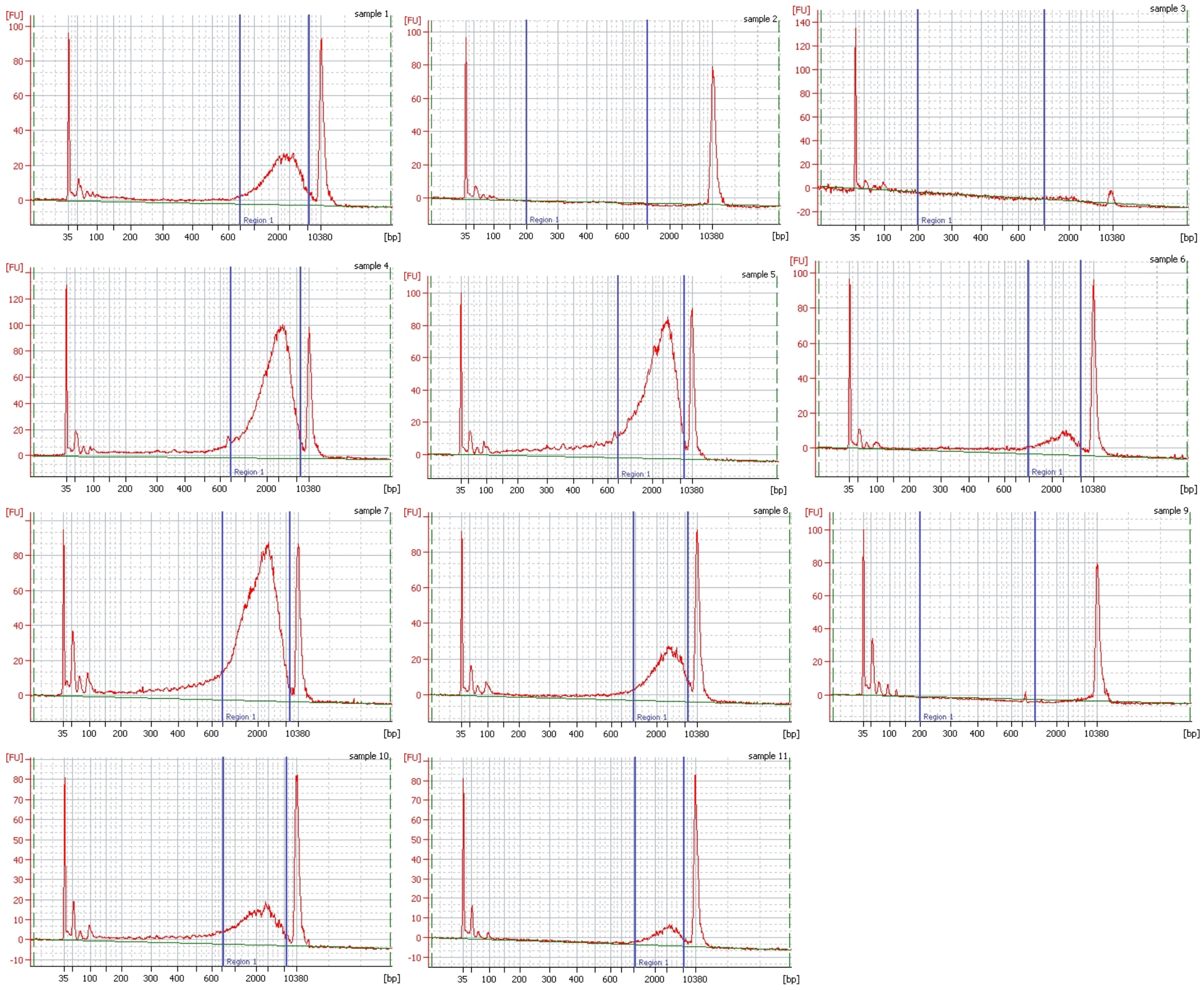
We used three independent methods to evaluate the size and the quality of the purified cDNA from the single cells as well as the bulk controls. We used 1µL of the purified WTA product to measure the fragment size distribution with the Agilent HS DNA BioAnalyzer. Figure 30 shows 11 profiles obtained on SC after WTA, describing good and bad profiles reflecting various yields obtained after the WTA protocol.

To estimate the library concentration, we used the Quant-iT™ PicoGreen® dsDNA Assay Kit (as per manufacturer’s recommendations, Life Technologies) using 1µL of the purified WTA product.

We finally estimated the single cell and the bulk control viability with the Taqman qPCR assay, using well described housekeeping genes (*B2M*, *ACTB* or *RPLP0*). We then compared results from the three tests in order to select the best single cell candidates for the library construction.

Figure 30:

Bioanalyzer results of 11 samples selected among the 96 SC A2B5 positive from tumor PLGG 22



2.4. Library construction procedure

We performed library preparation of 48 A2B5 positive cells and 48 A2B5 negative cells (on a same 96-well plate) using a double indexing strategy according to the Nextera XT sequencing-library construction protocol (Illumina). This protocol performs a first step sharing the double strand DNA and followed by attaching a 8bp index AA sequence at the 3’ and 5’ end of the small strand DNA before performing PCR amplification. We used 1.2 μ L of the diluted PCR product from SMART RT-PCR for the library construction (as per manufacturer’s recommandations).

We then pooled 2.5 μ L from each well containing the product of the amplified library into a single 1.5mL tube and purified a last time the cDNA using the DNA SPRI beads (Agencourt AMPure XP). After three washing steps with ethanol 80% we allowed the beads to dry at room temperature for approximately 10 minutes and eluted the DNA with 30μl of TE buffer and repeated the DNA SPRI beads cleanup a last time. We finally measured the average sized of the library obtained using the Agilent HS DNA BioAnalyzer.

2.5. Sequencing and data analyses

We sequenced the libraries on a Illumina sequencer (HiSeq 2500, 2 x 25 cycles) using one lane for the 96 pooled single cells and another line for the pooled bulk controls of the corresponding tumor. After demultiplexing the FASTQ files, we matched each of the paired-end reads with its mate and aligned the reads to reference genome hg19 using TopHat[349]. Gene-expression profiles were then assessed using a built Picard tool inside the Broad Institute Firehose environment. Z-scores were generated for each gene across samples. Self Organizing Maps[326], hierarchical clustering[202], Comparative Marker Selection Analysis (CMS)[327,328], Support Vector Machine (SVM) algorithm and Gene-set Enrichment Analysis (GSEA)[206,329] were performed using GenePattern[330]

(Broad Institute, http://genepattern.broadinstitute.org). Principal Component Analysis (PCA) plots were generated in Matlab. For CMS analyses, genes that described p<0.01, FDR<0.1 and fold-change > 500 were considered significant. For GSEA analysis, gene-sets that showed p<0.01 and FDR<0.1 were only considered significant. For those two later analyses, we used non z-scored matrix.

3. Results

3.1 Profiling of glioneuronal markers in PLGGs

We first analyzed by flow cytometry the fraction of PLGG cells expressing various glioneuronal markers. We profiled in total 26 PLGG tumors from different ages, locations, and histological subtypes as well as various *BRAF* genomic statuses (Table 12 and Figure 31) and tested different glioneuronal markers. We used glioneuroanl Ab available on the market and expressed at the membrane of cells in order to allow further potential nucleic acid extraction. Most of the cells tested were obtained from freshly dissociated PLGG tumors (20/26 tumors, 77%). Six samples (23%) were from frozen cells and one case was an autopsy.

We observed important variability in terms of the fraction of positive glioneuronal staining across the PLGG tumors. We were able to assess the expression of the early glial progenitor marker A2B5 in 21/26 tumors (81% of tumors) and found a wide range of A2B5 expression (mean 27%, range 1-88%). Among the 17 tumors analyzed for the astrocyte marker GLAST (63%), we observed similar important variability in the fraction of positive PLGG cells (mean 28%, range 0-90%). O4 was tested in 10 tumors (37%)

and was expressed between 0.5-30% of the cells (mean 7%). FGFR1 was expressed at a lower amplitude, 0.5-30% of the PLGG cells (mean 7%). The neuronal marker N-CAM was globally poorly expressed in the 4 tumors analyzed (mean 1% of positive cells).

For the single-cell sorting and sequencing experiment we decided to focus our analysis using the anti-A2B5 Ab, as we have optimized the use this Ab with a positive control (cultured mouse neural stem cells, staining the A2B5 Ab between 10-20%, Figure 32A-C) and a negative control (cultured 3T3-cells, Figure 32D and E).

Table 12:

Summary of the glioneuronal profling in the 26 PLGG tumors analyzed by flow cytometry The three tumors highlighted in orange are the ones that have been processed for single cell sorting and RNA-sequencencing

PLGG ID	Age at surgery (years)	Histo	Location	BRAF status	Fresh/ frozen	A2B5 +	N-CAM	GLAST	O4	FGFR1
PLGG 1	5	GG	ST	ND	fresh	76	3	73	3	ND
PLGG 2	0.6	PMA	ST	V600E neg	fresh	1	2	1.5	0,5	ND
PLGG 3	1.4	PA	IT	BRAF DUP	fresh	88	0	48	3	ND
PLGG 3	1.4	PA	IT	BRAF DUP	frozen	50	0	35	5	ND
PLGG 4	1.9	PA	IT	no BRAF alteration	frozen	15	ND	83	13	75
PLGG 5	2.5	PMA	ST	V600E neg	frozen	ND	ND	76	4	ND
PLGG 6	1	PMA	ST	BRAF DUP	fresh	ND	ND	3	6	ND
PLGG 7	2.5	LGG NOS	IT	V600E neg	frozen	30	ND	2	30	60
PLGG 8	5.5	LGG NOS	IT	V600E neg	fresh	ND	ND	1	1,5	ND
PLGG 9	13	DNT	ST	ND	fresh	9	ND	0.5	2	21
PLGG 10	6	PA recurrent	ST	ND	frozen	ND	ND	1.5	2	ND
PLGG 11	13.1	DNT	ST	ND	fresh	2	ND	ND	ND	10
PLGG 12	10	DNT	ST	ND	fresh	63	ND	ND	ND	ND
PLGG 13	5	PA	IT	BRAF DUP	fresh	11	ND	22	ND	ND
PLGG 14	3.4	PA	IT	BRAF DUP	fresh	8	ND	90	ND	80
PLGG 15	10.8	DA	ST	IDH1+	fresh	0	ND	2	ND	1,8
PLGG 16	16	DNT	ST	ND	fresh	0	ND	2	ND	0
PLGG 17	14.5	GG	ST	no BRAF alteration	frozen	18.5	ND	7.5	ND	0
PLGG 18	5	PA	IT	BRAF DUP	fresh	13	ND	ND	ND	ND
PLGG 19	6.4	LGG NOS	ST	no BRAF alteration	fresh	6.5	ND	ND	ND	ND
PLGG 20	10.3	DNT	ST	no BRAF alteration	fresh	6.5	ND	ND	ND	ND
PLGG 21	13.1	PA recurrent	IT	BRAF DUP	fresh	35	ND	ND	ND	ND
PLGG 22	16	GG recurrent	ST	V600E mut	fresh	6	ND	0	ND	ND
PLGG 23	6.1	PA	IT	ND	fresh	47	ND	ND	ND	ND
PLGG 24	4	LGG NOS	ST	ND	autopsia	ND	ND	ND	ND	ND
PLGG 25	7	LGG NOS recurrent	IT	BRAF DUP	fresh	40	ND	ND	ND	ND
PLGG 26		PLGG			fresh	65	ND	ND	ND	ND

Figure 31:

Graphical representation of flow cytometry results from 26 tumors

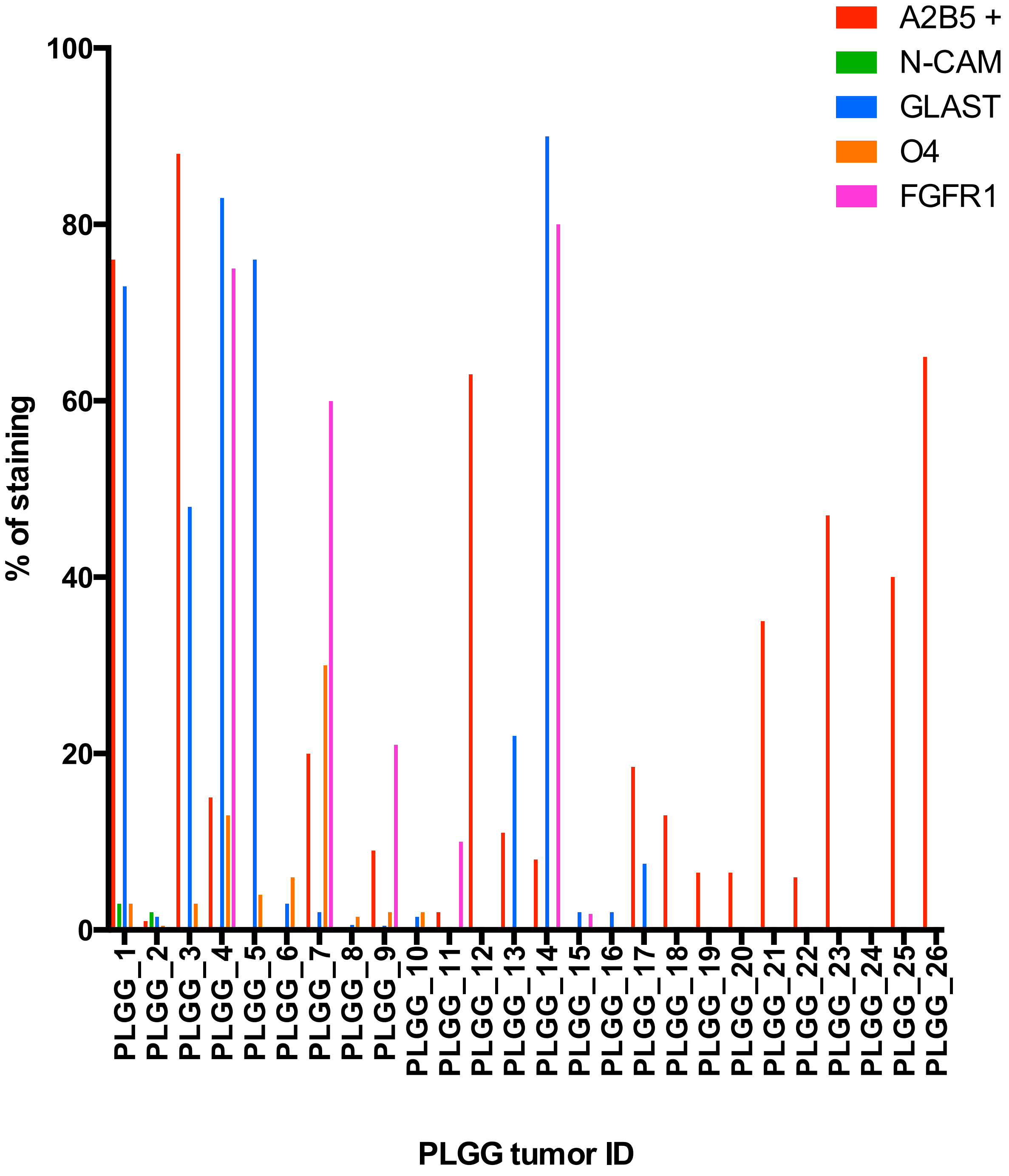
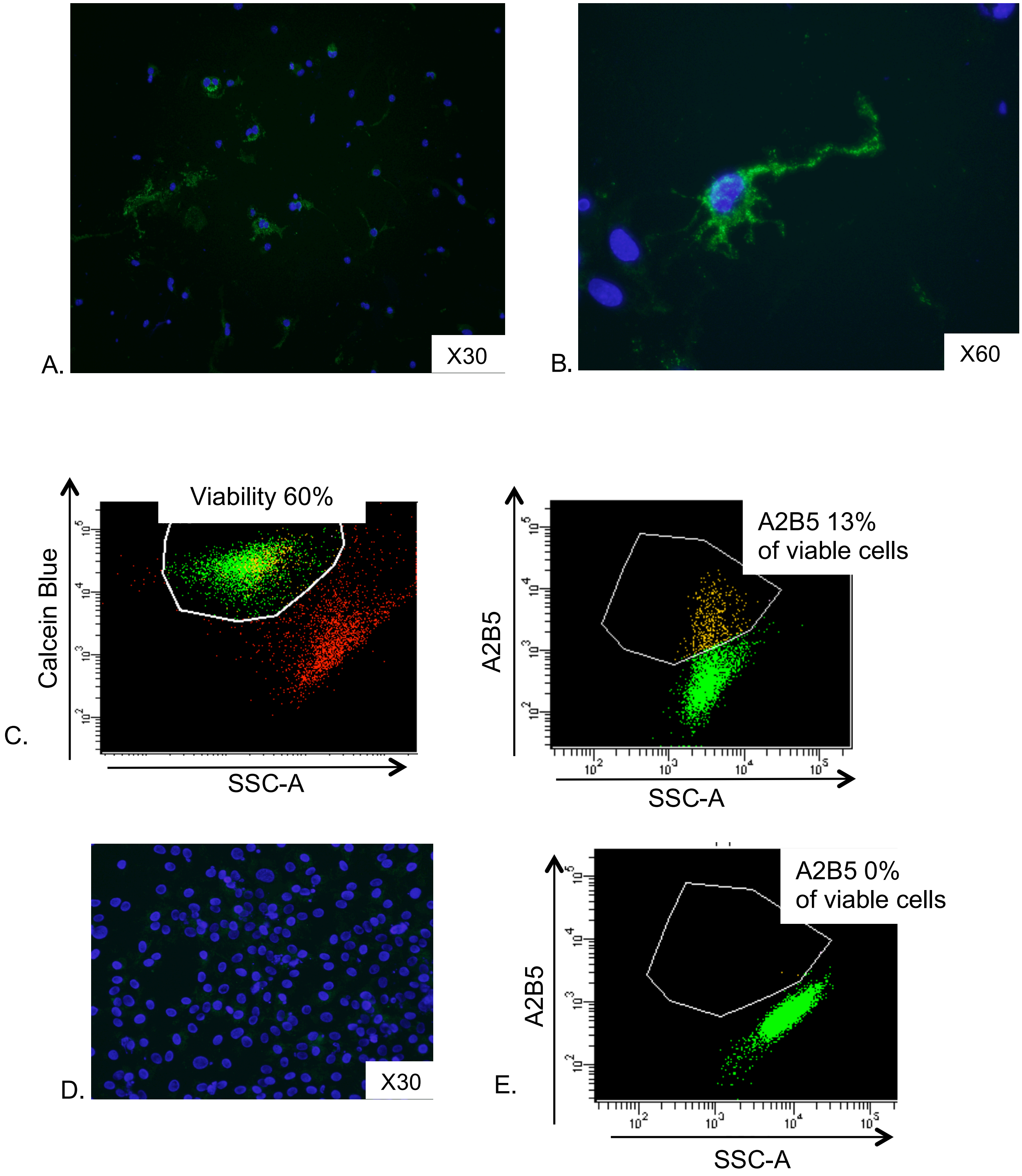


Figure 32:
Immunofluorescence and flow cytometry profiles of the positive (mouse neural stem cells, A, B and C) and negative controls (NIH 3T3 cells, D and E) for the A2B5 antibody



3.2. Single-cell profiling results

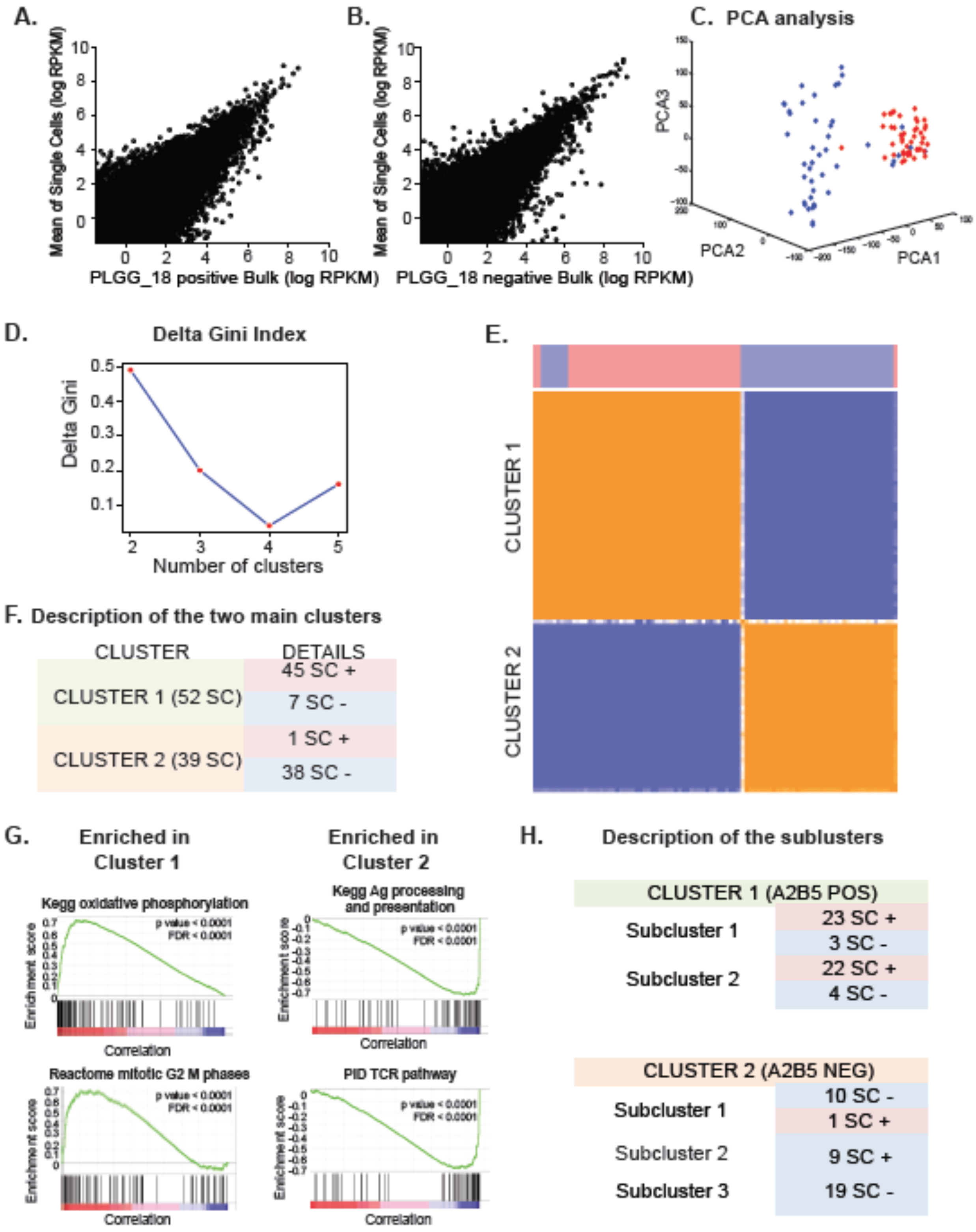
Among the 26 PLGGs that we analyzed by flow cytometry, we sorted 6 tumors at a single cell level by collecting separately A2B5 positive and negative cells. Among those 6 PLGGs, we processed the single cells RNA until the library preparation step for sequencing in 3 tumors. We then sequenced the RNA of these cells in order to analyze their genome-wide transcriptome.

The first tumor analyzed (PLGG 18) had 46 A2B5 positive and 46 A2B5 negative single cells (SC) that met these criteria. We observed for both A2B5 positive and negative SC a good correlation between the mean gene expression of the SC to the gene expression in the corresponding bulk control (Figure 33A and 33B). PCA revealed that the A2B5 positive SC clustered separately from the A2B5 negative cells (Fig 33C). SOM consensus clustering showed that two clusters best segregated the 91 SC (Figure 33D). Strikingly, the clusters obtained by SOM were significantly strongly enriched with A2B5 positive and negative cells, respectively (Figure 33F and 33E, $p<0.0001$).

We then performed comparative marker selection (CMS) and GSEA analyses comparing both clusters in order to identify the molecular differences between the A2B5 positive and negative cells. We found 709 genes significantly overexpressed in the cluster 1 dominated by the A2B5 positive cells and 312 genes significantly overexpressed in cluster 2 enriched with the A2B5 negative cells. Among the top 100 genes that describe the highest fold-change in cluster 1 relative to cluster 2, we found genes involved in intracellular metabolic regulation such as *GNG3* (guanine nucleotide binding protein 3), *STK32A* (serine-threonine kinase 32A), *LDLRAD3* (low-density lipoprotein receptor class A domain containing 3) and *RRP9* (ribosomal RNA processing 9) and genes coding proteins involved in cytoskeletal organization such as *MARCKSL1* (myristoylated alanie-rich C-kinase substrate-like 1), *ECM1* (extracellular matrix 1) and *KIF1A* (kinesin family member 1) (Table 13A). Among the top 10% genes that describe the highest fold-change in cluster 2, dominated by the A2B5 relative to cluster 1, we found genes essentially genes involved in inflammatory mechanisms, such as *CCRL2* (chemokine receptor-like 2), essential for the recruitment of immune cells to the site of inflammation, *TLR-7* (toll-like receptor 7) mediating the production of cytokines and *HPGDS* (hematopoietic prostaglandin D synthase) playing a role in the production of prostanoids in the immune system and mast cells (Table 13B).

The same molecular patterns were found in GSEA between cluster 1 and 2. Among the 151 gene-sets significantly enriched in the A2B5 positive cells (cluster 1), we found that 58/151 (38%) of the gene-sets were represented by metabolic process and mitosis related pathways (such as KEGG OXIDATIVE PHOSPHORYLATION, REACTOME MITOTIC G2 M PHASES) and 21/151 (14%) were represented by cell mobility and intracellular communications (such as KEGG GAP JUNCTION, REACTOME COLLAGEN FORMATION, KEGG ECM RECEPTOR INTERACTION or REACTOME CELL CELL JUNCTION ORGANIZATION, Table 14A). 68/90 (75%) of the gene-sets enriched in cluster 2 were related to inflammatory pathways (Table 14B and Figure 33G). We then wanted to explore the degree of heterogeneity within each subcluster. SOM consensus clustering performed on the cluster 1 SC revealed two major groups whereas cluster 2 segregated in 3 groups (Figure 33H), suggesting that the heterogeneity within the A2B5 negative cells is significantly higher than the A2B5 positive SC.

Figure 33: Results of PLGG 18 single-cell RNA-sequencing experiment



F. Description of the two main clusters

CLUSTER	DETAILS
CLUSTER 1 (52 SC)	45 SC +
	7 SC -
CLUSTER 2 (39 SC)	1 SC +
	38 SC -

H. Description of the sublusters

CLUSTER 1 (A2B5 POS)	
Subcluster 1	23 SC +
	3 SC -
Subcluster 2	22 SC +
	4 SC -

CLUSTER 2 (A2B5 NEG)	
Subcluster 1	10 SC -
	1 SC +
Subcluster 2	9 SC +
Subcluster 3	19 SC -

Figure 33 (legend)

A-B, graphical representation of the correlation between the average gene expression of SC relative to the bulk gene expression for the positive A2B5 cells (A) and negative cells (B). C, PCA analysis showing the three first components of the SC A2B5 positive (red) and negative (blue). D, graphical representation of the delta-gini index variation obtained by SOM clustering, E, heat-map of the SOM clustering (red represents A2B5 SC positive cells, blue the A2B5 negative cells). F, summary table of the distribution of the different SC by SOM clustering. G, GSEA enrichment plots significantly enriched in cluster 1 and 2. H, summary table of the distribution of the different sublclusters of SC by SOM clustering

Table 13: Genes significantly overexpressed in cluster 1 (A) and 2 (B)

A.

Genes upregulated in Cluster 1	FDR	Fold change
GNG3	5,51E-04	28225
C1orf198	5,51E-04	15586
STK32A	5,51E-04	11705
RP11-342M3.5	5,51E-04	8476
MARCKSL1	5,51E-04	8467
LDLRAD3	5,51E-04	8370
NOVA1	5,51E-04	7727
ECM1	5,51E-04	7241
RRP9	5,51E-04	5605
KIF1A	5,51E-04	5434
BHLHB9	5,51E-04	5291
NSDHL	5,51E-04	5280
CDR1	5,51E-04	5188
KIF5C	5,51E-04	5093
ACTL6A	5,51E-04	4924
RP3-398G3.3	5,51E-04	4924
RP11-161M6.2	5,51E-04	4632
IFIT3	5,51E-04	4491
CHAD	5,51E-04	4450
ALDH1A3	5,51E-04	4330
RP11-547I7.2	5,51E-04	4099
USP30	5,51E-04	4009
GRM5-AS1	5,51E-04	3946
IFI27L2	5,51E-04	3924
RP11-480D4.3	5,51E-04	3800
EPN2	5,51E-04	3782
RP5-1177M21.1	5,51E-04	3750
SPATA6L	5,51E-04	3562
GRIK1	5,51E-04	3548
C8orf46	5,51E-04	3464
CCDC28B	5,51E-04	3408
IGSF9B	5,51E-04	3368
JAM3	5,51E-04	3363
MORN2	5,51E-04	3343
RP3-412A9.11	5,51E-04	3325
PABPC5	5,51E-04	3318
TSPAN12	5,51E-04	3296
FLRT3	5,51E-04	3251
FSTL1	5,51E-04	3250
TMEM231	5,51E-04	3235
RP11-548C21.1	5,51E-04	3143
DNALI1	5,51E-04	3116
TMEM132C	5,51E-04	3112

SWI5	5,51E-04	3075
CPE	5,51E-04	3041
BBS10	5,51E-04	3038
AP3M2	5,51E-04	2959
RP11-508N12.4	5,51E-04	2860
LMCD1	5,51E-04	2835
BNIP3P1	5,51E-04	2771
GLIPR1L2	5,51E-04	2716
DIRAS3	5,51E-04	2698
BAMBI	5,51E-04	2688
CDO1	5,51E-04	2685
BTBD8	5,51E-04	2677
RP4-788L13.1	5,51E-04	2675
KLHL32	5,51E-04	2663
AMPH	5,51E-04	2652
ARC	5,51E-04	2652
C12orf29	5,51E-04	2630
PNMAL1	5,51E-04	2610
SORCS3	5,51E-04	2595
AEBP1	5,51E-04	2589
BBS7	5,51E-04	2571
TAGLN3	5,51E-04	2568
TRMT61B	5,51E-04	2533
THBS2	5,51E-04	2519
ZCCHC24	5,51E-04	2505
NRSN2	5,51E-04	2487
CTD-2033A16.3	5,51E-04	2450
ZNF521	5,51E-04	2448
POP7	5,51E-04	2436
C14orf132	5,51E-04	2418
PRPF31	5,51E-04	2399
TIAM2	5,51E-04	2356
Y RNA	5,51E-04	2337
POMGNT1	5,51E-04	2305
sept-08	5,51E-04	2288
GPD1	5,51E-04	2288
APBA2	5,51E-04	2255
MTSS1L	5,51E-04	2208
AC015936.3	5,51E-04	2206
ASAP3	5,51E-04	2188
KLHL23	5,51E-04	2173
CHST3	5,51E-04	2157
NKAIN4	5,51E-04	2156
TXNL4A	5,51E-04	2117
CDCA7L	5,51E-04	2112
SLC24A2	5,51E-04	2111
AKAP12	5,51E-04	2103
TECPR2	5,51E-04	2097
MASP1	5,51E-04	2090
MLLT11	5,51E-04	2084

ACAN	5,51E-04	2050
SHISA4	5,51E-04	2048
DHCR24	5,51E-04	2045
snoU13	5,51E-04	2040
PRRX1	5,51E-04	1979
ATCAY	5,51E-04	1977
NAP1L3	5,51E-04	1964

B.

Genes upregulated in Cluster 2	FDR	Fold change
CCRL2	5,51E-04	30825
ACSL5	5,51E-04	13643
LGALS9	5,51E-04	12737
FOLR2	5,51E-04	10297
TLR7	5,51E-04	9619
HPGDS	5,51E-04	8571
VTRNA1-3	5,51E-04	6975
C1orf162	5,51E-04	6496
HAMP	5,51E-04	6054
U6	5,51E-04	5802
LILRA1	5,51E-04	5791
TNFAIP3	5,51E-04	5418
TRBJ2-1	5,51E-04	5225
AP000908.1	5,51E-04	5217
RP11-1094M14.3	5,51E-04	5063
AC005232.1	5,51E-04	4837
MIR335	5,51E-04	4813
AC090282.1	5,51E-04	4645
FCGBP	5,51E-04	4478
TRAJ42	5,51E-04	4478
RP11-489O18.1	5,51E-04	4168
ABI3	5,51E-04	4084
PIK3R5	5,51E-04	4058
CRYBB1	5,51E-04	4056
TNF	5,51E-04	3842
AC006953.1	5,51E-04	3829
TRAV1-2	5,51E-04	3787
RP11-300J18.2	5,51E-04	3711
EMR1	5,51E-04	3397
CEACAM21	5,51E-04	3395
TRAV12-2	5,51E-04	3349
CD300C	5,51E-04	3315
CTD-2616J11.3	5,51E-04	3253
snoU13	5,51E-04	2984
S100A4	5,51E-04	2964
PLCG2	5,51E-04	2860
TRAV36DV7	5,51E-04	2853
SAMD9L	5,51E-04	2840
AC069363.1	5,51E-04	2762
TBXAS1	5,51E-04	2738
SNORD119	5,51E-04	2735
RNASE2	5,51E-04	2683
CALHM2	5,51E-04	2676
LILRA2	5,51E-04	2647
CD48	5,51E-04	2645

RP11-493L12.2	5,51E-04	2642
TRBJ2-4	5,51E-04	2555
PTGER4	5,51E-04	2462
RP11-750H9.5	5,51E-04	2450
VSIG4	5,51E-04	2407
U4	5,51E-04	2393
ITPRIP	5,51E-04	2377
NFKBID	5,51E-04	2363
GPR65	5,51E-04	2345
TNFSF13	5,51E-04	2341
CD163L1	5,51E-04	2332
RN5S226	5,51E-04	2331
AREG	5,51E-04	2326
NCKAP1L	5,51E-04	2315
RP4-612C19.2	5,51E-04	2295
MS4A4A	5,51E-04	2269
CMKLR1	5,51E-04	2267
CTD-2616J11.2	5,51E-04	2233
IL4R	5,51E-04	2210
LILRB4	5,51E-04	2195
RN5S498	5,51E-04	2175
TRBV20-1	5,51E-04	2165
TRAJ16	5,51E-04	2070
TRAJ26	5,51E-04	1988
RILPL2	5,51E-04	1957
AL118508.1	5,51E-04	1946
NAPSB	5,51E-04	1885
CSF2RA	5,51E-04	1816
PTAFR	5,51E-04	1776
CXorf21	5,51E-04	1764
DLEU7-AS1	5,51E-04	1763
ITGAM	5,51E-04	1740
TNFSF18	5,51E-04	1735
FMNL1	5,51E-04	1734
ADORA3	5,51E-04	1725
GS1-526D21.5	5,51E-04	1712
CD300LB	5,51E-04	1706
IRAK3	5,51E-04	1705
FILIP1L	5,51E-04	1688
SIGLEC10	5,51E-04	1675
TRAJ9	5,51E-04	1648
RNASE3	5,51E-04	1620
F11R	5,51E-04	1612
CD84	5,51E-04	1578
STAT6	5,51E-04	1557
MRC1L1	5,51E-04	1545
NCF2	5,51E-04	1539
SH3TC1	5,51E-04	1483
TLR4	5,51E-04	1470
RP11-394B5.2	5,51E-04	1429
RP5-1073O3.7	5,51E-04	1389
AC010518.3	5,51E-04	1386
C1orf38	5,51E-04	1364
SNORD58C	5,51E-04	1356
F13A1	5,51E-04	1345

Table 14: Gene-sets enriched in cluster 1 (A) and 2 (B) in PLGG 18.

A.

NAME	FDR
KEGG OXIDATIVE PHOSPHORYLATION	0,00
REACTOME GLUCOSE METABOLISM	0,00
KEGG PARKINSONS DISEASE	0,00
KEGG HUNTINGTONS DISEASE	0,00
KEGG ALZHEIMERS DISEASE	0,00
REACTOME MITOCHONDRIAL PROTEIN IMPORT	0,00
REACTOME RESPIRATORY ELECTRON TRANSPORT	0,00
REACTOME TCA CYCLE AND RESPIRATORY ELECTRON TRANSPORT	0,00
REACTOME RESPIRATORY ELECTRON TRANSPORT ATP SYNTHESIS BY CHEMIOSMOTIC COUPLING AND HEAT PRODUCTION BY UNCOUPLING PROTEINS	0,00
REACTOME ABORTIVE ELONGATION OF HIV1 TRANSCRIPT IN THE ABSENCE OF TAT	0,00
KEGG PROPANOATE METABOLISM	0,00
KEGG CITRATE CYCLE TCA CYCLE	0,00
REACTOME TRANSMISSION ACROSS CHEMICAL SYNAPSES	0,00
REACTOME UNBLOCKING OF NMDA RECEPTOR GLUTAMATE BINDING AND ACTIVATION	0,00
REACTOME NEUROTRANSMITTER RECEPTOR BINDING AND DOWNSTREAM TRANSMISSION IN THE POSTSYNAPTIC CELL	0,00
KEGG VALINE LEUCINE AND ISOLEUCINE DEGRADATION	0,00
REACTOME ACTIVATION OF KAINATE RECEPTORS UPON GLUTAMATE BINDING	0,00
REACTOME ELONGATION ARREST AND RECOVERY	0,00
REACTOME TRAFFICKING OF AMPA RECEPTORS	0,00
KEGG PYRUVATE METABOLISM	0,00
KEGG BUTANOATE METABOLISM	0,00
REACTOME MICRORNA MIRNA BIOGENESIS	0,00
KEGG RNA POLYMERASE	0,00
REACTOME REGULATORY RNA PATHWAYS	0,00
REACTOME PEROXISOMAL LIPID METABOLISM	0,00
REACTOME GLYCOLYSIS	0,00
REACTOME RECRUITMENT OF MITOTIC CENTROSOME PROTEINS AND COMPLEXES	0,00
REACTOME NEUROTRANSMITTER RELEASE CYCLE	0,00
KEGG GLYCINE SERINE AND THREONINE METABOLISM	0,00
REACTOME GABA SYNTHESIS RELEASE REUPTAKE AND DEGRADATION	0,00
REACTOME LOSS OF NLP FROM MITOTIC CENTROSOMES	0,00
REACTOME NCAM1 INTERACTIONS	0,00
PID NCADHERINPATHWAY	0,00
PID P75NTRPATHWAY	0,00
REACTOME ADHERENS JUNCTIONS INTERACTIONS	0,00
KEGG GLYCOLYSIS GLUCONEOGENESIS	0,00
REACTOME GLYCOGEN BREAKDOWN GLYCOGENOLYSIS	0,00
REACTOME MRNA SPLICING MINOR PATHWAY	0,00
REACTOME CREB PHOSPHORYLATION THROUGH THE ACTIVATION OF CAMKII	0,00
KEGG BIOSYNTHESIS OF UNSATURATED FATTY ACIDS	0,01
REACTOME BRANCHED CHAIN AMINO ACID CATABOLISM	0,01
REACTOME A TETRASACCHARIDE LINKER SEQUENCE IS REQUIRED FOR GAG SYNTHESIS	0,01
REACTOME FORMATION OF RNA POL II ELONGATION COMPLEX	0,01
REACTOME GLUCONEOGENESIS	0,01
ST DIFFERENTIATION PATHWAY IN PC12 CELLS	0,01
REACTOME TRAFFICKING OF GLUR2 CONTAINING AMPA RECEPTORS	0,01
REACTOME NEURONAL SYSTEM	0,01
KEGG UBIQUITIN MEDIATED PROTEOLYSIS	0,01
REACTOME FORMATION OF THE HIV1 EARLY ELONGATION COMPLEX	0,01
REACTOME INHIBITION OF INSULIN SECRETION BY ADRENALINE NORADRENALINE	0,01
KEGG CARDIAC MUSCLE CONTRACTION	0,01
REACTOME REGULATION OF INSULIN SECRETION	0,01

KEGG PURINE METABOLISM	0,01
REACTOME METABOLISM OF AMINO ACIDS AND DERIVATIVES	0,01
KEGG PENTOSE PHOSPHATE PATHWAY	0,01
REACTOME NETRIN1 SIGNALING	0,01
KEGG ALANINE ASPARTATE AND GLUTAMATE METABOLISM	0,01
REACTOME TRNA AMINOACYLATION	0,01
REACTOME METABOLISM OF CARBOHYDRATES	0,01
PID NETRIN PATHWAY	0,01
REACTOME PYRUVATE METABOLISM AND CITRIC ACID TCA CYCLE	0,01
REACTOME MITOTIC G2 G2 M PHASES	0,01
REACTOME FORMATION OF TUBULIN FOLDING INTERMEDIATES BY CCT TRIC	0,01
PID P38 MKK3 6PATHWAY	0,01
BIOCARTA NOS1 PATHWAY	0,01
REACTOME ACTIVATION OF NMDA RECEPTOR UPON GLUTAMATE BINDING AND POSTSYNAPTIC EVENTS	0,01
REACTOME INSULIN SYNTHESIS AND PROCESSING	0,01
KEGG FRUCTOSE AND MANNOSE METABOLISM	0,01
PID NOTCH PATHWAY	0,01
PID BMPPATHWAY	0,01
REACTOME INTEGRATION OF ENERGY METABOLISM	0,01
KEGG AMINOACYL TRNA BIOSYNTHESIS	0,01
PID SYNDECAN 3 PATHWAY	0,01
REACTOME CITRIC ACID CYCLE TCA CYCLE	0,02
REACTOME INTERACTION BETWEEN L1 AND ANKYRINS	0,02
REACTOME CHONDROITIN SULFATE DERMATAN SULFATE METABOLISM	0,02
ST WNT CA2 CYCLIC GMP PATHWAY	0,02
ST JNK MAPK PATHWAY	0,02
REACTOME NCAM SIGNALING FOR NEURITE OUT GROWTH	0,02
REACTOME EXTRACELLULAR MATRIX ORGANIZATION	0,02
REACTOME ANTIGEN PROCESSING UBIQUITINATION PROTEASOME DEGRADATION	0,02
REACTOME PREFOLDIN MEDIATED TRANSFER OF SUBSTRATE TO CCT TRIC	0,02
KEGG PEROXISOME	0,02
KEGG TRYPTOPHAN METABOLISM	0,02
REACTOME PROTEIN FOLDING	0,02
REACTOME PYRUVATE METABOLISM	0,02
REACTOME GENERIC TRANSCRIPTION PATHWAY	0,02
ST GRANULE CELL SURVIVAL PATHWAY	0,02
REACTOME CELL CELL JUNCTION ORGANIZATION	0,02
REACTOME CHONDROITIN SULFATE BIOSYNTHESIS	0,02
KEGG ECM RECEPTOR INTERACTION	0,03
REACTOME INWARDLY RECTIFYING K CHANNELS	0,03
PID BETACATENIN NUC PATHWAY	0,03
REACTOME CHOLESTEROL BIOSYNTHESIS	0,03
PID A6B1 A6B4 INTEGRIN PATHWAY	0,03
REACTOME ION CHANNEL TRANSPORT	0,03
REACTOME RAS ACTIVATION UOPN CA2 INFUX THROUGH NMDA RECEPTOR	0,03
REACTOME SMOOTH MUSCLE CONTRACTION	0,03
REACTOME HEPARAN SULFATE HEPARIN HS GAG METABOLISM	0,03
REACTOME GLYCOSAMINOGLYCAN METABOLISM	0,03
REACTOME COLLAGEN FORMATION	0,03
KEGG PROTEIN EXPORT	0,03
REACTOME RNA POL III TRANSCRIPTION	0,03
PID AR PATHWAY	0,03
KEGG PYRIMIDINE METABOLISM	0,03
KEGG LONG TERM POTENTIATION	0,03
REACTOME RNA POL III TRANSCRIPTION TERMINATION	0,03
REACTOME METABOLISM OF LIPIDS AND LIPOPROTEINS	0,03
REACTOME FATTY ACID TRIACYLGLYCEROL AND KETONE BODY METABOLISM	0,03
REACTOME BILE ACID AND BILE SALT METABOLISM	0,03

REACTOME G BETA GAMMA SIGNALLING THROUGH PLC BETA	0,04
REACTOME GABA RECEPTOR ACTIVATION	0,04
REACTOME MUSCLE CONTRACTION	0,04
REACTOME BASE EXCISION REPAIR	0,04
REACTOME AQUAPORIN MEDIATED TRANSPORT	0,04
PID PS1PATHWAY	0,04
BIOCARTA GPCR PATHWAY	0,04
KEGG GAP JUNCTION	0,04
REACTOME G PROTEIN ACTIVATION	0,05
KEGG MELANOMA	0,05
REACTOME REGULATION OF WATER BALANCE BY RENAL AQUAPORINS	0,05
REACTOME OPIOID SIGNALLING	0,05
KEGG VASOPRESSIN REGULATED WATER REABSORPTION	0,05
PID AR TF PATHWAY	0,05
REACTOME GABA B RECEPTOR ACTIVATION	0,05
PID TRKRPATHWAY	0,05
REACTOME MRNA SPLICING	0,05
KEGG PPAR SIGNALING PATHWAY	0,06
REACTOME NUCLEAR SIGNALING BY ERBB4	0,06
PID TAP63PATHWAY	0,06
REACTOME CELL JUNCTION ORGANIZATION	0,06
REACTOME DEVELOPMENTAL BIOLOGY	0,06
REACTOME RESPONSE TO ELEVATED PLATELET CYTOSOLIC CA2	0,06
KEGG RNA DEGRADATION	0,06
PID AURORA B PATHWAY	0,06
PID FOXOPATHWAY	0,06
REACTOME P75 NTR RECEPTOR MEDIATED SIGNALLING	0,07
REACTOME AXON GUIDANCE	0,07
REACTOME RNA POL II TRANSCRIPTION	0,07
BIOCARTA AGR PATHWAY	0,07
ST FAS SIGNALING PATHWAY	0,07
REACTOME TRANSCRIPTION	0,07
PID ERA GENOMIC PATHWAY	0,07
WNT SIGNALING	0,08
PID ECADHERIN STABILIZATION PATHWAY	0,08
KEGG WNT SIGNALING PATHWAY	0,08
REACTOME METABOLISM OF PROTEINS	0,08
KEGG MELANOGENESIS	0,09
REACTOME BIOLOGICAL OXIDATIONS	0,09
REACTOME PHOSPHOLIPID METABOLISM	0,10
REACTOME POTASSIUM CHANNELS	0,10

B.

NAME	FDR
KEGG LEISHMANIA INFECTION	0,00
KEGG ALLOGRAFT REJECTION	0,00
KEGG INTESTINAL IMMUNE NETWORK FOR IGA PRODUCTION	0,00
KEGG AUTOIMMUNE THYROID DISEASE	0,00
REACTOME GENERATION OF SECOND MESSENGER MOLECULES	0,00
KEGG HEMATOPOIETIC CELL LINEAGE	0,00
KEGG TYPE I DIABETES MELLITUS	0,00
BIOCARTA TH1TH2 PATHWAY	0,00
KEGG ANTIGEN PROCESSING AND PRESENTATION	0,00
KEGG GRAFT VERSUS HOST DISEASE	0,00
PID TCR PATHWAY	0,00
PID IL12 2PATHWAY	0,00
REACTOME TCR SIGNALING	0,00
REACTOME CHEMOKINE RECEPTORS BIND CHEMOKINES	0,00
KEGG VIRAL MYOCARDITIS	0,00
REACTOME COSTIMULATION BY THE CD28 FAMILY	0,00
REACTOME PD1 SIGNALING	0,00
KEGG ASTHMA	0,00
REACTOME IMMUNOREGULATORY INTERACTIONS BETWEEN A LYMPHOID AND A NON LYMPHOID CELL	0,00
REACTOME INTERFERON GAMMA SIGNALING	0,00
BIOCARTA NKT PATHWAY	0,00
KEGG NATURAL KILLER CELL MEDIATED CYTOTOXICITY	0,00
BIOCARTA DC PATHWAY	0,00
BIOCARTA CTLA4 PATHWAY	0,00
REACTOME PHOSPHORYLATION OF CD3 AND TCR ZETA CHAINS	0,00
BIOCARTA LAIR PATHWAY	0,00
REACTOME PEPTIDE LIGAND BINDING RECEPTORS	0,00
REACTOME DOWNSTREAM TCR SIGNALING	0,00
KEGG PRIMARY IMMUNODEFICIENCY	0,00
PID IL23PATHWAY	0,00
REACTOME DEFENSINS	0,00
REACTOME CLASS A1 RHODOPSIN LIKE RECEPTORS	0,00
BIOCARTA NO2IL12 PATHWAY	0,00
PID IL12 STAT4PATHWAY	0,00
PID CD8TCRPATHWAY	0,00
KEGG FC GAMMA R MEDIATED PHAGOCYTOSIS	0,00
SA MMP CYTOKINE CONNECTION	0,00
REACTOME IL 2 SIGNALING	0,00
BIOCARTA INFLAM PATHWAY	0,00
BIOCARTA TOB1 PATHWAY	0,00
KEGG B CELL RECEPTOR SIGNALING PATHWAY	0,00
REACTOME IL RECEPTOR SHC SIGNALING	0,00
PID PI3KCIPATHWAY	0,00
PID CXCR4 PATHWAY	0,00
REACTOME IL 3 5 AND GM CSF SIGNALING	0,00
REACTOME INNATE IMMUNE SYSTEM	0,00
KEGG SYSTEMIC LUPUS ERYTHEMATOSUS	0,00
PID FCER1PATHWAY	0,01
REACTOME CELL SURFACE INTERACTIONS AT THE VASCULAR WALL	0,01
REACTOME GPVI MEDIATED ACTIVATION CASCADE	0,01
KEGG CYTOKINE CYTOKINE RECEPTOR INTERACTION	0,01
PID IL27PATHWAY	0,01
REACTOME CYTOKINE SIGNALING IN IMMUNE SYSTEM	0,01
REACTOME MHC CLASS II ANTIGEN PRESENTATION	0,01

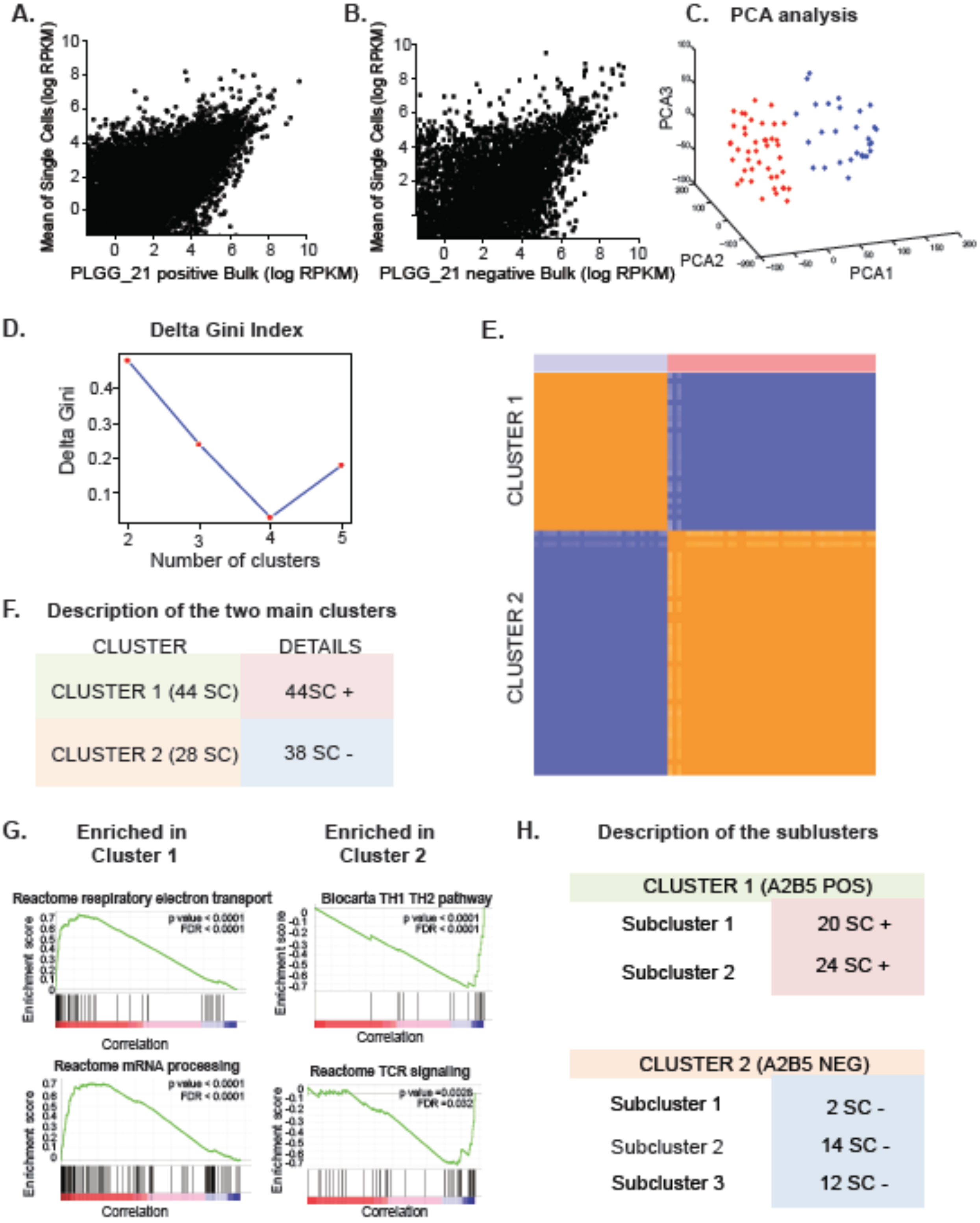
PID BCR 5PATHWAY	0,01
REACTOME INTERFERON SIGNALING	0,01
KEGG NOD LIKE RECEPTOR SIGNALING PATHWAY	0,02
KEGG TOLL LIKE RECEPTOR SIGNALING PATHWAY	0,02
REACTOME GPCR LIGAND BINDING	0,03
KEGG CHEMOKINE SIGNALING PATHWAY	0,03
KEGG LEUKOCYTE TRANSENDOTHELIAL MIGRATION	0,04
PID TXA2PATHWAY	0,04
REACTOME TOLL RECEPTOR CASCADES	0,05
BIOCARTA IL17 PATHWAY	0,00
BIOCARTA CYTOKINE PATHWAY	0,00
REACTOME G ALPHA I SIGNALLING EVENTS	0,07
BIOCARTA IL7 PATHWAY	0,01
PID CMYB PATHWAY	0,06
BIOCARTA CSK PATHWAY	0,00
ST INTERLEUKIN 4 PATHWAY	0,01
PID IL2 STAT5PATHWAY	0,01
BIOCARTA IL10 PATHWAY	0,01
KEGG FC EPSILON RI SIGNALING PATHWAY	0,02
KEGG CELL ADHESION MOLECULES CAMS	0,02
ST T CELL SIGNAL TRANSDUCTION	0,04
KEGG T CELL RECEPTOR SIGNALING PATHWAY	0,01
REACTOME SIGNALING BY ILS	0,03
BIOCARTA NKCELLS PATHWAY	0,01
REACTOME THE ROLE OF NEF IN HIV1 REPLICATION AND DISEASE PATHOGENESIS	0,00
KEGG JAK STAT SIGNALING PATHWAY	0,10
BIOCARTA MPR PATHWAY	0,04
BIOCARTA NTHI PATHWAY	0,02
REACTOME NUCLEOTIDE BINDING DOMAIN LEUCINE RICH REPEAT CONTAINING RECEPTOR NLR SIGNALING PATHWAYS	0,03
KEGG CYTOSOLIC DNA SENSING PATHWAY	0,04
BIOCARTA IL2RB PATHWAY	0,06
SIG IL4RECEPTOR IN B LYPHOCYTES	0,05
REACTOME NOD1 2 SIGNALING PATHWAY	0,06
PID PTP1BPATHWAY	0,06
PID INTEGRIN2 PATHWAY	0,02
BIOCARTA CDC42RAC PATHWAY	0,01

The second tumor analyzed (PLGG 21) contained 44 A2B5 positive and 28 A2B5 negative SC that each exhibited expression of at least 1500 genes. Similarly to PLGG 19, we observed for both A2B5 positive and negative SC a good correlation between the mean gene expression of the SC to the gene expression in the corresponding bulk control (Figure 34A and 34B). PCA revealed that the A2B5 positive SC clustered separately from the A2B5 negative cells (Figure 34C). SOM consensus clustering showed that two clusters best segregated the 82 SC (Figure 34D). Strikingly, we observed that the clusters obtained by SOM segregated strongly A2B5 positive and negative SC, respectively (Figure 34F and 34E, p<0.0001).

Using CMS and GSEA analyses, we found that 3,420 genes and 283 gene-sets were significantly differentially expressed between the A2B5 positive and negative SC. Among the top 100 genes significantly overexpressed in the A2B5 positive cells relative to the negative SC and harboring the highest fold-change, we noticed a predominance of genes mostly involved in metabolic pathways, such as *PLPP* (plasmolipin), considered as a myelin protein involved in human hereditary demyelinating neuropathies, *NDUFA1* (NADH dehydrogenase (ubiquinone) 1 alpha subcomplex), involved in the proton translocation process in the respiratory chain, *ALDOC* (aldolase C, fructose-bisphosphate), a glycolytic enzyme specifically expressed in the hippocampus and Purkinje cells of the brain and *COX8A* (cytochrome c oxidase subunit VIIIA), involved in the catalytic function of the terminal steps in the respiratory chain (Table 15). Conversely, among the top 100 genes significantly overexpressed in the A2B5 negative SC relative to the positive SC, we found a enrichment in genes involved in inflammatory pathways, such as *C1QB* (complement component 1, q subcomponent, B chain), a major constituent of the human complement subcomponent C1q), *CCL3L3* (chemokine ligand 3-like 3), *HLA-DQB1* (major histocompatibility complex, class II, DQ beta 1), *CD69* and *RFPL4A* (ret finger protein-like 4A) (Table 15).

Similar patterns were also observed in the GSEA results. The three first gene-sets significantly enriched in the cluster 1 were represented by respiratory electron transport (Figure 34G and Table 16). We also noticed a significant enrichment in NOTCH pathways gene-sets in the A2B5 positive cells (7 NOTCH related gene-sets were enriched, $p<0.012$). 9/14 (64%) gene-sets significantly enriched in the A2B5 negative population were directly related to inflammatory pathways. We then wanted to explore the degree of heterogeneity within each subgroup of SC. SOM consensus clustering performed on the A2B5 positive SC revealed two major groups whereas the A2B5 negative SC segregated in 3 groups (Figure 34H), suggesting that the heterogeneity within the A2B5 negative cells is significantly higher than the A2B5 positive SC in this tumor.

Figure 34: Results of PLGG 21 single-cell RNA-sequencing experiment



F. Description of the two main clusters

CLUSTER	DETAILS
CLUSTER 1 (44 SC)	44SC +
CLUSTER 2 (28 SC)	38 SC -

H. Description of the sublusters

CLUSTER 1 (A2B5 POS)	
Subcluster 1	20 SC +
Subcluster 2	24 SC +

CLUSTER 2 (A2B5 NEG)	
Subcluster 1	2 SC -
Subcluster 2	14 SC -
Subcluster 3	12 SC -

Figure 34 (legend):

A-B, graphical representation of the correlation between the average gene expression of SC relative to the bulk gene expression for the positive A2B5 cells (A) and negative cells (B). C, PCA analysis showing the three first components of the SC A2B5 positive (red) and negative (blue). D, graphical representation of the delta-gini index variation obtained by SOM clustering, E, heat-map of the SOM clustering (red represents A2B5 SC positive cells, blue the A2B5 negative cells). F, summary table of the distribution of the different SC by SOM clustering. G, GSEA enrichment plots significantly enriched in cluster 1 and 2. H, summary table of the distribution of the different sublclusters of SC by SOM clustering

Table 15: List of the top 100 genes that are significantly differentially expressed between A2B5 positive cells (POS) and negative cells (NEG) with the fold change of gene expression.

Overexpressed in	gene	FDR	Fold-change
POS	PLLP	0,0	50429,2
POS	AL132780.1	0,0	48890,6
POS	RP11-701H24.3	0,0	34673,1
POS	NDUFA1	0,0	32654,1
POS	ALDOC	0,0	27847,4
POS	AC012379.1	0,0	24540,7
POS	AL590762.1	0,0	23715,0
POS	RP3-398G3.3	0,0	22231,1
POS	NDUFB7	0,0	22040,3
POS	SNX22	0,0	20967,8
POS	COX8A	0,0	20374,7
POS	MRPS18A	0,0	20075,9
POS	BEX4	0,0	19319,9
POS	RP11-566K8.1	0,0	19264,8
POS	COPS6	0,0	18834,6
POS	NDUFB11	0,0	18260,1
POS	PSMB5	0,0	17889,7
POS	RP11-512N4.2	0,0	16823,3
POS	PEPD	0,0	16815,1
POS	RP11-599J14.2	0,0	16511,0
POS	PSMC4	0,0	15986,2
POS	NNMT	0,0	15963,8
POS	BEX1	0,0	15361,1
POS	RABAC1	0,0	14655,1
POS	RP11-21N7.2	0,0	14426,1
POS	RP11-728F11.6	0,0	14314,1
POS	EIF3G	0,0	14293,1
POS	SLC25A35	0,0	14161,5
POS	TMEM179B	0,0	13936,6
POS	CTD-2206G10.2	0,0	13917,7
POS	MAGEH1	0,0	13759,2
POS	HIST1H3A	0,0	13610,5
POS	ATP5J	0,0	13536,1
POS	AC004797.1	0,0	13171,0
POS	CNBP	0,0	13083,1
POS	NDUFB9	0,0	13045,6
POS	SNORA11	0,0	12904,4
POS	TCEB2	0,0	12899,3
POS	NDUFA6	0,0	12703,6
POS	MT1M	0,0	12188,6
POS	UBA3	0,0	11784,5
POS	HIRIP3	0,0	11756,3
POS	FAM127A	0,0	11737,6
POS	OAT	0,0	11591,0
POS	PRDX2	0,0	11554,6
POS	DCXR	0,0	11552,7
POS	HIGD2A	0,0	11434,5
POS	SLC25A4	0,0	11366,2
POS	RP11-354M1.3	0,0	11329,4
POS	RP3-525N10.2	0,0	10852,7
POS	SARS	0,0	10672,6

POS	RP4-604K5.1	0,0	10650,3
POS	EDF1	0,0	10474,3
POS	RP11-342M3.5	0,0	10282,7
POS	HINT1	0,0	10244,8
POS	SSBP1	0,0	10196,8
POS	MRPL51	0,0	10141,4
POS	RP3-330M21.5	0,0	10113,6
POS	BEX2	0,0	10075,1
POS	AC006195.2	0,0	9885,3
POS	C6orf108	0,0	9851,4
POS	ACTRT2	0,0	9843,1
POS	RP11-216M21.1	0,0	9828,1
POS	LRRC14	0,0	9734,5
POS	AC138744.2	0,0	9685,6
POS	AC009948.3	0,0	9610,2
POS	NDUFA13	0,0	9524,8
POS	RP11-230C9.1	0,0	9360,1
POS	OR51D1	0,0	9159,4
POS	S100A10	0,0	8887,5
POS	CHST12	0,0	8876,7
POS	ZMAT5	0,0	8834,9
POS	PSMD8	0,0	8808,6
POS	SEPP1	0,0	8747,2
POS	FIBP	0,0	8694,3
POS	PRNP	0,0	8620,8
POS	RP11-801I18.1	0,0	8578,8
POS	RPS15	0,0	8558,4
POS	C16orf80	0,0	8552,6
POS	APOA1BP	0,0	8494,1
POS	RP11-474L23.3	0,0	8479,2
POS	EFNA1	0,0	8376,0
POS	GALT	0,0	8351,8
POS	RCBTB1	0,0	8346,9
POS	APEH	0,0	8345,4
POS	TUBG2	0,0	8131,1
POS	DHPS	0,0	8127,9
POS	FAM127B	0,0	8113,8
POS	RP11-453M1.1	0,0	8088,6
POS	PEF1	0,0	8066,4
POS	RP11-63M22.2	0,0	8043,6
POS	SNX3	0,0	7951,6
POS	TH1L	0,0	7807,7
POS	FAM96B	0,0	7794,6
POS	MIR593	0,0	7792,9
POS	HIGD1B	0,0	7756,2
POS	RP3-388M5.8	0,0	7727,8
POS	BRP44L	0,0	7606,1
POS	RP11-447H19.3	0,0	7564,4
POS	AL445309.1	0,0	7533,7
NEG	RN5S151	0,0	114386,0
NEG	C1QB	0,0	46165,0
NEG	AC112165.1	0,0	45406,6
NEG	MIR30E	0,0	32716,1
NEG	CCL3L3	0,0	26970,7
NEG	OR8Q1P	0,0	22624,7
NEG	Y RNA	0,0	21421,8
NEG	APOC1	0,0	21140,8
NEG	AC022201.5	0,0	20234,2
NEG	U6	0,0	19247,0
NEG	CYBB	0,0	18433,4
NEG	OR13C3	0,0	15062,2
NEG	AL162430.1	0,0	14237,6

NEG	AC108696.1	0,0	14015,9
NEG	AC073352.1	0,0	13635,8
NEG	RP11-2A1.1	0,0	12488,6
NEG	HPDL	0,0	12222,6
NEG	RP11-299L17.3	0,0	11632,0
NEG	Y RNA	0,0	11503,6
NEG	RP13-128O4.3	0,0	10936,3
NEG	RNASE1	0,0	10097,8
NEG	RP11-456A14.1	0,0	10097,6
NEG	ADORA3	0,0	9666,2
NEG	SNORA51	0,0	9250,3
NEG	RN5S98	0,0	9200,9
NEG	FAM71E1	0,0	8879,5
NEG	HLA-DQB1	0,0	8012,9
NEG	CD69	0,0	7345,0
NEG	RFPL4A	0,0	7336,3
NEG	RP11-1085N6.5	0,0	7254,9
NEG	C11orf52	0,0	7152,5
NEG	CXCR4	0,0	7054,4
NEG	SRGN	0,0	6947,7
NEG	VSIG4	0,0	6834,7
NEG	SNORA81	0,0	6815,4
NEG	XX-CR54.1	0,0	6511,9
NEG	AIF1	0,0	6466,1
NEG	CCL3	0,0	5780,3
NEG	U6	0,0	5670,9
NEG	MS4A6A	0,0	5415,7
NEG	RGS1	0,0	5240,0
NEG	U6	0,0	5129,0
NEG	U6atac	0,0	4954,6
NEG	MIR548E	0,0	4864,4
NEG	RP11-25K21.1	0,0	4647,5
NEG	CTD-3195I5.1	0,0	4640,1
NEG	P2RY12	0,0	4635,2
NEG	AC007163.3	0,0	4509,2
NEG	snoU13	0,0	4345,3
NEG	CCL4	0,0	4183,2
NEG	RP11-147K6.2	0,0	4052,7
NEG	RP11-505P4.7	0,0	4032,8
NEG	RP11-867G2.5	0,0	4023,7
NEG	PRICKLE4	0,0	4020,9
NEG	RP11-379P15.1	0,0	3892,4
NEG	OR4F6	0,0	3880,1
NEG	RP11-667M19.1	0,0	3856,4
NEG	Y RNA	0,0	3785,8
NEG	OR52S1P	0,0	3506,9
NEG	AC005005.1	0,0	3423,9
NEG	RP11-64K12.4	0,0	3267,1
NEG	C2orf70	0,0	3263,8
NEG	C3	0,0	3257,4
NEG	RP11-445H22.4	0,0	3155,7
NEG	PARVG	0,0	3118,9
NEG	ZAR1L	0,0	3061,7
NEG	RP11-930O11.1	0,0	3045,1
NEG	SNORA79	0,0	2881,7
NEG	RN5S226	0,0	2779,4
NEG	Y RNA	0,0	2729,5
NEG	RP11-404P21.1	0,0	2407,9
NEG	TAAR5	0,0	2401,6
NEG	OR10V3P	0,0	2344,6
NEG	AC015987.1	0,0	2324,5
NEG	RP11-66H6.3	0,0	2257,9

NEG	U6	0,0	2218,9
NEG	RP11-49O14.2	0,0	2122,8
NEG	FBXO46	0,0	2016,8
NEG	CTSS	0,0	1966,8
NEG	ATP8B5P	0,0	1914,8
NEG	AC140120.1	0,0	1900,4
NEG	RP11-52I18.1	0,0	1839,3
NEG	RNU5A-6P	0,0	1741,2
NEG	IL8	0,0	1651,6
NEG	TNFAIP3	0,0	1591,5
NEG	C12orf35	0,0	1591,1
NEG	CD2	0,0	1562,2
NEG	TRNAQ36P	0,0	1562,0
NEG	LAPTM5	0,0	1502,3
NEG	XAGE1E	0,0	1499,5
NEG	U6	0,0	1368,1
NEG	SNORD56	0,0	1367,4
NEG	REM1	0,0	1318,4
NEG	AL049831.1	0,0	1216,8
NEG	CEBPB	0,0	1211,4
NEG	RP11-108K3.1	0,0	1140,6
NEG	BX255972.1	0,0	1138,3
NEG	SPINK14	0,0	996,9
NEG	TRAV38-2DV8	0,0	986,2
NEG	CTD-2316B1.2	0,0	969,1

Table 16: GSEA results for cluster 1 (A) and 2 (B) in PLGG 21 tumor.

A.

NAME	FDR
REACTOME RESPIRATORY ELECTRON TRANSPORT	0,00
REACTOME RESPIRATORY ELECTRON TRANSPORT ATP SYNTHESIS BY CHEMIOSMOTIC COUPLING AND HEAT PRODUCTION BY UNCOUPLING PROTEINS	0,00
REACTOME TCA CYCLE AND RESPIRATORY ELECTRON TRANSPORT	0,00
KEGG PARKINSONS DISEASE	0,00
KEGG OXIDATIVE PHOSPHORYLATION	0,00
KEGG HUNTINGTONS DISEASE	0,00
KEGG ALZHEIMERS DISEASE	0,00
REACTOME MRNA SPLICING	0,00
BIOCARTA PROTEASOME PATHWAY	0,00
REACTOME NGF SIGNALLING VIA TRKA FROM THE PLASMA MEMBRANE	0,00
REACTOME REGULATION OF ORNITHINE DECARBOXYLASE ODC	0,00
REACTOME PROCESSING OF CAPPED INTRON CONTAINING PRE MRNA	0,00
REACTOME HIV INFECTION	0,00
REACTOME MRNA PROCESSING	0,00
REACTOME DOWNSTREAM SIGNALING EVENTS OF B CELL RECEPTOR BCR	0,00
REACTOME ER PHAGOSOME PATHWAY	0,00
KEGG ENDOCYTOSIS	0,00
REACTOME LATE PHASE OF HIV LIFE CYCLE	0,00
REACTOME PREFOLDIN MEDIATED TRANSFER OF SUBSTRATE TO CCT TRIC	0,00
KEGG VIBRIO CHOLERAE INFECTION	0,00
PID REG GR PATHWAY	0,00
REACTOME MRNA SPLICING MINOR PATHWAY	0,00
REACTOME SIGNALING BY THE B CELL RECEPTOR BCR	0,00
REACTOME MRNA CAPPING	0,00
REACTOME SIGNALING BY WNT	0,00
BIOCARTA MEF2D PATHWAY	0,00
REACTOME ACTIVATION OF KAINATE RECEPTORS UPON GLUTAMATE BINDING	0,00
KEGG PEROXISOME	0,00
REACTOME AUTODEGRADATION OF CDH1 BY CDH1 APC C	0,00
REACTOME REGULATION OF MRNA STABILITY BY PROTEINS THAT BIND AU RICH ELEMENTS	0,00
REACTOME CROSS PRESENTATION OF SOLUBLE EXOGENOUS ANTIGENS ENDOSOMES	0,00
REACTOME GLUCOSE METABOLISM	0,00
REACTOME RNA POL II TRANSCRIPTION	0,00
REACTOME HIV LIFE CYCLE	0,00
PID ERBB1 DOWNSTREAM PATHWAY	0,00
REACTOME RNA POL II PRE TRANSCRIPTION EVENTS	0,00
PID MTOR 4PATHWAY	0,00
PID ECADHERIN STABILIZATION PATHWAY	0,00
REACTOME FORMATION OF RNA POL II ELONGATION COMPLEX	0,00
REACTOME SIGNALING BY ERBB4	0,00
BIOCARTA GPCR PATHWAY	0,00
REACTOME VIF MEDIATED DEGRADATION OF APOBEC3G	0,00
REACTOME FORMATION OF THE HIV1 EARLY ELONGATION COMPLEX	0,00
BIOCARTA NOS1 PATHWAY	0,00
REACTOME DOWNSTREAM SIGNAL TRANSDUCTION	0,00
REACTOME PEROXISOMAL LIPID METABOLISM	0,00
KEGG PROTEASOME	0,00
KEGG SPLICEOSOME	0,00
REACTOME REGULATION OF APOPTOSIS	0,00
REACTOME FORMATION OF TRANSCRIPTION COUPLED NER TC NER REPAIR COMPLEX	0,00
REACTOME NUCLEOTIDE EXCISION REPAIR	0,00
KEGG BIOSYNTHESIS OF UNSATURATED FATTY ACIDS	0,00
REACTOME RNA POL II TRANSCRIPTION PRE INITIATION AND PROMOTER OPENING	0,00
PID HDAC CLASSIII PATHWAY	0,00
REACTOME ACTIVATION OF NF KAPPAB IN B CELLS	0,00
REACTOME P53 INDEPENDENT G1 S DNA DAMAGE CHECKPOINT	0,00
REACTOME MITOCHONDRIAL PROTEIN IMPORT	0,00
REACTOME SIGNALING BY NOTCH	0,00
REACTOME SCF BETA TRCP MEDIATED DEGRADATION OF EMI1	0,00
PID TELOMERASEPATHWAY	0,00
REACTOME SCFSKP2 MEDIATED DEGRADATION OF P27 P21	0,00
REACTOME SIGNALLING BY NGF	0,00

REACTOME DARPP 32 EVENTS	0,00
REACTOME HOST INTERACTIONS OF HIV FACTORS	0,00
REACTOME NOTCH1 INTRACELLULAR DOMAIN REGULATES TRANSCRIPTION	0,00
REACTOME GLYCOLYSIS	0,00
KEGG PROPANOATE METABOLISM	0,00
REACTOME DESTABILIZATION OF MRNA BY AUF1 HNRNP D0	0,00
REACTOME FORMATION OF INCISION COMPLEX IN GG NER	0,00
KEGG VALINE LEUCINE AND ISOLEUCINE DEGRADATION	0,00
REACTOME CYCLIN E ASSOCIATED EVENTS DURING G1 S TRANSITION	0,00
REACTOME ANTIGEN PROCESSING UBIQUITINATION PROTEASOME DEGRADATION	0,00
KEGG PROTEIN EXPORT	0,00
KEGG CARDIAC MUSCLE CONTRACTION	0,00
REACTOME ABORTIVE ELONGATION OF HIV1 TRANSCRIPT IN THE ABSENCE OF TAT	0,00
REACTOME GLUCAGON SIGNALING IN METABOLIC REGULATION	0,00
REACTOME CDK MEDIATED PHOSPHORYLATION AND REMOVAL OF CDC6	0,00
REACTOME CLASS I MHC MEDIATED ANTIGEN PROCESSING PRESENTATION	0,00
PID TRKRPATHWAY	0,00
KEGG ADHERENS JUNCTION	0,00
REACTOME ACTIVATION OF CHAPERONE GENES BY XBP1S	0,00
REACTOME AUTODEGRADATION OF THE E3 UBIQUITIN LIGASE COP1	0,00
REACTOME DIABETES PATHWAYS	0,00
REACTOME TRANSCRIPTION COUPLED NER TC NER	0,00
KEGG GLIOMA	0,00
REACTOME MEMBRANE TRAFFICKING	0,00
BIOCARTA MPR PATHWAY	0,00
REACTOME MAPK TARGETS NUCLEAR EVENTS MEDIATED BY MAP KINASES	0,00
PID HDAC CLASSI PATHWAY	0,00
REACTOME CHOLESTEROL BIOSYNTHESIS	0,00
REACTOME APC C CDC20 MEDIATED DEGRADATION OF MITOTIC PROTEINS	0,00
BIOCARTA MTA3 PATHWAY	0,00
KEGG INSULIN SIGNALING PATHWAY	0,00
KEGG LYSINE DEGRADATION	0,00
KEGG GAP JUNCTION	0,00
REACTOME NEUROTRANSMITTER RECEPTOR BINDING AND DOWNSTREAM TRANSMISSION IN THE POSTSYNAPTIC CELL	0,00
REACTOME TRANSMISSION ACROSS CHEMICAL SYNAPSES	0,00
REACTOME SIGNALING BY ERBB2	0,00
KEGG RNA POLYMERASE	0,00
KEGG LONG TERM POTENTIATION	0,00
REACTOME METABOLISM OF CARBOHYDRATES	0,00
BIOCARTA CARM ER PATHWAY	0,00
REACTOME PYRUVATE METABOLISM AND CITRIC ACID TCA CYCLE	0,00
REACTOME PROCESSING OF CAPPED INTRONLESS PRE MRNA	0,00
REACTOME ACTIVATION OF NMDA RECEPTOR UPON GLUTAMATE BINDING AND POSTSYNAPTIC EVENTS	0,00
PID RAC1 REG PATHWAY	0,00
KEGG CITRATE CYCLE TCA CYCLE	0,00
REACTOME TRANSPORT OF MATURE MRNA DERIVED FROM AN INTRONLESS TRANSCRIPT	0,00
BIOCARTA NDKDYNAMIN PATHWAY	0,00
REACTOME SMOOTH MUSCLE CONTRACTION	0,00
REACTOME SIGNALING BY FGFR IN DISEASE	0,00
KEGG UBIQUITIN MEDIATED PROTEOLYSIS	0,00
BIOCARTA EIF PATHWAY	0,00
REACTOME NUCLEAR SIGNALING BY ERBB4	0,00
BIOCARTA AKAPCENTROSOME PATHWAY	0,00
KEGG OOCYTE MEIOSIS	0,00
KEGG PYRIMIDINE METABOLISM	0,00
REACTOME FATTY ACID TRIACYLGLYCEROL AND KETONE BODY METABOLISM	0,00
REACTOME CDT1 ASSOCIATION WITH THE CDC6 ORC ORIGIN COMPLEX	0,00
REACTOME BMAL1 CLOCK NPAS2 ACTIVATES CIRCADIAN EXPRESSION	0,00
REACTOME ASSEMBLY OF THE PRE REPLICATIVE COMPLEX	0,00
PID BETACATENIN NUC PATHWAY	0,00
REACTOME REGULATION OF WATER BALANCE BY RENAL AQUAPORINS	0,00
REACTOME PRE NOTCH EXPRESSION AND PROCESSING	0,00
REACTOME OPIOID SIGNALLING	0,00
REACTOME METABOLISM OF AMINO ACIDS AND DERIVATIVES	0,00
KEGG GLYCOLYSIS GLUCONEOGENESIS	0,00
REACTOME CIRCADIAN REPRESSION OF EXPRESSION BY REV ERBA	0,00
REACTOME PI3K EVENTS IN ERBB2 SIGNALING	0,00
BIOCARTA CALCINEURIN PATHWAY	0,00
REACTOME MITOTIC G1 G1 S PHASES	0,00

REACTOME CIRCADIAN CLOCK	0,00
PID PDGFRBPATHWAY	0,00
PID ERA GENOMIC PATHWAY	0,00
REACTOME FORMATION OF TUBULIN FOLDING INTERMEDIATES BY CCT TRIC	0,00
REACTOME METABOLISM OF NON CODING RNA	0,00
REACTOME ANTIGEN PROCESSING CROSS PRESENTATION	0,00
REACTOME DOWNSTREAM SIGNALING OF ACTIVATED FGFR	0,00
SIG INSULIN RECEPTOR PATHWAY IN CARDIAC MYOCYTES	0,00
PID HES HEYPATHWAY	0,00
KEGG RENAL CELL CARCINOMA	0,00
REACTOME SIGNALING BY PDGF	0,00
REACTOME PRE NOTCH TRANSCRIPTION AND TRANSLATION	0,00
REACTOME TRANSFERRIN ENDOCYTOSIS AND RECYCLING	0,00
REACTOME METABOLISM OF LIPIDS AND LIPOPROTEINS	0,00
KEGG NUCLEOTIDE EXCISION REPAIR	0,00
REACTOME SIGNALING BY NOTCH1	0,00
REACTOME P53 DEPENDENT G1 DNA DAMAGE RESPONSE	0,00
REACTOME TRAFFICKING OF AMPA RECEPTORS	0,00
REACTOME DNA REPAIR	0,00
REACTOME AXON GUIDANCE	0,00
REACTOME GLUCONEOGENESIS	0,00
BIOCARTA CREB PATHWAY	0,00
REACTOME APOPTOSIS	0,00
REACTOME CA DEPENDENT EVENTS	0,00
REACTOME SIGNALING BY EGFR IN CANCER	0,00
REACTOME CELL CYCLE CHECKPOINTS	0,00
REACTOME PKA MEDIATED PHOSPHORYLATION OF CREB	0,00
REACTOME PLC BETA MEDIATED EVENTS	0,00
PID NOTCH PATHWAY	0,00
PID BMPPATHWAY	0,00
REACTOME ELONGATION ARREST AND RECOVERY	0,00
REACTOME PI3K AKT ACTIVATION	0,00
REACTOME INTEGRATION OF ENERGY METABOLISM	0,00
REACTOME REGULATION OF INSULIN SECRETION	0,00
REACTOME SIGNALING BY FGFR	0,00
REACTOME RORA ACTIVATES CIRCADIAN EXPRESSION	0,00
REACTOME GENERIC TRANSCRIPTION PATHWAY	0,00
REACTOME MRNA 3 END PROCESSING	0,00
REACTOME PIP3 ACTIVATES AKT SIGNALING	0,00
REACTOME S PHASE	0,00
BIOCARTA BCR PATHWAY	0,00
BIOCARTA GATA3 PATHWAY	0,00
PID IL6 7PATHWAY	0,00
REACTOME GLOBAL GENOMIC NER GG NER	0,00
KEGG CYSTEINE AND METHIONINE METABOLISM	0,00
REACTOME APC C CDH1 MEDIATED DEGRADATION OF CDC20 AND OTHER APC C CDH1 TARGETED PROTEINS IN LATE MITOSIS EARLY G1	0,00
PID TGFBRPATHWAY	0,00
REACTOME REGULATORY RNA PATHWAYS	0,00
BIOCARTA ERK PATHWAY	0,00
REACTOME INSULIN RECEPTOR RECYCLING	0,00
BIOCARTA RHO PATHWAY	0,00
REACTOME REGULATION OF INSULIN SECRETION BY GLUCAGON LIKE PEPTIDE1	0,00
REACTOME SIGNALLING TO ERKS	0,00
BIOCARTA PPARA PATHWAY	0,00
REACTOME RECRUITMENT OF MITOTIC CENTROSOME PROTEINS AND COMPLEXES	0,00
REACTOME CLEAVAGE OF GROWING TRANSCRIPT IN THE TERMINATION REGION	0,00
REACTOME GAB1 SIGNALOSOME	0,00
REACTOME METABOLISM OF RNA	0,00
REACTOME SYNTHESIS SECRETION AND DEACYLATION OF GHRELIN	0,00
KEGG NOTCH SIGNALING PATHWAY	0,00
REACTOME REGULATION OF MITOTIC CELL CYCLE	0,00
REACTOME TRAFFICKING OF GLUR2 CONTAINING AMPA RECEPTORS	0,00
KEGG MISMATCH REPAIR	0,00
REACTOME G1 S TRANSITION	0,00
REACTOME M G1 TRANSITION	0,00
REACTOME DAG AND IP3 SIGNALING	0,00
REACTOME NEURONAL SYSTEM	0,00
PID PS1PATHWAY	0,00
REACTOME MITOTIC G2 G2 M PHASES	0,00

REACTOME DEVELOPMENTAL BIOLOGY	0,00
REACTOME PROTEIN FOLDING	0,00
REACTOME CELL CYCLE	0,00
REACTOME ORC1 REMOVAL FROM CHROMATIN	0,00
REACTOME INTERACTIONS OF VPR WITH HOST CELLULAR PROTEINS	0,00
REACTOME TRANSPORT OF MATURE TRANSCRIPT TO CYTOPLASM	0,00
PID ERBB1 INTERNALIZATION PATHWAY	0,00
REACTOME TRANSCRIPTION	0,00
PID HIF1 TFPATHWAY	0,00
PID MET PATHWAY	0,00
KEGG VASOPRESSIN REGULATED WATER REABSORPTION	0,00
REACTOME SIGNALING BY INSULIN RECEPTOR	0,00
REACTOME MUSCLE CONTRACTION	0,00
SA TRKA RECEPTOR	0,00
REACTOME METABOLISM OF PROTEINS	0,00

B.

NAME	FDR
BIOCARTA TH1TH2 PATHWAY	0,00
KEGG ASTHMA	0,00
BIOCARTA NKT PATHWAY	0,00
REACTOME PHOSPHORYLATION OF CD3 AND TCR ZETA CHAINS	0,00
REACTOME PD1 SIGNALING	0,00
KEGG INTESTINAL IMMUNE NETWORK FOR IGA PRODUCTION	0,00
REACTOME CHEMOKINE RECEPTORS BIND CHEMOKINES	0,01
REACTOME GENERATION OF SECOND MESSENGER MOLECULES	0,01
REACTOME DOWNSTREAM TCR SIGNALING	0,01

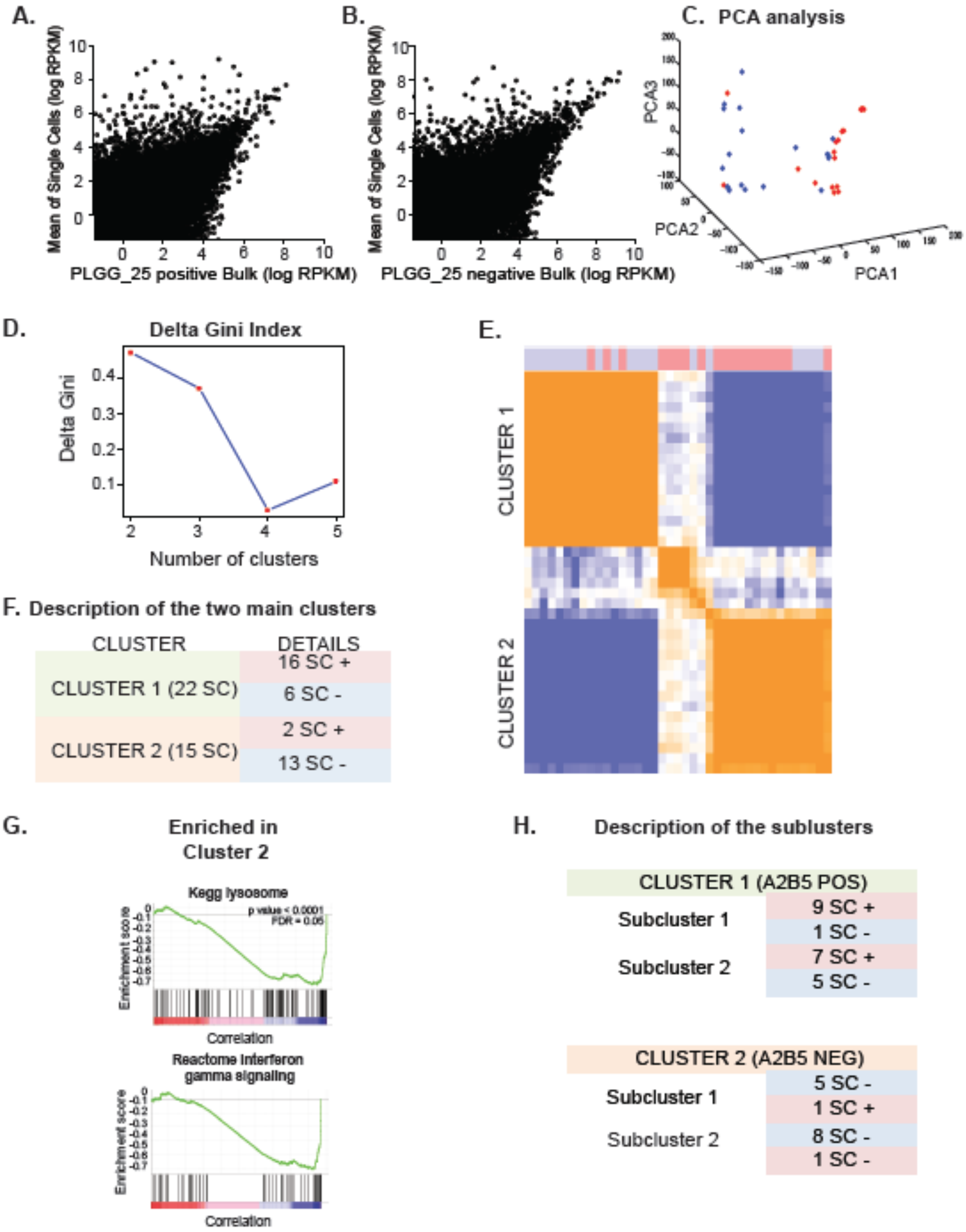
The third tumor analyzed (PLGG 25) contained 22 A2B5 positive SC and 15 A2B5 negative SC that exhibited expression of at least 1500 genes. Similarly to PLGG 19 and PLGG 22, we observed for both A2B5 positive and negative SC a good correlation between the mean gene expression of the SC to the gene expression in the corresponding bulk control (Figure 35A and 35B).

Although the segregation between A2B5 positive and negative SC was not as striking as the two previous tumors, PCA showed relatively distinct distributions between the two clusters for the A2B5 positive SC and the A2B5 negative cells (Fig 35C). SOM consensus clustering performed on the 37 SC showed that 2 clusters distinguished the best the matrix (Figure 35D). We observed a significant enrichment of the A2B5 positive and negative cells in cluster 1 and 2, respectively (Figure 35F and 35E, $p<0.0001$).

Using CMS and GSEA analyses, we found that 1,637 genes and 19 gene-sets were significantly differentially expressed between cluster 1 and 2. Beside the pseudogenes and small RNA molecules, we found *S100A9* (S100 calcium binding protein A9) involved in the regulation of a number of cellular processes such as cell cycle progression and differentiation, *CLEC18C* and *CLEC18A* (C-type lectin domain family 18) that belong to the superfamily of proteins containing C-type lectin-like domains proteins, harboring various cell functions such as cell trafficking[350] that are among the top 100 genes overexpressed in the cluster 1 (Table 17A). Among the top 100 genes overexpressed in cluster 2, we found *APOC1* (apolipoprotein C-I) activated in the transformation of monocytes into macrophages, *EVI2B* (ecotropic viral integration site 2B) overexpressed in cells derived from NF1 neurofibromas[351] and *SYNDIG1* (synapse differentiation inducing 1) involved in the development of the central nervous system[352] (Table 17B).

GSEA analysis revealed that 13/18 gene-sets (72%) significantly enriched in the cluster 2 were involved in inflammatory processes (Figure 35G and Table 18). We performed SOM consensus clustering on the cluster 1 and 2 separately and observed that both clusters segregated preferentially in 2 groups (Figure 35H).

Figure 35: Results of PLGG 25 single-cell RNA-sequencing experiment.



F. Description of the two main clusters

CLUSTER	DETAILS
CLUSTER 1 (22 SC)	16 SC +
	6 SC -
CLUSTER 2 (15 SC)	2 SC +
	13 SC -

Description of the sublusters

CLUSTER 1 (A2B5 POS)	
Subcluster 1	9 SC +
	1 SC -
Subcluster 2	7 SC +
	5 SC -

CLUSTER 2 (A2B5 NEG)	
Subcluster 1	5 SC -
	1 SC +
Subcluster 2	8 SC -
	1 SC -

Figure 35 (legend):

A-B, graphical representation of the correlation between the average gene expression of SC relative to the bulk gene expression for the positive A2B5 cells (A) and negative cells (B). C, PCA analysis showing the three first components of the SC A2B5 positive (red) and negative (blue). D, graphical representation of the delta-gini index variation obtained by SOM clustering, E, heat-map of the SOM clustering (red represents A2B5 SC positive cells, blue the A2B5 negative cells). F, summary table of the distribution of the different SC by SOM clustering. G, GSEA enrichment plots significantly enriched in cluster 2. H, summary table of the distribution of the different sublclusters of SC by SOM clustering

Table 17: Comparative marker selection results in A2B5 positive (A) and negative (B) cells for PLGG 25 tumor (top 100 genes with highest fold-change are represented)

A.

overexpressed	gene	FDR	Fold-change
pos	RN5S87	0,00	827228
pos	KRTAP21-3	0,00	795687
pos	AL132855.1	0,00	738701
pos	7SK	0,01	364228
pos	AC105383.1	0,00	236174
pos	MRPL20P1	0,00	185028
pos	S100A9	0,00	150278
pos	RN5S447	0,00	148862
pos	AC073352.1	0,00	132915
pos	OR5R1	0,00	109973
pos	AC114491.2	0,00	96716
pos	snoU13	0,00	96247
pos	TRNAK42P	0,00	62721
pos	RP11-411D10.1	0,00	60619
pos	AC103783.1	0,00	57426
pos	AC022027.1	0,00	39078
pos	RP11-115C10.1	0,00	38046
pos	RP11-353N14.4	0,00	37676
pos	DDX18P6	0,00	30109
pos	RP11-312P12.2	0,00	29363
pos	CLEC18B	0,00	29169
pos	Y RNA	0,00	28435
pos	Y RNA	0,00	27373
pos	SNORD112	0,00	26066
pos	RP11-95K23.7	0,00	25218
pos	GCOM1	0,00	25023
pos	snoU13	0,00	24775
pos	AC024367.1	0,00	23170
pos	AC000111.4	0,00	22689
pos	CLEC18C	0,00	21097
pos	AC126365.7	0,00	20799
pos	Y RNA	0,00	20733
pos	Y RNA	0,00	20530
pos	RNU7-55P	0,00	20180
pos	U6atac	0,00	20019
pos	AL162415.1	0,00	19628
pos	AL606830.1	0,01	19279
pos	snR39B	0,00	18133
pos	CLEC18A	0,00	17822
pos	RP11-5N11.5	0,00	17527
pos	Y RNA	0,00	17289
pos	TRIM47	0,01	17126
pos	CD52	0,00	17123
pos	U3	0,00	16794
pos	U6	0,00	16513
pos	RP11-379H8.1	0,00	16345
pos	AC104456.2	0,00	16042
pos	AC127391.3	0,00	15019
pos	U6	0,00	14735
pos	TRNAK37P	0,00	14129
pos	U6	0,00	13882

pos	TRBJ2-2	0,00	13329
pos	Y RNA	0,00	13105
pos	ZBTB20-AS2	0,00	12740
pos	RP11-365D9.1	0,00	12613
pos	Y RNA	0,00	12512
pos	Y RNA	0,00	12379
pos	snoU13	0,00	12296
pos	snoU13	0,00	12266
pos	snoU13	0,00	12030
pos	RP11-12D24.6	0,00	11934
pos	U6	0,00	11693
pos	UBQLN4P1	0,00	11454
pos	TMIE	0,00	11391
pos	SNORD42	0,00	11073
pos	MIR297	0,00	10576
pos	AC018607.3	0,01	10176
pos	RNU7-23P	0,00	10090
pos	LCK	0,01	10023
pos	Y RNA	0,00	9926
pos	RP11-554D15.4	0,00	9890
pos	CSRP2	0,00	9689
pos	AC026992.1	0,00	9474
pos	CTD-2542O7.2	0,00	9455
pos	AL022344.4	0,00	9342
pos	AC011453.1	0,00	9330
pos	UBL7	0,00	9304
pos	RP11-85G21.2	0,00	9196
pos	ELL2P2	0,00	9014
pos	RNU7-26P	0,00	9009
pos	AP003402.1	0,00	8937
pos	DGCR6L	0,00	8883
pos	OR4A40P	0,00	8767
pos	AC026406.1	0,00	8719
pos	TRNAQ41P	0,00	8649
pos	SNORA44	0,00	8624
pos	PIK3CA	0,00	8493
pos	RBM22P3	0,00	8395
pos	SNORD60	0,00	8330
pos	AC092966.1	0,00	8112
pos	CTD-2530H12.2	0,00	8036
pos	Y RNA	0,00	7882
pos	AP002982.1	0,00	7861
pos	AC012485.2	0,00	7778
pos	snoU13	0,00	7702
pos	TMEM194B	0,00	7666
pos	AC004673.1	0,00	7481
pos	RP11-390D11.2	0,00	7466
pos	SNORD113	0,00	7275
pos	Y RNA	0,00	7192

B.

overexpressed	gene	FDR	Fold-change
neg	APOC1	0,00	35141
neg	TMEM11	0,00	14192
neg	SYNDIG1	0,00	14084
neg	MIR616	0,00	11837
neg	U6	0,00	10385
neg	EVI2B	0,01	10099
neg	CYB561D2	0,00	9738
neg	SPA17P1	0,00	8357
neg	RP11-407G23.3	0,00	7584
neg	RP11-96D1.7	0,00	7488
neg	AC019131.1	0,00	7391
neg	DOK1	0,01	6602
neg	FCGR1B	0,00	6397
neg	U6	0,00	5983
neg	CTD-2647E9.3	0,00	5868
neg	SRA1	0,00	5623
neg	SHC1P1	0,00	5368
neg	BHLHE41	0,01	5354
neg	ADPGK	0,00	4930
neg	RP1-20B21.4	0,00	4642
neg	AC011754.1	0,01	4491
neg	EDDM3A	0,01	4478
neg	DLEU7-AS1	0,00	4438
neg	RP11-423H2.1	0,00	4242
neg	RP11-63M22.2	0,00	4233
neg	PCDH12	0,00	4201
neg	OR6A2	0,00	4185
neg	AF240627.2	0,00	4117
neg	AC091177.1	0,00	4110
neg	GPN3	0,01	3967
neg	FAN1	0,01	3960
neg	SIGLEC9	0,01	3928
neg	IL10RB	0,00	3911
neg	ACSF3	0,00	3667
neg	BLCAP	0,01	3652
neg	7SK	0,00	3646
neg	C1QC	0,00	3397
neg	RP1-257A7.4	0,00	3362
neg	SNORD13	0,00	3331
neg	FTSJ2	0,00	3298
neg	ZNF846	0,00	3258
neg	RABIF	0,00	3144
neg	RP13-644M16.3	0,00	3024
neg	EFTUD1	0,00	2942
neg	7SK	0,00	2928
neg	S100A11P1	0,00	2896
neg	TMEM109	0,00	2839
neg	MTIF3	0,01	2790
neg	U6	0,00	2749
neg	RN5S477	0,00	2641
neg	RP11-641C17.4	0,00	2619
neg	7SK	0,00	2614
neg	RP11-181C21.4	0,01	2597
neg	snoU13	0,01	2570
neg	AC006483.5	0,00	2557
neg	snoU13	0,00	2541
neg	RP11-950C14.8	0,00	2531
neg	AL592166.1	0,00	2504
neg	RP11-160E2.11	0,00	2459
neg	S100A11P2	0,00	2440

neg	RP11-25J3.2	0,00	2415
neg	CEBPG	0,00	2357
neg	ELMO1-AS1	0,00	2357
neg	RP11-685N10.1	0,00	2324
neg	SNRPA1	0,01	2280
neg	RN5S194	0,00	2244
neg	MIR29B2	0,00	2227
neg	RP3-417O22.1	0,00	2214
neg	RP5-886K2.3	0,00	2176
neg	OR2AG1	0,00	2133
neg	RP11-452F19.3	0,01	2112
neg	HPGDS	0,01	2084
neg	RP11-536L15.1	0,00	2078
neg	RP4-760C5.3	0,00	2064
neg	AC079140.1	0,00	2058
neg	CTD-2272D18.1	0,02	2047
neg	Y RNA	0,00	2039
neg	RP4-686C3.7	0,00	2008
neg	RP5-881L22.5	0,00	1982
neg	CTD-2001E22.1	0,01	1966
neg	AC073063.10	0,01	1937
neg	RP11-49O14.2	0,00	1900
neg	ATP8B4	0,00	1886
neg	RAB1C	0,00	1871
neg	RP13-192B19.2	0,00	1871
neg	AL137798.1	0,00	1854
neg	RP11-313E4.1	0,00	1842
neg	RN5S89	0,00	1836
neg	RAB15	0,00	1775
neg	RP11-582E3.2	0,02	1740
neg	LY96	0,01	1726
neg	TDGF1P6	0,01	1725
neg	COG6	0,01	1625
neg	5S rRNA	0,02	1608
neg	AC004979.2	0,02	1608
neg	AC093484.4	0,01	1599
neg	U4atac	0,01	1514
neg	SNORA20	0,00	1511
neg	RP11-438F14.3	0,02	1489
neg	PCNA	0,01	1475

Table 18: GSEA results for cluster 1 (A) and 2 (B) in PLGG 25 tumor.

A.

NAME	FDR
REACTOME CLASS B 2 SECRETIN FAMILY RECEPTORS	0,20
REACTOME A TETRASACCHARIDE LINKER SEQUENCE IS REQUIRED FOR GAG SYNTHESIS	0,24

B.

NAME	FDR
KEGG VIRAL MYOCARDITIS	0,00
KEGG AUTOIMMUNE THYROID DISEASE	0,00
KEGG LEISHMANIA INFECTION	0,00
KEGG ASTHMA	0,00
KEGG ALLOGRAFT REJECTION	0,00
KEGG TYPE I DIABETES MELLITUS	0,01
KEGG ANTIGEN PROCESSING AND PRESENTATION	0,01
KEGG INTESTINAL IMMUNE NETWORK FOR IGA PRODUCTION	0,01

3.3. A2B5 positive cells appear as a homogeneous population across different PLGG tumors

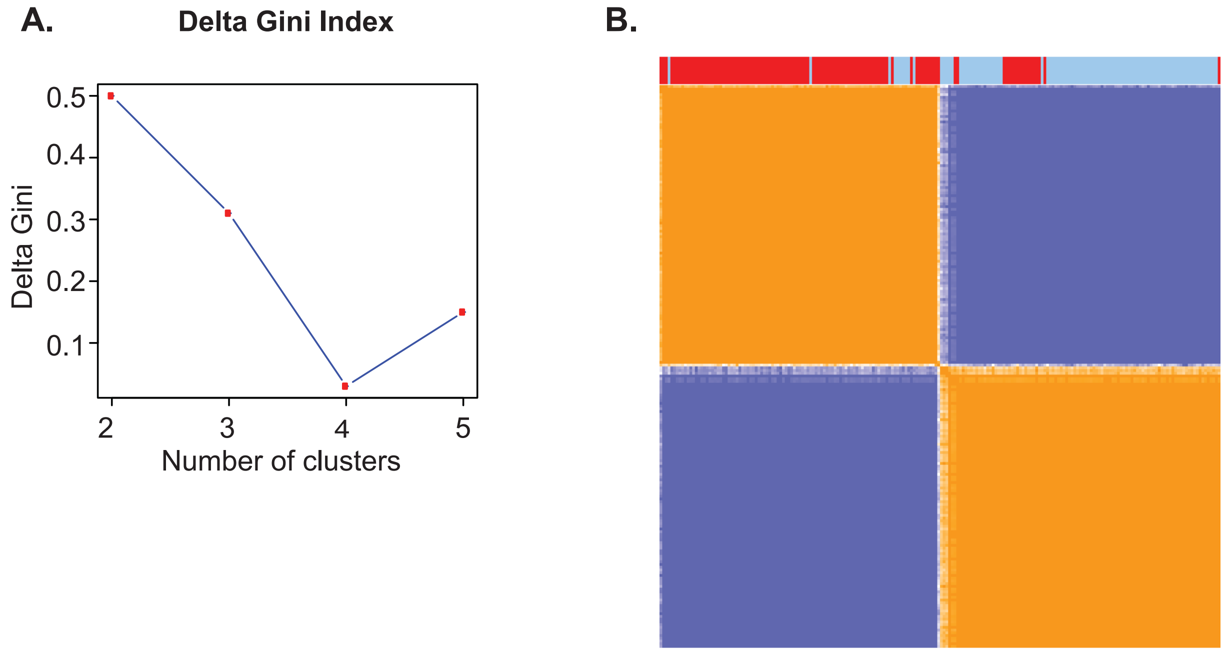
In order to evaluate whether the molecular patterns of A2B5 positive and negative cells were reproducible across different tumors we performed SOM consensus clustering including all the SC originated from the 3 tumors analyzed. We observed that the SC clustered preferentially in two groups (Figure 36A) and that the clusters were significantly enriched in A2B5 positive and negative cells, respectively ($p<0.0001$, Figure 36B). We further performed SVM analysis to evaluate how predictive was the molecular profile of A2B5 positive and negative SC across different individuals. We observed that 100% of the positive A2B5 SC were successfully predicted by the algorithm.

Figure 36:

SOM clustering results performed on all the SC across the 3 tumors tested

A. Delta Gini Index graph obtained by SOM analysis including all the SC from the 3 tumors

B. Heat-map of the SOM clustering (legend: red represents A2B5 SC positive cells, blue the A2B5 negative cells).



3.4. Discussion

In the first part of this section we presented the profiling of 26 PLGG tumors from different ages, locations, histological subtypes as well as BRAF genomic statuses using different markers of glioneuronal differentiation. Besides N-CAM that was uniformly poorly expressed across the tumors tested, we noticed a high degree of heterogeneity of expression of the different markers tested, as it has been described[353,354]. Although our cohort of tumors was composed with different histological subtypes, we could not distinguish distinct level of expression of GLAST, A2B5, FGFR1 or O4 according to a specific tumor type or location. The fact that we observed such important variability of the level of expression of the different glial differentiation markers can be related either by a lack of specificity of the different antibodies used or by an extreme heterogeneity of the populations of cells inside the tumors between different individuals.

We decided to perform the sorting experiment for SC RNA-sequencing using the A2B5 Ab as we had a higher confidence of the specificity of the staining of the Ab. Indeed for every sorting experiment we used mouse neural stem cells that stained the A2B5 Ab by flow cytometry as well as immunohistochemistry between 10-20%, which is similar to what has been previously reported[355]. We also used NIH 3T3 cells that systematically did not show any positive staining for the Ab as a negative control. Furthermore, A2B5 is well-recognized as a cancer stem cell marker. Several studies have shown that A2B5 isolated cells from human high grade gliomas were able to generate highly infiltrative tumors, suggesting that this population might be important in the initiation and the maintenance of tumors[356].

We then showed that we were able to dissociate PLGG tumors and sort successfully A2B5 positive and negative SC to perform genome wide RNA-sequencing analysis. Single-cell RNA-sequencing has recently been described in normal human and bacterial cells[357-359] as well as several cancer models such as prostate[360] and higher grade gliomas[361]. This represents a new opportunity to study in detail, cell by cell, the degree of heterogeneity of PLGGs and understand what major patterns are expressed in the different cells.

Unsupervised hierarchical clustering performed successively on the 3 tumors revealed strong similarities in terms of the significant clustering between positive and negative SC. Indeed, all cells originated from the 3 distinct tumors clustered preferentially in 2 major clusters characterized for one by A2B5 positive cells and for the other cluster the A2B5 negative cells. We showed that A2B5 positive cells were enriched in various metabolic processes pathways as well as cell-to-cell interaction pathways, compared to the A2B5 negative SC. Additionally, we observed that in the PLGG 18 tumor cells the expression of MKI67IP, associated with mitosis and cell cycle progression, was significantly higher in the A2B5 positive population relative to the negative population ($p< 0.0001$, Mann-Whitney test).

Even when combining the 3 tumors cells originated from different individuals, we were able to observe the similar pattern. This suggest that the A2B5 positive cells can be considered as a distinct population of cells among PLGGs, and especially PAs, as 2/3 tumors tested were PAs.

These results suggest that the results are accurate enough to reliably detect differences between cell types across different tumors and experiments.

One next important question is to understand how the BRAF duplicated cells are defined in our SC. Indeed, the 3 tumors tested have been diagnosed by FISH as BRAF duplicated. Based on the RNA-seq results, we have not been able to distinguish the SC that harbor the BRAF duplication, as the coverage of the gene was not sufficient. We are currently sequencing the transcriptome of the bulk tumors in order to identify the exact BRAF rearrangement type in order to validate by classical RT-PCR the mutation in our SC cDNAs left before library preparation. This information will allow us to observe whether the BRAF duplicated SC are preferentially A2B5 positive cells or not.

The other important finding was the importance of immunological perturbations enriched in the cluster dominated by the A2B5 negative cells across all 3 tumors. This suggests that within the tumors, there is a high degree of immunologic perturbation related probably to an important inflammatory cell infiltration within the tumor that might activate several cascades responsible of the tumor behavior.

Looking individually at the different subclusters of SC, we also discovered that the clusters dominated by the A2B5 negative cells were segregating in three distinct groups, whereas the A2B5 positive cells were only clustering in two groups. This pattern has not been observed in the last tumor, probably due to a lack of power (the number of single cells expressing at least 1500 genes was significantly reduced compared to the two first tumors analyzed).

This altogether suggests that PLGGs are characterized by a ‘stem cell-like’ population of A2B5 positive cells that might play a role in the initiation and the maintenance of PLGG tumors. The other cells are more heterogeneous and are highly characterized by an activation of inflammatory processes that are activated probably in response of the tumor growth but also causing itself damages to the normal CNS.

Limitations of the current study are represented by the bias of the sorting upfront. The cells are sorted based on their viability by a double staining strategy by flow cytometry. It is possible that not all the cells representative of the tumors are sorted and therefore are limiting the exhaustive landscape of the tumors cells. The other bias might be related to the tumor dissociation process that may cause damage to cells, especially glial cells. One last bias might be related by the RNA-sequencing protocol. We observed a non-negligible number of single cells that harbor a low number of genes expressed with poor overall coverage. This might be due to the cDNA and PCR steps that are not uniformly performed and generating preferential amplifications of regions, whereas other regions are not covered enough. All those steps require further closer optimization of the current protocol in order to increase the yield of the cDNA and PCR steps across the genome.

Chapter 5: General discussion and conclusion

We have explored the alterations of both DNA and RNA of PLGGs. A strength of this study is the size of the cohorts analyzed. Both the sequencing meta-analysis as well as the gene expression profiling study benefit from a large number of samples. The statistical power to detect rare events across samples is therefore partially reached and enables further discovery of new insights in the biology of PLGGs such as MYB-QKI recurrent alterations in AGs or specific immunologic alterations is the ST PAs. This work has been possible only through multi-institutional efforts. As we understand even more that PLGGs are very heterogeneous tumors, this effort of collecting and profiling more tumors is still needed in order to apprehend better the characteristics of each tumor subtypes.

Another important aspect that is highlighted in this work is the exploration of new paths through innovative techniques to apprehend the biology of PLGGs. The single cell expression profiling in PAs illustrates this effort. This study shows for the first time reliable results of transcriptomic patterns of SCs in PLGGs. These results represent a great opportunity to benefit such technological advances in order to study at a single cell level the biological patterns that are dominant in tumors and driving those cells. We have showed that A2B5 positive cells represent a distinct population in the tumor. The identification that *BRAF* duplication in specifically described in that population of cells would further reinforce the hypothesis that A2B5 positive cells are the “cancer cells” within the tumors. The expression results generated in our experiments would then represent a precious resource to explore the molecular patterns of the PA cancer cells in comparison to the other cells represented in the tumor. Additionally, the biology of A2B5 positive cells needs to be explored further. After FACS sorting A2B5 positive cells, they can be cultured *in vitro* and also be implanted as xenograft in order to explore their potential oncogenic properties. Furthermore, cultured A2B5 positive cells could represent a precious resource to test candidate drugs in order to accelerate the discovery of efficient therapies against AGs. In parallel of functional developments, further efforts are needed to improve the quality and the reliability of the results in order to built strong biological conclusions.

This work underlines also the desperate need to link closely clinical features and genomic profiles of PLGGs in order to provide a better global definition of what is a PLGG. The latest WHO classification of PLGGs is still only based on clinical features of the cells identified in the tumor. The last decade has brought a new dimension in the field: genetic alterations. *BRAF* duplications and point mutation is certainly dominating the landscape of PLGGs alterations. This work is confirming and highlighting the fact that other genetic events such as *FGFR1* duplication and mutation, *MYB/MYBL1* alteration, *IDH1* mutation, *MYB-QKI* rearrangement need to be considered as actors in the physiopathology of those tumors. Based on morphological features, PLGGs are often difficult to classify into a specific diagnosis group. The contribution of genetic features will therefore help clinicians to better define the type of low-grade tumor. Indeed, such as *BRAF* duplication identified in a posterior fossa tumor speaks strongly in favor of a PA, the discovery that *MYB-QKI* alteration is only represented in AG tumors might represent a strong biological feature to define AGs and might be used as a diagnostic test in routine clinic.

Another aspect highlighted by our study is the relative low number of DNA alterations observed in PLGG tumors. Furthermore, those genetic alterations (such as *BRAF-KIAA* duplications, *FGFR1* alterations, *MYB* alterations) have not been described in the literature as transforming cells into a cancer cell, part due to a desperate lack of relevant model. Those few alterations need to be further explored in better models to understand their exact contribution in the physiopathology of PLGGs. Alternatively, the fact that those genetic alterations are yet not defined as clear drivers in *in vitro* and *in vivo* models argues in favor of another level of biological regulation that interplays in the biology of PLGGs. Epigenetic and environmental influences can certainly be considered as important contributors in the physiopathology of PLGGs. Our bulk gene expression study has shown that specific molecular patterns characterize distinct tumor subtypes (histology and location mainly), suggesting that distinct biological perturbations are causing distinct tumor phenotypes. However we were not able to define clear patterns between the different PLGG tumor types, partially due to the mixed cell populations represented in the bulk expression profiling. To that extent, the SC profiling work opens new avenues of research as it allows dissecting the expression patterns of the “tumor cells” (ie presumably A2B5 cells) relative to “non-tumor cells” and thus understanding which molecular perturbations are acting in the transformation of PLGG tumor cells.

References

1. Louis DN, Ohgaki H, Wiestler OD, et al. The 2007 WHO classification of tumours of the central nervous system. Acta neuropathologica 2007;114:97-109.
2. Dolecek TA, Propp JM, Stroup NE, Kruchko C. CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2005-2009. Neuro-oncology 2012;14 Suppl 5:v1-49.
3. Tung JN, Tsao TY, Tai CJ, Yeh KT, Cheng YW, Jiang MC. Distribution of lysosome-associated membrane proteins-1 and -2, and cathepsin D in eosinophilic granular bodies: possible relationship to cyst development in pilocytic astrocytomas. The Journal of international medical research 2010;38:1354-64.
4. Berhouma M, Jemel H, Kchir N. Calcified pilocytic astrocytoma of the medulla mimicking a brainstem "stone". Pathologica 2008;100:408-10.
5. Takei H YS, Wong KK, Mehta V, Chintagumpala M, Dauser RC, Lau CC, Adesina AM. Expression of oligodendroglial differentiation markers in pilocytic astrocytomas identifies two clinical subsets and shows a significant correlation with proliferation index and progression free survival. J Neurooncol 2008;86:183-90.
6. Otero JJ RD, Vandenberg S. OLIG2 is differentially expressed in pediatric astrocytic and in ependymal neoplasms. J Neurooncol 2011;104:423-38.
7. Tanaka Y SA, Ishiuchi S, Nakazato Y. Diversity of glial cell components in pilocytic astrocytoma. Neuropathology 2008 Aug;28(4):399-407 Epub 2008 Feb 26 2008;28:399-407.
8. Ligon KL, Huillard E, Mehta S, et al. Olig2-regulated lineage-restricted pathway controls replication competence in neural stem cells and malignant glioma. Neuron 2007;53:503-17.
9. Gullotta F, Schindler F, Schmutzler R, Weeks-Seifert A. GFAP in brain tumor diagnosis: possibilities and limitations. Pathol Res Pract 1985;180:54-60.
10. Pfister S, Witt O. Pediatric gliomas. Recent results in cancer research Fortschritte der Krebsforschung Progres dans les recherches sur le cancer 2009;171:67-81.
11. Zentner J, Wolf HK, Ostertun B, et al. Gangliogliomas: clinical, radiological, and histopathological findings in 51 patients. J Neurol Neurosurg Psychiatry 1994;57:1497-502.
12. Daumas-Duport C, Scheithauer BW, Chodkiewicz JP, Laws ER, Jr., Vedrenne C. Dysembryoplastic neuroepithelial tumor: a surgically curable tumor of young patients with intractable partial seizures. Report of thirty-nine cases. Neurosurgery 1988;23:545-56.
13. Park JY, Suh YL, Han J. Dysembryoplastic neuroepithelial tumor. Features distinguishing it from oligodendroglioma on cytologic squash preparations. Acta Cytol 2003;47:624-9.
14. Lellouch-Tubiana A, Boddaert N, Bourgeois M, et al. Angiocentric neuroepithelial tumor (ANET): a new epilepsy-related clinicopathological entity with distinctive MRI. Brain pathology 2005;15:281-6.
15. Buccoliero AM, Castiglione F, Degl'innocenti DR, et al. Angiocentric glioma: clinical, morphological, immunohistochemical and molecular features in three pediatric cases. Clin Neuropathol 2013;32:107-13.

16. Forbes JA, Mobley BC, O'Lynnger TM, et al. Pediatric cerebellar pilomyxoid-spectrum astrocytomas. J Neurosurg Pediatr 2011;8:90-6.
17. Amatya VJ, Akazawa R, Sumimoto Y, Takeshima Y, Inai K. Clinicopathological and immunohistochemical features of three pilomyxoid astrocytomas: comparative study with 11 pilocytic astrocytomas. Pathology international 2009;59:80-5.
18. Machen SK, Prayson RA. Cyclin D1 and MIB-1 immunohistochemistry in pilocytic astrocytomas: a study of 48 cases. Human pathology 1998;29:1511-6.
19. Prayson RA, Khajavi K, Comair YG. Cortical architectural abnormalities and MIB1 immunoreactivity in gangliogliomas: a study of 60 patients with intracranial tumors. Journal of neuropathology and experimental neurology 1995;54:513-20.
20. Prayson RA, Morris HH, Estes ML, Comair YG. Dysembryoplastic neuroepithelial tumor: a clinicopathologic and immunohistochemical study of 11 tumors including MIB1 immunoreactivity. Clinical neuropathology 1996;15:47-53.
21. Johannessen AL, Torp SH. The clinical value of Ki-67/MIB-1 labeling index in human astrocytomas. Pathology oncology research : POR 2006;12:143-7.
22. Neder L, Colli BO, Machado HR, Carlotti CG, Jr., Santos AC, Chimelli L. MIB-1 labeling index in astrocytic tumors--a clinicopathologic study. Clinical neuropathology 2004;23:262-70.
23. Roessler K BA, Jezan H, Ba-Ssalamah A, Slavc I, Czech T, Budka H. Proliferative activity as measured by MIB-1 labeling index and long-term outcome of cerebellar juvenile pilocytic astrocytomas. J Neurooncol 2002;58:141-6.
24. Kayaselcuk F, Zorludemir S, Gumurduhu D, Zeren H, Erman T. PCNA and Ki-67 in central nervous system tumors: correlation with the histological type and grade. Journal of neuro-oncology 2002;57:115-21.
25. Paixao Becker A, de Oliveira RS, Saggioro FP, Neder L, Chimelli LM, Machado HR. In pursuit of prognostic factors in children with pilocytic astrocytomas. Child's nervous system : ChNS : official journal of the International Society for Pediatric Neurosurgery 2010;26:19-28.
26. Bowers DC GL, Kapur P, Reisch JS, Mulne AF, Shapiro KN, Elterman RD, Winick NJ, Margraf LR. Study of the MIB-1 labeling index as a predictor of tumor progression in pilocytic astrocytomas in children and adolescents. J Clin Oncol 2003;21:2968-73.
27. Margraf LR GL, Butt Y, Raghunathan N, Bowers DC. Proliferative and metabolic markers in incompletely excised pediatric pilocytic astrocytomas--an assessment of 3 new variables in predicting clinical outcome.
. Neuro Oncol 2011;13:767-74.
28. Cummings TJ, Provenzale JM, Hunter SB, et al. Gliomas of the optic nerve: histological, immunohistochemical (MIB-1 and p53), and MRI analysis. Acta neuropathologica 2000;99:563-70.
29. Faria C, Miguens J, Antunes JL, et al. Genetic alterations in a papillary glioneuronal tumor. J Neurosurg Pediatr 2008;1:99-102.
30. Czech T, Slavc I, Aichholzer M, et al. Proliferative activity as measured by MIB-1 labeling index and long-term outcome of visual pathway astrocytomas in children. J Neurooncol 1999;42:143-50.
31. Dirven CM KJ, Mooij JJ, Molenaar WM. The proliferative potential of the pilocytic astrocytoma: the relation between MIB-1 labeling and clinical and neuro-radiological follow-up. J Neurooncol 1998;37:9-16.

32. Hemmati HD NI, Lazareff JA, Masterman-Smith M, Geschwind DH, Bronner-Fraser M, Kornblum HI. Cancerous stem cells can arise from pediatric brain tumors. Proceedings of the National Academy of Sciences of the United States of America 2003;100:15178-83.
33. CBTRUS. 2012 CBTRUS Statistical Report: Primary Brain and Central Nervous System Tumors Diagnosed in the United States in 2004-20082012.
34. Bleyer WA. Epidemiologic impact of children with brain tumors. Child's nervous system : ChNS : official journal of the International Society for Pediatric Neurosurgery 1999;15:758-63.
35. Kohler BA, Ward E, McCarthy BJ, et al. Annual report to the nation on the status of cancer, 1975-2007, featuring tumors of the brain and other nervous system. Journal of the National Cancer Institute 2011;103:714-36.
36. Kaatsch P, Rickert CH, Kuhl J, Schuz J, Michaelis J. Population-based epidemiologic data on brain tumors in German children. Cancer 2001;92:3155-64.
37. Bauchet L, Rigau V, Mathieu-Daude H, et al. Clinical epidemiology for childhood primary central nervous system tumors. J Neurooncol 2009;92:87-98.
38. Gjerris F, Agerlin N, Borgesen SE, et al. Epidemiology and prognosis in children treated for intracranial tumours in Denmark 1960-1984. Child's nervous system : ChNS : official journal of the International Society for Pediatric Neurosurgery 1998;14:302-11.
39. Pinho RS, Andreoni S, Silva NS, et al. Pediatric central nervous system tumors: a single-center experience from 1989 to 2009. J Pediatr Hematol Oncol 2011;33:605-9.
40. Rickert CH, Paulus W. Epidemiology of central nervous system tumors in childhood and adolescence based on the new WHO classification. Child's nervous system : ChNS : official journal of the International Society for Pediatric Neurosurgery 2001;17:503-11.
41. Radner H, Blumcke I, Reifenberger G, Wiestler OD. [The new WHO classification of tumors of the nervous system 2000. Pathology and genetics]. Pathologe 2002;23:260-83.
42. Duffau H, Capelle L. Preferential brain locations of low-grade gliomas. Cancer 2004;100:2622-6.
43. Sievert AJ, Fisher MJ. Pediatric low-grade gliomas. Journal of child neurology 2009;24:1397-408.
44. Freeman CR, Farmer JP, Montes J. Low-grade astrocytomas in children: evolving management strategies. International journal of radiation oncology, biology, physics 1998;41:979-87.
45. Rosemberg S, Fujiwara D. Epidemiology of pediatric tumors of the nervous system according to the WHO 2000 classification: a report of 1,195 cases from a single institution. Child's nervous system : ChNS : official journal of the International Society for Pediatric Neurosurgery 2005;21:940-4.
46. Peters O, Gnekow AK, Rating D, Wolff JE. Impact of location on outcome in children with low-grade oligodendroglioma. Pediatric blood & cancer 2004;43:250-6.
47. Luyken C, Blumcke I, Fimmers R, Urbach H, Wiestler OD, Schramm J. Supratentorial gangliogliomas: histopathologic grading and tumor recurrence in 184 patients with a median follow-up of 8 years. Cancer 2004;101:146-55.
48. Fernandez C, Figarella-Branger D, Girard N, et al. Pilocytic astrocytomas in children: prognostic factors--a retrospective study of 80 cases. Neurosurgery 2003;53:544-53; discussion 54-5.

49. Scheinemann K, Bartels U, Huang A, et al. Survival and functional outcome of childhood spinal cord low-grade gliomas. Clinical article. J Neurosurg Pediatr 2009;4:254-61.
50. Robertson PL, Allen JC, Abbott IR, Miller DC, Fidel J, Epstein FJ. Cervicomedullary tumors in children: a distinct subset of brainstem gliomas. Neurology 1994;44:1798-803.
51. Rodriguez FJ, Perry A, Gutmann DH, et al. Gliomas in neurofibromatosis type 1: a clinicopathologic study of 100 patients. Journal of neuropathology and experimental neurology 2008;67:240-9.
52. Guillamo JS, Creange A, Kalifa C, et al. Prognostic factors of CNS tumours in Neurofibromatosis 1 (NF1): a retrospective study of 104 patients. Brain : a journal of neurology 2003;126:152-60.
53. Korf BR. Malignancy in neurofibromatosis type 1. The oncologist 2000;5:477-85.
54. Hernaiz Driever P, von Hornstein S, Pietsch T, et al. Natural history and management of low-grade glioma in NF-1 children. Journal of neuro-oncology 2010;100:199-207.
55. Listernick R, Louis DN, Packer RJ, Gutmann DH. Optic pathway gliomas in children with neurofibromatosis 1: consensus statement from the NF1 Optic Pathway Glioma Task Force. Annals of neurology 1997;41:143-9.
56. Blazo MA, Lewis RA, Chintagumpala MM, Frazier M, McCluggage C, Plon SE. Outcomes of systematic screening for optic pathway tumors in children with Neurofibromatosis Type 1. Am J Med Genet A 2004;127A:224-9.
57. Al-Saleem T, Wessner LL, Scheithauer BW, et al. Malignant tumors of the kidney, brain, and soft tissues in children and young adults with the tuberous sclerosis complex. Cancer 1998;83:2208-16.
58. Hirose T, Scheithauer BW, Lopes MB, et al. Tuber and subependymal giant cell astrocytoma associated with tuberous sclerosis: an immunohistochemical, ultrastructural, and immunoelectron and microscopic study. Acta neuropathologica 1995;90:387-99.
59. van Slegtenhorst M, de Hoogt R, Hermans C, et al. Identification of the tuberous sclerosis gene TSC1 on chromosome 9q34. Science 1997;277:805-8.
60. Bilginer B, Narin F, Oguz KK, Uzun S, Soylemezoglu F, Akalan N. Benign cerebellar pilocytic astrocytomas in children. Turkish neurosurgery 2011;21:22-6.
61. Gelabert-Gonzalez M, Amo JM, Arcos Algaba A, et al. [Intracranial gangliogliomas. A review of a series of 20 patients]. Neurologia 2011;26:405-15.
62. Prayson RA. Tumours arising in the setting of paediatric chronic epilepsy. Pathology 2010;42:426-31.
63. Alexiou GA, Varela M, Sfakianos G, Prodromou N. Benign lesions accompanied by intractable epilepsy in children. Journal of child neurology 2009;24:697-700.
64. De Rose M, Luzi M, Trignani R, et al. Cingulate epilepsy in a child with a low-grade glioma. Child's nervous system : ChNS : official journal of the International Society for Pediatric Neurosurgery 2009;25:1507-11.
65. Gaggero R, Consales A, Fazzini F, et al. Epilepsy associated with supratentorial brain tumors under 3 years of life. Epilepsy Res 2009;87:184-9.
66. Huber J, Sovinz P, Lackner H, Mokry M, Eder H, Urban C. Diencephalic syndrome: a frequently delayed diagnosis in failure to thrive. Klin Padiatr 2007;219:91-4.

67. Waga S, Shimizu T, Sakakura M. Diencephalic syndrome of emaciation (Russell's syndrome). Surg Neurol 1982;17:141-6.
68. Perilongo G, Carollo C, Salviati L, et al. Diencephalic syndrome and disseminated juvenile pilocytic astrocytomas of the hypothalamic-optic chiasm region. Cancer 1997;80:142-6.
69. Czyzyk E, Jozwiak S, Roszkowski M, Schwartz RA. Optic pathway gliomas in children with and without neurofibromatosis 1. Journal of child neurology 2003;18:471-8.
70. Rondinelli PI, Osorio CA, Cohen MP, Novaes PE. Unusual dissemination patterns of low-grade astrocytomas in childhood. Arq Neuropsiquiatr 2008;66:45-9.
71. Civitello LA, Packer RJ, Rorke LB, Siegel K, Sutton LN, Schut L. Leptomeningeal dissemination of low-grade gliomas in childhood. Neurology 1988;38:562-6.
72. Perilongo G, Garre ML, Giangaspero F. Low-grade gliomas and leptomeningeal dissemination: a poorly understood phenomenon. Child's nervous system : ChNS : official journal of the International Society for Pediatric Neurosurgery 2003;19:197-203.
73. Hukin J SJ, Cohen H, Velasquez L, Zagzag D, Allen J. Leptomeningeal dissemination at diagnosis of pediatric low-grade neuroepithelial tumors. Neuro Oncol 2003;5:188-96.
74. Mazloom A, Hodges JC, Teh BS, Chintagumpala M, Paulino AC. Outcome of patients with pilocytic astrocytoma and leptomeningeal dissemination. International journal of radiation oncology, biology, physics 2012;84:350-4.
75. von Hornstein S, Kortmann RD, Pietsch T, et al. Impact of chemotherapy on disseminated low-grade glioma in children and adolescents: Report from the HIT-LGG 1996 trial. Pediatric blood & cancer 2011;56:1046-54.
76. Porto L, Kieslich M, Franz K, et al. Spectroscopy of untreated pilocytic astrocytomas: do children and adults share some metabolic features in addition to their morphologic similarities? Child's nervous system : ChNS : official journal of the International Society for Pediatric Neurosurgery 2010;26:801-6.
77. Strong JA HHJ, Brown MT, Debatin JF, Friedman HS, Oakes WJ, Tien R. Pilocytic astrocytoma: correlation between the initial imaging features and clinical aggressiveness. AJR Am J Roentgenol 1993;161:369-72.
78. Lee YY, Van Tassel P, Bruner JM, Moser RP, Share JC. Juvenile pilocytic astrocytomas: CT and MR characteristics. AJR Am J Roentgenol 1989;152:1263-70.
79. Seo HS, Kim JH, Lee DH, et al. Nonenhancing intramedullary astrocytomas and other MR imaging features: a retrospective study and systematic review. AJNR Am J Neuroradiol 2010;31:498-503.
80. Komotar RJ, Zacharia BE, Sughrue ME, et al. Magnetic resonance imaging characteristics of pilomyxoid astrocytoma. Neurol Res 2008;30:945-51.
81. Chang SC, Lai PH, Chen WL, et al. Diffusion-weighted MRI features of brain abscess and cystic or necrotic brain tumors: comparison with conventional MRI. Clin Imaging 2002;26:227-36.
82. Fisher PG, Tihan T, Goldthwaite PT, et al. Outcome analysis of childhood low-grade astrocytomas. Pediatric blood & cancer 2008;51:245-50.
83. Perilongo G, Moras P, Carollo C, et al. Spontaneous partial regression of low-grade glioma in children with neurofibromatosis-1: a real possibility. Journal of child neurology 1999;14:352-6.
84. Rozen WM JS, Lo PA. Spontaneous regression of low-grade gliomas in pediatric patients without neurofibromatosis. Pediatr Neurosurg 2008;44:324-8.

85. Gunny RS, Hayward RD, Phipps KP, Harding BN, Saunders DE. Spontaneous regression of residual low-grade cerebellar pilocytic astrocytomas in children. Pediatric radiology 2005;35:1086-91.
86. Schmandt SM, Packer RJ, Vezina LG, Jane J. Spontaneous regression of low-grade astrocytomas in childhood. Pediatr Neurosurg 2000;32:132-6.
87. Kernan JC, Horgan MA, Piatt JH, D'Agostino A. Spontaneous involution of a diencephalic astrocytoma. Pediatr Neurosurg 1998;29:149-53.
88. Kornreich L, Blaser S, Schwarz M, et al. Optic pathway glioma: correlation of imaging findings with the presence of neurofibromatosis. AJNR Am J Neuroradiol 2001;22:1963-9.
89. Pollack IF, Shultz B, Mulvihill JJ. The management of brainstem gliomas in patients with neurofibromatosis 1. Neurology 1996;46:1652-60.
90. Milstein JM, Geyer JR, Berger MS, Bleyer WA. Favorable prognosis for brainstem gliomas in neurofibromatosis. J Neurooncol 1989;7:367-71.
91. Stokland T, Liu JF, Ironside JW, et al. A multivariate analysis of factors determining tumor progression in childhood low-grade glioma: a population-based cohort study (CCLG CNS9702). Neuro-oncology 2010;12:1257-68.
92. Pollack IF, Claassen D, al-Shboul Q, Janosky JE, Deutsch M. Low-grade gliomas of the cerebral hemispheres in children: an analysis of 71 cases. Journal of neurosurgery 1995;82:536-47.
93. Fouladi M, Hunt DL, Pollack IF, et al. Outcome of children with centrally reviewed low-grade gliomas treated with chemotherapy with or without radiotherapy on Children's Cancer Group high-grade glioma study CCG-945. Cancer 2003;98:1243-52.
94. Smoots DW, Geyer JR, Lieberman DM, Berger MS. Predicting disease progression in childhood cerebellar astrocytoma. Child's nervous system : ChNS : official journal of the International Society for Pediatric Neurosurgery 1998;14:636-48.
95. El Khashab M GL, Margraf L, Koral K, Nejat F, Swift D, Weprin B, Bowers DC. Predictors of tumor progression among children with gangliogliomas. Clinical article. J Neurosurg Pediatr 2009;3:461-6.
96. Bowers DC, Mulne AF, Weprin B, Bruce DA, Shapiro K, Margraf LR. Prognostic factors in children and adolescents with low-grade oligodendrogliomas. Pediatr Neurosurg 2002;37:57-63.
97. Opocher E, Kremer LC, Da Dalt L, et al. Prognostic factors for progression of childhood optic pathway glioma: a systematic review. European journal of cancer 2006;42:1807-16.
98. Ater JL, Zhou T, Holmes E, et al. Randomized study of two chemotherapy regimens for treatment of low-grade glioma in young children: a report from the Children's Oncology Group. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2012;30:2641-7.
99. Opocher E, Kremer LC, Da Dalt L, et al. Prognostic factors for progression of childhood optic pathway glioma: a systematic review. Eur J Cancer 2006;42:1807-16.
100. Bandopadhayay P, Bergthold G, London WB, et al. Long-term outcome of 4,040 children diagnosed with pediatric low-grade gliomas: An analysis of the Surveillance Epidemiology and End Results (SEER) database. Pediatric blood & cancer 2014.
101. Parsa CF GS. Juvenile pilocytic astrocytomas do not undergo spontaneous malignant transformation: grounds for designation as hamartomas. Br J Ophthalmol 2008;92:40-6.

102. Zoeller GK, Brathwaite CD, Sandberg DI. Malignant transformation of an optic pathway glioma without prior radiation therapy. J Neurosurg Pediatr 2010;5:507-10.
103. Unal E, Koksal Y, Cimen O, Paksoy Y, Tavli L. Malignant glioblastomatous transformation of a low-grade glioma in a child. Child's nervous system : ChNS : official journal of the International Society for Pediatric Neurosurgery 2008;24:1385-9.
104. Qaddoumi I, Sultan I, Gajjar A. Outcome and prognostic features in pediatric gliomas: a review of 6212 cases from the Surveillance, Epidemiology, and End Results database. Cancer 2009;115:5761-70.
105. Wisoff JH, Sanford RA, Heier LA, et al. Primary neurosurgery for pediatric low-grade gliomas: a prospective multi-institutional study from the Children's Oncology Group. Neurosurgery 2011;68:1548-54; discussion 54-5.
106. Benesch M EH, Sovinz P, Raith J, Lackner H, Moser A, Urban C. Residual or recurrent cerebellar low-grade glioma in children after tumor resection: is re-treatment needed? A single center experience from 1983 to 2003. Pediatr Neurosurg 2006;42:159-64.
107. Laws ER, Jr., Taylor WF, Clifton MB, Okazaki H. Neurosurgical management of low-grade astrocytoma of the cerebral hemispheres. Journal of neurosurgery 1984;61:665-73.
108. Uliel-Sibony S, Kramer U, Fried I, Fattal-Valevski A, Constantini S. Pediatric temporal low-grade glial tumors: epilepsy outcome following resection in 48 children. Child's nervous system : ChNS : official journal of the International Society for Pediatric Neurosurgery 2011;27:1413-8.
109. Teo C, Siu TL. Radical resection of focal brainstem gliomas: is it worth doing? Child's nervous system : ChNS : official journal of the International Society for Pediatric Neurosurgery 2008;24:1307-14.
110. Ramina R CNM, Fernandes YB, Borges G, Honorato DC, Arruda WO. Intrinsic tectal low grade astrocytomas: is surgical removal an alternative treatment? Long-term outcome of eight cases. Arq Neuropsiquiatr 2005;63:40-5.
111. Moghrabi A, Friedman HS, Burger PC, Tien R, Oakes WJ. Carboplatin treatment of progressive optic pathway gliomas to delay radiotherapy. Journal of neurosurgery 1993;79:223-7.
112. Aquino VM, Fort DW, Kamen BA. Carboplatin for the treatment of children with newly diagnosed optic chiasm gliomas: a phase II study. J Neurooncol 1999;41:255-9.
113. Gururangan S, Cavazos CM, Ashley D, et al. Phase II study of carboplatin in children with progressive low-grade gliomas. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2002;20:2951-8.
114. Hsu TR, Wong TT, Chang FC, et al. Responsiveness of progressive optic pathway tumors to cisplatin-based chemotherapy in children. Child's nervous system : ChNS : official journal of the International Society for Pediatric Neurosurgery 2008;24:1457-61.
115. Beaty O, 3rd, Berg S, Blaney S, et al. A phase II trial and pharmacokinetic study of oxaliplatin in children with refractory solid tumors: a Children's Oncology Group study. Pediatric blood & cancer 2010;55:440-5.
116. Friedman HS, Krischer JP, Burger P, et al. Treatment of children with progressive or recurrent brain tumors with carboplatin or iproplatin: a Pediatric Oncology Group randomized phase II study. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 1992;10:249-56.

117. Packer RJ, Lange B, Ater J, et al. Carboplatin and vincristine for recurrent and newly diagnosed low-grade gliomas of childhood. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 1993;11:850-6.
118. Packer RJ, Ater J, Allen J, et al. Carboplatin and vincristine chemotherapy for children with newly diagnosed progressive low-grade gliomas. Journal of neurosurgery 1997;86:747-54.
119. Demaerel P, de Ruyter N, Casteels I, Renard M, Uyttebroeck A, van Gool S. Visual pathway glioma in children treated with chemotherapy. Eur J Paediatr Neurol 2002;6:207-12.
120. Gnekow AK KR, Pietsch T, Emser A. Low grade chiasmatic-hypothalamic glioma-carboplatin and vincristin chemotherapy effectively defers radiotherapy within a comprehensive treatment strategy -- report from the multicenter treatment study for children and adolescents with a low grade glioma -- HIT-LGG 1996 -- of the Society of Pediatric Oncology and Hematology (GPOH). Klin Padiatr 2004;216:331-42.
121. Ronghe M, Hargrave D, Bartels U, et al. Vincristine and carboplatin chemotherapy for unresectable and/or recurrent low-grade astrocytoma of the brainstem. Pediatric blood & cancer 2010;55:471-7.
122. Sawamura Y, Kamoshima Y, Kato T, Tajima T, Tsubaki J. Chemotherapy with cisplatin and vincristine for optic pathway/hypothalamic astrocytoma in young children. Jpn J Clin Oncol 2009;39:277-83.
123. Massimino M, Spreafico F, Cefalo G, et al. High response rate to cisplatin/etoposide regimen in childhood low-grade glioma. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2002;20:4209-16.
124. Massimino M, Spreafico F, Riva D, et al. A lower-dose, lower-toxicity cisplatin-etoposide regimen for childhood progressive low-grade glioma. J Neurooncol 2010;100:65-71.
125. Ater JL, Zhou T, Holmes E, et al. Randomized study of two chemotherapy regimens for treatment of low-grade glioma in young children: a report from the Children's Oncology Group. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2012;30:2641-7.
126. Castello MA SA, Padula A, Varrasso G, Properzi E, Trasimeni G, Operamolla P, Gualdi GF, Clerico A. Does chemotherapy have a role in low-grade astrocytoma management? A report of 13 cases. Med Pediatr Oncol 1995;25:102-8.
127. Yu DY, Dahl GV, Shames RS, Fisher PG. Weekly dosing of carboplatin increases risk of allergy in children. J Pediatr Hematol Oncol 2001;23:349-52.
128. Lazzareschi I, Ruggiero A, Riccardi R, Attina G, Colosimo C, Lasorella A. Hypersensitivity reactions to carboplatin in children. J Neurooncol 2002;58:33-7.
129. Lafay-Cousin L, Sung L, Carret AS, et al. Carboplatin hypersensitivity reaction in pediatric patients with low-grade glioma: a Canadian Pediatric Brain Tumor Consortium experience. Cancer 2008;112:892-9.
130. Lancaster DL, Hoddes JA, Michalski A. Tolerance of nitrosurea-based multiagent chemotherapy regime for low-grade pediatric gliomas. J Neurooncol 2003;63:289-94.
131. Prados MD, Edwards MS, Rabbitt J, Lamborn K, Davis RL, Levin VA. Treatment of pediatric low-grade gliomas with a nitrosourea-based multiagent chemotherapy regimen. Journal of neuro-oncology 1997;32:235-41.
132. Mishra KK, Squire S, Lamborn K, et al. Phase II TPDCV protocol for pediatric low-grade hypothalamic/chiasmatic gliomas: 15-year update. J Neurooncol 2010;100:121-7.

133. Walter AW, Gajjar A, Reardon DA, et al. Tamoxifen and carboplatin for children with low-grade gliomas: a pilot study at St. Jude Children's Research Hospital. J Pediatr Hematol Oncol 2000;22:247-51.
134. Jakacki RI, Bouffet E, Adamson PC, et al. A phase 1 study of vinblastine in combination with carboplatin for children with low-grade gliomas: a Children's Oncology Group phase 1 consortium study. Neuro-oncology 2011;13:910-5.
135. Laithier V, Grill J, Le Deley MC, et al. Progression-free survival in children with optic pathway tumors: dependence on age and the quality of the response to chemotherapy--results of the first French prospective study for the French Society of Pediatric Oncology. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2003;21:4572-8.
136. Bruggers CS, Greene D. A phase 2 feasibility study of sequential, dose intensive chemotherapy to treat progressive low-grade gliomas in children. J Pediatr Hematol Oncol 2007;29:602-7.
137. Lee MJ, Ra YS, Park JB, et al. Effectiveness of novel combination chemotherapy, consisting of 5-fluorouracil, vincristine, cyclophosphamide and etoposide, in the treatment of low-grade gliomas in children. J Neurooncol 2006;80:277-84.
138. Scheinemann K, Bartels U, Tsangaris E, et al. Feasibility and efficacy of repeated chemotherapy for progressive pediatric low-grade gliomas. Pediatric blood & cancer 2010.
139. Nicholson HS, Kretschmar CS, Krailo M, et al. Phase 2 study of temozolomide in children and adolescents with recurrent central nervous system tumors: a report from the Children's Oncology Group. Cancer 2007;110:1542-50.
140. Gururangan S, Fisher MJ, Allen JC, et al. Temozolomide in children with progressive low-grade glioma. Neuro-oncology 2007;9:161-8.
141. Khaw SL, Coleman LT, Downie PA, Heath JA, Ashley DM. Temozolomide in pediatric low-grade glioma. Pediatric blood & cancer 2007;49:808-11.
142. Bartels U BS, Carret AS, Crooks B, Hukin J, Johnston D, Silva M, Strother D, Wilson B, Zelcer S, Eisenstat D, Sung L, Bouffet E. The use and effectiveness of temozolomide in children with central nervous system tumours: a survey from the Canadian Paediatric Brain Tumour Consortium. Curr Oncol 2011;18:e19-24.
143. Lafay-Cousin L, Holm S, Qaddoumi I, et al. Weekly vinblastine in pediatric low-grade glioma patients with carboplatin allergic reaction. Cancer 2005;103:2636-42.
144. Bouffet E, Jakacki R, Goldman S, et al. Phase II study of weekly vinblastine in recurrent or refractory pediatric low-grade glioma. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2012;30:1358-63.
145. Singh G, Wei XC, Hader W, Chan JA, Bouffet E, Lafay-Cousin L. Sustained Response to Weekly Vinblastine in 2 Children With Pilomyxoid Astrocytoma Associated With Diencephalic Syndrome. J Pediatr Hematol Oncol 2012.
146. Kadota RP, Kun LE, Langston JW, et al. Cyclophosphamide for the treatment of progressive low-grade astrocytoma: a Pediatric Oncology Group phase II Study. J Pediatr Hematol Oncol 1999;21:198-202.
147. Pons MA, Finlay JL, Walker RW, Puccetti D, Packer RJ, McElwain M. Chemotherapy with vincristine (VCR) and etoposide (VP-16) in children with low-grade astrocytoma. J Neurooncol 1992;14:151-8.

148. Sumer T, Freeman AI, Cohen M, Bremer AM, Thomas PR, Sinks LF. Chemotherapy in recurrent noncystic low-grade astrocytomas of the cerebrum in children. J Surg Oncol 1978;10:45-54.
149. Rosenstock JG, Evans AE, Schut L. Response to vincristine of recurrent brain tumors in children. Journal of neurosurgery 1976;45:135-40.
150. Sutton LN, Molloy PT, Sernyak H, et al. Long-term outcome of hypothalamic/chiasmatic astrocytomas in children treated with conservative surgery. Journal of neurosurgery 1995;83:583-9.
151. Heideman RL, Douglass EC, Langston JA, et al. A phase II study of every other day high-dose ifosfamide in pediatric brain tumors: a Pediatric Oncology Group Study. J Neurooncol 1995;25:77-84.
152. McCowage G, Tien R, McLendon R, et al. Successful treatment of childhood pilocytic astrocytomas metastatic to the leptomeninges with high-dose cyclophosphamide. Med Pediatr Oncol 1996;27:32-9.
153. Disabato JA HM, Strain JD, Fleitz JM, Foreman NK. Successful use of intracavitary bleomycin for low-grade astrocytoma tumor cyst. Pediatr Neurosurg 1999;31:246-50.
154. Blaney SM, Phillips PC, Packer RJ, et al. Phase II evaluation of topotecan for pediatric central nervous system tumors. Cancer 1996;78:527-31.
155. Dreyer ZE, Kadota RP, Stewart CF, et al. Phase 2 study of idarubicin in pediatric brain tumors: Pediatric Oncology Group study POG 9237. Neuro-oncology 2003;5:261-7.
156. Warren KE, Goldman S, Pollack IF, et al. Phase I trial of lenalidomide in pediatric patients with recurrent, refractory, or progressive primary CNS tumors: Pediatric Brain Tumor Consortium study PBTC-018. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2011;29:324-9.
157. Packer RJ, Jakacki R, Horn M, et al. Objective response of multiply recurrent low-grade gliomas to bevacizumab and irinotecan. Pediatric blood & cancer 2009;52:791-5.
158. Couec ML, Andre N, Thebaud E, et al. Bevacizumab and irinotecan in children with recurrent or refractory brain tumors: toxicity and efficacy trends. Pediatric blood & cancer 2012;59:34-8.
159. Gururangan S, Fangusaro J, Poussaint TY, et al. Efficacy of bevacizumab plus irinotecan in children with recurrent low-grade gliomas--a Pediatric Brain Tumor Consortium study. Neuro-oncology 2013.
160. Grill J, Couanet D, Cappelli C, et al. Radiation-induced cerebral vasculopathy in children with neurofibromatosis and optic pathway glioma. Ann Neurol 1999;45:393-6.
161. Armstrong GT, Conklin HM, Huang S, et al. Survival and long-term health and cognitive outcomes after low-grade glioma. Neuro-oncology 2011;13:223-34.
162. Dirks PB, Jay V, Becker LE, et al. Development of anaplastic changes in low-grade astrocytomas of childhood. Neurosurgery 1994;34:68-78.
163. Collet-Solberg PF, Sernyak H, Satin-Smith M, et al. Endocrine outcome in long-term survivors of low-grade hypothalamic/chiasmatic glioma. Clin Endocrinol (Oxf) 1997;47:79-85.
164. Merchant TE, Kun LE, Wu S, Xiong X, Sanford RA, Boop FA. Phase II trial of conformal radiation therapy for pediatric low-grade glioma. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2009;27:3598-604.

165. Merchant TE, Conklin HM, Wu S, Lustig RH, Xiong X. Late effects of conformal radiation therapy for pediatric patients with low-grade glioma: prospective evaluation of cognitive, endocrine, and hearing deficits. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2009;27:3691-7.
166. Nishihori T, Shirato H, Aoyama H, et al. Three-dimensional conformal radiotherapy for astrocytic tumors involving the eloquent area in children and young adults. J Neurooncol 2002;60:177-83.
167. Saran FH, Baumert BG, Khoo VS, et al. Stereotactically guided conformal radiotherapy for progressive low-grade gliomas of childhood. International journal of radiation oncology, biology, physics 2002;53:43-51.
168. Marcus KJ, Goumnerova L, Billett AL, et al. Stereotactic radiotherapy for localized low-grade gliomas in children: final results of a prospective trial. International journal of radiation oncology, biology, physics 2005;61:374-9.
169. Roberge D SL, Olivier A, Leblanc R, Podgorsak E. Hypofractionated stereotactic radiotherapy for low grade glioma at McGill University: long-term follow-up. Technol Cancer Res Treat 2006;5:1-8.
170. Henderson MA, Fakiris AJ, Timmerman RD, Worth RM, Lo SS, Witt TC. Gamma knife stereotactic radiosurgery for low-grade astrocytomas. Stereotact Funct Neurosurg 2009;87:161-7.
171. Wang LW, Shiau CY, Chung WY, et al. Gamma Knife surgery for low-grade astrocytomas: evaluation of long-term outcome based on a 10-year experience. Journal of neurosurgery 2006;105 Suppl:127-32.
172. Kida Y, Kobayashi T, Mori Y. Gamma knife radiosurgery for low-grade astrocytomas: results of long-term follow up. Journal of neurosurgery 2000;93 Suppl 3:42-6.
173. Hug EB, Muenter MW, Archambeau JO, et al. Conformal proton radiation therapy for pediatric low-grade astrocytomas. Strahlenther Onkol 2002;178:10-7.
174. Kuhlthau KA, Pulsifer MB, Yeap BY, et al. Prospective study of health-related quality of life for children with brain tumors treated with proton radiotherapy. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2012;30:2079-86.
175. Cagnetta A, Adamia S, Acharya C, et al. Role of genotype-based approach in the clinical management of adult acute myeloid leukemia with normal cytogenetics. Leukemia research 2014;38:649-59.
176. Wassie EA, Itzykson R, Lasho TL, et al. Molecular and prognostic correlates of cytogenetic abnormalities in chronic myelomonocytic leukemia: a Mayo Clinic-French Consortium Study. American journal of hematology 2014.
177. Pearson PL. Historical development of analysing large-scale changes in the human genome. Cytogenetic and genome research 2006;115:198-204.
178. Beroukhim R, Getz G, Nghiemphu L, et al. Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. Proceedings of the National Academy of Sciences of the United States of America 2007;104:20007-12.
179. Beroukhim R, Mermel CH, Porter D, et al. The landscape of somatic copy-number alteration across human cancers. Nature 2010;463:899-905.
180. Northcott PA, Shih DJ, Peacock J, et al. Subgroup-specific structural variation across 1,000 medulloblastoma genomes. Nature 2012;488:49-56.

181. Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. Nat Rev Genet 2010;11:685-96.
182. Strausberg RL, Levy S, Rogers YH. Emerging DNA sequencing technologies for human genomic medicine. Drug discovery today 2008;13:569-77.
183. Kim DW, Nam SH, Kim RN, Choi SH, Park HS. Whole human exome capture for high-throughput sequencing. Genome / National Research Council Canada = Genome / Conseil national de recherches Canada 2010;53:568-74.
184. Mamanova L, Coffey AJ, Scott CE, et al. Target-enrichment strategies for next-generation sequencing. Nature methods 2010;7:111-8.
185. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25:2078-9.
186. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research 2010;20:1297-303.
187. Raphael BJ. Chapter 6: Structural variation and medical genomics. PLoS computational biology 2012;8:e1002821.
188. Schatz MC, Delcher AL, Salzberg SL. Assembly of large genomes using second-generation sequencing. Genome research 2010;20:1165-73.
189. Li H, Ruan J, Durbin R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome research 2008;18:1851-8.
190. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754-60.
191. Homer N, Merriman B, Nelson SF. BFAST: an alignment tool for large scale genome resequencing. PloS one 2009;4:e7767.
192. Hudson AM, Yates T, Li Y, et al. Discrepancies in Cancer Genomic Sequencing Highlight Opportunities for Driver Mutation Discovery. Cancer research 2014.
193. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nature biotechnology 2013;31:213-9.
194. Drier Y, Lawrence MS, Carter SL, et al. Somatic rearrangements across cancer reveal classes of samples with distinct patterns of DNA breakage and rearrangement-induced hypermutability. Genome research 2013;23:228-35.
195. Lawrence MS, Stojanov P, Mermel CH, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature 2014.
196. Liu X, Wang J, Chen L. Whole-exome sequencing reveals recurrent somatic mutation networks in cancer. Cancer letters 2013;340:270-6.
197. Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature 2013;499:214-8.
198. Cancer Genome Atlas Research N. Integrated genomic analyses of ovarian carcinoma. Nature 2011;474:609-15.
199. Pavlidis P, Wapinski I, Noble WS. Support vector machine classification on the web. Bioinformatics 2004;20:586-7.
200. Li M, Reilly C. Assessing the quality of hybridized RNA in Affymetrix GeneChips using linear regression. Journal of biomolecular techniques : JBT 2008;19:122-8.
201. Rapaport F, Zinovyev A, Dutreix M, Barillot E, Vert JP. Classification of microarray data using gene networks. BMC bioinformatics 2007;8:35.

202. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proceedings of the National Academy of Sciences of the United States of America 1998;95:14863-8.
203. Tavazoie S, Hughes JD, Campbell MJ, Cho RJ, Church GM. Systematic determination of genetic network architecture. Nature genetics 1999;22:281-5.
204. McLachlan GJ, Flack LK, Ng SK, Wang K. Clustering of gene expression data via normal mixture models. Methods in molecular biology 2013;972:103-19.
205. Smolkin M, Ghosh D. Cluster stability scores for microarray data in cancer studies. BMC bioinformatics 2003;4:36.
206. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America 2005;102:15545-50.
207. Abate F, Zairis S, Ficarra E, et al. Pegasus: a comprehensive annotation and prediction tool for detection of driver gene fusions in cancer. BMC systems biology 2014;8:97.
208. Torres-Garcia W, Zheng S, Sivachenko A, et al. PRADA: pipeline for RNA sequencing data analysis. Bioinformatics 2014;30:2224-6.
209. Kalari KR, Nair AA, Bhavsar JD, et al. MAP-RSeq: Mayo Analysis Pipeline for RNA sequencing. BMC bioinformatics 2014;15:224.
210. Finotello F, Di Camillo B. Measuring differential gene expression with RNA-seq: challenges and strategies for data analysis. Briefings in functional genomics 2014.
211. Bhattacharjee MB, Armstrong DD, Vogel H, Cooley LD. Cytogenetic analysis of 120 primary pediatric brain tumors and literature review. Cancer genetics and cytogenetics 1997;97:39-53.
212. Bigner SH, McLendon RE, Fuchs H, McKeever PE, Friedman HS. Chromosomal characteristics of childhood brain tumors. Cancer genetics and cytogenetics 1997;97:125-34.
213. Rey JA, Bello MJ, de Campos JM, Kusak ME, Moreno S. Chromosomal composition of a series of 22 human low-grade gliomas. Cancer genetics and cytogenetics 1987;29:223-37.
214. Griffin CA, Long PP, Carson BS, Brem H. Chromosome abnormalities in low-grade central nervous system tumors. Cancer genetics and cytogenetics 1992;60:67-73.
215. Neumann E KD, Norman MG, Steinbok P, Cochrane DD, Goddard K. Cytogenetic analysis of 109 pediatric central nervous system tumors. Cancer genetics and cytogenetics 1993;71:40-9.
216. White FV, Anthony DC, Yunis EJ, Tarbell NJ, Scott RM, Schofield DE. Nonrandom chromosomal gains in pilocytic astrocytomas of childhood. Human pathology 1995;26:979-86.
217. Zattara-Cannoni H GD, Lena G, Dufour H, Choux M, Grisoli F, Vagner-Capodano AM. Are juvenile pilocytic astrocytomas benign tumors? A cytogenetic study in 24 cases. Cancer Genet Cytogenet 1998;104:157-60.
218. Karnes PS, Tran TN, Cui MY, et al. Cytogenetic analysis of 39 pediatric central nervous system tumors. Cancer genetics and cytogenetics 1992;59:12-9.
219. Neumann E, Kalousek DK, Norman MG, Steinbok P, Cochrane DD, Goddard K. Cytogenetic analysis of 109 pediatric central nervous system tumors. Cancer genetics and cytogenetics 1993;71:40-9.

220. Willert JR, Daneshvar L, Sheffield VC, Cogen PH. Deletion of chromosome arm 17p DNA sequences in pediatric high-grade and juvenile pilocytic astrocytomas. Genes Chromosomes Cancer 1995;12:165-72.
221. Sawyer JR, Thomas JR, Teo C. Low-grade astrocytoma with a complex four-breakpoint inversion of chromosome 8 as the sole cytogenetic aberration. Cancer genetics and cytogenetics 1995;83:168-71.
222. Orr LC, Fleitz J, McGavran L, Wyatt-Ashmead J, Handler M, Foreman NK. Cytogenetics in pediatric low-grade astrocytomas. Med Pediatr Oncol 2002;38:173-7.
223. Pfister S, Janzarik WG, Remke M, et al. BRAF gene duplication constitutes a mechanism of MAPK pathway activation in low-grade astrocytomas. The Journal of clinical investigation 2008;118:1739-49.
224. Bar EE LA, Tihan T, Burger PC, Eberhart CG. Frequent gains at chromosome 7q34 involving BRAF in pilocytic astrocytoma. J Neuropathol Exp Neurol 2008;67:878-87.
225. Jacob K AS, Sollier C, Faury D, Sader E, Montpetit A, Serre D, Hauser P, Garami M, Bognar L, Hanzely Z, Montes JL, Atkinson J, Farmer JP, Bouffet E, Hawkins C, Tabori U, Jabado N. Duplication of 7q34 is specific to juvenile pilocytic astrocytomas and a hallmark of cerebellar and optic pathway tumours. British journal of cancer 2009;101:722-33.
226. Sievert AJ, Jackson EM, Gai X, et al. Duplication of 7q34 in pediatric low-grade astrocytomas detected by high-density single-nucleotide polymorphism-based genotype arrays results in a novel BRAF fusion gene. Brain pathology 2009;19:449-58.
227. Ramkissoon LA, Horowitz PM, Craig JM, et al. Genomic analysis of diffuse pediatric low-grade gliomas identifies recurrent oncogenic truncating rearrangements in the transcription factor MYBL1. Proceedings of the National Academy of Sciences of the United States of America 2013;110:8188-93.
228. Jones DT, Hutter B, Jager N, et al. Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. Nature genetics 2013;45:927-32.
229. Zhang J, Wu G, Miller CP, et al. Whole-genome sequencing identifies genetic alterations in pediatric low-grade gliomas. Nature genetics 2013;45:602-12.
230. Ballester R, Marchuk D, Boguski M, et al. The NF1 locus encodes a protein functionally related to mammalian GAP and yeast IRA proteins. Cell 1990;63:851-9.
231. Yunoue S, Tokuo H, Fukunaga K, et al. Neurofibromatosis type I tumor suppressor neurofibromin regulates neuronal differentiation via its GTPase-activating protein function toward Ras. The Journal of biological chemistry 2003;278:26958-69.
232. Deshmukh H, Yeh TH, Yu J, et al. High-resolution, dual-platform aCGH analysis reveals frequent HIPK2 amplification and increased expression in pilocytic astrocytomas. Oncogene 2008;27:4745-51.
233. Hoischen A, Ehrler M, Fassunke J, et al. Comprehensive characterization of genomic aberrations in gangliogliomas by CGH, array-based CGH and interphase FISH. Brain pathology 2008;18:326-37.
234. Forshew T, Tatevossian RG, Lawson AR, et al. Activation of the ERK/MAPK pathway: a signature genetic defect in posterior fossa pilocytic astrocytomas. J Pathol 2009;218:172-81.
235. Lin A, Rodriguez FJ, Karajannis MA, et al. BRAF alterations in primary glial and glioneuronal neoplasms of the central nervous system with identification of 2 novel

KIAA1549:BRAF fusion variants. Journal of neuropathology and experimental neurology 2012;71:66-72.
236. Jones DT KS, Liu L, Pearson DM, Ichimura K, Collins VP. Oncogenic RAF1 rearrangement and a novel BRAF mutation as alternatives to KIAA1549:BRAF fusion in activating the MAPK pathway in pilocytic astrocytoma. Oncogene 2009;28:2119-23.
237. Cin H, Meyer C, Herr R, et al. Oncogenic FAM131B-BRAF fusion resulting from 7q34 deletion comprises an alternative mechanism of MAPK pathway activation in pilocytic astrocytoma. Acta neuropathologica 2011;121:763-74.
238. Lawson AR, Hindley GF, Forshew T, et al. RAF gene fusion breakpoints in pediatric brain tumors are characterized by significant enrichment of sequence microhomology. Genome research 2011;21:505-14.
239. Jones DT, Kocialkowski S, Liu L, et al. Tandem duplication producing a novel oncogenic BRAF fusion gene defines the majority of pilocytic astrocytomas. Cancer research 2008;68:8673-7.
240. Maldonado JL, Fridlyand J, Patel H, et al. Determinants of BRAF mutations in primary melanomas. Journal of the National Cancer Institute 2003;95:1878-90.
241. Li WQ, Kawakami K, Ruszkiewicz A, Bennett G, Moore J, Iacopetta B. BRAF mutations are associated with distinctive clinical, pathological and molecular features of colorectal cancer independently of microsatellite instability status. Mol Cancer 2006;5:2.
242. Takahashi Y, Mori J, Kami M. BRAF mutations in hairy-cell leukemia. The New England journal of medicine 2011;365:960-1; author reply 1-2.
243. Yeo YH, Byrne NP, Counelis GJ, Perry A. Adult with cerebellar anaplastic pilocytic astrocytoma associated with BRAF V600E mutation and p16 loss. Clinical neuropathology 2013;32:159-64.
244. Schiffman JD HJ, VandenBerg SR, Flaherty P, Polley MY, Yu M, Fisher PG, Rowitch DH, Ford JM, Berger MS, Ji H, Gutmann DH, James CD. Oncogenic BRAF mutation with CDKN2A inactivation is characteristic of a subset of pediatric malignant astrocytomas. Cancer research 2010;70:512-9.
245. Dougherty MJ, Santi M, Brose MS, et al. Activating mutations in BRAF characterize a spectrum of pediatric low-grade gliomas. Neuro-oncology 2010;12:621-30.
246. Dias-Santagata D, Lam Q, Vernovsky K, et al. BRAF V600E mutations are common in pleomorphic xanthoastrocytoma: diagnostic and therapeutic implications. PloS one 2011;6:e17948.
247. Schindler G, Capper D, Meyer J, et al. Analysis of BRAF V600E mutation in 1,320 nervous system tumors reveals high mutation frequencies in pleomorphic xanthoastrocytoma, ganglioglioma and extra-cerebellar pilocytic astrocytoma. Acta neuropathologica 2011;121:397-405.
248. MacConaill LE, Campbell CD, Kehoe SM, et al. Profiling critical cancer gene mutations in clinical tumor samples. PloS one 2009;4:e7887.
249. Craig Horbinski MNN, Jill M. Hagenkord, Ronald L. Hamilton,, Pollack aIF. Interplay among BRAF, p16, p53, and MIB1 in pediatric low-grade gliomas. Neuro Oncol 2012.
250. Raabe EH LK, Kim JM, Meeker A, Mao XG, Nikkhah G, Maciaczyk J, Kahlert U, Jain D, Bar E, Cohen KJ, Eberhart CG. BRAF activation induces transformation and then senescence in human neural stem cells: a pilocytic astrocytoma model. Clinical

cancer research : an official journal of the American Association for Cancer Research 2011;17:3590-9.
251. Sharma MK, Zehnbauer BA, Watson MA, Gutmann DH. RAS pathway activation and an oncogenic RAS mutation in sporadic pilocytic astrocytoma. Neurology 2005;65:1335-6.
252. Janzarik WG KC, Loges NT, Olbrich H, Klein C, Schäfer T, Scheurlen W, Roggendorf W, Weiller C, Niemeyer C, Korinthenberg R, Pfister S, Omran H. Further evidence for a somatic KRAS mutation in a pilocytic astrocytoma. Neuropediatrics 2007;38:61-3.
253. Kaul A, Chen YH, Emnett RJ, Dahiya S, Gutmann DH. Pediatric glioma-associated KIAA1549:BRAF expression regulates neuroglial cell growth in a cell type-specific and mTOR-dependent manner. Genes & development 2012;26:2561-6.
254. Li X, Newbern JM, Wu Y, et al. MEK Is a Key Regulator of Gliogenesis in the Developing Brain. Neuron 2012;75:1035-50.
255. Tabori U, Rienstein S, Dromi Y, et al. Epidermal growth factor receptor gene amplification and expression in disseminated pediatric low-grade gliomas. Journal of neurosurgery 2005;103:357-61.
256. Berman DM, Karhadkar SS, Hallahan AR, et al. Medulloblastoma growth inhibition by hedgehog pathway blockade. Science 2002;297:1559-61.
257. Sarangi A, Valadez JG, Rush S, Abel TW, Thompson RC, Cooper MK. Targeted inhibition of the Hedgehog pathway in established malignant glioma xenografts enhances survival. Oncogene 2009;28:3468-76.
258. Rush SZ, Abel TW, Valadez JG, Pearson M, Cooper MK. Activation of the Hedgehog pathway in pilocytic astrocytomas. Neuro-oncology 2010;12:790-8.
259. Sie M, de Bont ES, Scherpen FJ, Hoving EW, den Dunnen WF. Tumour vasculature and angiogenic profile of paediatric pilocytic astrocytoma; is it much different from glioblastoma? Neuropathology and applied neurobiology 2010;36:636-47.
260. Tatevossian RG, Tang B, Dalton J, et al. MYB upregulation and genetic aberrations in a subset of pediatric low-grade gliomas. Acta neuropathologica 2010;120:731-43.
261. Clappier E CW, Kalota A, Crinquette A, Cayuela JM, Dik WA, Langerak AW, Montpellier B, Nadel B, Walrafen P, Delattre O, Aurias A, Leblanc T, Dombret H, Gewirtz AM, Baruchel A, Sigaux F, Soulier J. The C-MYB locus is involved in chromosomal translocation and genomic duplications in human T-cell acute leukemia (T-ALL), the translocation defining a new T-ALL subtype in very young children. Blood 2007 ;110:1251-61.
262. Lahortiga I, De Keersmaecker K, Van Vlierberghe P, et al. Duplication of the MYB oncogene in T cell acute lymphoblastic leukemia. Nature genetics 2007;39:593-5.
263. Kauraniemi P, Hedenfalk I, Persson K, et al. MYB oncogene amplification in hereditary BRCA1 breast cancer. Cancer research 2000;60:5323-8.
264. Wallrapp C M-PF, Solinas-Toldo S, Lichter P, Friess H, Büchler M, Fink T, Adler G, Gress TM. Characterization of a high copy number amplification at 6q24 in pancreatic cancer identifies c-myb as a candidate oncogene. Cancer research 1997;57:3135-9.
265. Rodriguez-Paredes M, Esteller M. Cancer epigenetics reaches mainstream oncology. Nat Med 2011;17:330-9.
266. Bird A. Perceptions of epigenetics. Nature 2007;447:396-8.
267. Costa FF. Non-coding RNAs, epigenetics and complexity. Gene 2008;410:9-17.

268. Corn PG KS, van Noesel MM, Esteller M, Compitello N, Baylin SB, Herman JG. Transcriptional silencing of the p73 gene in acute lymphoblastic leukemia and Burkitt's lymphoma is associated with 5' CpG island methylation. Cancer Res 1999;59:3352-6.
269. Kuerbitz SJ, Pahys J, Wilson A, Compitello N, Gray TA. Hypermethylation of the imprinted NNAT locus occurs frequently in pediatric acute leukemia. Carcinogenesis 2002;23:559-64.
270. Wagner KJ, Cooper WN, Grundy RG, et al. Frequent RASSF1A tumour suppressor gene promoter methylation in Wilms' tumour and colorectal cancer. Oncogene 2002;21:7277-82.
271. Choy KW, Pang CP, To KF, Yu CB, Ng JS, Lam DS. Impaired expression and promotor hypermethylation of O6-methylguanine-DNA methyltransferase in retinoblastoma tissues. Invest Ophthalmol Vis Sci 2002;43:1344-9.
272. Wong IH, Chan J, Wong J, Tam PK. Ubiquitous aberrant RASSF1A promoter methylation in childhood neoplasia. Clinical cancer research : an official journal of the American Association for Cancer Research 2004;10:994-1002.
273. Jones DT, Jager N, Kool M, et al. Dissecting the genomic complexity underlying medulloblastoma. Nature 2012;488:100-5.
274. Kieran MW, Roberts CW, Chi SN, et al. Absence of oncogenic canonical pathway mutations in aggressive pediatric rhabdoid tumors. Pediatric blood & cancer 2012;59:1155-7.
275. Hasselblatt M, Isken S, Linge A, et al. High-resolution genomic analysis suggests the absence of recurrent genomic alterations other than SMARCB1 aberrations in atypical teratoid/rhabdoid tumors. Genes Chromosomes Cancer 2013;52:185-90.
276. Schwartzentruber J, Korshunov A, Liu XY, et al. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature 2012;482:226-31.
277. Yu J, Deshmukh H, Gutmann RJ, et al. Alterations of BRAF and HIPK2 loci predominate in sporadic pilocytic astrocytoma. Neurology 2009;73:1526-31.
278. Melendez B, Fiano C, Ruano Y, Hernandez-Moneo JL, Mollejo M, Martinez P. BCR gene disruption in a pilomyxoid astrocytoma. Neuropathology : official journal of the Japanese Society of Neuropathology 2006;26:442-6.
279. Hawkins C, Walker E, Mohamed N, et al. BRAF-KIAA1549 fusion predicts better clinical outcome in pediatric low-grade astrocytoma. Clinical cancer research : an official journal of the American Association for Cancer Research 2011;17:4790-8.
280. Lin A, Rodriguez FJ, Karajannis MA, et al. BRAF alterations in primary glial and glioneuronal neoplasms of the central nervous system with identification of 2 novel KIAA1549: BRAF fusion variants. Journal of neuropathology and experimental neurology 2012;71:66-72.
281. Jacob K, Quang-Khuong DA, Jones DT, et al. Genetic aberrations leading to MAPK pathway activation mediate oncogene-induced senescence in sporadic pilocytic astrocytomas. Clinical cancer research : an official journal of the American Association for Cancer Research 2011;17:4650-60.
282. Reddy JP, Li Y. Oncogene-induced senescence and its role in tumor suppression. J Mammary Gland Biol Neoplasia 2011;16:247-56.
283. Horbinski C NM, Hagenkord JM, Hamilton RL, Pollack IF. Interplay among BRAF, p16, p53, and MIB1 in pediatric low-grade gliomas. Neuro Oncol 2012;14:777-89.

284. Dahiya S, Haydon DH, Alvarado D, Gurnett CA, Gutmann DH, Leonard JR. BRAF(V600E) mutation is a negative prognosticator in pediatric ganglioglioma. Acta neuropathologica 2013;125:901-10.
285. Makar AB, McMartin KE, Palese M, Tephly TR. Formate assay in body fluids: application in methanol poisoning. Biochem Med 1975;13:117-26.
286. Smith JS PA, Borell TJ, Lee HK, O'Fallon J, Hosek SM, Kimmel D, Yates A, Burger PC, Scheithauer BW, Jenkins RB. Alterations of chromosome arms 1p and 19q as predictors of survival in oligodendrogliomas, astrocytomas, and mixed oligoastrocytomas. J Clin Oncol 2000;18:636-45.
287. Nakamura M, Shimada K, Ishida E, et al. Molecular pathogenesis of pediatric astrocytic tumors. Neuro-oncology 2007;9:113-23.
288. Suri V JP, Agarwal S, Pathak P, Sharma MC, Sharma V, Shukla S, Somasundaram K, Mahapatra AK, Kale SS, Sarkar C. Molecular profile of oligodendrogliomas in young patients. Neuro Oncol 2011;13:1099-106.
289. Kreiger PA, Okada Y, Simon S, Rorke LB, Louis DN, Golden JA. Losses of chromosomes 1p and 19q are rare in pediatric oligodendrogliomas. Acta neuropathologica 2005;109:387-92.
290. Huang H, Okamoto Y, Yokoo H, et al. Gene expression profiling and subgroup identification of oligodendrogliomas. Oncogene 2004;23:6012-22.
291. Pollack IF, Finkelstein SD, Burnham J, et al. Age and TP53 mutation frequency in childhood malignant gliomas: results in a multi-institutional cohort. Cancer research 2001;61:7404-7.
292. Litofsky NS, Hinton D, Raffel C. The lack of a role for p53 in astrocytomas in pediatric patients. Neurosurgery 1994;34:967-72; discussion 72-3.
293. Patt S, Gries H, Giraldo M, et al. p53 gene mutations in human astrocytic brain tumors including pilocytic astrocytomas. Human pathology 1996;27:586-9.
294. Paulus W, Lisle DK, Tonn JC, et al. Molecular genetic alterations in pleomorphic xanthoastrocytoma. Acta neuropathologica 1996;91:293-7.
295. Yan H, Parsons DW, Jin G, et al. IDH1 and IDH2 mutations in gliomas. The New England journal of medicine 2009;360:765-73.
296. Miwa T HY, Sasaki H, Ezaki T, Yoshida K, Kawase T. Single-copy gain of chromosome 1q is a negative prognostic marker in pediatric nonependymal, nonpilocytic gliomas. Neurosurgery 2011;68:206-12.
297. Padovani L, Colin C, Fernandez C, et al. Search for distinctive markers in DNT and cortical grade II glioma in children: same clinicopathological and molecular entities? Curr Top Med Chem 2012;12:1683-92.
298. Lawrence MS, Stojanov P, Mermel CH, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature 2014;505:495-501.
299. Jones DT, Hutter B, Jager N, et al. Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. Nature genetics 2013;45:927-32.
300. Chiang DY, Getz G, Jaffe DB, et al. High-resolution mapping of copy-number alterations with massively parallel sequencing. Nature methods 2009;6:99-103.
301. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome biology 2011;12:R41.
302. Chapman MA, Lawrence MS, Keats JJ, et al. Initial genome sequencing and analysis of multiple myeloma. Nature 2011;471:467-72.

303. Eisenhardt AE, Olbrich H, Roring M, et al. Functional characterization of a BRAF insertion mutant associated with pilocytic astrocytoma. International journal of cancer Journal international du cancer 2011;129:2297-303.
304. Clark SS, McLaughlin J, Crist WM, Champlin R, Witte ON. Unique forms of the abl tyrosine kinase distinguish Ph1-positive CML from Ph1-positive ALL. Science 1987;235:85-8.
305. Garraway LA, Widlund HR, Rubin MA, et al. Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. Nature 2005;436:117-22.
306. Wang M, Tihan T, Rojiani AM, et al. Monomorphous angiocentric glioma: a distinctive epileptogenic neoplasm with features of infiltrating astrocytoma and ependymoma. Journal of neuropathology and experimental neurology 2005;64:875-81.
307. Preusser M, Hoischen A, Novak K, et al. Angiocentric glioma: report of clinico-pathologic and genetic findings in 8 cases. The American journal of surgical pathology 2007;31:1709-18.
308. Tatevossian RG, Tang B, Dalton J, et al. MYB upregulation and genetic aberrations in a subset of pediatric low-grade gliomas. Acta neuropathologica 2010;120:731-43.
309. Baker SJ, Ma'ayan A, Lieu YK, et al. B-myb is an essential regulator of hematopoietic stem cell and myeloid progenitor cell development. Proceedings of the National Academy of Sciences of the United States of America 2014;111:3122-7.
310. Costa AF, Altemani A, Garcia-Inclan C, et al. Analysis of MYB oncogene in transformed adenoid cystic carcinomas reveals distinct pathways of tumor progression. Laboratory investigation; a journal of technical methods and pathology 2014;94:692-702.
311. Doukhanine E, Gavino C, Haines JD, Almazan G, Richard S. The QKI-6 RNA binding protein regulates actin-interacting protein-1 mRNA stability during oligodendrocyte differentiation. Molecular biology of the cell 2010;21:3029-40.
312. Radomska KJ, Halvardson J, Reinius B, et al. RNA-binding protein QKI regulates Glial fibrillary acidic protein expression in human astrocytes. Human molecular genetics 2013;22:1373-82.
313. Zhao Y, Zhang G, Wei M, et al. The tumor suppressing effects of QKI-5 in prostate cancer: a novel diagnostic and prognostic protein. Cancer biology & therapy 2014;15:108-18.
314. Hall MP, Nagel RJ, Fagg WS, et al. Quaking and PTB control overlapping splicing regulatory networks during muscle cell differentiation. Rna 2013;19:627-38.
315. Bian Y, Wang L, Lu H, et al. Downregulation of tumor suppressor QKI in gastric cancer and its implication in cancer prognosis. Biochemical and biophysical research communications 2012;422:187-93.
316. Chen AJ, Paik JH, Zhang H, et al. STAR RNA-binding protein Quaking suppresses cancer via stabilization of specific miRNA. Genes & development 2012;26:1459-72.
317. Huang FW, Hodis E, Xu MJ, Kryukov GV, Chin L, Garraway LA. Highly recurrent TERT promoter mutations in human melanoma. Science 2013;339:957-9.
318. Gutmann DH, Hedrick NM, Li J, Nagarajan R, Perry A, Watson MA. Comparative gene expression profile analysis of neurofibromatosis 1-associated and sporadic pilocytic astrocytomas. Cancer research 2002;62:2085-91.

319. Sharma MK MD, Reifenberger G, Perry A, Leonard JR, Aldape KD, Albin MG, Emnett RJ, Loeser S, Watson MA, Nagarajan R, Gutmann DH. Distinct genetic signatures among pilocytic astrocytomas relate to their brain region origin. Cancer research 2007;67:890-900.
320. Tchoghandjian A, Fernandez C, Colin C, et al. Pilocytic astrocytoma of the optic pathway: a tumour deriving from radial glia cells with a specific gene signature. Brain 2009;132:1523-35.
321. Tian Y, Rich BE, Vena N, et al. Detection of KIAA1549-BRAF fusion transcripts in formalin-fixed paraffin-embedded pediatric low-grade gliomas. J Mol Diagn 2011;13:669-77.
322. Fan JB, Yeakley JM, Bibikova M, et al. A versatile assay for high-throughput gene expression profiling on universal array matrices. Genome research 2004;14:878-85.
323. Hoshida Y, Villanueva A, Kobayashi M, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. The New England journal of medicine 2008;359:1995-2004.
324. Workman C, Jensen LJ, Jarmer H, et al. A new non-linear normalization method for reducing variability in DNA microarray experiments. Genome biology 2002;3:research0048.
325. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 2007;8:118-27.
326. Tamayo P, Slonim D, Mesirov J, et al. Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. Proceedings of the National Academy of Sciences of the United States of America 1999;96:2907-12.
327. Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999;286:531-7.
328. Gould J, Getz G, Monti S, Reich M, Mesirov JP. Comparative gene marker selection suite. Bioinformatics 2006;22:1924-5.
329. Mootha VK, Lindgren CM, Eriksson KF, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nature genetics 2003;34:267-73.
330. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP. GenePattern 2.0. Nature genetics 2006;38:500-1.
331. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 2002;30:207-10.
332. BrainSpan: Atlas of the Developing Human Brain [Internet]. Funded by ARRA Awards 1RC2MH089921-01, 1RC2MH090047-01, and 1RC2MH089929-01. © 2011. Available from: http://developinghumanbrain.org.
333. Carpenter AE, Jones TR, Lamprecht MR, et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome biology 2006;7:R100.
334. Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. Statistics in medicine 1990;9:811-8.
335. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics 2011;27:1739-40.
336. Lambert SR, Witt H, Hovestadt V, et al. Differential expression and methylation of brain developmental genes define location-specific subsets of pilocytic astrocytoma. Acta neuropathologica 2013;126:291-301.

337. Bacon CM, Petricoin EF, 3rd, Ortaldo JR, et al. Interleukin 12 induces tyrosine phosphorylation and activation of STAT4 in human lymphocytes. Proceedings of the National Academy of Sciences of the United States of America 1995;92:7307-11.
338. Sugimoto N, Nakahira M, Ahn HJ, et al. Differential requirements for JAK2 and TYK2 in T cell proliferation and IFN-gamma production induced by IL-12 alone or together with IL-18. Eur J Immunol 2003;33:243-51.
339. Cogswell JP, Zeleznik-Le N, Ting JP. Transcriptional regulation of the HLA-DRA gene. Crit Rev Immunol 1991;11:87-112.
340. Matsushita K, Takenouchi T, Kobayashi S, et al. HLA-DR antigen expression in colorectal carcinomas: influence of expression by IFN-gamma in situ and its association with tumour progression. British journal of cancer 1996;73:644-8.
341. Lovig T, Andersen SN, Thorstensen L, et al. Strong HLA-DR expression in microsatellite stable carcinomas of the large bowel is associated with good prognosis. British journal of cancer 2002;87:756-62.
342. Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 2012;366:883-92.
343. Saunders NA, Simpson F, Thompson EW, et al. Role of intratumoural heterogeneity in cancer drug resistance: molecular and clinical perspectives. EMBO molecular medicine 2012;4:675-84.
344. Barrett CL, Schwab RB, Jung H, et al. Transcriptome sequencing of tumor subpopulations reveals a spectrum of therapeutic options for squamous cell lung cancer. PloS one 2013;8:e58714.
345. Duechler M, Peczek L, Zuk K, Zalesna I, Jeziorski A, Czyz M. The heterogeneous immune microenvironment in breast cancer is affected by hypoxia-related genes. Immunobiology 2014;219:158-65.
346. Swanton C. Intratumor heterogeneity: evolution through space and time. Cancer research 2012;72:4875-82.
347. Sottoriva A, Spiteri I, Piccirillo SG, et al. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. Proceedings of the National Academy of Sciences of the United States of America 2013;110:4009-14.
348. Trombetta JJ, Gennert D, Lu D, Satija R, Shalek AK, Regev A. Preparation of Single-Cell RNA-Seq Libraries for Next Generation Sequencing. Current protocols in molecular biology / edited by Frederick M Ausubel [et al] 2014;107:4 22 1-4 17.
349. Trapnell C, Roberts A, Goff L, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature protocols 2012;7:562-78.
350. Zelensky AN, Gready JE. The C-type lectin-like domain superfamily. The FEBS journal 2005;272:6179-217.
351. Kaufmann D, Gruener S, Braun F, et al. EVI2B, a gene lying in an intron of the neurofibromatosis type 1 (NF1) gene, is as the NF1 gene involved in differentiation of melanocytes and keratinocytes and is overexpressed in cells derived from NF1 neurofibromas. DNA and cell biology 1999;18:345-56.
352. Lovero KL, Blankenship SM, Shi Y, Nicoll RA. SynDIG1 promotes excitatory synaptogenesis independent of AMPA receptor trafficking and biophysical regulation. PloS one 2013;8:e66171.

353. Rebetz J, Tian D, Persson A, et al. Glial progenitor-like phenotype in low-grade glioma and enhanced CD133-expression and neuronal lineage differentiation potential in high-grade glioma. PloS one 2008;3:e1936.
354. Auvergne RM, Sim FJ, Wang S, et al. Transcriptional differences between normal and glioma-derived glial progenitor cells identify a core set of dysregulated genes. Cell reports 2013;3:2127-41.
355. Panchision DM, Chen HL, Pistollato F, Papini D, Ni HT, Hawley TS. Optimized flow cytometric analysis of central nervous system tissue reveals novel functional relationships among cells expressing CD133, CD15, and CD24. Stem cells 2007;25:1560-70.
356. Tchoghandjian A, Baeza N, Colin C, et al. A2B5 cells from human glioblastoma have cancer stem cell properties. Brain pathology 2010;20:211-21.
357. Shalek AK, Satija R, Adiconis X, et al. Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature 2013;498:236-40.
358. Hashimshony T, Wagner F, Sher N, Yanai I. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. Cell reports 2012;2:666-73.
359. Flynn JM, Spusta SC, Rosen CJ, Melov S. Single cell gene expression profiling of cortical osteoblast lineage cells. Bone 2013;53:174-81.
360. Cann GM, Gulzar ZG, Cooper S, et al. mRNA-Seq of single prostate cancer circulating tumor cells reveals recapitulation of gene expression and pathways found in prostate cancer. PloS one 2012;7:e49144.
361. Patel AP, Tirosh I, Trombetta JJ, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science 2014;344:1396-401.

APPENDIX I – Résumé en Français

CONTEXTE SCIENTIFIQUE ET DEFINITIONS

Les gliomes de bas-grade représentent la tumeur cérébrale la plus fréquente chez l'enfant. Leur définition actuelle est basée sur l'aspect morphologique des tumeurs. Elles sont caractérisées par un très large spectre de sous-types tumoraux, très hétérogènes tant cliniquement que radiologiquement.

Leur définition actuelle est principalement basée sur des critères histologiques, en fonction de la prédominance et de l'architecture des sous-types cellulaires décrits dans la tumeur. Il existe deux principales classifications histologiques couramment utilisées ; la première est définie par l'Organisation Mondiale de la Santé (OMS) et la seconde est celle définie par la classification de Sainte-Anne, plus fréquemment utilisée en France. Elles sont dites de bas-grade car elles correspondent à un grade I ou II selon la classification OMS. Ces classifications ont l'avantage de répertorier les multiples tumeurs en différentes catégories afin d'harmoniser leur appellation à travers le monde. Leur principale limite est qu'elles se basent uniquement sur l'architecture cellulaire au microscope. Ces classifications souffrent ainsi d'un manque de consensualisme et de précision due à la variabilité interindividuelle lors de l'appellation d'une tumeur dans une certaine catégorie.

Les astrocytomes pilocytiques (PA) sont les gliomes de bas-grade les plus fréquents chez l'enfant. Dans la majorité des cas, les PA sont aisément diagnostiqués car leur architecture tumorale est relativement typique et caractérisée par une la présence de fibres de Rosenthal, d'une prolifération vasculaire et de corps éosinophiles fréquemment localisés en bordure des régions kystiques qui sont caractéristiques de ces tumeurs. Elles sont caractérisées par une expression constante de Glial Fibrillary Acid Protein (GFAP) en immunohistochimie. Elles expriment également typiquement le marqueur de différenciation OLIG2 et la myelin basic protein (MBP).

Les gangliogliomes (GG) sont également des gliomes de grade I et se caractérisent typiquement par une prolifération de neurones binuclés, d'une inflammation périvasculaire, de calcifications et de kystes.

Les tumeurs neuroépithéliales dysembryoplastiques (DNT) et les gliomes angiocentriques (AG ou ANET) sont des tumeurs de grade I qui ont été récemment décrites et reconnues comme entité histologique à part. Les DNT se caractérisent par des cellules apparentées aux oligodendrocytes, GFAP négatives associées à des neurones immergés dans un environnement éosinophile. Les AGs décrits initialement par Lellouch-Tubiana et al. sont composés d'astrocytomes fusiformes et bipolaires fixant la GFAP se disposant classiquement en palissade autour des structures vasculaires.

Les gliomes de grade II selon l'OMS sont définis essentiellement par les astrocytomes diffus (DA), les astrocytomes pilomixoides (PMA), les xanthoastrocytomes pleomorphic (PXA) et les tumeurs oligodendrogliales (OD).

Globalement, l'index de prolifération cellulaire MIB est faible avec moins de 10%.

Sur le plan clinique, les gliomes de bas-grade ont un comportement tumoral intriguant. Leur croissance tumorale est lentement progressive, marquée par des phases de croissance puis de « quiescence » probable. Leur traitement actuel est principalement basé sur la chirurgie et la chimiothérapie. La radiothérapie tend à être moins utilisée compte tenu des complications à long terme. Malgré une excellente survie globale dans la population pédiatrique, ces tumeurs rechutent ou progressent fréquemment et sont responsables d'une comorbidité importante liée aux différents traitements utilisés. Malgré un manque de preuve formelle sur le plan scientifique, ces tumeurs ont une tendance à entrer dans une phase de quiescence définitive lors de la transition à l'âge adulte, expliquant ainsi leur excellente survie globale. Les raisons physiopathologiques à l'origine de ce phénomène d'arrêt de croissance tumorale demeurent inconnues et interpellent les scientifiques.

Les progrès récents de la biologie moléculaire et de la génomique, avec l'expansion des techniques de séquençage à haut débit et le développement de plateformes permettant l'exploration à haute résolution du génome tumoral ont permis d'approfondir la biologie de nombreux types de cancers. Dans les gliomes de bas-grade de l'enfant, ces techniques modernes nous permettent premièrement d'approfondir considérablement les connaissances sur la biologie de ces tumeurs afin d'enrichir et compléter leur classification morphologique actuelle. Elles permettent également d'appréhender des questions plus fondamentales comme celle du type cellulaire initial à l'origine de ces tumeurs multiples et des voies de différenciation cellulaires responsables du développement puis du ralentissement de la croissance tumorale.

CHAPITRE 1 : REVUE DE LA LITTERATURE

Dans le premier chapitre, ce travail résume de façon détaillée les connaissances actuelles sur la prise en charge clinique de ces tumeurs. Il décrit ensuite toutes les différentes technologies d'exploration de l'ADN tumoral (détection des altérations de nombre, des mutations ponctuelles, des insertions, délétions, réarrangements chromosomiques) et de l'ARN tumoral (étude du transcriptome par méthodes array ou par séquençage direct de l'ARN tumoral) ainsi que les outils bioinformatiques actuellement disponibles pour exploiter l'ensemble de ces ressources.

Sur la base d'une revue de la bibliographie, ce travail décrit l'ensemble des altérations génétiques qui ont été récemment découvertes grâce aux progrès des techniques de séquençage à haut débit.

Sur le plan génétique, plusieurs facteurs influencent la probabilité de découvrir de nouvelles altérations potentiellement oncogéniques. La fréquence des mutations au sein d'un sous-type tumoral, le taux de mutation somatique tumoral, et surtout le nombre de tumeurs utilisées sont importants. L'importance de la cohorte de tumeurs étudiée est donc un aspect primordial. Compte tenu de leur faible prévalence à travers le monde, l'accessibilité à un grand nombre de gliomes de bas-grade pédiatriques représente un obstacle important dans les études de séquençage du génome. Par conséquent, les collaborations et le partage des résultats sont nécessaires afin d'augmenter la cohorte des tumeurs et de bénéficier de conditions optimales pour découvrir de nouvelles altérations génétiques à l'origine de ces tumeurs.

Ce travail bénéficie directement de cet effort de collecter un maximum de tissu et d'information sur le séquençage tumoral afin de réaliser une analyse approfondie des altérations génomiques de l'ADN et l'ARN des gliomes de bas-grade pédiatriques.

CHAPITRE 2 : ETUDE DES ALTERATIONS DE L'ADN

Dans le deuxième chapitre, notre travail se base sur la description des altérations de l'ADN des gliomes de bas-grade de l'enfant à partir des données de séquençage à haut débit récemment publiées par des équipes de chercheurs aux USA (Washington University et St-Jude Hospital, Tenesse) et en Europe (DKFZ, Heidelberg) ainsi que des tumeurs nouvellement séquencées au Dana-Farber Cancer Institute et au Children's Hospital of Philadelphia et des tumeurs originaires de l'Hôpital Necker à Paris. Nous avons analysé au total le génome de 169 gliomes de bas-grade de l'enfant, comprenant 138 whole genome sequencing (WGS) et 53 RNA-sequencing (parmi lesquelles 22 tumeurs ont également bénéficié de whole exome sequencing (WES)).

Le taux médian de mutation dans les régions exoniques était de 1.02×10^{-5} mutations /base. Nous avons pu confirmer que les mutations du gène *BRAF* sont les plus fréquentes. 98 tumeurs (57%) présentent un des 7 différents isoformes de duplication de *BRAF*. 12 tumeurs présentent une mutation *BRAF* (V600E, insertion p.599 599T>TT, insertion VLR à la position 506 et une mutation ponctuelle p.E451D). Les altérations du gène *FGFR1* ont été retrouvées dans 12 tumeurs (5 tumeurs avec une mutation faux-sens, 4 duplications dans la région exonique du gène et 3 duplication-inversions entre *FGFR1* et *TACC1*). Nous avons également identifié un réarrangement chromosomique entre *FGFR3* et *TACC4* dans un PA. Les réarrangements du gène codant pour le récepteur à la Tyrosine kinase *NTRK2* ont été retrouvés dans 4 tumeurs. Deux de ces tumeurs présentent un réarrangement identique entre l'exon 15 du gène *NTRK2* et l'exon 6 de *QKI*. Nous avons également identifié un réarrangement entre *NTRK3* et *NAV6* dans une tumeur PXA analysée par RNA-sequencing.

NF1 était muté dans 5 tumeurs.

Nous avons observé des mutations dans les gènes codant pour les histones (*H3F3A* et *HIST1H3B*) dans 5 tumeurs. 80% de ces mutations ont également été décrites dans les gliomes de haut-grade pédiatriques.

Notre travail présente pour la première fois un réarrangement chromosomique prédominant de façon quasi exclusive dans les AGs. Parmi les 6 AGs analysés dans notre cohorte, nous avons découvert que 5 présentaient une translocation entre le gène *MYB* et *QKI.* Au total nous avons décrit 4 différents types de réarrangements entre MYB et QKI, résultant tous en une perte de l'extrémité 3' end de MYB et 5' de QKI. Deux de ces tumeurs ont pu être analysées par WGS ayant permis d'identifier de façon précise la cartographie des cassures chromosomiques. Dans la première tumeur, la cassure chromosomique se situe entre l'exon 9 et 10 de *MYB* et entre l'exon 4 et 5 de *QKI*. La seconde tumeur décrit un réarrangement complexe entre le gène *EYA4*, *MYB* et *QKI* résultant à la perte du dernier exon de *MYB* accolé aux exons 5-7 de *QKI*. Les quatre premiers exons de *QKI* sont accolés aux 11 premiers exons de *EYA4.* De façon intéressante, *MYB* (activateur transcriptionnel) et *QKI* (protéine se liant à l'ARN) ont été décrits dans d'autres études génétiques comme potentiels oncogène et gène suppresseur de tumeur, respectivement.

Des études complémentaires dans divers modèles *in vitro* and *in vivo* réalisées principalement par Bandopahhayay et Ramkisson, deux collaboratrices dans le laboratoire, ont permis d'appréhender les répercussions fonctionnelles de ce réarrangement particulier. Elles ont pu décrire que le réarrangement entre *MYB* et *QKI* entrainait entre autre trois potentiels mécanismes à l'origine des mécanismes d'oncogenèse ; la création d'une protéine oncogénique MYB-QKI, l'augmentation de l'expression de MYB en partie lié à la translocation des enhancers présents à l'extrémité 3' de *QKI* proches des domaines de transactivation de *MYB* et enfin par une perte hétérozygote de *QKI*.

CHAPITRE 3 : ETUDE DES ALTERATIONS DE L'ARN DANS LE BULK TUMORAL

Dans le troisième chapitre, ce travail décrit les résultats d'analyses des profils d'expression génétiques de 151 gliomes de bas grade pédiatriques extraits à partir de tissu conservé en paraffine. Les tumeurs pédiatriques sélectionnées reflètent en partie l'hétérogénéité clinique et génétique avec une bonne représentativité des différents sous-types histologiques. Nous avons pu analyser le profil d'expression de 6100 gènes de 76 PAs, 17 GGs, 11 DAs, 12 DNTs, 2 ODs et 33 tumeurs NOS (histologie indifférenciée). 48% de ces tumeurs étaient localisées dans des régions supratentorielles (cortex cérébral, région hypothalamo-chiasmatiques) et 52% provenaient de la fosse postérieure. Nous avons également pu déterminer le statut *BRAF* (91% pour la duplication *BRAF-KIAA1549* et 62% pour la mutation V600E) des tumeurs. Nous disposions également d'une répartition homogène des âges au sein de la population pédiatrique avec 13 tumeurs développées dans la petite enfance (entre 0 et 18 mois), 89 tumeurs dans l'enfance (entre 19 mois et 11 ans) et 49 tumeurs dans l'adolescence (entre 12 et 19 ans). Nous avons pu mettre en évidence des différences moléculaires en fonction de leur sous-type histologique, de la localisation tumorale et de leur statut *BRAF*.

Grâce aux différentes techniques de clustering non supervisées, nous avons identifié que les 151 gliomes de bas-grade pédiatriques se distribuaient majoritairement en 3 groupes moléculaires. Le premier cluster moléculaire était significativement composé de DA et DNT et d'autres tumeurs supratentotorielles. Ce cluster était enrichi avec des tumeurs portant la mutation V600E. Le deuxième cluster était significativement enrichi de PAs développés dans la fosse postérieure comportant la duplication *BRAF-KIAA1549*. Le troisième cluster, groupe mixte, se composait majoritairement de PA, GG et NOS provenant des régions supra ou infratentorielles. L'étude de la survie sans événement des patients portant ces tumeurs n'a pas permis d'identifier de différence significative en fonction des trois clusters moléculaires.

En comparant les profils d'expression des différents sous-types histologiques, nous avons observé que certains sous-types histologiques divergeaient, comme les PA des DA ou les PA des DNT, alors que d'autres sous-types ne présentaient pas de différences moléculaires significatives, comme les PA et les GG.

Nous avons également observé que les tumeurs développées dans les régions supratentorielles divergeaient clairement des tumeurs naissant dans la fosse postérieure. De façon complémentaire, nous avons pu démontrer que les PAs développés dans les régions corticales surexprimaient notamment un grand nombre de gènes impliqués dans les voies de signalisation de l'inflammation, par rapport aux PAs de la fosse postérieure. Ces résultats suggèrent que selon la localisation tumorale, les patterns d'activation moléculaires divergent au sein d'une même entité histologique.

Nous avons également observé certaines différences moléculaires entre les tumeurs primitives et les tumeurs lors de la rechute ainsi qu'entre les tumeurs *BRAF* V600E mutées et *BRAF-KIAA1549* dupliquées.

Enfin, nous n'avons pas observé de différences moléculaires significatives entre les tumeurs développées dans la petite enfance avec celles diagnostiquées dans l'enfance ou l'adolescence. De façon opposée, en comparant le profil d'expression génétique de 20 gliomes de bas-grade de l'adulte nous avons observé que celles-ci divergent de façon importante avec celles développées chez l'enfant.

CHAPITRE 4 : ETUDE DES ALTERATIONS DE L'ARN A L'ECHELLE D'UNE CELLULE UNIQUE

Un des obstacles de l'analyse du transcriptome en cancérologie est la contamination des cellules tumorales par du tissu conjonctif normal. Dans le dernier volet de ce travail, nous avons pu tester la faisabilité d'isoler successivement par cytométrie en flux des cellules uniques issues de la dissociation tumorale provenant d'une exérèse chirurgicale récente. Afin de distinguer deux sous-types de population cellulaire au sein de la tumeur dissociée, nous avons initialement testé le taux d'expression dans les tumeurs fraichement dissociées de plusieurs marqueurs membranaires de différentiation neuro-gliale, comme N-CAM, A2B5, GLAST, O4, FGFR1. Après avoir optimisé les contrôles positifs et négatifs pour ces anticorps membranaires, nous avons privilégié l'utilisation de l'anticorps A2B5 qui nous semblait le plus robuste. Nous avons pu ensuite dissocier trois PAs et trier à l'échelle de la cellule unique des populations de cellules A2B5 positives et négatives.

En parallèle, nous avons bénéficié de l'expertise technologique du Broad Institute qui nous a permis d'optimiser les étapes d'extraction d'ARN, de génération de cDNA suivi d'amplification du cDNA et de la préparation des librairies en vue d'une analyse transcriptomique à haut-débit de l'ARN d'une cellule unique A2B5 positive et négative. Notre hypothèse de travail était que ces deux sous-populations cellulaires exprimaient des profils différents. Nous souhaitions ensuite étudier leur degré d'hétérogénéité au sein de chaque population cellulaire (A2B5 positives ou négatives) et entre ces deux sous-populations.

L'analyse des données de RNA-sequencing des cellules A2B5 positives et négatives issues des trois PAs nous a permis de confirmer notre hypothèse que ces deux sous-populations cellulaires divergeaient significativement. Nous avons pu décrire de façon exhaustive les gènes et voies de signalisations divergentes entre ces deux sous-populations dans les trois PAs analysés.

Nous avons observé que le degré d'hétérogénéité semblait plus important pour la population cellulaire A2B5 négatif et que les voies de signalisation de l'inflammation prédominaient fréquemment dans ces cellules, en comparaison avec celles exprimant A2B5. Des études complémentaires de validation sont nécessaires pour consolider ces premiers résultats.

De façon complémentaire, il serait nécessaire d'identifier parmi les cellules uniques étudiées celles qui revêtent les altérations génétiques de BRAF afin d'observer si celles-ci sont majoritairement caractérisées par l'expression membranaire de A2B5. Ces résultats pourront fortement suggérer que les cellules exprimant le marqueur de différentiation A2B5 représentent la fraction de cellules cancéreuses responsable de la croissance et du développement tumoral.

Ces premiers résultats soulignent l'intérêt d'exploiter des nouvelles technologies de pointe pour servir de base à l'étude des caractéristiques biologiques des cellules tumorales ouvrant de nouvelles perspectives sur la compréhension fine des mécanismes moléculaires à l'origine des tumeurs.

APPENDIX II – Other published papers during PhD

1) BET-bromodomain inhibition of MYC-amplified Medulloblastoma

Bandopadhayay P*, Bergthold G*, Nguyen B, Schubert S, Gholamin S, Tang Y, Bolin S, Schumacher SE, Zeid R, Masoud S, Yu F, Vue N, Gibson WJ, Paolella BR, Mitra SS, Cheshier S, Qi J, Liu KW, Wechsler-Reya RJ, Weiss WA, Swartling FJ, Kieran MW, Bradner JE, Beroukhim R*, Cho YJ* ; (*co-authors)

Clin Cancer Res. 2014 Feb 15;20(4):912-25

Clinical Cancer Research

Cancer Therapy: Preclinical

BET Bromodomain Inhibition of *MYC*-Amplified Medulloblastoma

Pratiti Bandopadhayay[1,3,5], Guillaume Bergthold[1,5], Brian Nguyen[6], Simone Schubert[6], Sharareh Gholamin[7], Yujie Tang[6], Sara Bolin[11], Steven E. Schumacher[1,5], Rhamy Zeid[2], Sabran Masoud[6], Furong Yu[6], Nujsaubnusi Vue[6], William J. Gibson[1,5], Brenton R. Paolella[1,5], Siddhartha S. Mitra[7], Samuel H. Cheshier[7], Jun Qi[2], Kun-Wei Liu[9], Robert Wechsler-Reya[9], William A. Weiss[10], Fredrik J. Swartling[11], Mark W. Kieran[3], James E. Bradner[2,5], Rameen Beroukhim[1,2,4,5], and Yoon-Jae Cho[6,7,8]

Abstract

Purpose: *MYC*-amplified medulloblastomas are highly lethal tumors. Bromodomain and extraterminal (BET) bromodomain inhibition has recently been shown to suppress MYC-associated transcriptional activity in other cancers. The compound JQ1 inhibits BET bromodomain-containing proteins, including BRD4. Here, we investigate BET bromodomain targeting for the treatment of *MYC*-amplified medulloblastoma.

Experimental Design: We evaluated the effects of genetic and pharmacologic inhibition of BET bromodomains on proliferation, cell cycle, and apoptosis in established and newly generated patient- and genetically engineered mouse model (GEMM)-derived medulloblastoma cell lines and xenografts that harbored amplifications of *MYC* or *MYCN*. We also assessed the effect of JQ1 on MYC expression and global MYC-associated transcriptional activity. We assessed the *in vivo* efficacy of JQ1 in orthotopic xenografts established in immunocompromised mice.

Results: Treatment of *MYC*-amplified medulloblastoma cells with JQ1 decreased cell viability associated with arrest at G_1 and apoptosis. We observed downregulation of MYC expression and confirmed the inhibition of MYC-associated transcriptional targets. The exogenous expression of MYC from a retroviral promoter reduced the effect of JQ1 on cell viability, suggesting that attenuated levels of MYC contribute to the functional effects of JQ1. JQ1 significantly prolonged the survival of orthotopic xenograft models of *MYC*-amplified medulloblastoma ($P < 0.001$). Xenografts harvested from mice after five doses of JQ1 had reduced the expression of *MYC* mRNA and a reduced proliferative index.

Conclusion: JQ1 suppresses MYC expression and MYC-associated transcriptional activity in medulloblastomas, resulting in an overall decrease in medulloblastoma cell viability. These preclinical findings highlight the promise of BET bromodomain inhibitors as novel agents for *MYC*-amplified medulloblastoma. *Clin Cancer Res; 20(4); 912–25. ©2013 AACR.*

Introduction

Medulloblastoma is the most common malignant brain tumor of childhood (1). Patients with local disease receive surgical resection, radiotherapy, and chemotherapy, with 5-year overall survival exceeding 80% (2). These treatments cause significant therapy-related morbidity, including disabling cognitive deficits (3), growth failure, and increased risk of secondary malignancies (4). However,

---

Authors' Affiliations: Departments of [1]Cancer Biology and [2]Medical Oncology, Dana-Farber Cancer Institute and Harvard Medical School; [3]Pediatric Neuro-Oncology, Department of Pediatric Oncology, Dana-Farber Cancer Institute and Division of Pediatric Hematology/Oncology, Boston Children's Hospital; [4]Center for Cancer Genome Characterization, Dana-Farber Cancer Institute, Boston; [5]The Broad Institute of MIT and Harvard, Cambridge, Massachusetts; Departments of [6]Neurology and Neurological Sciences and [7]Neurosurgery, Stanford University School of Medicine; [8]Stanford Cancer Institute, Stanford University Medical Center, Stanford; [9]Tumor Initiation and Maintenance Program, NCI-Designated Cancer Center, Sanford-Burnham Medical Research Institute, La Jolla; [10]Department of Neurology, Pediatrics, and Neurosurgery, University of California, San Francisco, California; and [11]Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Rudbeck Laboratory, Uppsala University, Uppsala, Sweden

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

P. Bandopadhayay, G. Bergthold, R. Beroukhim, and Y.-J. Cho contributed equally to this work.

Corresponding Authors: Yoon-Jae Cho, Department of Neurology and Neurological Sciences, Stanford University School of Medicine, 1201 Welch Road, MSLS Building, Room P213, Stanford, CA 94305. Phone: 617-755-0298; Fax: 650-723-7299; E-mail: yjcho1@stanford.edu; and Rameen Beroukhim, Department of Cancer Biology, Dana-Farber Cancer Institute, Harvard Medical School, 450 Brookline Avenue, SM 1022C, Boston, MA 02215. Phone: 617-582-7941; Fax: 617-394-2896; E-mail: Rameen_Beroukhim@dfci.harvard.edu

doi: 10.1158/1078-0432.CCR-13-2281

Translational Relevance

Collectively, *MYC*, *MYCN*, and *MYCL1* are the most commonly amplified oncogenes in medulloblastoma, and are associated with a dismal prognosis. The recent development of strategies to block MYC activity through the inhibition of bromodomain and extraterminal (BET) bromodomain proteins represents a possible novel therapeutic strategy for these tumors. Here, we report that JQ1, a potent inhibitor of BET bromodomain proteins, results in both reduced cell proliferation and prominent apoptosis using *in vitro* models *of MYC*-amplified medulloblastoma, and prolongs survival in xenograft models. We confirm effective downregulation of MYC-related pathways with JQ1 and suppression of the expression of MYC. We also show reduced cell proliferation with JQ1 treatment of cells derived from MYCN-driven tumors harvested from a genetically engineered mouse model. BET bromodomain inhibition, therefore, represents a novel therapeutic strategy for children with *MYC*-amplified medulloblastoma. These data support further evaluation in early-phase clinical trials.

despite intensive chemotherapy and radiotherapy, the overall survival of "high-risk" patients remains dismal, with 10-year overall survival rates as low as 20% (5). Thus, tremendous impetus exists for the development of more effective treatments, based on known molecular targets, in medulloblastoma.

Medulloblastoma is a genetically heterogeneous disease, composed of molecular subtypes characterized by differing transcriptional signatures, genomic alterations, and clinical courses (6–9). The current consensus is of at least four distinct subtypes, including Wingless (WNT), Sonic Hedgehog (SHH), and groups 3 and 4 (10). Group 3 medulloblastomas have the worst prognosis, and are commonly metastatic and refractory to standard therapy, with 10-year overall survival rates of 39% (5, 6, 10). Amplifications of one of three members of the MYC family of genes (*MYC*, *MYCN*, and *MYCL1*) are found in several subtypes. Group 3 tumors are often associated with amplification of *MYC* (11), which is the most frequently observed amplification observed across multiple cancer types (12). Group 3 tumors without *MYC* amplification are often characterized by overexpression of MYC (6) or amplification of *MYCN* (11). *MYCL1* amplifications have been reported in a few SHH tumor cases whereas SHH and group 4 tumors are enriched with amplifications of *MYCN* (11).

MYC and other transcription factors complicit in cancer are poor targets for small molecule inhibition. Alternative strategies target MYC through epigenetic modulation of *MYC* transcription itself or of MYC target genes (13, 14). In particular, bromodomain and extraterminal (BET)–containing proteins, which recognize and engage side-chain acetylated lysine on open chromatin to facilitate gene transcription (15), have been identified as novel targets for small molecule development (13). The evidence that transcription of *MYC* and *MYCN* and subsequent activation of their downstream transcriptional programs can be targeted by BET bromodomain inhibition (13, 16, 17) presents a novel therapeutic strategy for patients with *MYC*-amplified medulloblastoma.

Materials and Methods

Ethics statement

Ethics approval was granted by the relevant human Institutional Review Board and/or animal ethics Institutional Animal Care and Use Committee (IACUC) research committees of Dana-Farber Cancer Institute (Boston, MA) and Stanford University (Stanford, CA).

Cell lines and culture

D283, D425, D458, and D556 were generously provided by Dr. Darrell Bigner (Duke University, Durham, NC). Daoy cells were obtained from the American Type Culture Collection. Cell lines were maintained in Dulbecco's Modified Eagle Medium (Gibco) supplemented with 10% FBS (100106; Benchmark) and 1% penicillin–streptomycin with 1% glutamine (Gibco). UW228, R256, R262, and R308 were a kind gift from Michael Bobola (University of Washington, Seattle, WA). MB002 cells were derived from an autopsy specimen of the leptomeningeal compartment from a child with metastatic, treatment-refractory (chemotherapy only) medulloblastoma. The MB002 primary tumor displayed histologic features of large-cell medulloblastoma and gene expression markers consistent with group 3 medulloblastoma (Supplementary Fig. S1A; ref.11). MB004 cells were derived from the primary surgical resection of a tumor in a child whose tumor recurred after therapy. The MB004 primary tumor displayed focal anaplasia and gene expression markers consistent with group 3 and 4 medulloblastomas (see Supplementary Fig. S1; ref. 6). *MYC* amplification in the MB002 and MB004 cells was confirmed with NanoString nCounter v2 Cancer CN Codeset, which estimates a copy number of 86 genes commonly amplified or deleted in cancer (Supplementary Table S1). Human neural stem cells were derived from subventricular zone tissue surgically excised during a functional hemispherectomy in a child with refractory seizures. MB002 and MB004 cells were maintained in culture media with 1:1 Dulbecco's Modified Eagle Medium (Gibco) and neural stem cell media (Gibco) supplemented with B27 (Gibco) EGF (02653, Stem Cell), fibroblast growth factor (GF003; Millipore), Heparin (07980, Stem Cell), and leukemia inhibitory factor (LIF; LIF1010; Millipore). The subventricular zone (SVZ)–derived neural stem cells were maintained similarly, with the exception of the LIF supplement. The MYC- and MYCN-driven medulloblastoma GEMM cell lines were derived and cultured as previously described (18, 19). Briefly, for *Myc*-amplified GEMM lines, cerebellar stem cells infected with *Myc* and *DNp53* retroviruses were transplanted into cerebella of NOD-SCID-IL2Rgamma$^{\text{null}}$ (NSG) mice. When mice became symptomatic, tumors were harvested and dissociated into single cell suspensions.

Patient-derived cell lines MB002 and MB004 were authenticated using sequence-tagged site (STS) fingerprinting. Cell lines obtained from the Bigner and Babola laboratories were authenticated using SNP250k or SNP6.0 arrays, which revealed copy-number alterations consistent with previously published karyotypes (12, 20).

Short hairpin RNA suppression

Lentiviral vectors encoding short hairpin RNAs (shRNA) specific for *BRD4*, *MYC*, and the control *LACZ* were obtained from The RNAi Consortium (Clones and sequence: shBRD4 TRCN0000021426, 5′CGTCCGATTGATGTTCTCCAA; shMYC TRCN0000039640, 5′CAGTTGAAACACAAACTTGAA; and shLacZ TRCN0000231726, 5′TGTTCGCATTATCCGAACCAT). Lentivirus was produced by the transfection of 293T cells with vectors encoding each shRNA (2 μg) with packaging plasmids encoding PSPAX2 and VSVG using Lipofectamine (Invitrogen, 56532). Lentivirus-containing supernatant was collected 48 and 72 hours after transfection, pooled, and stored at −80°C. Cells were infected (a ratio of 1:4 virus media) in polybrene-containing media (2.5 μg/mL), and incubated overnight. Cells were selected in puromycin (2.0 μg/mL) starting 48 hours after infection.

Overexpression of MYC in D283 for MYC rescue experiments

293T cells were transfected with 2 μg retroviral pBabe expression vectors (empty vector or pBabeMYC) with packaging plasmids encoding gag-pol and VSVG using Lipofectamine. Retrovirus containing supernatant was collected 48 and 72 hours after transfection, and was pooled and stored at −80°C. D283 ells were infected (a ratio of 1:4 virus media) in polybrene-containing media (2.5 μg/mL). Treatment with 1 μmol/L JQ1R or JQ1S was commenced 6 hours after infection. Cell viability was assessed 24 hours after treatment with JQ1.

Cell viability assays following treatment with JQ1 or shRNA suppression

To assess responsiveness to JQ1, 1,000 cells were plated in 96-well plates in serial dilutions of either JQ1R or JQ1S, in triplicate. Cell viability was measured by assessing ATP content at 0, 24, 48, 72, 96, and 120 hours using Cell Titre-Glo (Promega) according to the manufacturer's instructions. Mean ± SD was calculated. Nonlinear dose–response curves were applied to the data using GraphPad Prism.

To assess the dependence of cells on BRD4 or MYC, cells were infected with lentiviral plasmids encoding shRNA. Forty-eight hours after infection, 1,000 cells were plated in each well of 96-well plates, in triplicate, in media containing puromycin (1 μg/mL). Cell viability was measured by assessing ATP content using Cell Titre-Glo (Promega), and results were normalized to baseline. Mean ± SD was calculated.

Flow cytometry

Cell-cycle analysis was performed by measuring DNA content by propidium iodide (PI)–stained cells treated with 1 μmol/L of JQ1R or JQ1S for 72 hours. Apoptosis was measured with Annexin V/PI staining. The Annexin V was labeled with Alexa Fluor (A13201; Invitrogen) and flow cytometry was performed per the manufacturer's guidelines.

Protein extraction and immunoblotting

MYC-amplified cells were lysed in boiling RIPA (radioimmunoprecipitation assay) lysis buffer containing protease and phosphatase inhibitors, and centrifuged at 13,000 × g for 10 minutes. For *MYCN*-amplified lines, Western blot analysis was performed as previously described (19) with the following modification: lysis buffer with 1% SDS. Supernatant was mixed with 4× SDS sample buffer, boiled for 10 minutes, and subjected to SDS-PAGE on 4% to 12% gradient gels. Blots were probed with antibodies against BRD4 (12183; Cell Signaling Technology), MYC (sc-764; Santa Cruz Biotechnology), MYCN (ab-16898; Abcam), β-tubulin (MAB 3408; Millipore), and actin (sc-1615; Santa Cruz).

RNA extraction and real-time reverse transcriptase PCR

RNA was extracted with the RNeasy Kit (Qiagen). cDNA was synthesized from 1 μg RNA using High Capacity RNA-to-cDNA kits (Applied Biosystems). Real-time reverse transcriptase (RT-PCR) was performed using SYBR Green master mix (Applied Biosystems). Cycling was performed as follows: 50°C for 2 minutes and 95°C for 10 minutes, followed by 40 cycles of 95°C for 15 seconds and 60°C for 30 seconds. This was followed with a dissociation stage of 95°C for 15 seconds, 60°C for 30 seconds, and 95°C for 15 seconds. Primers for BRD4, MYC, and β-actin are listed in Supplementary Table S2. Samples were amplified in triplicate and data were analyzed using the $\Delta\Delta C_T$ method.

Copy-number analysis

Relative copy-number estimates were generated from published Affymetrix SNP 6.0 data for 1,073 tumors (11) using comparison data from 131 normal samples and an analytic pipeline described in detail elsewhere (Tabak and colleagues, in preparation). Briefly, signal intensities for each probe were normalized to uniform intensity values and merged to form probe set–level values using SNPFileCreator, a Java implementation of dChip (21, 22). Marker-level intensities were calibrated to DNA copy-number levels using Birdseed for single-nucleotide polymorphism (SNP) markers (23) and using the results of experiments with cell lines with varied copy numbers of the X chromosome for copy-number probes (Tabak and colleagues, in preparation). Regions of frequent germline copy-number variation were identified using a large bank of normal tissue samples and excluded from the data (Tabak and colleagues, in preparation). Noise was reduced by applying tangent normalization (12), followed by circular binary segmentation (24, 25). Data were mean centered for each sample. Amplifications were defined as greater than a relative copy number of 2.4. For samples with a relative copy number of 2.4 to 3, we applied the ABSOLUTE algorithm (26), and confirmed that each of these samples had an absolute copy number of greater than three copies.

Genome-wide expression analysis

Previously published microarray expression and copy-number data (11) were obtained from the Gene Expression Omnibus (GEO; GSE37385 and GSE37382). The expression data were obtained using the GEOImporter module in GenePattern. *Z* scores of gene expression values of genes within samples were calculated.

For analyses of gene sets enriched among samples exhibiting high expression of MYC family members (z score > 1 for *MYC*, z score > 0.85 for *MYCL1*, and z score > 1.5 for *MYCN*), gene set enrichment analysis (GSEA; refs. 27, 28) was performed using the C2 canonical pathway (CP) gene sets and seven additional gene sets from The Molecular Signatures Database (MSigDB) that represent MYC activation signatures (Supplementary Table S2). Gene sets with a nominal *P* value of less then 0.05 were considered significant.

To examine the effect of JQ1 on global gene transcription, cell lines were treated with JQ1R or JQ1S (1 μmol/L for 24 hours) and RNA was extracted. Gene expression profiles were assayed using Affymetrix Human Gene 1.0 ST microarrays (Affymetrix). Affymetrix CEL files were normalized using Robust Multi-Array average (RMA) (29). Expression-array data have been deposited in the GEO portal under the accession number GSE51020. Comparative marker selection analysis (30) between JQ1S- and JQ1R-treated cells was performed in GenePattern using the default settings.

The recently described JQ1 consensus signature (16) was applied to the gene expression profiles using the 52 genes identified as being significantly differentially expressed following treatment with JQ1. Agglomerative hierarchical clustering was performed using pairwise complete linkage and a Pearson correlation metric across both samples and genes.

To identify gene sets differentially expressed following treatment with JQ1, GSEA was performed using the same customized C2 (CP) gene sets (MSigDB) with the seven additional MYC activation gene sets. Gene sets with a nominal *P* value of less then 0.05 were considered significant.

In vivo experiments

In vivo efficacy studies were performed in accordance with protocols approved by the IACUC at Stanford University. Briefly, MB002 cells were transduced with a GFP-luciferase lentiviral expression construct and FACS sorted to obtain 30,000 GFP-luciferase–positive cells that were then injected with stereotaxic guidance into the cerebella of 4- to 6-week-old NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ mice (The Jackson Laboratory). To confirm engraftment, mice were administered D-luciferin (75 mg/kg; Promega) and imaged on a Xenogen IVIS2000 (PerkinElmer) 14 days after injection. Mice were randomized into treatment and control groups ($n = 5$ mice/group) and administered JQ1-S (50 mg/kg in 1:10 solution of DMSO:10% cyclodextrin) or vehicle alone (1:10 solution of DMSO:10% cyclodextrin), daily via intraperitoneal injection, until euthanasia was required. Tumor growth was monitored by in vivo imaging systems (IVIS) imaging at 14 and 21 days of treatment. Statistical significance for Kaplan–Meier analysis was determined by the log-rank (Mantel–Cox) test.

For immunohistochemistry analysis of xenografted medulloblastomas, 4- to 6-week-old NSG mice received intracerebellar injections of MB002 cells (30,000 cells) and were administered JQ1-S (50 mg/kg twice daily; $n = 3$) or vehicle ($n = 3$), for five doses and then euthanized. Brains were carefully dissected and either frozen in RNAlater (Qiagen) or preserved in 4% paraformaldehyde and subsequently embedded in paraffin. RNA was extracted from frozen cerebellum using the RNeasy Kit (Qiagen) as per the manufacturer's instructions.

Immunohistochemistry

JQ1- and vehicle-treated MB002 xenografts were harvested, rinsed in PBS, and fixed in 4% paraformaldehyde overnight at 4°C. Then, 5-μm-thick sections were mounted on poly-D-lysine–coated slides and treated with xylene, followed by several changes of graded alcohol. Antigen retrieval was performed by application of citrate buffer pH 6.00 for 20 minutes. Slides were then incubated with anti-Ki67 (Lab Vision; SP6 RM-9106-S, lot 9106S1210D) overnight at 4°C. Cells were washed with several changes of PBS, and secondary antibody conjugated to horseradish peroxidase was applied and detected using the Dako Envision Kit for 3,3′-diaminobenzidine.

Immunofluorescence

Primary medulloblastoma cells (MB002) were cultured in 12-well plates at a density of 1×10^5 cells per well and treated with vehicle, (S)-JQ1 (500 nmol/L and 1 μmol/L for 6 and 12 hours). Cells were centrifuged at 1,000 rpm for 5 minutes, washed in PBS, and mechanically dissociated for 5 minutes at 37°C. Single cell suspensions were transferred to coverslips precoated with poly-L-lysine (10 μg/mL in double-distilled water, catalog number P6516; Sigma-Aldrich) and then fixed with 4% paraformaldehyde for 15 minutes. Immunolabeling was carried out with the antibody anti-BAD (C-7; 1:400, SC8044; Santa Cruz), detected by Cy3-conjugated secondary Donkey anti-mouse antibody (1:200; JacksonImmuno Research) and visualized by confocal fluorescence microscopy (Leica DM5500 B; Leica Microsystems).

For statistical analysis, *P* values were calculated using the Fisher, *t* tests, or Pearson, as appropriate. ANOVA with correction was used for the comparison of multiple groups.

Results

Medulloblastomas exhibit several indicators of MYC pathway activation

We evaluated indicators of *MYC* pathway activation using integrated data sets of genome-wide copy-number estimates from 1,071 medulloblastomas and corresponding gene expression data for 282 medulloblastomas (11). These tumors comprised 79 Wnt subgroup medulloblastomas, 265 SHH subgroup tumors, 168 group 3 medulloblastomas, and 313 group 4 tumors, as determined by exon array or nanoString analysis (11). Subtype designation was not available for 245 tumors. We evaluated copy numbers and expression of all three MYC family members (*MYC*, *MYCN*, and *MYCL1*) and nine signatures of *MYC* pathway activation obtained from the Gene Set Enrichment Database.

We found that 23% of all medulloblastomas exhibit amplifications of one or more of *MYC*, *MYCN*, or *MYCL1* (9%, 12%, and 2% of all tumors, respectively; Fig. 1A). Group 3 tumors exhibit amplifications of MYC family members approximately three times as frequently (44% of cases) as the other subtypes (13%, 19%, and 16% among Wnt, SHH and group 4 tumors, respectively). This enrichment was most profound for *MYC*; group 3 tumors accounted for 70% of medulloblastomas exhibiting *MYC* amplifications. Amplifications of *MYCN* were observed slightly more often in SHH and group 4 tumors (38% and 35% of all *MYCN* amplifications, respectively) than in group 3 tumors (10%).

Amplifications of *MYC* and *MYCN* were anticorrelated ($P < 0.05$; Fig. 1B), suggesting that they have similar functional effects. Anticorrelated genetic alterations often indicate functional redundancies (31–33) because redundant alterations are not required by the same tumor. Although there was a trend toward anticorrelation between amplification of *MYCL1* and either *MYC* or *MYCN*, this did not reach significance ($P = 0.8$ and 0.8, respectively), perhaps because *MYCL1* amplifications were observed so infrequently.

Amplifications of each MYC family member were also associated with increased expression of its gene transcript ($P < 0.0001$ in all cases). Indeed, high expression of *MYCN* or *MYCL1* tends to be found exclusively in medulloblastomas with amplification of those genes. High *MYC* expression, however, is often present in samples without *MYC* amplification, suggesting alternative mechanisms for increased *MYC* expression. Increased expression of *MYC* and *MYCN* was anticorrelated ($P = 0.007$; Fig. 1B), as were *MYC* and *MYCN* amplifications. Increased expression of MYCL1 trended toward a similar anticorrelation with the other MYC family members, but did not reach statistical significance.

Amplification and overexpression of each MYC family member was also associated with increased expression of genes upregulated by *MYC* ($P < 0.05$ in all cases; Fig. 1C). We assessed *MYC* pathway activation by summing the expression levels of 68 previously published genes known to be upregulated by *MYC* (14). We obtained similar results with GSEA using nine signatures of *MYC* activation (Supplementary Table S2) and comparing tumors with high expression of any MYC to tumors with low expression of all MYCs. Among the nine *MYC* activation signatures present in our gene sets, six were significantly associated with high *MYC* expression, five with *MYCN*, and four with *MYCL1*, respectively (Supplementary Table S3).

Medulloblastomas that exhibited indicators of *MYC* activation also exhibited high expression of genes observed to be downregulated with treatment with the BET bromodomain inhibitor JQ1 (16). Genes downregulated by JQ1 were previously identified in tumors from multiple lineages, including multiple myeloma, leukemia, and neuroblastoma (16). We found a positive correlation between the expression of genes that are targeted by JQ1 (and thus opposite to the signature following treatment with JQ1) and MYC activation signatures in medulloblastoma (Fig. 1D). We also observed positive correlations between expression of genes targeted by JQ1 and amplifications of *MYC* ($P = 0.003$) and *MYCN* ($P < 0.05$). The finding that medulloblastomas with indicators of *MYC* activation exhibit gene expression profiles that are opposite to the signature of JQ1 is not surprising, because the JQ1 signature has already been shown to reflect the downregulation of MYC activity (16). Nevertheless, these findings raise the hypothesis that treatment with JQ1 may limit MYC activity and suppress the proliferation of MYC-driven medulloblastomas.

JQ1 reduces cell proliferation in *MYC*- and *MYCN*-amplified medulloblastoma cells

We examined the efficacy of treatment with JQ1S in six patient-derived medulloblastoma cell lines documented to have *MYC* amplification relative to five non-*MYC*–amplified medulloblastoma lines, and human neural stem cells (34, 35). Western immunoblotting of the patient-derived cell lines confirmed increased expression of *MYC* in lines with *MYC* amplification (Supplementary Fig. S1B). No patient-derived *MYCN*-amplified medulloblastoma cell lines have been generated to date; therefore, we evaluated JQ1 activity in the setting of *Mycn* amplification using tumor cells derived from recently developed mouse models of group 4 *MYNC*–amplified medulloblastomas (18, 36). We also evaluated JQ1 activity in a murine model of group 3 MYC–driven medulloblastomas. The activity of JQ1 was initially assessed by comparing proliferation rates in the presence of the active stereoisomer of JQ1 (JQ1S) to proliferation rates in the presence of an inactive stereoisomer, JQ1R (13).

In all *MYC*-amplified patient-derived cell lines, treatment with JQ1S for 48 hours at doses less than 1 μmol/L resulted in 57% to 69% reduction in cell viability compared with treatment with JQ1R (Fig. 2A). Each *MYC*-amplified line had an IC_{50} of less than 1 μmol/L. In contrast, cell viability of non-*MYC*–amplified cell lines (Fig. 2B) and neurons derived from the SVZ (Fig. 2C) was not substantially altered by treatment with either JQ1S or JQ1R. Cells with *MYC* amplification exhibited up to 5-fold increases in cell

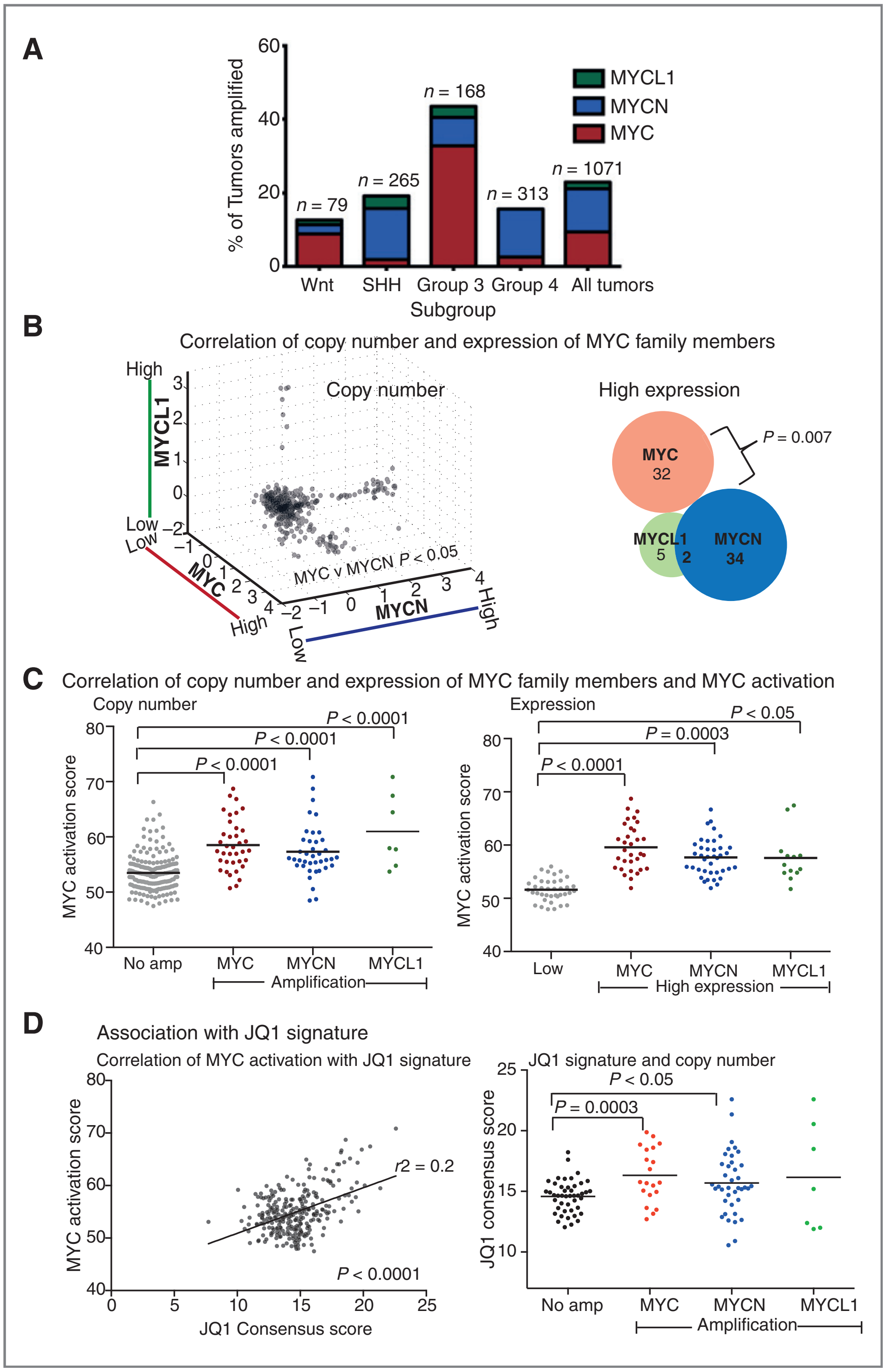


Figure 1. Amplifications of the MYC isoforms *MYC*, *MYCN*, or *MYCL1* are common and mutually exclusive in medulloblastoma. A, fractions of tumors in the Wnt, SHH, group 3, and group 4 subtypes of medulloblastoma with amplifications of *MYC*, *MYCN*, or *MYCL1* in 1,071 medulloblastomas. B, left, scatter plot depicting copy numbers of different MYC isoforms in 1,071 medulloblastoma samples. *P* value depicts significant anticorrelation between *MYC* and *MYCN*. Right, Venn diagram showing medulloblastomas with high expression of MYC (*z* score > 1), MYCN (*z* score > 1.5), or MYCL1 (*z* score > 0.85). *P* value depicts significant anticorrelation between MYC and MYCN. C, left, correlation of MYC activation scores with copy numbers of MYC isoforms and correlation of MYC activation scores with gene expression of MYC isoforms. *P* values depict significant differences in MYC activation scores. ANOVA *P* value < 0.0001. D, left, association between MYC activation and JQ1 consensus score. *P* value depicts significant correlation as determined by the Pearson correlation test. Right, the association between JQ1 consensus score and MYC amplifications. *P* values depict significant differences in JQ1 scores. ANOVA *P* value < 0.05.

number by day 4 of treatment with JQ1R, but showed no increase in cell numbers with JQ1S treatment over the same time period (Supplementary Fig. S2A).

Treatment with JQ1S also reduced the proliferation of cells derived from a murine model of group 3–4 medulloblastoma with *MYCN* overexpression. In these cells, treatment with JQ1 reduced viability by 75% compared with treatment with JQ1R (Fig. 2D). Medulloblastoma cells from a mouse model with *MYC* overexpression exhibited an even greater (91%) sensitivity to treatment with JQ1S (Fig. 2D). Taken together, these data show that JQ1 is efficacious in reducing cell viability in medulloblastoma cell lines driven by *MYC* and also in cells from a murine model of MYCN-driven medulloblastoma, with

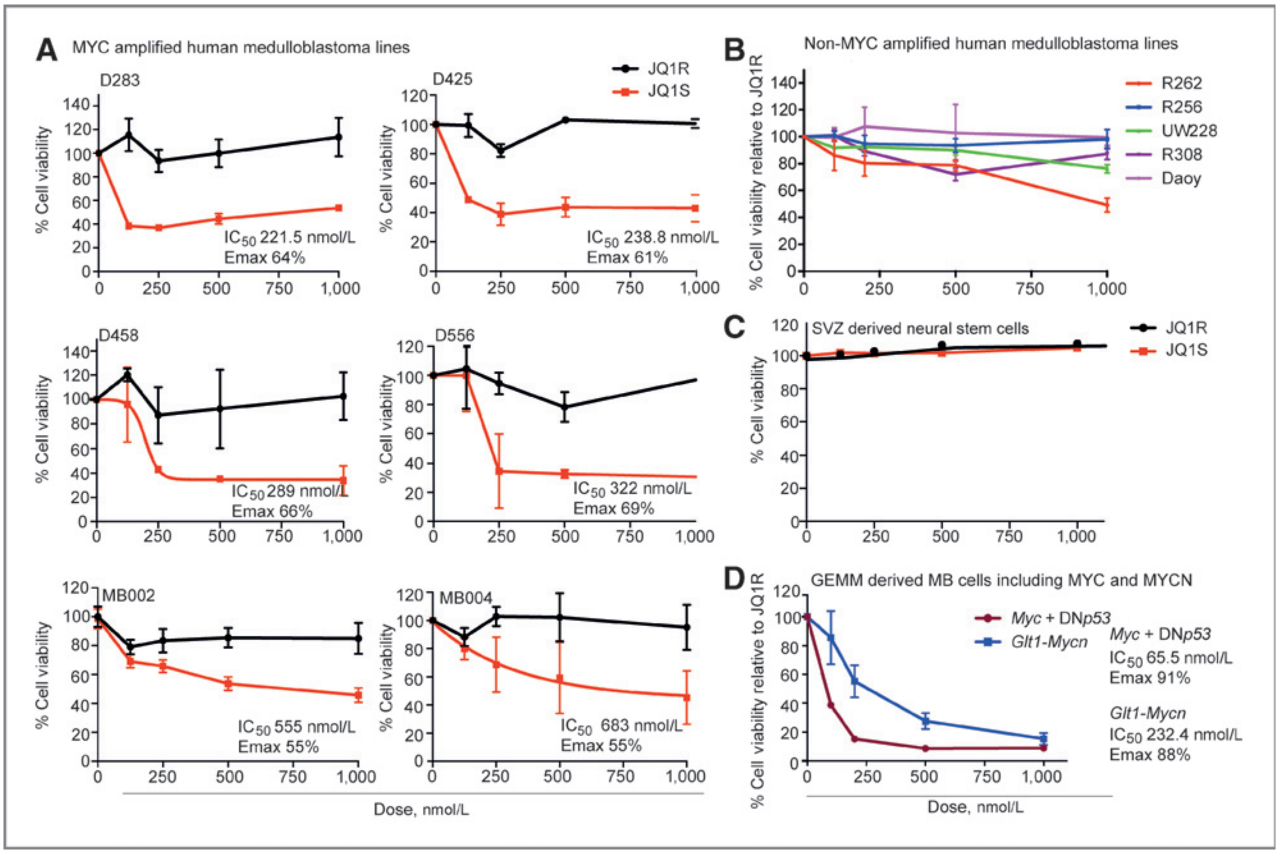


Figure 2. JQ1 reduces cell proliferation in *MYC*-amplified medulloblastoma cell lines and murine models of *MYC*- and *MYCN*-amplified medulloblastoma. A, dose–response curves of patient-derived *MYC*-amplified medulloblastoma cell lines treated with JQ1S and JQ1R for 48 hours. The percentage of cell viability is relative to untreated cells. Values, mean ± SD. B, dose–response curves of patient-derived non-*MYC*–amplified medulloblastoma cell lines treated with JQ1S for 72 hours. The percentage of cell viability is relative to untreated cells. Values, mean ± SD. C, dose–response curves of neural stem cells derived from the SVZ treated with JQ1S and JQ1R. Values, mean ± SD. D, dose–response curves of cancer cells from genetically engineered mice models of group 3 and 4 medulloblastoma with *Myc* or *Mycn* overexpression, after treatment with JQ1S. Values, mean ± SD.

minimal effect observed in non-*MYC*–amplified medulloblastoma cells or neural stem cells.

JQ1 reduces cell viability by inducing G_1 arrest and apoptosis in *MYC*-amplified medulloblastoma cell lines

Treatment with JQ1S significantly affected cell cycling of the MYC-driven cell lines. We profiled cell cycling by flow-cytometry measurement of PI-stained cells treated for 72 hours with either 1 μmol/L of JQ1S or JQ1R. In six patient-derived *MYC*-amplified cell lines, and one *Mycn*-driven GEMM-derived cell line, we observed that treatment with JQ1S resulted in G_1 arrest, with an increased percentage of cells in G_1 and reduction of cells in S phase (Fig. 3A). Across all *MYC*-amplified lines, treatment with JQ1S reduced the number of cells in S phase by approximately 50% compared with JQ1R (21% ± 3% vs. 43% ± 5%; $P < 0.001$) and increased the number of cells in G_1 (73 ± 4% vs. 51 ± 6%; $P < 0.001$; Fig. 3B).

Treatment with JQ1S also induced apoptosis in patient-derived *MYC*-amplified cells. We assessed the induction of apoptosis in cells treated with 1 μmol/L JQ1S or JQ1R for 72 hours by flow cytometry analysis of Annexin V/PI staining (Fig. 3C). MB004 was observed to have an increase in the number of necrotic cells, whereas in all other cell lines, there was a clear increase in apoptosis noted. When the results of all cell lines were pooled (including MB004), JQ1S almost tripled the number of apoptotic cells (8% ± 2% vs. 22.5% ± 4%, $P = 02$; Fig. 3D). As an independent correlation of induction of apoptosis, we treated MB002 cells with JQ1 and observed induction of the apoptotic protein, BAD, within 6 hours of treatment (Fig. 3D. right panel). These data suggest that JQ1S reduces the cell viability of *MYC*-driven medulloblastoma cell lines by inducing G_1 arrest and apoptosis.

BRD4 suppression attenuates expression of MYC in medulloblastoma cells

We hypothesized that the effects of JQ1 in the medulloblastoma cells were mediated in large part by the decreased activity of MYC through inhibition of BRD4. BRD4 has been shown to be a cofactor for MYC-dependent transcription in multiple cell types (13), and JQ1 has previously been shown to have higher affinity for binding domains of BRD4 than other bromodomain-containing proteins (13).

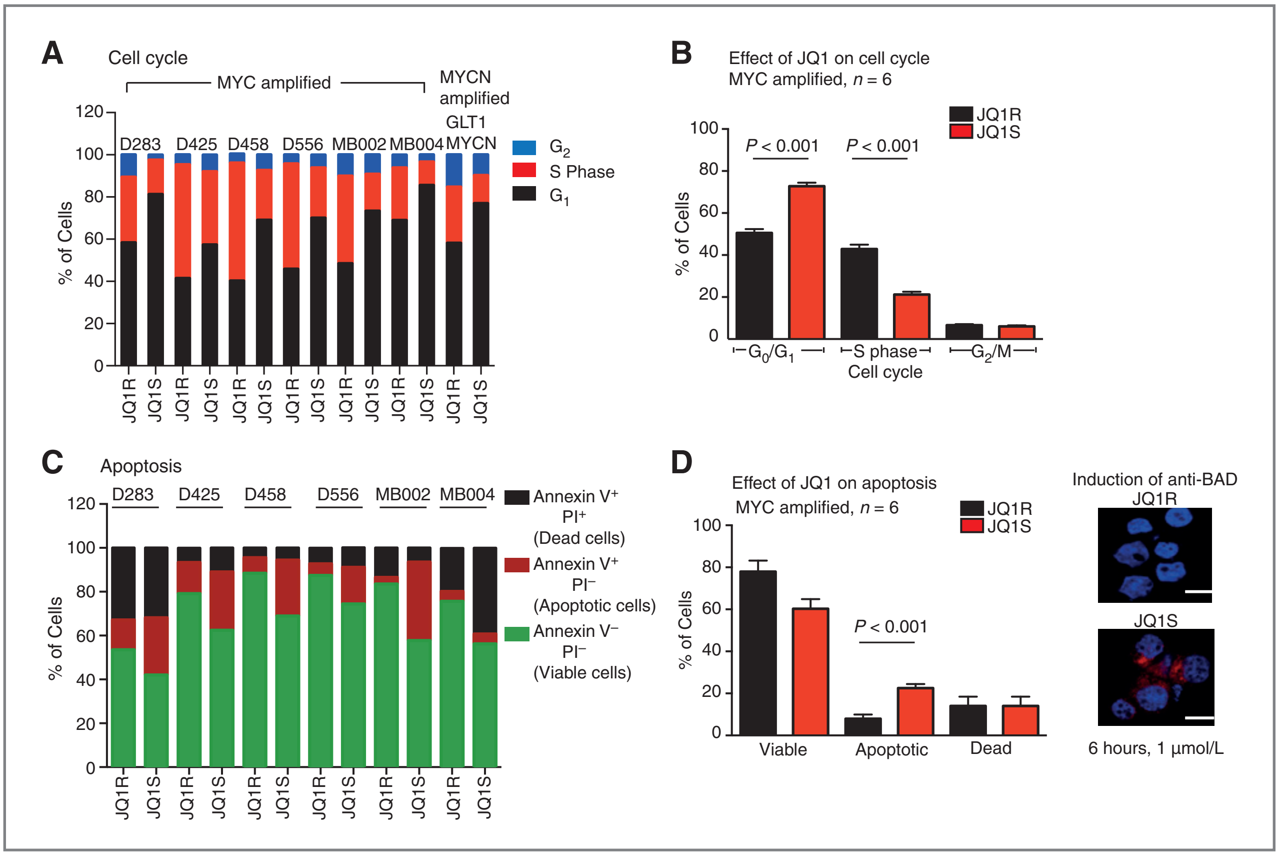


Figure 3. JQ1 induces G_1 arrest and apoptosis in patient-derived *MYC*-amplified medulloblastoma cell lines. A, cell-cycle analysis of *MYC*-amplified patient medulloblastoma cell lines and one *MYCN*-amplified cell line derived from a GEMM treated with 1 μmol/L of JQ1S and JQ1R for 72 hours. B, pooled cell-cycle analysis of *MYC*-amplified medulloblastoma cell lines treated with 1 μmol/L of JQ1S and JQ1R for 72 hours. Values, mean ± SD. ANOVA *P* value < 0.001. C, Annexin V/PI apoptosis assays of *MYC*-amplified medulloblastoma cell lines treated with 1 μmol/L of JQ1S and JQ1R for 72 hours. Viable cells are negative for Annexin V and PI, apoptotic cells are positive for Annexin V and negative for PI, and dead cells are positive for both Annexin V and PI. D, pooled Annexin V/PI apoptosis analysis of the six *MYC*-amplified medulloblastoma cell lines following treatment with 1 μmol/L of JQ1S and JQ1R for 72 hours. Values, mean ± SD. Right, induction of the apoptotic protein, BAD, in the MB002 cell line, 6 hours following treatment with 1 μmol/L of JQ1S. Scale, 20 μm.

We, therefore, suppressed the expression of either *MYC* or *BRD4* in each of the six *MYC*-amplified patient-derived cell lines using *MYC*- and *BRD4*-directed shRNAs. We measured the proliferation and compared the results with controls treated with LacZ-targeted shRNAs. Suppression of MYC and BRD4 protein levels was confirmed by immunoblotting (Fig. 4A and Supplementary Fig. S3).

In all cell lines, suppression of either MYC or BRD4 resulted in greater than 50% reductions in proliferation relative to cells treated with LacZ-targeted shRNAs (Fig. 4B). We observed absolute decreases in cell numbers among four lines with both MYC suppression and BRD4 suppression.

In all cell lines, suppression of BRD4 was also associated with a reduction of MYC mRNA, by an average of 45% relative to the LacZ-suppressed controls ($P < 0.0001$; Fig. 4C). We also observed a reduction in MYC protein levels in cells following suppression of BRD4 (Fig. 4A and Supplementary Fig. S3). This was most pronounced in the D425 and D556 cell lines. Suppression of BRD4 has previously been shown to decrease the MYC expression in other cell types (13). The finding that BRD4 suppression resulted in attenuation of *MYC* mRNA and protein suggested that JQ1 also exerted its effects in part through suppression of MYC.

JQ1 suppresses MYC activation pathways and expression of MYC and MYCN themselves

To determine the transcriptional effects of JQ1 treatment, we performed genome-wide expression profiling of five patient-derived *MYC*-amplified cell lines treated with 1 μmol/L of either JQ1S or JQ1R for 24 hours. We identified 43 genes that were significantly downregulated by JQ1S (false discovery rate (FDR) < 0.1; Supplementary Table S4). No genes were significantly upregulated at this significance threshold.

The JQ1S-treated cells exhibited significant enrichment of a signature of JQ1 treatment derived from multiple myeloma, leukemia, and neuroblastoma cells ($P < 0.0001$; Fig. 5A; 16), and significant attenuation of MYC activity ($P < 0.05$ in all cases; Fig. 5B). An unbiased screen of pathways altered by treatment with JQ1S (using the gene set enrichment algorithm and the C2 (CP) set of signatures)

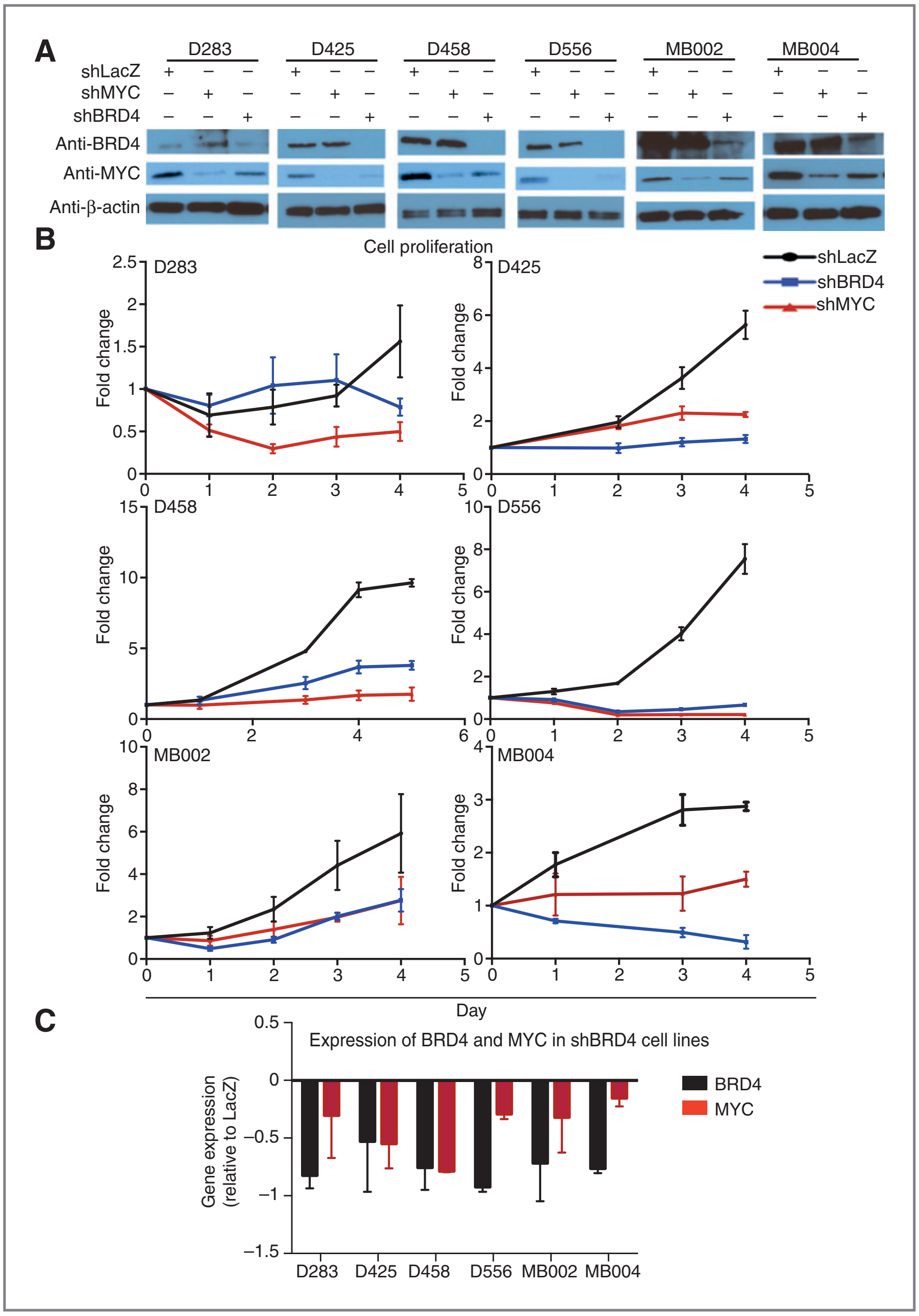


Figure 4. Suppression of BRD4 in patient-derived *MYC*-amplified medulloblastoma cell lines suppresses proliferation and MYC expression. A, immunoblots of BRD, MYC, and β-actin loading controls following infection with shRNAs targeting BRD4, MYC, or LacZ controls. B, cell viability following infection with shRNAs targeting BRD4, MYC, or LacZ. Cell proliferation of *MYC*-amplified medulloblastoma cell lines following suppression of BRD4, MYC, or LacZ. Error bars, mean ± SD of three replicates per condition. C, BRD4 and MYC mRNA levels determined 72 hours after BRD4 suppression in the indicated cell lines. Values, mean ± SD of three replicates per condition, normalized to expression in LacZ controls.

indicated significant alteration of 46 pathways ($P < 0.05$; Supplementary Table S5). Four of these 46 pathways represented MYC activation signatures (Fig. 5B), and were downregulated following treatment with JQ1S (compared with cells treated with JQ1R).

In all patient-derived *MYC*-amplified cell lines, treatment with 1 μmol/L of JQ1S also led to decreased expression of MYC itself. Levels of *MYC* mRNA declined by an average of 46% (range, 29%–78%; $P < 0.0001$; Fig. 5C), and these changes were associated with decreased expression of the MYC protein. We also examined the expression of *MYC* mRNA in three cell lines treated with lower doses of JQ1S, and found attenuation of *MYC* mRNA expression at doses of JQ1S as low as 125 nmol/L (Supplementary Fig. S2B). Treatment of Mycn-overexpressing cells derived from tumors from the Glt1 *Mycn* GEMM model with 0.5 μmol/L of JQ1 for 72 hours was associated with the suppression of *Mycn* mRNA and protein expression (Fig. 5C).

We next overexpressed MYC from an exogenous retroviral promoter in D283 cells and treated them with JQ1 to assess whether forced expression of MYC could rescue

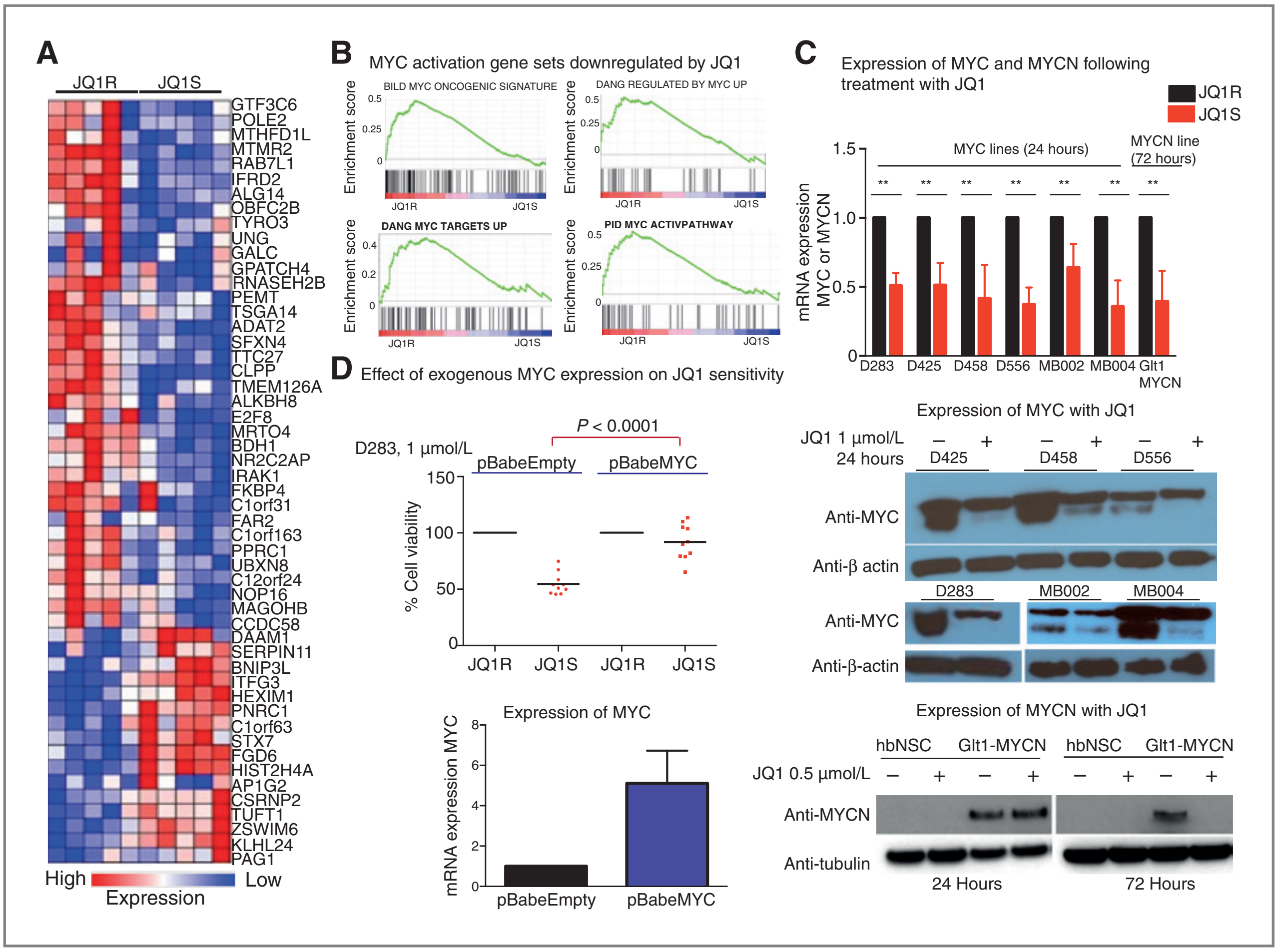


Figure 5. JQ1 treatment downregulates the transcription of MYC, MYCN, and MYC activating pathways. A, expression of genes in the JQ1 consensus signature in *MYC*-amplified medulloblastoma cell lines following a 24-hour treatment with 1 μmol/L of JQ1R or JQ1S. Red indicates increased gene expression and blue indicates reduced gene expression. B, GSEA of gene sets downregulated among genes in medulloblastoma cell lines treated with JQ1S. C, expression of MYC and MYCN mRNA and protein following the treatment with JQ1R or JQ1S. *MYC*-amplified cells were treated with 1 μmol/L of JQ1R or JQ1S for 24 hours and MYCN-derived GEM lines were treated with 0.5 μmol/L of JQ1S over the indicated time course. mRNA expression values are normalized to housekeeping control and expression is depicted relative to JQ1R controls. Values, mean ± SD of three replicates per condition. **, $P < 0.05$. Immunoblots for MYC and β-actin loading controls following the 24-hour treatment with 1 μmol/L of JQ1S in the indicated cell lines or for MYCN and β-tubulin loading controls treated with 0.5 μmol/L of JQ1S over the indicated time course are shown. Lysates from hindbrain neural stems cells are shown as a negative control for MYCN expression. D, cell viability of D283 infected with pBabeEmpty or pBabeMYC, which were treated with 1 μmol/L of JQ1R or JQ1S 6 hours after infection. Cell viability is shown at 24 hours following treatment with JQ1R or JQ1S. Bars, mean of 10 replicates (individual values shown as scatter plot). Right, quantitative PCR for mRNA expression of MYC in D283 cells infected with retrovirus for pBabeEmpty or pBabeMYC. Expression is shown relative to β-actin controls and normalized to expression for pBabeEmpty. Values, mean ± SD of six replicates.

the cells from the effects of JQ1 (Fig. 5D). In cells infected with the empty pBabe retroviral vector, we observed reduced cell viability following treatment with JQ1S. In contrast, in cells infected with pBabe MYC, we observed minimal effect on cell viability following treatment with JQ1S.

Taken together, these results confirm that BET bromodomain inhibition with JQ1 results in the downregulation of MYC and MYC activation pathways, resulting in reduced cell proliferation. JQ1 treatment also reduces MYC and MYCN expression in patient-derived *MYC*-amplified medulloblastoma cell lines and in cells derived from a *MYCN-driven* medulloblastoma mouse model. Forced expression of MYC from an exogenous promoter rescues D283 cells from the effects of JQ1S.

Treatment with JQ1 prolongs survival of *MYC*-amplified medulloblastoma xenografts

We examined the efficacy of JQ1 *in vivo*, using orthotopic xenografts generated by intracerebellar injections of MB002 cells. Compared with vehicle control, mice undergoing daily treatment with JQ1 (50 mg/kg/d) exhibited a significantly prolonged survival (Fig. 6A) with a slower rate of tumor growth at 14 and 21 days postinjection as indicated by bioilluminescence (Fig. 6B). This was statistically significant on day 14 (*P*, 0.03). A cohort of

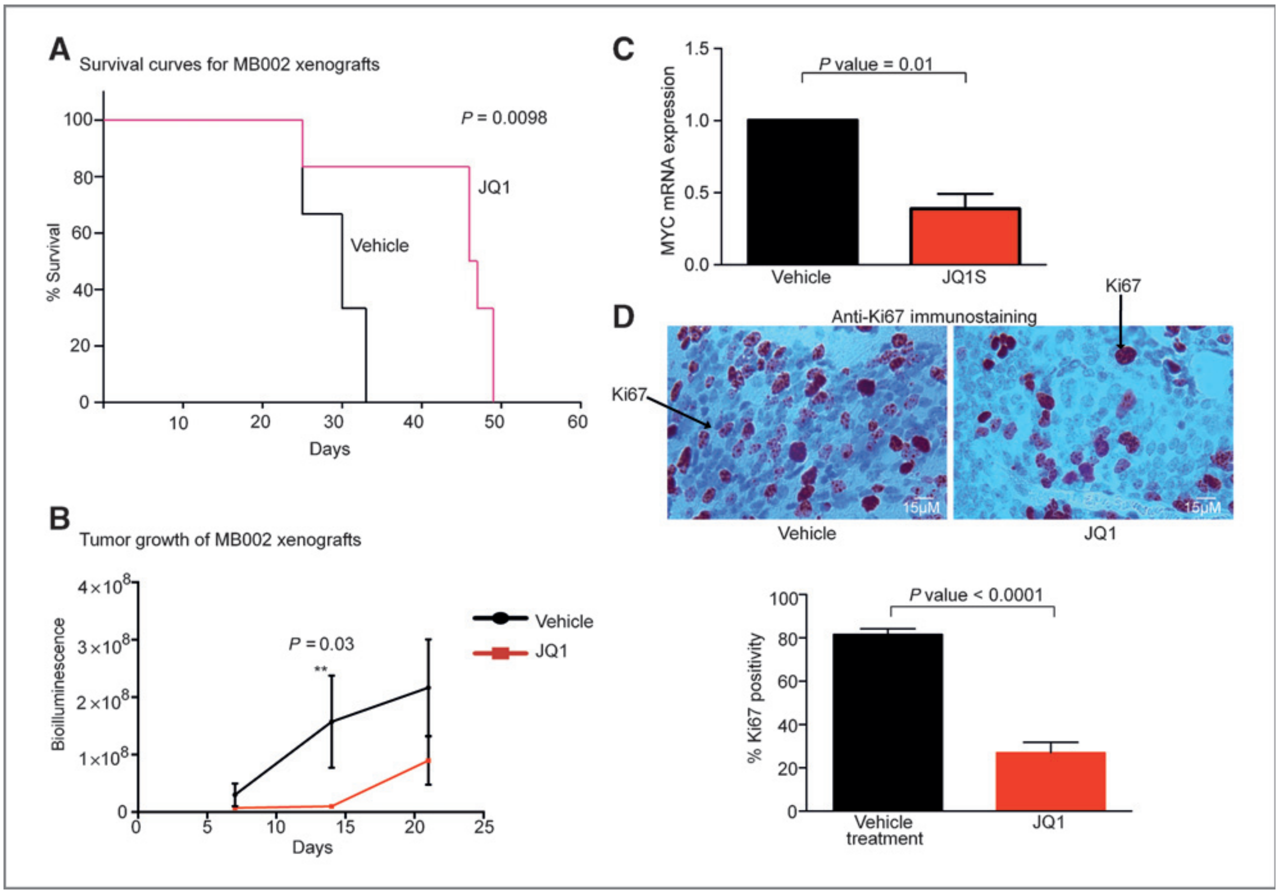


Figure 6. JQ1 prolongs survival of xenograft models of group 3 medulloblastoma. A, Kaplan–Meier curves of mice with intracranial xenograft injections of MB002 treated with JQ1 or vehicle. B, rate of tumor growth in NSG mice with orthotopic MB002 xenografts following treatment with JQ1 as monitored by bioilluminescence detection. C, quantitative PCR of MYC mRNA levels (normalized to β-actin) in cerebellar xenografts of MB002 following treatment with either vehicle control or JQ1. Values, mean ± SD of six replicate measurements. D, representative images of Ki67 staining of MB002 xenografts treated with vehicle control and JQ1 (top). The percentage of Ki67 positivity in tumors from mice treated with either vehicle control ($n = 2$) or JQ1 ($n = 3$). Values, mean ± SD of three replicate measurements for each animal.

mice was sacrificed following five doses of JQ1 or vehicle treatment. We extracted RNA from the cerebellum of vehicle- or JQ1-treated mice and assessed the expression of *MYC* mRNA (Fig. 6C). We observed significant downregulation of *MYC* mRNA expression following treatment with JQ1. Ki67 immunostaining confirmed a reduced proliferative index in tumors treated with JQ1 compared with vehicle controls (82% vehicle vs. 27% JQ1 treated, $P < 0.0001$; Fig. 6D).

Discussion

Our data provide the first preclinical evidence for a potential therapeutic role of BET bromodomain inhibition for *MYC*-amplified medulloblastoma. Furthermore, we show redundancy between amplification of the different *MYC* genes. We observed responsiveness to JQ1 therapy of all patient cell lines harboring *MYC* amplifications, and in a murine cell line of a *Mycn*-driven medulloblastoma model.

MYC-amplified medulloblastomas are characteristically resistant to conventional treatments, including radiotherapy and chemotherapy. In this study, the use of JQ1 as a single agent was able to reduce cell proliferation and tumor growth and prolong the survival of mice with intracranial xenografts. Thus, our data provide rationale for the development of clinical trials to assess the role of these agents in the treatment of patients with medulloblastoma.

Ongoing work, however, is required to determine optimal strategies to incorporate these agents into upfront therapy for children with newly diagnosed *MYC*-amplified medulloblastoma. Specifically, further work is required to assess strategies to combine BET bromodomain inhibition with other treatment modalities such as radiotherapy and chemotherapy. Specific predictors of sensitivity and resistance to JQ1 remain to be elucidated.

The anticorrelation between amplifications of *MYC* and *MYCN*, and the correlations between these amplifications (and amplifications of *MYCL1*) and expression of genes upregulated by MYC, indicate that these amplifications are associated with increased MYC activity. However, some samples without *MYC* amplification had high expression of *MYC* or of other genes upregulated by MYC. This observation suggests that other factors are also involved in the regulation of MYC activation pathways, and that some tumors without *MYC* amplification may benefit from therapeutics targeting MYC, including BET bromodomain inhibition.

Although we observed minimal responsiveness of the non-MYC–amplified medulloblastoma cell lines to JQ1, it is possible that these lines may not represent the full spectrum of non-*MYC*--amplified medulloblastomas. For example, none of our non-*MYC*–amplified lines included either the Wnt or SHH subtypes. We are unable to speculate whether these subtypes may also have sensitivity to treatment with BET bromodomain inhibition.

Importantly, the inhibition of MYC family members and activation pathways may result in downregulation of other cell-signaling pathways in tumors that do not harbor amplification of MYC isoforms. This is particularly relevant in the SHH subtype. Both *MYC* (37) and *MYCN* (38–40) have been reported to interact with and regulate transcription factors involved in SHH signaling, raising the possibility that inhibition of MYC activation may also result in inhibition of the SHH pathway, and other pathways important in the pathogenesis of medulloblastoma.

We show that *MYC*-amplified medulloblastomas can be targeted by epigenetic modulation of oncogenes. BRD4 is an epigenetic reader that binds to ε-N-lysine acetylation motifs (41). It is increasingly recognized that genomic alterations may result in global epigenetic dysregulation in pediatric brain tumors (42–44). Indeed, group 3 medulloblastomas have been found to harbor somatic copy-number alterations of genes involved in chromatin modeling (45, 46). These alterations frequently affect genes of modifiers of H3K27 methylation (46). Thus, it is possible that targeting other chromatin modifiers may also have activity in this group of tumors. Further work is required to define the interplay between methylation and acetylation of these histone marks to help determine whether the combination of different classes of epigenetic modulators may improve their efficacy.

We were unable to obtain any patient-derived medulloblastoma cell lines that harbor amplifications of either *MYCN* or *MYCL1*. Our integrative analysis of copy number and expression profiles suggest that both isoforms are associated with increased expression of MYC activation pathways, suggesting that BET bromodomain inhibition may be effective in tumors that harbor these amplifications. Indeed, BET bromodomain inhibition has been shown to suppress MYCN expression in patient-derived neuroblastoma cell lines (16). We have shown efficacy of JQ1 treatment in a cell line generated from a *MYCN*-driven mouse model; however, ideally this should be validated in patient-derived *MYCN*-amplified medulloblastoma cell lines.

In summary, we show that BET bromodomain inhibition results in the suppression of MYC activation pathways in MYC-driven medulloblastoma models, resulting in reduced cell viability, induction of G_1 arrest, and apoptosis with prolongation of survival in xenograft models. These data provide a rationale for early-phase clinical trials for BET bromodomain inhibitors for children with this aggressive disease.

Disclosure of Potential Conflicts of Interest

R. Beroukhim is a consultant/advisory board member of Novartis.

The Dana-Farber Cancer Institute has licensed drug-like derivatives of JQ1 prepared in the Bradner laboratory to Tensha Therapeutics for clinical translation as cancer therapeutics. Dana-Farber and Dr. Bradner have been provided minority equity shares in Tensha. Dr Qi has a consultant agreement with Tensha Therapeutics.

Authors' Contributions

Conception and design: P. Bandopadhayay, G. Bergthold, J. Qi, R. Wechsler-Reya, M.W. Kieran, R. Beroukhim, Y.-J. Cho
Development of methodology: P. Bandopadhayay, S. Gholamin, S. Bolin, Y.-J. Cho
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): P. Bandopadhayay, G. Bergthold, B. Nguyen, S. Gholamin, Y. Tang, R. Zeid, S. Masoud, N. Vue, W.J. Gibson, S. Mitra, S. Cheshier, K.-W. Liu, F.J. Swartling, Y.-J. Cho
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): P. Bandopadhayay, G. Bergthold, Y. Tang, S. Bolin, S.E. Schumacher, R. Zeid, S. Masoud, N. Vue, B.R. Paolella, S. Mitra, S. Cheshier, R. Wechsler-Reya, F.J. Swartling, J.E. Bradner, R. Beroukhim, Y.-J. Cho
Writing, review, and/or revision of the manuscript: P. Bandopadhayay, G. Bergthold, B.R. Paolella, S. Cheshier, R. Wechsler-Reya, W.A. Weiss, M.W. Kieran, J.E. Bradner, R. Beroukhim, Y.-J. Cho
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): P. Bandopadhayay, G. Bergthold, B. Nguyen, S. Schubert, S. Gholamin, F. Yu, J. Qi, R. Beroukhim, Y.-J. Cho
Study supervision: P. Bandopadhayay, S. Mitra, R. Beroukhim, Y.-J. Cho

Acknowledgments

The authors thank Vida Shokoohi and John Coller of the Stanford Functional Genomics Core, Christopher DeHeer and Karen Krantz at NanoString Technologies, Inc., Steve Avolicino at Histo-Tec Laboratory, and John Daley and Suzan Lazo-Kallanian of the DFCI Flow Cytometry Core Facility for their technical assistance. The authors also thank Drs. Scott Pomeroy and Tenley Archer, Boston Children's Hospital (Boston, MA), for their assistance and sharing of reagents.

Grant Support

This work was financially supported by the following sources: St. Baldrick's Foundation Scholar Award (Y.-J. Cho); Beirne Faculty Scholar Endowment (Y.-J. Cho); NIH U01-CA176287 (Y.-J. Cho and W.A. Weiss); Stanford Center for Children's Brain Tumors (Y.-J. Cho and S. Cheshier); Pediatric Low-Grade Astrocytoma Foundation (P. Bandopadhayay, G. Bergthold, M.W. Kieran, and R. Beroukhim); Friends of DFCI (P. Bandopadhayay); the Sontag Foundation (R. Beroukhim); Gray Matters Foundation (R. Beroukhim); Stop&Shop Pediatric Brain Tumor Program (P. Bandopadhayay and M.W. Kieran); the Mill Foundation for Kids (M.W. Kieran); Men's Collaborative for Women's Cancers (J.E. Bradner and R. Beroukhim); Damon-Runyon Cancer Research Foundation (J. Qi and J.E. Bradner); Nuovo-Soldati Foundation (G. Bergthold); Philippe Foundation (G. Bergthold); R01-CA159859 (W.A. Weiss and R. Wechsler-Reya); NIH R01-CA148699 (W.A. Weiss); R01-CA133091 (W.A. Weiss); Swedish Childhood Cancer Foundation (S. Bolin and F.J. Swartling); the Swedish Cancer Society (S. Bolin and F.J. Swartling); the Swedish Research Council (S. Bolin and F.J. Swartling); the Swedish Society of Medicine (S. Bolin and F.J. Swartling); the Swedish Brain Foundation (S. Bolin and F.J. Swartling); éke Wiberg's Foundation (S. Bolin and F.J. Swartling); and the Association for International Cancer Research (S. Bolin and F.J. Swartling). R. Wechsler-Reya is the recipient of a Leadership Award (LA1-01747) from the California Institute of Regenerative Medicine (San Francisco, CA).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received August 19, 2013; revised November 6, 2013; accepted November 26, 2013; published OnlineFirst December 2, 2013.

References

1. Kieran MW, Walker D, Frappaz D, Prados M. Brain tumors: from childhood through adolescence into adulthood. J Clin Oncol 2010;28:4783–9.
2. Packer RJ, Gajjar A, Vezina G, Rorke-Adams L, Burger PC, Robertson PL, et al. Phase III study of craniospinal radiation therapy followed by adjuvant chemotherapy for newly diagnosed average-risk medulloblastoma. J Clin Oncol 2006;24:4202–8.
3. Mulhern RK, Palmer SL, Merchant TE, Wallace D, Kocak M, Brouwers P, et al. Neurocognitive consequences of risk-adapted therapy for childhood medulloblastoma. J Clin Oncol 2005;23:5511–9.
4. Packer RJ, Zhou T, Holmes E, Vezina G, Gajjar A. Survival and secondary tumors in children with medulloblastoma receiving radiotherapy and adjuvant chemotherapy: results of Children's Oncology Group trial A9961. Neuro Oncol 2013;15:97–103.
5. Kool M, Korshunov A, Remke M, Jones DT, Schlanstein M, Northcott PA, et al. Molecular subgroups of medulloblastoma: an international meta-analysis of transcriptome, genetic aberrations, and clinical data of WNT, SHH, Group 3, and Group 4 medulloblastomas. Acta Neuropathol 2012;123:473–84.
6. Cho YJ, Tsherniak A, Tamayo P, Santagata S, Ligon A, Greulich H, et al. Integrative genomic analysis of medulloblastoma identifies a molecular subgroup that drives poor clinical outcome. J Clin Oncol 2011;29:1424–30.
7. Northcott PA, Korshunov A, Witt H, Hielscher T, Eberhart CG, Mack S, et al. Medulloblastoma comprises four distinct molecular variants. J Clin Oncol 2011;29:1408–14.
8. Kool M, Koster J, Bunt J, Hasselt NE, Lakeman A, van Sluis P, et al. Integrated genomics identifies five medulloblastoma subtypes with distinct genetic profiles, pathway signatures and clinicopathological features. PLoS ONE 2008;3:e3088.
9. Thompson MC, Fuller C, Hogg TL, Dalton J, Finkelstein D, Lau CC, et al. Genomics identifies medulloblastoma subgroups that are enriched for specific genetic alterations. J Clin Oncol 2006;24:1924–31.
10. Taylor MD, Northcott PA, Korshunov A, Remke M, Cho YJ, Clifford SC, et al. Molecular subgroups of medulloblastoma: the current consensus. Acta Neuropathol 2012;123:465–72.
11. Northcott PA, Shih DJ, Peacock J, Garzia L, Morrissy AS, Zichner T, et al. Subgroup-specific structural variation across 1,000 medulloblastoma genomes. Nature 2012;488:49–56.
12. Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, et al. The landscape of somatic copy-number alteration across human cancers. Nature 2010;463:899–905.
13. Delmore JE, Issa GC, Lemieux ME, Rahl PB, Shi J, Jacobs HM, et al. BET bromodomain inhibition as a therapeutic strategy to target c-Myc. Cell 2011;146:904–17.
14. Raeder MB, Birkeland E, Trovik J, Krakstad C, Shehata S, Schumacher S, et al. Integrated genomic analysis of the 8q24 amplification in endometrial cancers identifies ATAD2 as essential to MYC-dependent cancers. PLoS ONE 2013;8:e54873.
15. Filippakopoulos P, Knapp S. The bromodomain interaction module. FEBS Lett 2012;586:2692–704.
16. Puissant A, Frumm SM, Alexe G, Bassil CF, Qi J, Chanthery YH, et al. Targeting MYCN in neuroblastoma by BET bromodomain inhibition. Cancer Discov 2013;3:308–23.
17. Herrmann H, Blatt K, Shi J, Gleixner KV, Cerny-Reiterer S, Mullauer L, et al. Small-molecule inhibition of BRD4 as a new potent approach to eliminate leukemic stem- and progenitor cells in acute myeloid leukemia AML. Oncotarget 2012;3:1588–99.
18. Pei Y, Moore CE, Wang J, Tewari AK, Eroshkin A, Cho YJ, et al. An animal model of MYC-driven medulloblastoma. Cancer Cell 2012;21:155–67.
19. Swartling FJ, Savov V, Persson AI, Chen J, Hackett CS, Northcott PA, et al. Distinct neural stem cell populations give rise to disparate brain tumors in response to N-MYC. Cancer Cell 2012;21:601–13.
20. Weeraratne SD, Amani V, Teider N, Pierre-Francois J, Winter D, Kye MJ, et al. Pleiotropic effects of miR-183 $\sim$ 96 $\sim$ 182 converge to regulate cell survival, proliferation and migration in medulloblastoma. Acta Neuropathol 2012;123:539–52.
21. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. Proc Natl Acad Sci U S A 2001;98:31–6.
22. Li C, Hung Wong W. Model-based analysis of oligonucleotide arrays: model validation, design issues and standard error application. Genome Biol 2001;2:RESEARCH0032.
23. Korn JM, Kuruvilla FG, McCarroll SA, Wysoker A, Nemesh J, Cawley S, et al. Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. Nat Genet 2008;40:1253–60.
24. Venkatraman ES, Olshen AB. A faster circular binary segmentation algorithm for the analysis of array CGH data. Bioinformatics 2007;23:657–63.
25. Olshen AB, Venkatraman ES, Lucito R, Wigler M. Circular binary segmentation for the analysis of array-based DNA copy number data. Biostatistics 2004;5:557–72.
26. Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, Zack T, et al. Absolute quantification of somatic DNA alterations in human cancer. Nat Biotechnol 2012;30:413–21.
27. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545–50.
28. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 2003;34:267–73.
29. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 2003;19:185–93.
30. Gould J, Getz G, Monti S, Reich M, Mesirov JP. Comparative gene marker selection suite. Bioinformatics 2006;22:1924–5.
31. Ciriello G, Cerami E, Sander C, Schultz N. Mutual exclusivity analysis identifies oncogenic network modules. Genome Res 2012;22:398–406.
32. Vaske CJ, Benz SC, Sanborn JZ, Earl D, Szeto C, Zhu J, et al. Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using PARADIGM. Bioinformatics 2010;26:i237–45.
33. Vandin F, Upfal E, Raphael BJ. Algorithms for detecting significantly mutated pathways in cancer. J Comput Biol 2011;18:507–22.
34. Friedman HS, Burger PC, Bigner SH, Trojanowski JQ, Wikstrand CJ, Halperin EC, et al. Establishment and characterization of the human medulloblastoma cell line and transplantable xenograft D283 Med. J Neuropathol Exp Neurol 1985;44:592–605.
35. Aldosari N, Wiltshire RN, Dutra A, Schrock E, McLendon RE, Friedman HS, et al. Comprehensive molecular cytogenetic investigation of chromosomal abnormalities in human medulloblastoma cell lines and xenograft. Neuro Oncol 2002;4:75–85.
36. Swartling FJ, Grimmer MR, Hackett CS, Northcott PA, Fan QW, Goldenberg DD, et al. Pleiotropic role for MYCN in medulloblastoma. Genes Dev 2010;24:1059–72.
37. Yoon JW, Gallant M, Lamm ML, Iannaccone S, Vieux KF, Proytcheva M, et al. Noncanonical regulation of the hedgehog mediator GLI1 by c-MYC in Burkitt lymphoma. Mol Cancer Res 2013;11:604–15.
38. Colvin Wanshura LE, Galvin KE, Ye H, Fernandez-Zapico ME, Wetmore C. Sequential activation of Snail1 and N-Myc modulates sonic hedgehog-induced transformation of neural cells. Cancer Res 2011;71:5336–45.
39. Kenney AM, Cole MD, Rowitch DH. Nmyc upregulation by sonic hedgehog signaling promotes proliferation in developing cerebellar granule neuron precursors. Development 2003;130:15–28.
40. Oliver TG, Grasfeder LL, Carroll AL, Kaiser C, Gillingham CL, Lin SM, et al. Transcriptional profiling of the Sonic hedgehog response: a critical role for N-myc in proliferation of neuronal precursors. Proc Natl Acad Sci U S A 2003;100:7331–6.
41. Filippakopoulos P, Picaud S, Mangos M, Keates T, Lambert JP, Barsyte-Lovejoy D, et al. Histone recognition and large-scale structural analysis of the human bromodomain family. Cell 2012;149:214–31.

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

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

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

45. Jones DT, Northcott PA, Kool M, Pfister SM. The role of chromatin remodeling in medulloblastoma. Brain Pathol 2013;23:193–9.

46. Dubuc AM, Remke M, Korshunov A, Northcott PA, Zhan SH, Mendez-Lago M, et al. Aberrant patterns of H3K4 and H3K27 histone lysine methylation occur across subgroups in medulloblastoma. Acta Neuropathol 2013;125:373–84.

2) Epigenetic targeting of Hedgehog pathway transcriptional output through BET bromodomain inhibition

Tang Y, Gholamin S, Schubert S, Willardson MI, Lee A, Bandopadhayay P, Bergthold G, Masoud S, Nguyen B, Vue N, Balansay B, Yu F, Oh S, Woo P, Chen S, Ponnuswami A, Monje M, Atwood SX, Whitson RJ, Mitr S, Cheshier SH, Qi J, Beroukhim R, Tang JY, Wechsler-Reya R, Oro AE, Link BA, Bradner JE, Cho YJ ;

Nature Med. 2014 Jul;20(7):732-40

Epigenetic targeting of Hedgehog pathway transcriptional output through BET bromodomain inhibition

Yujie Tang[1,2], Sharareh Gholamin[2], Simone Schubert[1,2], Minde I Willardson[3], Alex Lee[4], Pratiti Bandopadhayay[5–8], Guillame Bergthold[5–7], Sabran Masoud[1], Brian Nguyen[1], Nujsaubnusi Vue[1], Brianna Balansay[1], Furong Yu[1,2], Sekyung Oh[1,2], Pamelyn Woo[1], Spenser Chen[1], Anitha Ponnuswami[1], Michelle Monje[1], Scott X Atwood[4], Ramon J Whitson[4], Siddhartha Mitra[2], Samuel H Cheshier[2], Jun Qi[9], Rameen Beroukhim[5,8,9], Jean Y Tang[4], Rob Wechsler-Reya[10], Anthony E Oro[4], Brian A Link[3], James E Bradner[5,8,9] & Yoon-Jae Cho[1,2,11]

Hedgehog signaling drives oncogenesis in several cancers, and strategies targeting this pathway have been developed, most notably through inhibition of Smoothened (SMO). However, resistance to Smoothened inhibitors occurs by genetic changes of Smoothened or other downstream Hedgehog components. Here we overcome these resistance mechanisms by modulating *GLI* transcription through inhibition of bromo and extra C-terminal (BET) bromodomain proteins. We show that BRD4 and other BET bromodomain proteins regulate *GLI* transcription downstream of SMO and suppressor of fused (SUFU), and chromatin immunoprecipitation studies reveal that BRD4 directly occupies *GLI1* and *GLI2* promoters, with a substantial decrease in engagement of these sites after treatment with JQ1, a small-molecule inhibitor targeting BRD4. Globally, genes associated with medulloblastoma-specific GLI1 binding sites are downregulated in response to JQ1 treatment, supporting direct regulation of GLI activity by BRD4. Notably, patient- and GEMM (genetically engineered mouse model)-derived Hedgehog-driven tumors (basal cell carcinoma, medulloblastoma and atypical teratoid rhabdoid tumor) respond to JQ1 even when harboring genetic lesions rendering them resistant to Smoothened antagonists. Altogether, our results reveal BET proteins as critical regulators of Hedgehog pathway transcriptional output and nominate BET bromodomain inhibitors as a strategy for treating Hedgehog-driven tumors with emerged or *a priori* resistance to Smoothened antagonists.

The Hedgehog (Hh) pathway is an evolutionarily conserved signaling axis that directs embryonic patterning through strict temporal and spatial regulation of cell proliferation and differentiation[1]. Developmental aberrations in Hh signaling result in dysmorphology, such as cyclopism, holoprosencephaly and limb deformity, when its output is absent or decreased[2] and in cancer predisposition, as is seen in nevoid basal cell carcinoma syndrome (Gorlin syndrome)[3], when its output is increased or unchecked[1,4].

In canonical Hh signaling, several morphogens (sonic hedgehog (SHH), Indian hedgehog (IHH) and desert hedgehog (DHH))[5,6] have been identified that bind to the multipass cell-surface receptor Patched (PTCH1)[1]. When not bound by Hh ligand, PTCH1 inhibits the G protein–coupled receptor, SMO[7]. Once bound by ligand, however, PTCH1 no longer inhibits SMO, allowing SMO to positively regulate mobilization of the otherwise latent zinc finger transcription factor GLI2, residing in the cilia, to the nucleus, where GLI2 transactivates the *GLI1* promoter[8–10]. GLI1 and GLI2 directly transactivate transcription of Hh target genes, several of which are involved in proliferation, such as *MYCN* and *CCND1* (ref. 11). GLI1 also serves to amplify the output of Hh signaling in a positive feedback loop by activating transcription of *GLI2*, albeit indirectly[12]. Ultimately, the transcriptional programs mediated by Hh signaling orchestrate an array of events based on cellular, temporal and spatial context, with perhaps the most phenotypically consequential event being an increase in cell proliferation.

Inappropriate activation of Hh signaling results in tumor formation in several tissue lineages, including skin, brain, muscle, breast and pancreas[13–15]. The tumors most commonly associated with aberrant Hh signaling are basal cell carcinoma (BCC) and medulloblastoma, given their prevalence in individuals with germline mutations in PTCH1 (Gorlin syndrome)[3,4]. However, the overwhelming majority of Hh-driven BCCs and medulloblastomas activate Hh signaling through sporadic somatic mutations in *PTCH1* or other components of the Hh pathway[14,16,17]. These include activating mutations in *SMO*

---

[1]Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, California, USA. [2]Department of Neurosurgery, Stanford University School of Medicine, Stanford, California, USA. [3]Department of Cell Biology, Neurobiology and Anatomy, Medical College of Wisconsin, Milwaukee, Wisconsin, USA. [4]Program in Epithelial Biology, Stanford University School of Medicine, Stanford, California, USA. [5]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [6]Pediatric Neuro-oncology, Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [7]Division of Pediatric Hematology/Oncology, Boston Children's Hospital, Boston, Massachusetts, USA. [8]Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA. [9]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [10]Sanford-Burnham Medical Research, La Jolla, California, USA. [11]Stanford Cancer Institute, Stanford University Medical Center, Stanford, California, USA. Correspondence should be addressed to Y.-J.C. (yjcho1@stanford.edu).

Received 18 March; accepted 29 May; published online 29 June 2014; doi:10.1038/nm.3613

or inactivating mutations in *SUFU*, which negatively regulates Hh output downstream of SMO[17,18]. Genomic amplification of *GLI2*, and more rarely *GLI1*, has also been reported and is associated with a more aggressive clinical course[16,19–21]. In addition, noncanonical activation of the Hh pathway can occur through loss of SMARCB1, a component of the SWI/SNF chromatin remodeling complex, which results in derepression of transcriptional activity at the *GLI1* locus in malignant rhabdoid tumors[22]. Similarly, the EWS-FLI fusion oncogene responsible for Ewing sarcoma has been shown to directly transactivate the *GLI1* promoter[23].

The identification of SMO as the main pharmacological target of cyclopamine[24], a natural compound found in wild corn lily (*Veratrum californicum*)[2], fostered the development of clinically optimized compounds with potent activity against SMO[25–27]. Some of these compounds have shown clinical efficacy against BCC, medulloblastoma and other cancers[28–30]. However, emergence of resistance and *a priori* resistance have been encountered[25,29,31], prompting investigations into alternate strategies targeting new sites on SMO and Hh pathway components downstream of SMO[32,33] or signaling pathways that cooperate with Hh activation in development and disease[25,34,35]. High-throughput screens have also identified scaffolds that regulate GLI processing and its translocation to or from the cilia and nucleus[36]. However, the effectiveness of these strategies against Hh-driven cancers with *MYCN* amplification, such as SHH-subtype medulloblastomas, is unclear, as MYCN appears to be epistatic to the targets of many of these drugs.

A new class of drugs targeting BET bromodomain proteins (BRD2–BRD4 and BRDT) was described recently[37]. Bromodomains recognize and bind to ε-*N*-lysine acetylation motifs on open chromatin, such as those found on K27 residues of H3 histone N-terminal tails[38,39]. The BET proteins also interact with the positive transcription elongation factor (P-TEFb)[40,41] and phosphorylate Ser2 of RNA polymerase II (PolII), facilitating gene transcription at 'super-enhancer' sites across the genome[42,43]. BRD-containing complexes that bind at these super-enhancer sites often localize to promoter regions of key transcription factors such as MYC, and disruption of these complexes by BET inhibitors has produced substantial responses in mice bearing xenografts of treatment-refractory cancers driven by MYC and other previously 'untargetable' oncogenes, with limited or no toxicity to normal tissues[44–47].

Here we aimed to identify whether inhibition of BET bromodomain proteins could provide a strategy for treating Hh-driven tumors, including those resistant to SMO antagonists. We provide evidence that BRD4 is a critical regulator of *GLI1* and *GLI2* transcription through direct occupancy of their promoters. Furthermore, we show that occupancy of *GLI1* and *GLI2* promoters by BRD4 and transcriptional activation at cancer-specific GLI promoter–binding sites are markedly inhibited by the BET inhibitor JQ1. In GEMM- and patient-derived tumors with constitutive Hh pathway activation, JQ1 effectively decreases tumor cell proliferation and viability *in vitro* and *in vivo*, even when genetic lesions conferring resistance to SMO inhibition (SMOi) are present. Notably, the inhibition of cell proliferation by JQ1 can be rescued by *GLI2* expression driven by a plasmid-based cytomegalovirus (CMV) promoter, which, in contrast to endogenous *GLI* promoters, is not under direct transcriptional regulation by BET proteins. In sum, our study identifies BET proteins as epigenetic regulators of Hedgehog transcriptional output and establishes a rationale for the use of BET inhibitors in cancers with evidence of Hh pathway activation.

RESULTS

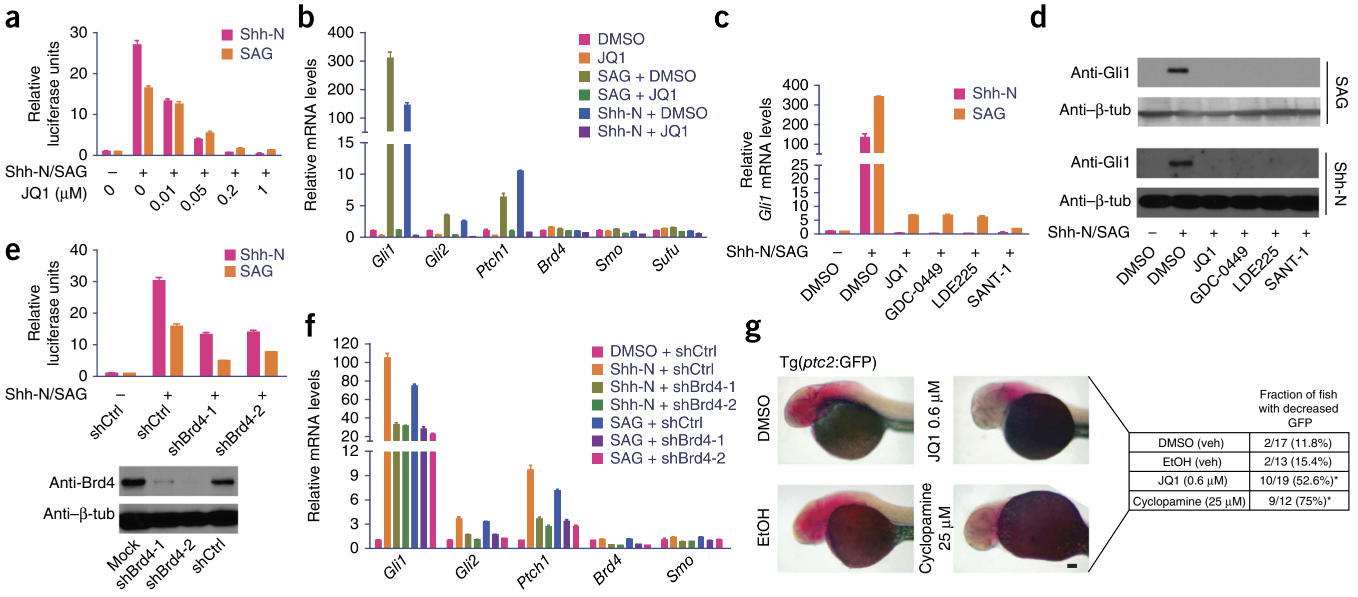
BRD4 is required for ligand-induced Hh transcriptional output

The BET protein BRD4 enhances the transcription of key genes involved in embryonic stem cell maintenance[42] and oncogenesis[43]. Therefore, we hypothesized that BRD4 is a transcriptional cofactor for Hh-responsive genes. In the mouse 3T3 cell–based Hh-Light2 reporter line containing a stably integrated *Gli*-luciferase reporter construct[48], ligand-induced activation of Hh-Light2 cells with either Shh-N conditioned medium (CM)[49] or Smoothened agonist (SAG)[48] resulted in an expected increase in *Gli1*-luciferase activity and *Gli1* mRNA levels, which were both potently inhibited by increasing doses of the BET inhibitor JQ1 (**Fig. 1a** and **Supplementary Fig. 1**). Upregulation of other Hh target genes such as *Ptch1* and *Gli2* was also inhibited by JQ1 (**Fig. 1b**). In contrast, *Smo* expression was modestly influenced, and expression of *Sufu* and *Brd4* was not substantially altered by JQ1 (**Fig. 1b**). Notably, the inhibition of *Gli1* expression by JQ1 equaled that by SMO inhibitors (GDC-0449, LDE225 or SANT-1) (**Fig. 1c**,**d**). Additionally, shRNA-mediated knockdown of *Brd4* in Hh-Light2 cells followed by Shh-N CM or SAG stimulation resulted in marked inhibition of ligand-induced *Gli*-luciferase activity and Hh target gene expression, directly supporting an essential role of Brd4 in Hh signaling (**Fig. 1e**,**f**).

To further assess inhibition of Hh transcriptional output by JQ1, we used zebrafish harboring a *ptc2*:GFP reporter transgene, a well-described canonical Hh pathway reporter in zebrafish[50,51]. Embryos exposed to JQ1 from 2 to 30 hours post fertilization (hpf) showed decreased expression of GFP mRNAs, similar to the results seen in cyclopamine-exposed fish (**Fig. 1g**). We also assessed whether JQ1 could revert abnormal phenotypes caused by aberrant Hh signaling in a temperature-sensitive transgenic fish line harboring an *hsp70l*:Shha–enhanced GFP (eGFP) transgene[51], which overexpresses Shh and produces a reliable and well-described dysgenic eye phenotype that often includes a ventral coloboma, a structural defect in the eye resulting from improper closure of gaps located between various eye structures during embryonic development[52,53]. As predicted, heat-shocked transgenic fish treated with vehicle alone (DMSO) developed abnormally shaped eyes with diminished diameter relative to their heat-shocked nontransgenic siblings (**Fig. 1h**). However, fish exposed to JQ1 immediately after heat shock trended toward more normal-appearing eyes with statistically significant increases in eye diameter, suggesting that BET inhibition countered the effects of aberrant Hh signaling *in vivo* in this model (**Fig. 1h**).

BRD4 regulates Hh signaling at Gli1 and Gli2 promoters

We next examined the effects of JQ1 on Hh signaling in *Sufu*$^{-/-}$ mouse embryonic fibroblasts (MEFs)[54] and Hh-Light2 cells overexpressing GLI2. SUFU positively regulates the degradation of GLI proteins[54], and thus loss of SUFU activity results in stabilization of GLI and constitutive Hh signaling downstream of SMO. As expected, we observed markedly increased *Gli1* mRNA and protein levels in *Sufu*$^{-/-}$ MEFs, which were substantially downregulated by JQ1 (**Fig. 2a**,**c** and **Supplementary Fig. 2a**,**b**). We also noted decreased transcription of *Gli2*, as well as *Smo* to a lesser extent, after JQ1 treatment, whereas *Brd4* mRNA levels remained unchanged (**Fig. 2a**). In stark contrast to JQ1 treatment, we observed little to no effect on *Gli* transcripts or Gli1 protein levels in *Sufu*$^{-/-}$ MEFs after treatment with the SMO inhibitors (LDE225, GDC-0449 or SANT-1) (**Fig. 2b**,**c**). Consistent with pharmacological inhibition of Brd4, shRNA-mediated knockdown of *Brd4* in *Sufu*$^{-/-}$ MEFs resulted in



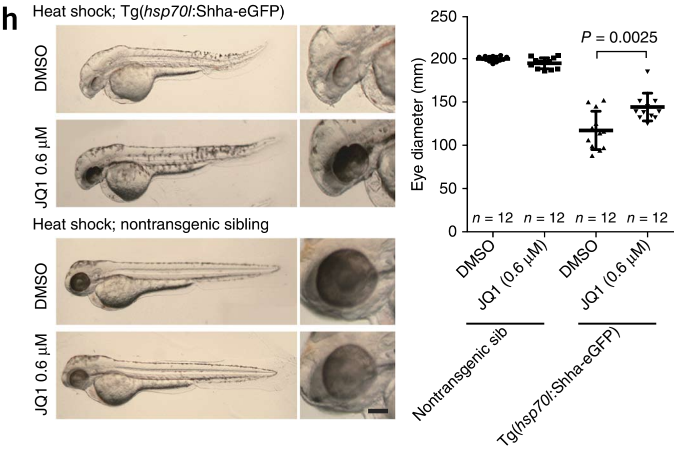


Figure 1 BRD4 is necessary for *GLI* transcription. (**a**) *Gli*-luciferase reporter activity in Hh-Light2 cells treated with Hh ligand (Shh-N CM or SAG) alone or in combination with increasing amount of JQ1. Data represent the mean of triplicates ± s.d. (**b**) qRT-PCR of Hh target genes (*Gli1*, *Gli2* and *Ptch1*), Hh pathway components (*Sufu* and *Smo*) and *Brd4* in Hh-Light2 cells treated with Hh ligand (Shh-N CM or SAG) alone or in combination with JQ1 (1 μM). Data represent the mean of triplicates ± s.d. (**c**) qRT-PCR of *Gli1* mRNA levels in Hh-Light2 cells treated with Hh ligand (Shh-N CM or SAG) alone or in combination with 1 μM of JQ1, GDC-0449, LDE225 or SANT-1. Data represent the mean of triplicates ± s.d. (**d**) Immunoblot detecting Gli1 expression in cell lysates from Hh-Light2 cells treated with Hh ligand (Shh-N CM or SAG) alone or in combination with 1 μM of JQ1, GDC-0449, LDE225 or SANT-1. An anti–β-tubulin (β-tub) immunoblot is shown as a loading control. The immunoblot shown represents a typical result from an experiment performed in duplicate. (**e**) *Gli*-luciferase reporter activity in Hh-Light2 cells treated with Hh ligand (Shh-N CM or SAG) in combination with shRNAs against *Brd4* (shBrd4-1 and shBrd4-2) or scrambled shRNA (shCtrl). Data represent the mean of quadruplicates ± s.d. (**f**) qRT-PCR of Hh target genes (*Gli1*, *Gli2* and *Ptch1*), Hh pathway components (*Sufu* and *Smo*) and *Brd4* after treatment with Hh ligand (Shh-N CM or SAG) alone or in combination with shBrd4-1, shBrd4-2 or shCtrl. Data represent the mean of triplicates ± s.d. (**g**) *In situ* hybridization detecting GFP mRNA levels in transgenic (Tg) zebrafish (*ptc2*:GFP) treated with JQ1 (0.6 μM), cyclopamine (25 μM) or vehicle (veh) controls (DMSO or EtOH). The fraction of zebrafish with decreased GFP expression is shown. Fisher's exact test was used for statistical analysis. $*P < 0.05$. Scale bar, 100 μm. (**h**) Images of a heat-shocked + Tg(*hsp70l*:Shha-eGFP) zebrafish or a nontransgenic sibling treated with JQ1 (0.6 μM) or DMSO. The eye diameter of each group ($n = 12$) was measured and is shown. Data represent the group means ± s.d. The *P* value shown was generated using Student's *t* test. Scale bar, 100 μm.

decreased *Gli1* and *Gli2* mRNA levels (**Fig. 2d**). It is worth noting that Brd4 knockdown did not abrogate GLI-luciferase activity or *Gli* expression as effectively as did JQ1 treatment. This result could be explained by incomplete knockdown of *Brd4*, or it could suggest that other BET proteins (all targets of JQ1) may also contribute to the transcriptional regulation of *Gli* genes. Indeed, knockdown of either Brd2 or Brd3 resulted in a substantial decrease of *Gli* mRNA levels in *Sufu*$^{-/-}$ MEFs (**Supplementary Fig. 2c**).

In Hh-Light2 cells, forced expression of full-length mouse Gli2 (hemagglutinin (HA)-Gli2-FL) or an N-terminally truncated active form of human GLI2 (Myc-GLI2-DN)[55] resulted in an increase in *Gli1* mRNA levels, which was inhibited by JQ1 but not SMO inhibitors (GDC-0449, LDE225 or SANT-1) (**Fig. 2e**). Notably, we did not observe any decrease in ectopic GLI2 expression driven by the CMV promoter expression construct after JQ1 treatment, in contrast to the marked decrease in endogenous *Gli* transcripts (**Figs. 1b** and **2f**). Additionally, upregulation of *Ptch1*, another Hh target gene, was not inhibited by JQ1, suggesting that not all Hh target genes are directly dependent on Brd4, as *Gli* genes themselves are (**Supplementary Fig. 2d**).

In *Sufu*$^{-/-}$ cells, JQ1 decreased *Gli1* and *Gli2* levels as early as 3 h after treatment, supporting a role for Brd4 as a transcriptional cofactor that directly regulates transactivation of *Gli* promoters (**Supplementary Fig. 2e**). Chromatin immunoprecipitation followed by quantitative PCR (ChIP-qPCR) using antibody to Brd4 of regions flanking the transcription start sites of *Gli1* and *Gli2* promoters confirmed increased Brd4 occupancy at both *Gli* promoters after SAG-mediated activation of Hh signaling in Hh-Light2 cells (**Fig. 2g,h**). Accordingly, ChIP-qPCR with antibody to PolII showed engagement of both *Gli* promoters by PolII after SAG stimulation. Notably, both Brd4 and PolII interactions at the *Gli* promoters were blocked by the addition of JQ1 (**Fig. 2g,h**). Similarly, in *Sufu*$^{-/-}$ MEFs, we observed increased baseline occupancy of *Gli* promoters by Brd4 and PolII relative to that in wild-type (WT) MEFs, which was markedly inhibited by JQ1 (**Fig. 2i,j**).

JQ1 inhibits Ptch-deficient medulloblastoma and BCC

We investigated the efficacy of JQ1 in Hh-driven tumors using cell lines derived from autochthonous medulloblastomas (SmoWT-MB and Med1-MB) arising in *Ptch*$^{+/-}$; *Trp53*$^{-/-}$ and *Ptch*$^{+/-}$; *lacZ* mice, respectively[32,56], and BCC (ASZ001)[57], also derived from *Ptch*$^{+/-}$ mice. JQ1 treatment resulted in marked downregulation of *Gli* mRNA and protein expression with little to no effect on *Smo*, *Sufu* or *Brd4* (**Fig. 3a–d** and **Supplementary Fig. 3a**). Again, we observed a rapid decrease of *Gli* gene expression after JQ1 treatment (as early as 3 h), supporting a direct effect of BET inhibition on *Gli* promoters

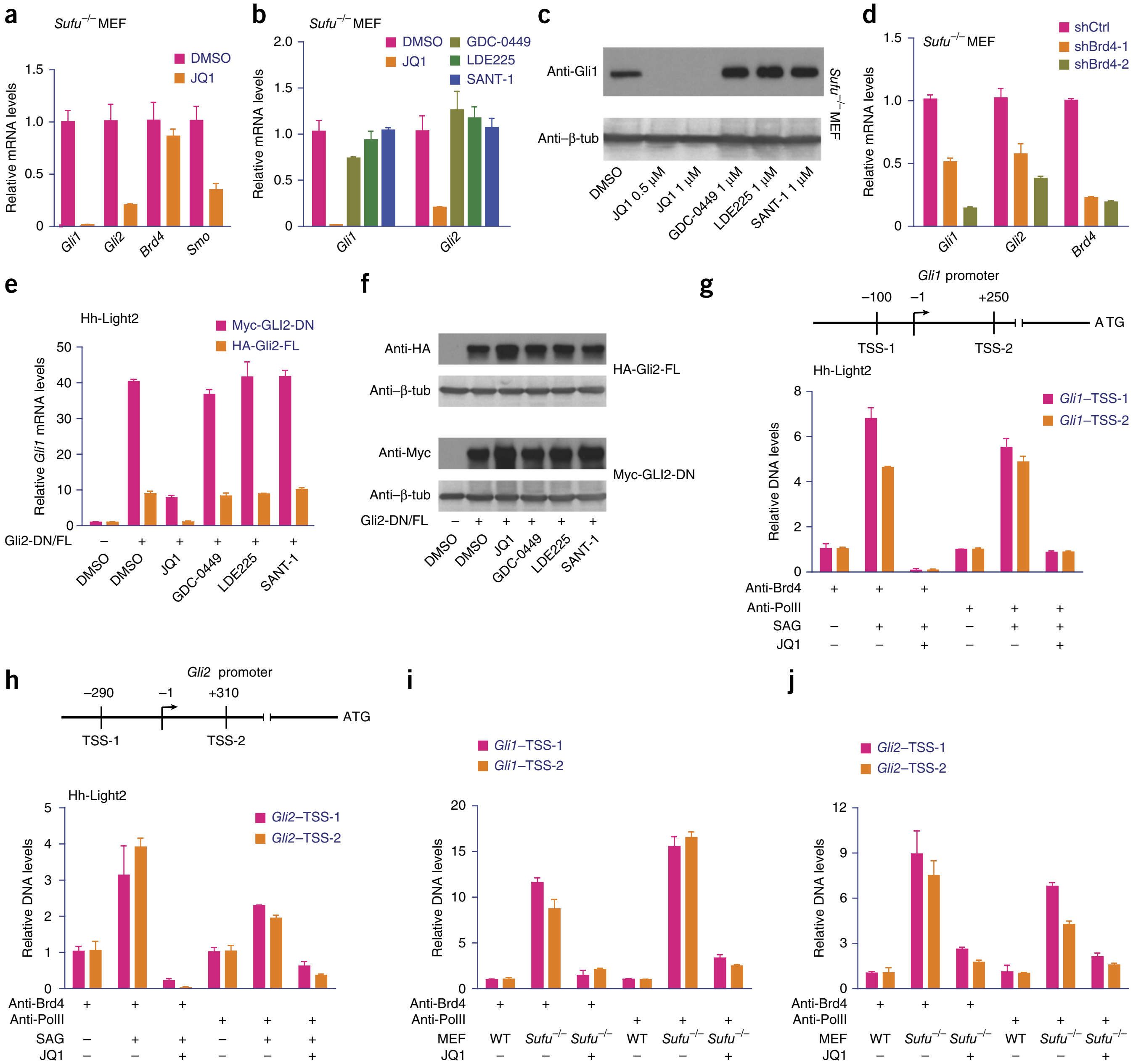


Figure 2 JQ1 inhibits Hh target gene activation at the level of the *GLI1* and *GLI2* promoters. (**a**) qRT-PCR showing *Gli1*, *Gli2*, *Smo* and *Brd4* mRNA levels in *Sufu*$^{-/-}$ MEFs treated with JQ1. Data represent the mean of triplicates ± s.d. (**b**) *Gli1* and *Gli2* mRNA levels in *Sufu*$^{-/-}$ MEFs treated with DMSO, JQ1, GDC-0449, LDE225 or SANT-1. Data represent the mean of triplicates ± s.d. (**c**) Immunoblot detecting GLI1 expression in cell lysates from *Sufu*$^{-/-}$ MEFs treated with DMSO, JQ1, GDC-0449, LDE225 or SANT-1. An anti–β-tubulin immunoblot is shown as a loading control. The immunoblots in **c** and **f** represent a typical result from each experiment performed in duplicate. (**d**) qRT-PCR showing *Gli1*, *Gli2* and *Brd4* mRNA levels in *Sufu*$^{-/-}$ cells expressing shBrd4-1, shBrd4-2 or shCtrl. Data represent the mean of triplicates ± s.d. (**e**) qRT-PCR showing *Gli1* mRNA levels in Hh-Light2 cells transiently transfected with HA-Gli2-FL or Myc-GLI2-DN and their responses to JQ1, GDC-0449, LDE225 or SANT-1. Data represent the mean of triplicates ± s.d. (**f**) Anti-HA and anti-Myc immunoblots on cell lysates from Hh-Light2 cells transfected with HA-Gli2-FL or Myc-GLI2-DN and treated with DMSO, JQ1, GDC-0449, LDE225 or SANT-1. An anti–β-tubulin immunoblot is shown as a loading control. (**g**–**j**) Schematic of regions flanking the *Gli1* and *Gli2* promoter transcription start sites (TSS) analyzed by ChIP-qPCR of Brd4 and PolII occupancies in Hh-Light2 cells treated with SAG and JQ1 (**g**,**h**) and in *Sufu*$^{-/-}$ MEFs treated with JQ1 (**i**,**j**). Data represent the mean of triplicates ± s.d. Except where indicated, cells were treated with 1 µM of JQ1, GDC-0449, LDE225 or SANT-1.

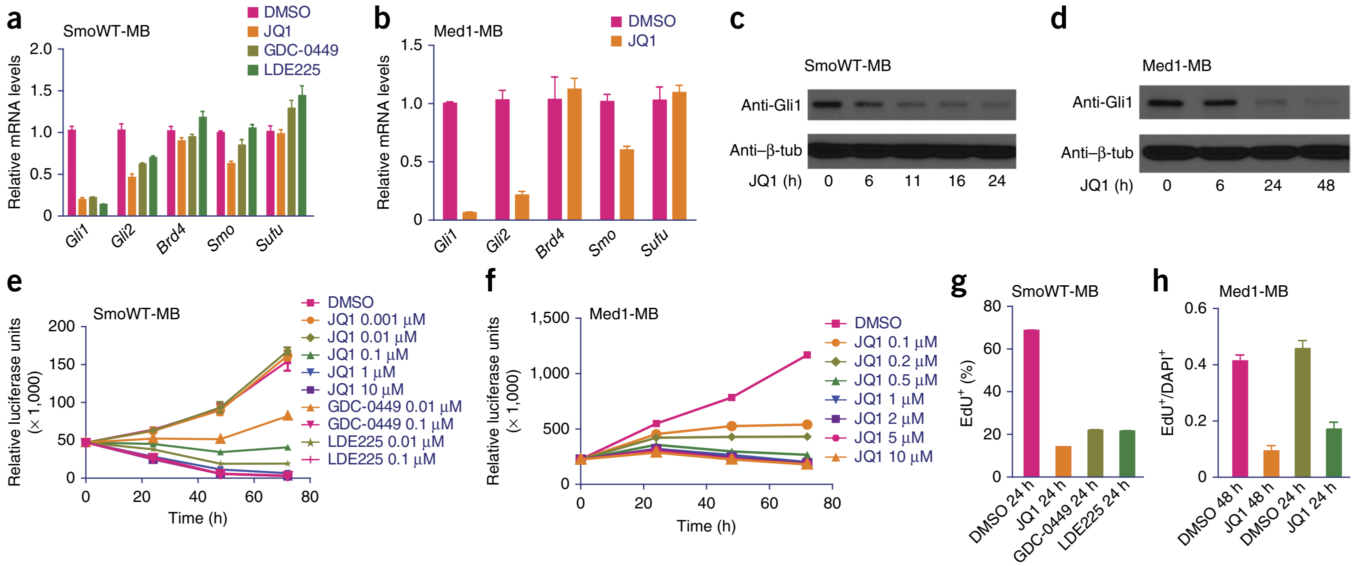


Figure 3 JQ1 inhibits Hh pathway activity and cell viability and proliferation in Ptch-mutated medulloblastoma (SmoWT-MB and Med1-MB) cells. (**a**,**b**) qRT-PCR of the expression of Hh pathway target genes (*Gli1* and *Gli2*), components (*Smo* and *Sufu*) and *Brd4* in SmoWT-MB and Med1-MB cells treated with JQ1 (1 µM), GDC-0449 (0.1 µM) or LDE225 (0.1 µM). Data represent the mean of triplicates ± s.d. (**c**,**d**) Immunoblots detecting Gli1 expression in response to JQ1 treatment over time. An anti–β-tubulin immunoblot is shown as a loading control. The immunoblots represent a typical result from each experiment performed in duplicate. (**e**,**f**) Cell viability detection over time with increasing doses of JQ1 or SMO inhibitors. Data represent the group means ± s.d. (**g**,**h**) Proliferative index in response to JQ1 (1 µM) or SMO inhibitors (GDC-0449 or LDE225 at 0.1 µM) as measured by EdU incorporation. Data represent the group means ± s.d.

(**Supplementary Fig. 3b**). Accordingly, ChIP-qPCR using antibodies to Brd4 and PolII showed potent inhibition of Brd4 and PolII occupancy at *Gli* promoters in all cell lines after exposure to JQ1 (**Supplementary Fig. 3c**,**d**).

In Med1-MB and SmoWT-MB cells, JQ1 treatment resulted in dose-responsive decreases in cell viability to a much greater extent than those observed in Hh-Light2 or *Sufu*$^{-/-}$ MEFs (**Supplementary Fig. 4a**). Potent growth inhibition was achieved (half-maximum inhibitory concentration (IC$_{50}$) ~50–150 nM; **Supplementary Fig. 4b**,**c**) with marked decreases of proliferation (**Fig. 3e–h**), induction of apoptosis (**Supplementary Fig. 4d**,**e**) and, in Med1-MB cells, an increased fraction of cells in G1 and a decreased fraction of cells transitioning through S phase (**Supplementary Fig. 4f**). Notably, in SmoWT-MB cells, the inhibitory effects of JQ1 on *Gli* expression, cell viability and proliferation were equivalent to those of SMO inhibitors (GDC-0449 or LDE225) (**Fig. 3a**,**e**,**g** and **Supplementary Figs. 3a** and **4d**), and these effects were enhanced when we exposed cells to both JQ1 and GDC-0449 in combination (**Supplementary Fig. 4g**).

Using microarray analysis, we assessed changes in global gene expression in JQ1-treated SmoWT-MB cells compared with DMSO- and GDC0449-treated cells. We observed a substantial overlap between significantly differentially expressed genes ($P < 0.0001$) or gene sets ($P < 0.0001$; **Supplementary Dataset**) by JQ1 and GDC0449 in both cell lines compared with DMSO-treated controls, including the anticipated GLI target genes *Gli2*, *Ptch1*, *Ccnd1*, *Ccnd2*, *Hhip* and *Cdk6* (**Supplementary Fig. 5a–c**). We next compared JQ1-induced gene expression profiles with gene sets derived from previously published ChIP-chip studies, which indexed gene promoters with Gli1-binding sites in normal granule neuron precursor cells (GNPs) and *Ptch*$^{+/-}$ medulloblastoma cells[58]. Specifically, we analyzed for enrichment of ChIP-chip peaks associated with GNPs, medulloblastoma, the overlap of both and peaks associated with GNPs alone or medulloblastoma alone (**Supplementary Table 1**). Gene set enrichment analysis (GSEA) revealed that only genes with Gli1 promoter–binding sites associated with medulloblastoma were significantly enriched ($P < 0.0001$) in JQ1-treated cells (**Supplementary Fig. 5d**). These results confirm the disruption of Gli1-mediated transcription by JQ1 and the preferential targeting of Gli1 transcriptional activity in tumor cells[43].

Ectopic GLI2 expression rescues growth inhibition by JQ1

We tested whether knockdown of Brd4 could phenocopy the effects of JQ1 in Hh-driven medulloblastoma cells. As expected, knockdown of Brd4 resulted in decreased *Gli* expression (**Fig. 4a**,**b**) and cell proliferation (**Fig. 4c**,**d**), suggesting that the inhibitory effect of JQ1 was through targeting of Brd4. Furthermore, to directly assess whether BET inhibition blocked cell proliferation in Hh-driven tumor cells through targeting of *Gli* transcription, we used plasmid-based expression of GLI2 (Myc-GLI2-DN; **Fig. 2f**) in SmoWT-MB cells and monitored its ability to rescue the inhibition of proliferation by JQ1 (**Fig. 4c**). Notably, ectopic expression of GLI2 inhibited the effects of JQ1 on 5-ethynyl-2′-deoxyuridine (EdU) incorporation, resulting in levels of EdU incorporation that were nearly equivalent to levels in DMSO-treated control cells (**Fig. 4e**,**f**). This result indicates that inhibition of proliferation by JQ1 is mediated largely through inhibition of *Gli* transcription and, intriguingly, that Brd4-independent transcriptional targets of Gli transcription factors are sufficient to overcome BET inhibition.

SMOi-resistant Hh-driven tumors are inhibited by JQ1

Given the documented mechanisms of resistance to current, clinically available SMO inhibitors[25,31] and the potential of BET inhibitors as a strategy to overcome this resistance, we examined the efficacy of JQ1 in Hh-driven cancers with either acquired or *a priori* resistance to SMO inhibitors (**Fig. 5a**). We analyzed the efficacy of JQ1 and SMO inhibitors (GDC-0449 and LDE225) against medulloblastoma cells

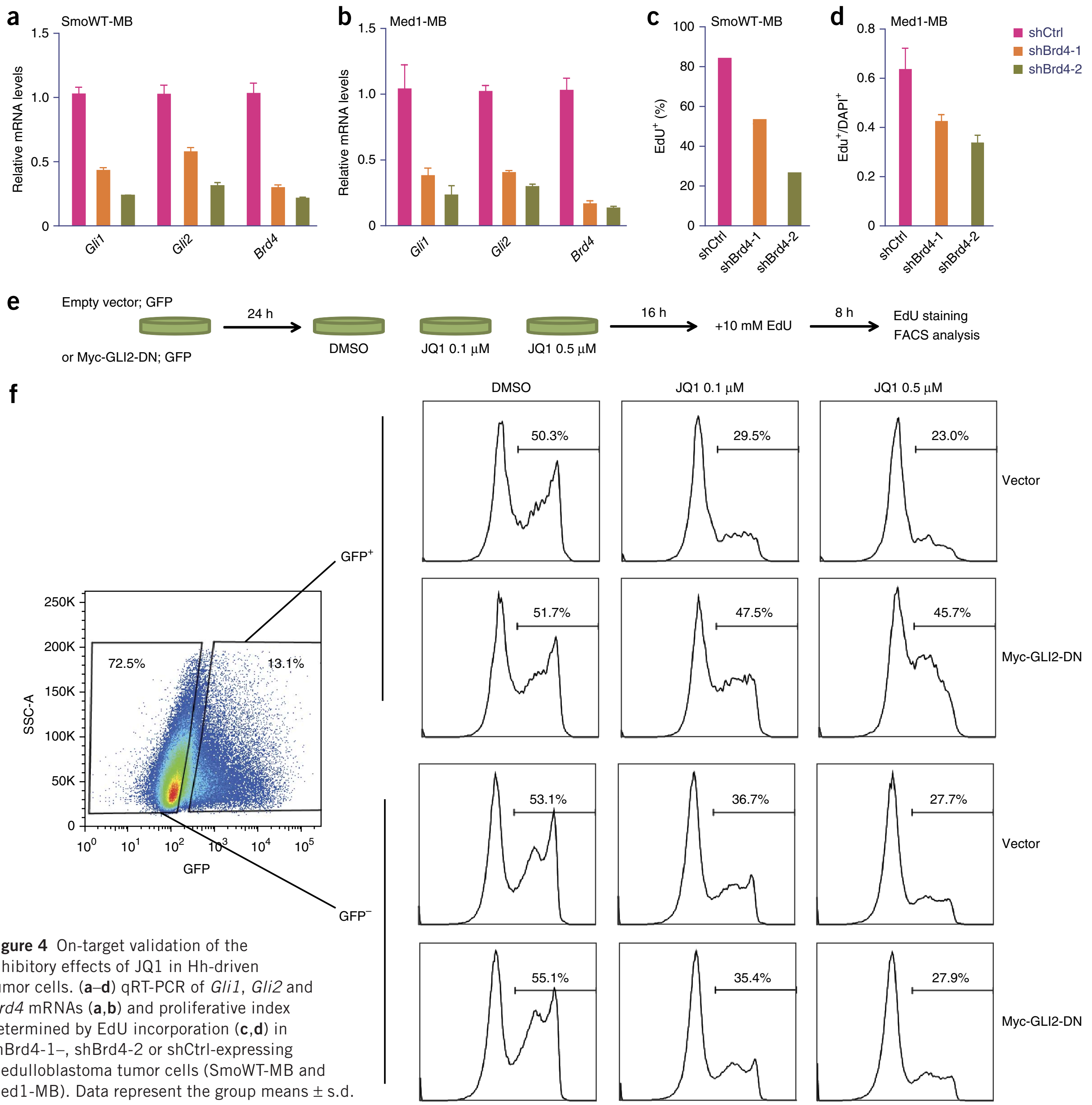


Figure 4 On-target validation of the inhibitory effects of JQ1 in Hh-driven tumor cells. (**a**–**d**) qRT-PCR of *Gli1*, *Gli2* and *Brd4* mRNAs (**a**,**b**) and proliferative index determined by EdU incorporation (**c**,**d**) in shBrd4-1–, shBrd4-2 or shCtrl-expressing medulloblastoma tumor cells (SmoWT-MB and Med1-MB). Data represent the group means ± s.d. (**e**) The schematic workflow of the GLI2-overexpressing rescue experiment in SmoWT-MB cells. (**f**) FACS analysis of EdU incorporation in SmoWT-MB cells transfected with empty vector or Myc-GLI2-DN followed by JQ1 treatment (0.1 or 0.5 µM). GFP-expressing plasmid was used for co-transfection to mark the transfected (GFP$^+$) cells. SSC-A, side scatter.

carrying an aspartate-to-glycine substitution at amino acid residue 477 in Smo that results in decreased sensitivity to SMO antagonists (SmoD477G-MB) (**Fig. 5b**)[32]; patient-derived *SUFU*-mutated primary SHH-subtype medulloblastoma cells (RCMB025); patient-derived primary atypical teratoid rhabdoid tumor (ATRT) cells (CHB_ATRT1 and SU_ATRT2) with derepression of *GLI1* transcription through loss of SMARCB1 (also called SNF5 or INI1) (ref. 22); and patient-derived *MYCN*-amplified primary SHH-subtype medulloblastoma cells (RCMB018). Cell viability (**Fig. 5b**–**f**, top), *Gli* and *GLI* levels (**Fig. 5b**–**f**, bottom) and EdU incorporation (**Supplementary Fig. 6a**–**c**) were markedly decreased in response to JQ1 in all of these cells, and we observed little or no effect with the SMO inhibitors GDC-0449 and LDE225. Additionally, we examined *Myc*, *MYC*, *Mycn* and *MYCN* expression in SmoWT-MB, SmoD477G-MB, RCMB025, CHB_ATRT1 and RCMB018 cells and found that *Mycn* and *MYCN* expression was consistently inhibited by JQ1 (**Fig. 5f**, bottom and **Supplementary Figs. 4h** and **6d**–**f**), suggesting that JQ1 targets at least two important driver oncogenes (GLI and MYCN) in these tumors.

In vivo inhibition of Hh-driven tumors by JQ1

To support a therapeutic role for BET inhibition in Hh-driven tumors, we assessed the *in vivo* efficacy of JQ1 against medulloblastomas and

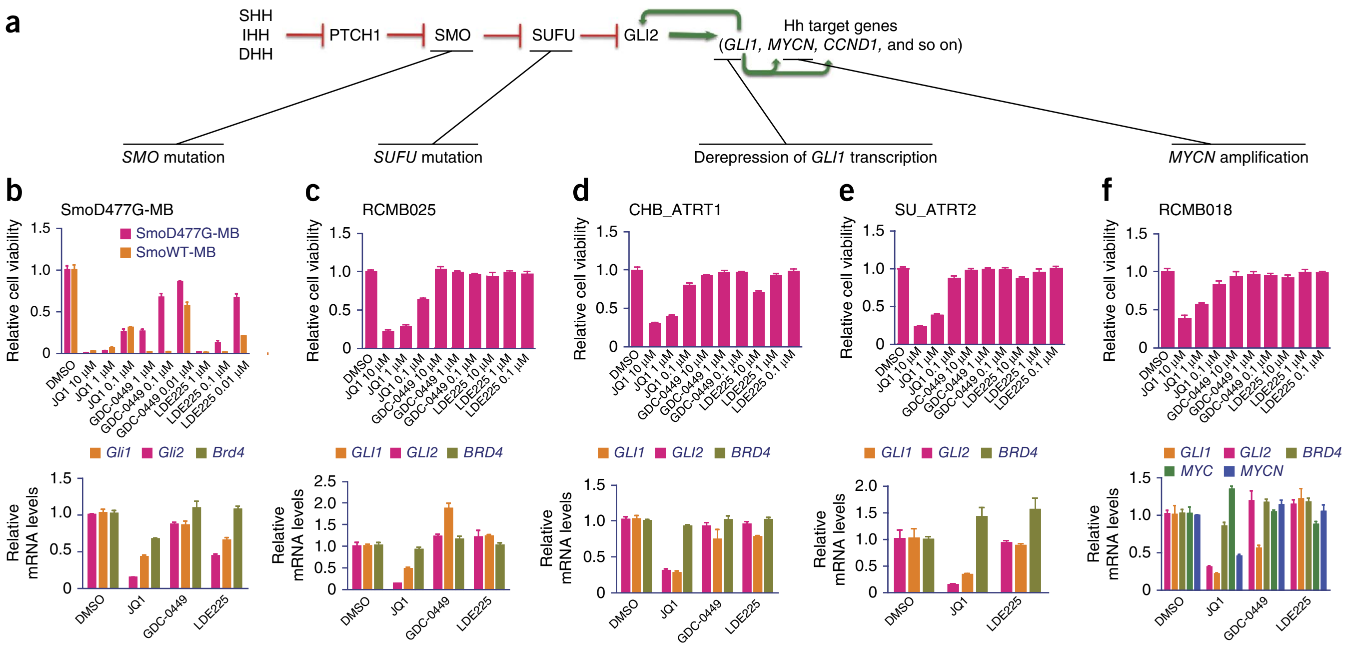
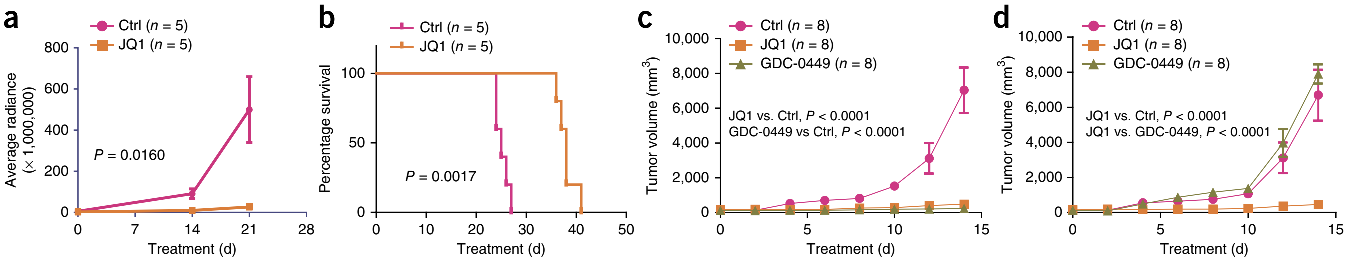


Figure 5 JQ1 inhibits Hh pathway activity and cell viability and proliferation in SMOi-resistant Hh-driven tumors. (**a**) Schematic depicting mechanisms of resistance to Smoothened antagonists in Hh-driven cancers. (**b**–**f**, top) Cell viability in SMOi-resistant medulloblastoma cells (SmoD477G-MB; **b**), patient-derived *SUFU* mutant medulloblastoma cells (RCMB025; **c**), patient-derived ATRT cells (CHB_ATRT1 and SU_ATRT2; **d**,**e**) and patient-derived *MYCN*-amplified medulloblastoma cells (RCMB018; **f**) treated with increasing doses of JQ1, GDC-0449 or LDE225. Data represent the group means ± s.d. (**b**–**f**, bottom) qRT-PCR of *Gli1*, *GLI1*, *Gli2*, *GLI2*, *Brd4* and *BRD4* (plus *MYC* and *MYCN* levels for RCMB018) in SmoD477G-MB (**b**), RCMB025 (**c**), CHB_ATRT1 (**d**), SU_ATRT2 (**e**) and RCMB018 (**f**) cells in response to JQ1 (1 µM), GDC-0449 or LDE225 (0.1 µM for SmoD477G-MB and 1 µM for the other groups). Data represent the mean of triplicates ± s.d.

BCCs. We treated flank and intracranial allografts of Med1-MB cells stably expressing a firefly luciferase reporter in immunodeficient NSG mice with either JQ1 (50 mg per kg body weight per day intraperitoneally (i.p.)) or vehicle control. We observed a significant reduction in flank tumor growth in JQ1-treated mice, as well as an increase in overall survival in JQ1-treated mice harboring intracranial allografts (**Fig. 6a**,**b** and **Supplementary Fig. 7a**). Additionally, we treated medulloblastoma flank allografts of SmoWT-MB or SmoD477G-MB cells with vehicle control, JQ1 (50 mg per kg body weight per day i.p.) or GDC-0449 (100 mg per kg body weight per day orally (p.o.)). We observed marked decreases in the growth of SmoD477G-MB flank allografts in response to JQ1 but not GDC-0449, whereas SmoWT-MB flank allografts responded to both GDC-0449 and JQ1 (**Fig. 6c**,**d**). To evaluate the efficacy of JQ1 against BCCs *in vivo*, we used an allograft model of *Ptch*$^{+/-}$; *K14-creER2*; *p53*$^{\text{flox/flox}}$–derived mouse BCC cells[34]. JQ1 treatment (50 mg per kg body weight per day i.p.)



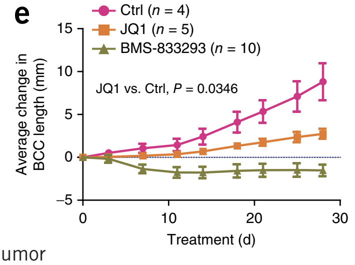


Figure 6 JQ1 inhibits Hh-driven tumor growth *in vivo*. (**a**,**b**) Med1-MB cells transduced with a lentiviral luciferase reporter were used for flank (**a**) or cerebellum (**b**) injections of NSG mice, which were then randomized for treatment with either JQ1 (50 mg per kg body weight daily i.p.) or vehicle. (**a**) Tumor growth of Med1-MB allografts assessed by IVIS imaging and presented as the average radiance. (**b**) Survival curve of the mice injected with Med1-MB cells in the cerebellum. (**c**,**d**) Tumor growth, assessed by caliper measurement, after SmoWT-MB (**c**) and SmoD477G-MB (**d**) cells were injected into the flanks of NSG mice (both flanks of each mouse) followed by treatment with JQ1 (50 mg per kg body weight daily i.p.), GDC-0449 (100 mg per kg body weight daily p.o.) or vehicle. (**e**) Tumor growth, assessed by caliper measurement, of SMOi-naive mouse BCC tumors generated under the dermis of NSG mice that were treated with JQ1 (50 mg per kg body weight daily i.p.), BMS-833293 (100 mg per kg body weight daily i.p.) or vehicle. Data on tumor growth represent the group means ± s.e.m. Two-way analysis of variance (ANOVA) was used for comparing tumor growth curves. Log-rank test was used for comparing survival curves. The results shown are from two separate experiments testing JQ1 and BMS-833293, independently, but are presented on the same graph.

resulted in significant growth inhibition of BCCs but was not as effective as the clinically optimized SMO inhibitor BMS-833293 (ref. 34) (**Fig. 6e**). Nonetheless, in all Hh-driven tumor models tested, we observed reduction of *Gli* mRNA levels after JQ1 treatment regardless of whether allografts were sensitive or resistant to SMO inhibition (SMOi) (**Supplementary Fig. 7b–f**). Together these results demonstrate *in vivo* efficacy of JQ1 against Hh-driven tumors, even those with acquired or *a priori* resistance to clinically available SMO inhibitors.

DISCUSSION

We have shown that BRD4 and other BET proteins are critical regulators of *GLI1* and *GLI2* transcription and that BET inhibition provides a new therapeutic strategy against Hh-driven tumors. Notably, as BET proteins regulate the far-downstream transcriptional output of Hh signaling, BET inhibition was effective against tumor cells that evade Smoothened antagonists through mutation of SMO or amplification of nodes downstream of SMO. Our study is clinically relevant for patients who have *a priori* resistance to SMO inhibitors and in cases in which the emergence of resistance develops after an initial response to such therapy. By acting directly on the *GLI1* and *GLI2* promoters, BET inhibition circumvents all SMOi resistance mechanisms that have been reported so far, which include mutations of SMO or SUFU or amplifications in *GLI2* or *MYCN*[16,17,20,25,31]. The response to JQ1 observed in *MYCN*-amplified SHH medulloblastoma cells (RCMB018), in terms of both decreased cell viability and *MYCN* levels, is similar to the results of a recent study showing the efficacy of BET inhibitors in *MYCN*-amplified neuroblastoma[46]. However, in Hh-driven tumors, it is likely that decreased MYCN levels in response to JQ1 treatment reflect the role of GLI in directly transactivating the *MYCN* promoter, in addition to the role of JQ1 in directly regulating expression of *Mycn* and *MYCN*.

Given the importance of Hh signaling in normal development, it will be essential to understand and anticipate potential toxicities of BET inhibitor therapies as they enter into clinical trials. We observed developmental anomalies at very high doses of JQ1 in our zebrafish studies (data not shown), consistent with those seen in *Brd4* heterozygous mice, which display a multitude of defects that overlap with cyclopamine-treated or Hh-deficient mice[59,60]. Of note, however, *Brd4* heterozygous mice develop craniofacial but not overt axial skeletal phenotypes[59], unlike cyclopamine-exposed embryos[2,60], suggesting lineage-specific differences of Hh pathway dependency on Brd4. Our finding that plasmid-driven GLI2 expression can rescue the proliferation defect induced by JQ1 supports the existence of GLI-responsive promoters that do not require BRD4 for their transactivation. Notably, such genes appear to be either individually or collectively sufficient to mediate part, if not all, of the oncogenic phenotype associated with Hh-GLI signaling.

Investigating how BRD4 regulates normal Hh-mediated biological processes and documenting BRD4-related changes that occur during Hh-mediated oncogenic transformation could potentially elucidate factors essential for tumor development that are independent of normal development. Our analysis of gene expression changes in JQ1-treated medulloblastoma supports observations by Lee *et al.*[58], who identified marked shifts in Gli1 occupancy across the genome in medulloblastoma compared to GNPs. An unbiased characterization of Brd4 binding across the genome in GNPs and medulloblastomas will clarify whether the genomic footprint of Brd4 overlaps with Gli occupancy in the oncogenic state relative to the normal developmental state. Related to this point, emerging evidence suggests BET proteins converge on super-enhancer sites across the genome and that these super-enhancer sites help transactivate promoters of key regulators of cellular identity in normal and pathogenic contexts[42,43]. Whether GLI transactivates super enhancer–related promoters and, accordingly, whether super-enhancer sites are positioned over *GLI* promoters is currently under active investigation.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Gene expression profiling data has been deposited into the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) with accession code GSE58185.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

This work was supported the St. Baldrick's Foundation Scholar Award (Y.-J.C.), a Beirne Faculty Scholar Endowment (Y.-J.C.), US National Institutes of Health (NIH) grant U01-CA176287 (Y.-J.C.), an Alex's Lemonade Stand Foundation Young Investigator Award (Y.T.), the Damon-Runyon Cancer Research Foundation (J.Q. and J.E.B.), NIH R01-CA159859 (R.W.-R.) and a pilot project grant from the Medical College of Wisconsin Cancer Center–Advancing a Healthier Wisconsin (B.A.L.). R.W.-R. is the recipient of a Leadership Award (LA1-01747) from the California Institute of Regenerative Medicine. We thank the Stanford Functional Genomics Facility (SFGF) and the Protein and Nucleic Acid (PAN) facility for their assistance in generating gene expression microarray data and reagents. We thank M. Scott (Stanford), C. Rudin (Memorial Sloan-Kettering Cancer Center), P. Beachy (Stanford), A. Sweet-Cordero (Stanford), P.-T. Chang (University of California, San Francisco), R. Karlstrom (University of Massachusetts, Amherst) and J.K. Chen (Stanford) for reagents, helpful suggestions and/or critical reading of the manuscript.

AUTHOR CONTRIBUTIONS

Y.T. and Y.-J.C. conceived the project and wrote the manuscript. Y.T., B.N., S.M., B.B., N.V., S.S., S.C., A.P., S.O. and F.Y. performed all molecular biology experiments. J.Q. and J.E.B. synthesized and supplied JQ1 for all studies. A.E.O., S.X.A., R.J.W., A.L. and J.Y.T. generated and prepared GEMM-derived BCC cells, and R.W.-R. generated and provided patient-derived medulloblastoma cells. P.B., G.B., R.B. and Y.-J.C. performed all informatics analyses. S.G., A.L., Y.T., S.M., P.W., M.M., S.H.C. and S.S. performed JQ1 mouse *in vivo* studies. M.I.W. and B.A.L. performed all zebrafish studies.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

---

1. Goodrich, L.V., Johnson, R.L., Milenkovic, L., McMahon, J.A. & Scott, M.P. Conservation of the hedgehog/patched signaling pathway from flies to mice: induction of a mouse patched gene by Hedgehog. *Genes Dev.* **10**, 301–312 (1996).
2. Binns, W., Shupe, J.L., Keeler, R.F. & James, L.F. Chronologic evaluation of teratogenicity in sheep fed *Veratrum californicum*. *J. Am. Vet. Med. Assoc.* **147**, 839–842 (1965).
3. Gorlin, R.J., Vickers, R.A., Kellen, E. & Williamson, J.J. Multiple nasal-cell nevi syndrome. An analysis of a syndrome consisting of multiple nevoid basal-cell carcinoma, jaw cysts, skeletal anomalies, medulloblastoma, and hyporesponsiveness to parathormone. *Cancer* **18**, 89–104 (1965).
4. Johnson, R.L. *et al.* Human homolog of patched, a candidate gene for the basal cell nevus syndrome. *Science* **272**, 1668–1671 (1996).
5. Echelard, Y. *et al.* Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. *Cell* **75**, 1417–1430 (1993).
6. Marigo, V. *et al.* Cloning, expression, and chromosomal location of SHH and IHH: two human homologues of the *Drosophila* segment polarity gene hedgehog. *Genomics* **28**, 44–51 (1995).
7. Alcedo, J., Ayzenzon, M., Von Ohlen, T., Noll, M. & Hooper, J.E. The *Drosophila* smoothened gene encodes a seven-pass membrane protein, a putative receptor for the hedgehog signal. *Cell* **86**, 221–232 (1996).

8. Huangfu, D. & Anderson, K.V. Cilia and Hedgehog responsiveness in the mouse. *Proc. Natl. Acad. Sci. USA* **102**, 11325–11330 (2005).
9. Haycraft, C.J. *et al.* Gli2 and Gli3 localize to cilia and require the intraflagellar transport protein polaris for processing and function. *PLoS Genet.* **1**, e53 (2005).
10. Liu, A., Wang, B. & Niswander, L.A. Mouse intraflagellar transport proteins regulate both the activator and repressor functions of Gli transcription factors. *Development* **132**, 3103–3111 (2005).
11. Oliver, T.G. *et al.* Transcriptional profiling of the Sonic hedgehog response: a critical role for N-myc in proliferation of neuronal precursors. *Proc. Natl. Acad. Sci. USA* **100**, 7331–7336 (2003).
12. Regl, G. *et al.* Human GLI2 and GLI1 are part of a positive feedback mechanism in basal cell carcinoma. *Oncogene* **21**, 5529–5539 (2002).
13. Mao, J. *et al.* A novel somatic mouse model to survey tumorigenic potential applied to the Hedgehog pathway. *Cancer Res.* **66**, 10171–10178 (2006).
14. Xie, J. *et al.* Mutations of the PATCHED gene in several types of sporadic extracutaneous tumors. *Cancer Res.* **57**, 2369–2372 (1997).
15. Thayer, S.P. *et al.* Hedgehog is an early and late mediator of pancreatic cancer tumorigenesis. *Nature* **425**, 851–856 (2003).
16. Cho, Y.J. *et al.* Integrative genomic analysis of medulloblastoma identifies a molecular subgroup that drives poor clinical outcome. *J. Clin. Oncol.* **29**, 1424–1430 (2011).
17. Pugh, T.J. *et al.* Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. *Nature* **488**, 106–110 (2012).
18. Taylor, M.D. *et al.* Mutations in SUFU predispose to medulloblastoma. *Nat. Genet.* **31**, 306–310 (2002).
19. ten Haaf, A. *et al.* Expression of the glioma-associated oncogene homolog (GLI) 1 in human breast cancer is associated with unfavourable overall survival. *BMC Cancer* **9**, 298 (2009).
20. Northcott, P.A. *et al.* Subgroup-specific structural variation across 1,000 medulloblastoma genomes. *Nature* **488**, 49–56 (2012).
21. Buczkowicz, P., Ma, J. & Hawkins, C. GLI2 is a potential therapeutic target in pediatric medulloblastoma. *J. Neuropathol. Exp. Neurol.* **70**, 430–437 (2011).
22. Jagani, Z. *et al.* Loss of the tumor suppressor Snf5 leads to aberrant activation of the Hedgehog-Gli pathway. *Nat. Med.* **16**, 1429–1433 (2010).
23. Zwerner, J.P. *et al.* The EWS/FLI1 oncogenic transcription factor deregulates GLI1. *Oncogene* **27**, 3282–3291 (2008).
24. Chen, J.K., Taipale, J., Cooper, M.K. & Beachy, P.A. Inhibition of Hedgehog signaling by direct binding of cyclopamine to Smoothened. *Genes Dev.* **16**, 2743–2748 (2002).
25. Buonamici, S. *et al.* Interfering with resistance to smoothened antagonists by inhibition of the PI3K pathway in medulloblastoma. *Sci. Transl. Med.* **2**, 51ra70 (2010).
26. Lee, M.J. *et al.* Hedgehog pathway inhibitor saridegib (IPI-926) increases lifespan in a mouse medulloblastoma model. *Proc. Natl. Acad. Sci. USA* **109**, 7859–7864 (2012).
27. Zhang, Y., Laterra, J. & Pomper, M.G. Hedgehog pathway inhibitor HhAntag691 is a potent inhibitor of ABCG2/BCRP and ABCB1/Pgp. *Neoplasia* **11**, 96–101 (2009).
28. LoRusso, P.M. *et al.* Phase I trial of hedgehog pathway inhibitor vismodegib (GDC-0449) in patients with refractory, locally advanced or metastatic solid tumors. *Clin. Cancer Res.* **17**, 2502–2511 (2011).
29. Rudin, C.M. *et al.* Treatment of medulloblastoma with hedgehog pathway inhibitor GDC-0449. *N. Engl. J. Med.* **361**, 1173–1178 (2009).
30. Von Hoff, D.D. *et al.* Inhibition of the hedgehog pathway in advanced basal-cell carcinoma. *N. Engl. J. Med.* **361**, 1164–1172 (2009).
31. Yauch, R.L. *et al.* Smoothened mutation confers resistance to a Hedgehog pathway inhibitor in medulloblastoma. *Science* **326**, 572–574 (2009).
32. Kim, J. *et al.* Itraconazole and arsenic trioxide inhibit Hedgehog pathway activation and tumor growth associated with acquired resistance to smoothened antagonists. *Cancer Cell* **23**, 23–34 (2013).
33. Kim, J., Lee, J.J., Gardner, D. & Beachy, P.A. Arsenic antagonizes the Hedgehog pathway by preventing ciliary accumulation and reducing stability of the Gli2 transcriptional effector. *Proc. Natl. Acad. Sci. USA* **107**, 13432–13437 (2010).
34. Atwood, S.X., Li, M., Lee, A., Tang, J.Y. & Oro, A.E. GLI activation by atypical protein kinase C ι/λ regulates the growth of basal cell carcinomas. *Nature* **494**, 484–488 (2013).
35. Kenney, A.M., Widlund, H.R. & Rowitch, D.H. Hedgehog and PI-3 kinase signaling converge on Nmyc1 to promote cell cycle progression in cerebellar neuronal precursors. *Development* **131**, 217–228 (2004).
36. Hyman, J.M. *et al.* Small-molecule inhibitors reveal multiple strategies for Hedgehog pathway blockade. *Proc. Natl. Acad. Sci. USA* **106**, 14132–14137 (2009).
37. Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468**, 1067–1073 (2010).
38. Dhalluin, C. *et al.* Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**, 491–496 (1999).
39. Jacobson, R.H., Ladurner, A.G., King, D.S. & Tjian, R. Structure and function of a human TAFII250 double bromodomain module. *Science* **288**, 1422–1425 (2000).
40. Jang, M.K. *et al.* The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II–dependent transcription. *Mol. Cell* **19**, 523–534 (2005).
41. Yang, Z. *et al.* Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol. Cell* **19**, 535–545 (2005).
42. Whyte, W.A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
43. Lovén, J. *et al.* Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**, 320–334 (2013).
44. Delmore, J.E. *et al.* BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell* **146**, 904–917 (2011).
45. Mertz, J.A. *et al.* Targeting MYC dependence in cancer by inhibiting BET bromodomains. *Proc. Natl. Acad. Sci. USA* **108**, 16669–16674 (2011).
46. Puissant, A. *et al.* Targeting MYCN in neuroblastoma by BET bromodomain inhibition. *Cancer Discov.* **3**, 308–323 (2013).
47. Bandopadhayay, P. *et al.* BET bromodomain inhibition of MYC-amplified medulloblastoma. *Clin. Cancer Res.* **20**, 912–925 (2014).
48. Chen, J.K., Taipale, J., Young, K.E., Maiti, T. & Beachy, P.A. Small molecule modulation of Smoothened activity. *Proc. Natl. Acad. Sci. USA* **99**, 14071–14076 (2002).
49. Maity, T., Fuse, N. & Beachy, P.A. Molecular mechanisms of Sonic hedgehog mutant effects in holoprosencephaly. *Proc. Natl. Acad. Sci. USA* **102**, 17026–17031 (2005).
50. Lewis, K.E., Concordet, J.P. & Ingham, P.W. Characterisation of a second patched gene in the zebrafish *Danio rerio* and the differential response of patched genes to Hedgehog signalling. *Dev. Biol.* **208**, 14–29 (1999).
51. Shen, M.C. *et al.* Heat-shock–mediated conditional regulation of hedgehog/gli signaling in zebrafish. *Dev. Dyn.* **242**, 539–549 (2013).
52. Macdonald, R. *et al.* Midline signalling is required for Pax gene regulation and patterning of the eyes. *Development* **121**, 3267–3278 (1995).
53. Lee, J. *et al.* An ENU mutagenesis screen in zebrafish for visual system mutants identifies a novel splice-acceptor site mutation in patched2 that results in colobomas. *Invest. Ophthalmol. Vis. Sci.* **53**, 8214–8221 (2012).
54. Chen, M.H. *et al.* Cilium-independent regulation of Gli protein function by Sufu in Hedgehog signaling is evolutionarily conserved. *Genes Dev.* **23**, 1910–1928 (2009).
55. Roessler, E. *et al.* A previously unidentified amino-terminal domain regulates transcriptional activity of wild-type and disease-associated human GLI2. *Hum. Mol. Genet.* **14**, 2181–2188 (2005).
56. Goodrich, L.V., Milenkovic, L., Higgins, K.M. & Scott, M.P. Altered neural cell fates and medulloblastoma in mouse patched mutants. *Science* **277**, 1109–1113 (1997).
57. Aszterbaum, M. *et al.* Ultraviolet and ionizing radiation enhance the growth of BCCs and trichoblastomas in patched heterozygous knockout mice. *Nat. Med.* **5**, 1285–1291 (1999).
58. Lee, E.Y. *et al.* Hedgehog pathway–regulated gene networks in cerebellum development and tumorigenesis. *Proc. Natl. Acad. Sci. USA* **107**, 9736–9741 (2010).
59. Houzelstein, D. *et al.* Growth and early postimplantation defects in mice deficient for the bromodomain-containing protein Brd4. *Mol. Cell. Biol.* **22**, 3794–3802 (2002).
60. Cooper, M.K., Porter, J.A., Young, K.E. & Beachy, P.A. Teratogen-mediated inhibition of target tissue response to Shh signaling. *Science* **280**, 1603–1607 (1998).

3) Genome Sequencing of SHH Medulloblastoma Predicts Genotype- Related Response to Smoothened-Inhibition

Kool M, Jones D, Jäger N, Northcott P, Pugh T, Hoverstadt V, Piro RM, Esparza LA, Markant SL, Remke M, Milde T, Bourdeaut F, Ryzhova M, Sturm D, Pfaff E, Stark S, Hutter S, Seker-Cin H, Johann P, Bender S, Schmidt C, Rausch T, Shih D, Reimand J, Sieber L, Wittmann A, Linke L, Witt H, Weber H, Zapatka M, König R, Beroukhim R, Bergthold G, van Sluis P, Volckmann R, Koster J, Versteeg R, Schmidt S, Wolf S, Lawerenz C, Bartholomae C, von Kalle C, Unterberg A, Herold-Mende C, Hofer S, Kulozik AE, von Deimling A, Scheurlen W, Felsberg J, Reifenberger G, Hasselblatt M, Crawford JR, Grant GA, Jabado N, Perry A, Cowdery C, Croul S, Zadeh G, Korbel JO, Doz F, Delattre O, Bader GD, McCabe MG, Collins VP, Kieran MK, Cho YJ, Pomeroy S, Brors B, Taylor MD, Schüller U, Korshunov A, Eils R, Wechsler-Reya R, Lichter P, Pfister SM ;

Cancer Cell. 2014 Mar 17;25(3):393-405

Cancer Cell
Article

Genome Sequencing of SHH Medulloblastoma Predicts Genotype-Related Response to Smoothened Inhibition

Marcel Kool,[1,*] David T.W. Jones,[1] Natalie Jäger,[2] Paul A. Northcott,[1] Trevor J. Pugh,[3] Volker Hovestadt,[4] Rosario M. Piro,[2] L. Adriana Esparza,[5] Shirley L. Markant,[5] Marc Remke,[6] Till Milde,[7] Franck Bourdeaut,[8,9] Marina Ryzhova,[10] Dominik Sturm,[1] Elke Pfaff,[1] Sebastian Stark,[1] Sonja Hutter,[1] Huriye Şeker-Cin,[1] Pascal Johann,[1] Sebastian Bender,[1] Christin Schmidt,[1] Tobias Rausch,[11] David Shih,[6] Jüri Reimand,[12] Laura Sieber,[1] Andrea Wittmann,[1] Linda Linke,[1] Hendrik Witt,[1,7] Ursula D. Weber,[4] Marc Zapatka,[4] Rainer König,[2,13,14] Rameen Beroukhim,[3,15,16] Guillaume Bergthold,[3,15,17] Peter van Sluis,[18] Richard Volckmann,[18] Jan Koster,[18] Rogier Versteeg,[18] Sabine Schmidt,[19] Stephan Wolf,[19] Chris Lawerenz,[20] Cynthia C. Bartholomae,[21] Christof von Kalle,[21] Andreas Unterberg,[21] Christel Herold-Mende,[21] Silvia Hofer,[22] Andreas E. Kulozik,[7] Andreas von Deimling,[23,24] Wolfram Scheurlen,[25] Jörg Felsberg,[26] Guido Reifenberger,[26] Martin Hasselblatt,[27] John R. Crawford,[28,29] Gerald A. Grant,[30,31] Nada Jabado,[32] Arie Perry,[33] Cynthia Cowdrey,[34] Sydney Croul,[35] Gelareh Zadeh,[35] Jan O. Korbel,[11] Francois Doz,[8,36] Olivier Delattre,[8,9] Gary D. Bader,[12] Martin G. McCabe,[37] V. Peter Collins,[38] Mark W. Kieran,[39] Yoon-Jae Cho,[40] Scott L. Pomeroy,[41] Olaf Witt,[42] Benedikt Brors,[2] Michael D. Taylor,[6] Ulrich Schüller,[43] Andrey Korshunov,[1,23,24] Roland Eils,[2] Robert J. Wechsler-Reya,[5,44] Peter Lichter,[4,44] and Stefan M. Pfister,[1,7,44] on behalf of the ICGC PedBrain Tumor Project

[1]Division of Pediatric Neurooncology, German Cancer Research Center (DKFZ), 69121 Heidelberg, Germany
[2]Division of Theoretical Bioinformatics, German Cancer Research Center (DKFZ), 69121 Heidelberg, Germany
[3]Broad Institute of MIT and Harvard, Cambridge, MA 02141, USA
[4]Division of Molecular Genetics, German Cancer Research Center (DKFZ), 69121 Heidelberg, Germany
[5]Sanford-Burnham Medical Research Institute, La Jolla, CA 92037, USA
[6]The Arthur and Sonia Labatt Brain Tumour Research Centre, Hospital for Sick Children, Toronto, ON M5G 1L7, Canada
[7]Department of Pediatric Oncology, Hematology and Immunology, University Hospital Heidelberg, 69120 Heidelberg, Germany
[8]Institut Curie, 75005 Paris, France
[9]Institut Curie/INSERM U830, 75248 Paris, France
[10]Department of Neuropathology, NN Burdenko Neurosurgical Institute, Moscow 125047, Russia
[11]European Molecular Biology Laboratory (EMBL), 69117 Heidelberg, Germany
[12]The Donnelly Centre, University of Toronto, Toronto, ON M5S 3E1, Canada
[13]Integrated Research and Treatment Center, Center for Sepsis Control and Care, Jena University Hospital, 07747 Jena, Germany
[14]Leibniz Institute for Natural Product Research and Infection Biology, Hans-Knöll-Institute (HKI), 07745 Jena, Germany
[15]Department of Cancer Biology, Dana Farber Cancer Institute, Boston, MA 02215, USA
[16]Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA
[17]UMR 8203, CNRS Vectorology and Anticancer Therapeutics, Gustave Roussy Cancer Institute, University Paris XI, 94805 Villejuif Cedex, France
[18]Department of Oncogenomics, Academic Medical Center, Amsterdam 1105 AZ, the Netherlands
[19]Genomics and Proteomics Core Facility, German Cancer Research Center (DKFZ), 69121 Heidelberg, Germany
[20]Data Management Facility, German Cancer Research Center (DKFZ), 69121 Heidelberg, Germany
[21]Division of Translational Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), 69121 Heidelberg, Germany
[22]Department of Oncology, University Hospital Zürich, 8006 Zürich, Switzerland
[23]Department of Neuropathology, University of Heidelberg, 69120 Heidelberg, Germany
[24]Clinical Cooperation Unit Neuropathology, German Cancer Research Center (DKFZ), 69121 Heidelberg, Germany
[25]Cnopf'sche Kinderklinik, Nürnberg Children's Hospital, 90419 Nürnberg, Germany
[26]Department of Neuropathology, Heinrich-Heine-University Düsseldorf, 40225 Düsseldorf, Germany
[27]Institute for Neuropathology, University Hospital Münster, 48149 Münster, Germany
[28]Departments of Pediatrics and Neurosciences, University of California San Diego, La Jolla, CA 92093
[29]Rady Children's Hospital, San Diego, CA 92123, USA
[30]Division of Pediatric Hematology/Oncology, Department of Pediatrics, Duke University Medical Center, Durham, NC 27710, USA
[31]Department of Surgery, Duke University Medical Center, Durham, NC 27710, USA
[32]Departments of Pediatrics and Human Genetics, McGill University Health Centre Research Institute, Montreal, QC H3H 1P3, Canada
[33]Departments of Pathology and Laboratory Medicine, University of California, San Francisco, San Francisco, CA 94143, USA
[34]Departments of Pathology and Neurological Surgery, Brain Tumor Research Center, University of California, San Francisco, San Francisco, CA 94143, USA
[35]Department of Neuropathology, The Arthur and Sonia Labatt Brain Tumour Research Centre, Toronto, ON M5G 1L7, Canada
[36]Université Paris Descartes, 75006 Paris, France
[37]Manchester Academic Health Science Centre, Manchester M13 9NT, UK
[38]Department of Pathology, University of Cambridge, Cambridge CB2 1QP, UK
[39]Pediatric Medical Neuro-Oncology, Dana-Farber Cancer Institute and Boston Children's Hospital, Boston, MA 02215, USA
[40]Department of Neurology and Neurosurgery, Stanford University School of Medicine, Stanford, CA 94305, USA
[41]Boston Children's Hospital and Harvard Medical School, Boston, MA 02115, USA

[42]CCU Pediatric Oncology, German Cancer Research Center (DKFZ), 69121 Heidelberg, Germany
[43]Center for Neuropathology and Prion Research, Ludwig-Maximilians-Universität, 81377 München, Germany
[44]These authors contributed equally to this work and are co-senior authors
*Correspondence: m.kool@dkfz.de
http://dx.doi.org/10.1016/j.ccr.2014.02.004

SUMMARY

Smoothened (SMO) inhibitors recently entered clinical trials for sonic-hedgehog-driven medulloblastoma (SHH-MB). Clinical response is highly variable. To understand the mechanism(s) of primary resistance and identify pathways cooperating with aberrant SHH signaling, we sequenced and profiled a large cohort of SHH-MBs (n = 133). SHH pathway mutations involved *PTCH1* (across all age groups), *SUFU* (infants, including germline), and *SMO* (adults). Children >3 years old harbored an excess of downstream *MYCN* and *GLI2* amplifications and frequent *TP53* mutations, often in the germline, all of which were rare in infants and adults. Functional assays in different SHH-MB xenograft models demonstrated that SHH-MBs harboring a *PTCH1* mutation were responsive to SMO inhibition, whereas tumors harboring an *SUFU* mutation or *MYCN* amplification were primarily resistant.

INTRODUCTION

Medulloblastoma (MB) comprises a collection of clinically and molecularly distinct tumor subgroups that arise either in the cerebellum or brainstem (Grammel et al., 2012; Louis et al., 2007; Taylor et al., 2012). In children, they comprise the most frequent embryonal brain tumor, whereas in adults the disease is relatively rare, accounting for less than 1% of all intracranial malignancies (Louis et al., 2007). Current therapy regimens including surgery, cranio-spinal radiotherapy, and chemotherapy, may cure 70%–80% of patients with MB. Most survivors, however, suffer from long-term sequelae because of the intensive treatment, demonstrating that less toxic treatments are urgently needed. Molecular analyses have shown that there are four major MB subgroups (WNT, Sonic Hedgehog [SHH], Group 3, and Group 4; Taylor et al., 2012). They are highly distinct in tumor cell histology and biology, and in addition show divergent clinical phenotypes such as patient demographics, tumor dissemination, and patient outcome (Kool et al., 2012; Northcott et al., 2012a; Taylor et al., 2012). Recent studies, largely focusing on pediatric MB, have utilized next-generation sequencing technologies to map the genomic landscape of MB and to identify novel driver mutations in each molecular subgroup (Jones et al., 2012; Northcott et al., 2012a, 2012b; Parsons et al., 2011; Pugh et al., 2012; Rausch et al., 2012; Robinson et al., 2012). Due to the infrequent occurrence of this disease in adulthood, little is known about the biology and genetics of MB in adults. This also explains why there are few prospective phase III trials for this age group. Most centers treat adult patients with MB either using glioblastoma protocols (which are largely ineffective) or, alternatively, using pediatric MB protocols, although toxicity profiles differ greatly between children and adults, leading to dose-limiting toxicity in a high proportion of adults treated on pediatric protocols (Brandes et al., 2009; Padovani et al., 2007; Spreafico et al., 2005).

Targeted therapy as an alternative treatment option for patients with MB is especially interesting for SHH-MBs. SHH pathway antagonists, primarily those inhibiting at the level of smoothened (SMO), are currently a major area of interest in the pharmaceutical industry because they can potentially be applied in multiple cancers with activated SHH signaling (Lin and Matsui, 2012). Some of these drugs are already in clinical trials for MB (Low and de Sauvage, 2010; Ng and Curran, 2011). SHH-MBs with alterations in downstream SHH pathway genes, however, such as *SUFU*, *GLI2*, or *MYCN*, may demonstrate primary resistance to SMO inhibition (Lee et al., 2007). Furthermore, as has been shown in both humans and mice, tumors may also rapidly acquire secondary resistance to treatment (Dijkgraaf et al., 2011; Rudin et al., 2009; Yauch et al., 2009), suggesting that such inhibitors might be ineffective as a curative option when administered as monotherapy. SHH-MBs present the most common subgroup in infants (≤3 years old) and adults (≥18 years old),

Significance

Our data show that most adults, but only half of the pediatric patients, with SHH-MB will likely respond to SMO inhibition as predicted by molecular analysis of the primary tumor and tested in the SHH xenografts, demonstrating that the next generation of SMO inhibitor trials should be based on these predictive biomarkers. Recurrent mutations in additional pathways suggest rational combination therapies including epigenetic modifiers and PI3K/AKT inhibitors, especially in adults. We also show that tumor predisposition (Gorlin syndrome and Li-Fraumeni syndrome) is highly prevalent in patients with SHH-MB. Each patient with SHH-MB, especially those 4–17 years old with LCA histology, should be tested for germline *TP53* mutations. Separate LFS-MB trials should be considered, sparing radiotherapy and excluding alkylating drugs.

whereas in children (4–17 years old) other subgroups are more prevalent (Kool et al., 2012). Transcriptome analyses and whole genome sequencing have already shown that SHH-MBs are quite heterogeneous (Northcott et al., 2011a; Rausch et al., 2012). Childhood SHH-MBs, for instance, are genetically distinct from those in infants, because they frequently harbor *TP53* mutations and as a result of chromothripsis, their genomes are often dramatically rearranged (Rausch et al., 2012). To preselect patients who might qualify for clinical trials using SMO antagonists or future combination therapies, a better understanding of the biology of SHH-MBs across different age groups is required. We have therefore sequenced the genomes of 133 cases of SHH-MB, including 50 adult and 83 pediatric cases. In addition, we analyzed the tumors for DNA methylation and gene expression.

RESULTS

SHH-MBs in Infants, Children, and Adults Are Genomically Distinct

Unsupervised *k*-means consensus cluster analysis of DNA methylation data (n = 129) identified two major clusters, mainly separating infant from childhood and adult SHH-MB tumors (Figure 1A, left panel). Unsupervised cluster analysis of gene expression data (n = 103) showed similar results, with the infant cases again being the most distinct (Figure 1A, right panel). GISTIC2 analysis of somatic copy number aberrations in all SHH-MB cases (n = 266) reported by MAGIC (Northcott et al., 2012b), however, showed that childhood SHH-MBs are very different from both infant and adult SHH-MBs (Figure 1B). Childhood SHH-MBs typically show much greater genomic instability and are characterized by frequent amplifications of oncogenes including *GLI2*, *MYCN*, and *PPM1D*, most likely due to underlying chromosome shattering (chromothripsis; Rausch et al., 2012).

Next-Generation Sequencing of SHH-MB

To determine the mutational landscape of SHH-MBs across age groups, we sequenced a large series of SHH-MB tumors from infants ($\leq$3 years old; n = 50), children (4–17 years old; n = 33), and adults ($\geq$18 years old; n = 50; Table 1; Table S1 available online). In the discovery cohort of 67 SHH-MBs, analyzed by whole genome or whole exome sequencing, we identified 1,090 nonsynonymous somatic single nucleotide variants (SNVs) and 155 small insertions or deletions (indels), 89 of which introduced translational frameshifts and 9 affected splice sites. In total, 1,054 genes were found to be somatically mutated in this discovery cohort, including 78 with alterations in more than one tumor. In the two replication cohorts (43 pediatric and 23 adults), we identified another 666 nonsynonymous SNVs and 76 indels. For the combined 133 SHH-MBs, we found mutations in 1,156 genes, 215 of which were recurrently altered. All coding somatic SNVs/indels identified are listed in Table S2.

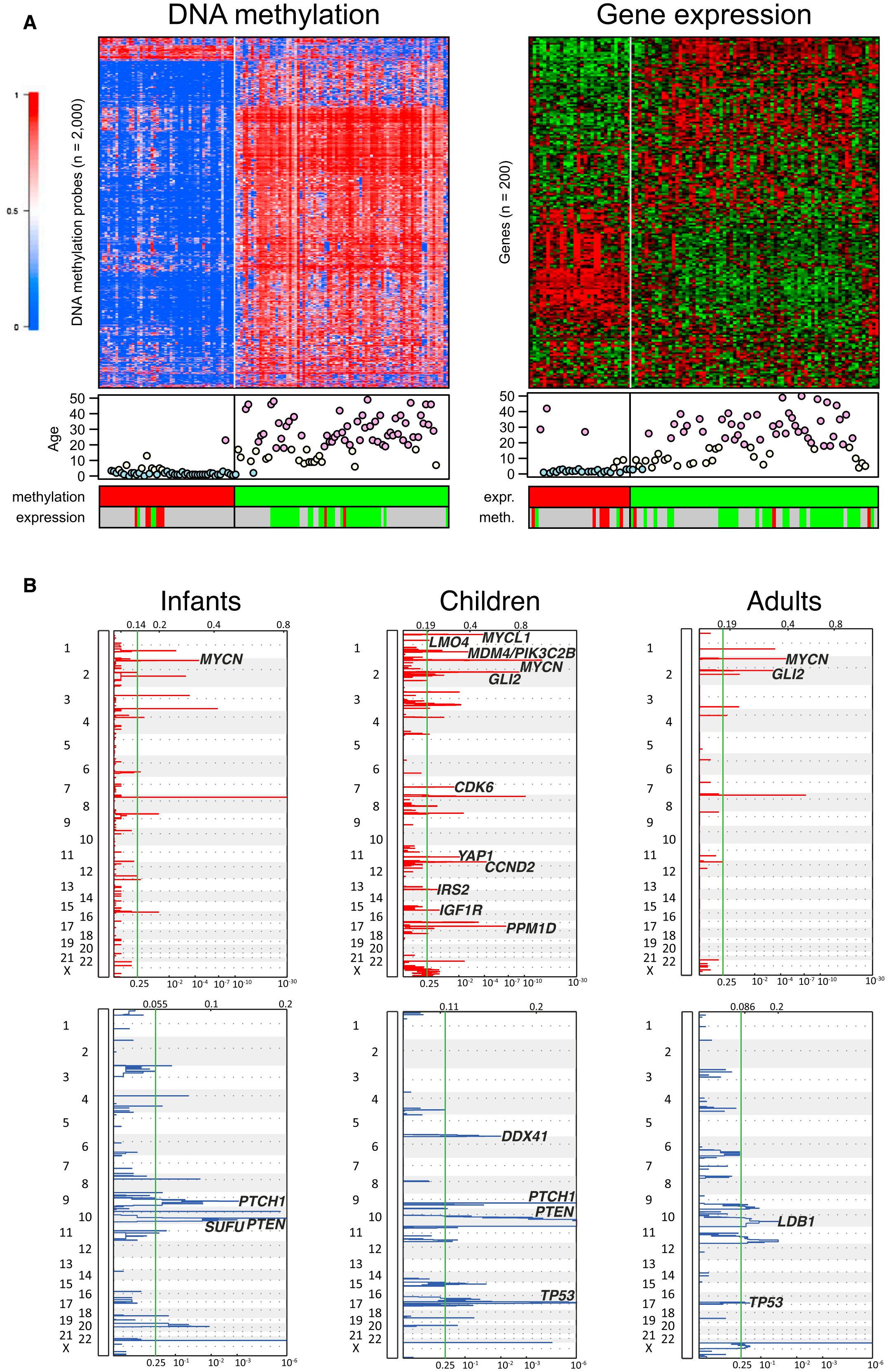
As previously reported (Jones et al., 2012), pediatric SHH-MBs harbored very few nonsynonymous SNVs (infants, 0–13, median 3.0; children [*TP53* wild-type], 1–26, median 9.5; Table S2; Figures 2A and 2B). Exceptions were the eight *TP53* mutated tumors in children, in this discovery cohort all between 9.5 and 14 years old, which harbored on average many more mutations (7–29, median 19.5). WGS data showed that adult SHH-MBs also contained many more nonsynonymous SNVs (9–48, median 25.0), in line with other adult solid tumors. The average number of small indels was also higher in adults (0–10, median 3.0) than in children (0–4, median 1.0) and infants (0–3, median 1.0). Interestingly, there was a much stronger correlation between somatic mutation rate and patient age, both genome-wide (r^2 = 0.58, p = 1.6×10^{-9}, Pearson's product moment correlation), and for coding mutations (r^2 = 0.62, p = 2.2×10^{-15}), than previously reported across all MB subgroups (Figures 2A and 2B; Jones et al., 2012). Assessment of mutation classes revealed a predominance of cytosine to thymine (C > T) transitions in a CpG context (likely due to deamination of methylated cytosines), as expected for an age-related background mutation pattern (Figures 2C and 2D; Welch et al., 2012). Interestingly, the C > T fraction in the *TP53* mutated cases appeared to be much lower, with a relatively higher proportion of cytosine to adenosine (C > A) transitions. Whether this can be explained by the *TP53* mutation itself remains elusive.

Mutations in the SHH Pathway

Overall, we detected mutations in known SHH pathway genes (116/133 cases; 87%), further substantiating the tumor-driving role of the SHH pathway in this medulloblastoma subgroup (Table S3). As expected, among the most frequently mutated genes were *PTCH1* (60 cases), *SMO* (19 cases), and *SUFU* (10 cases), all mutually exclusive (Figure 3A; Figures S1A–S1C). In addition, we found two *PTCH1* and six *SUFU* mutations in the germlines of eight pediatric patients, including two twin brothers with an identical small indel in *SUFU* (Table S3). The second replication cohort (for which germline controls were unavailable) contained another two cases from twin brothers both with the same inactivating *SUFU* mutation, strongly suggesting that this was also a germline event. For all other samples in this replication cohort, it remains unknown whether any of the identified *PTCH1* or *SUFU* mutations were germline events. Interestingly, while *PTCH1* mutations were found at roughly equal frequency in infants (42.0%), children (36.4%), and adults (54.0%), *SMO* mutations were highly enriched in adult patients (15/19 mutations; p = 1.8×10^{-4}), while *SUFU* mutations were almost exclusively found in infants $\leq$3 years old (16/18 mutations; p = 8.4×10^{-6}). Mutations in *SMO* and *SUFU* were absent or extremely rare in children (4–17 years old; Figures 3A and 3B). Instead, they frequently harbored *TP53* mutations (16/33 children; p = 1.2×10^{-11}), all found in children between 8 and 17 years old. The *TP53* mutations were mutually exclusive with *PTCH1* mutations but often co-occurred with amplifications of *GLI2* (p = 2.5×10^{-6}) and *MYCN* (p = 2.8×10^{-8}), three events that were rare in infants, young children, and adults (Figures 3A and 3B). In addition, we identified four cases, including three children with a *TP53* mutation, with an amplification of the *SHH* gene. These results show that activating mutations in the SHH pathway are detectable in almost all SHH-MBs, but the type of mutation and targeted genes are largely variable in the different age groups (Figure 3C).

Large Cell/Anaplastic Histology and 17p Loss Are Strongly Associated with TP53 Mutated SHH-MBs

Losses of 9q, 10q, and/or 17p are the most common copy number aberrations associated with SHH-MBs (Kool et al., 2012). All



(legend on next page)

Table 1. SHH-MB Patient Cohorts

Cohort	Number of Patients
Whole genome sequencing[a]	n = 45
Infants[b]	5
Children[c]	13
Adults[d]	27
Whole exome sequencing[a]	n = 22
Infants	13
Children	9
Adults	0
Targeted sequencing 2734 genes[a] 12	n = 12
Infants	7
Children	5
Adults	0
Targeted sequencing 400 genes[e]	n = 54
Infants	25
Children	6
Adults	23
Immunohistochemistry	n = 155
Infants	31
Children	54
Adults	70

See also Tables S1, S2, S3, and S4.

[a] Tumor-normal pairs were sequenced.

[b] Infants: 0–3 years of age.

[c] Children: 4–17 years of age.

[d] Adults: ≥ 18 years of age.

[e] Only tumors were sequenced.

three were most frequent in childhood cases, with 17p loss highly enriched in *TP53* mutant cases (14/17 had 17p loss; $p = 7.8 \times 10^{-8}$; Figures 3A and 3B). Histology was also unequally distributed between the three age groups, with most large cell/anaplastic (LCA) cases found in childhood (15/21; $p = 4.1 \times 10^{-9}$). Thirteen of these 15 had a *TP53* mutation. Nodular/desmoplastic MB variants were most prevalent in infant cases. Moreover, all four MBs with extensive nodularity (MBEN) were found in infants (Figures 3A and 3B). In contrast to a recent report (Brugières et al., 2012), which was, however, reporting on a larger number of MBEN MBs, only 1/4 MBEN cases in our series had an *SUFU* mutation, while two harbored a *PTCH1*, and one displayed an *SMO* mutation (Figure 3A).

TERT Promoter Mutations Are Highly Recurrent in Adult SHH-MBs

Recently, several groups have reported that *TERT* promoter mutations that drive telomerase activity are frequently found in various cancers, including medulloblastoma, of mainly adult patients (Killela et al., 2013; Koelsche et al., 2013; Remke et al., 2013). Two mutually exclusive hotspot mutations in the promoter region have been reported: C228T and the less frequent C250T. Using our WGS data and data from the replication cohort in which the *TERT* promoter region was analyzed by PCR and Sanger sequencing (Remke et al., 2013), we found that indeed these mutations almost exclusively and with high frequency occur in adult SHH-MBs (Table S1). Strikingly, almost all adult patients for which we had data available had a somatic *TERT* promoter mutation (43/44, 98%; 40 had the C228T mutation and 3 had the C250T mutation). In contrast, in infants and children, only 3/24 (13%) and 3/14 (21%) SHH-MBs, respectively, had a *TERT* mutation (five C228T and one C250T).

DDX3X and Chromatin Modifiers Are Frequently Mutated in Adult SHH-MBs

Other genes previously reported as being recurrently mutated in pediatric SHH-MBs (*MLL2*, *BCOR*, and *LDB1*) were also found in adult SHH-MBs (Figures 4A–4C). Interestingly, however, we identified several recurrent mutations in adult SHH-MBs that were completely absent or very rare in pediatric SHH-MBs, including *BRPF1*, *KIAA0182*, *TCF4*, *CREBBP*, *NEB*, *LRP1B*, *PIK3CA*, *FBXW7*, *KDM3B*, *XPO1*, *PRKAR1A*, and *PDE4D* (Figures 4A–4C; Figures S1D–S1I). Another striking example is the gene encoding the RNA helicase DDX3X, which was mutated in 27 adult SHH-MBs (54%) and only 6 pediatric MBs (7.2%, $p = 4.5 \times 10^{-9}$). *DDX3X* was among the new genes identified in recent sequencing studies of pediatric MB (Jones et al., 2012; Pugh et al., 2012; Robinson et al., 2012). Notably, whereas mutations were found in 50% of WNT-MBs in children (Northcott et al., 2012a), few *DDX3X* mutations were seen in SHH-MBs in these studies (Pugh et al., 2012; Robinson et al., 2012). All identified mutations affected one of the two helicase domains with no difference in their distribution between WNT- and SHH-MBs (Figure S1D). Interestingly, mutations affecting the SWI-SNF complex, also mainly found in the WNT-MBs in children (Jones et al., 2012; Northcott et al., 2012a; Pugh et al., 2012; Robinson et al., 2012), were also frequently seen in adult SHH-MBs.

Pathway analyses, performed separately for the three age groups, showed marked differences in altered processes. In infant cases, developmental processes and DNA/histone methylation are prominently affected. Both in children and in adults, chromatin organization is also affected, but especially in adults many more chromatin modifiers and/or transcription regulators were additionally altered, as well as different and larger gene sets involved in brain development (Figure S2 and Table S4). Remarkably, most of the mutations in chromatin modifiers in adults were found to be mutually exclusive (Figure 4D). Interestingly, some of these mutations in chromatin

Figure 1. Genetic and Epigenetic Differences between SHH-MBs from Infants, Children, and Adults

(A) Cluster analysis of DNA methylation and gene expression data of SHH-MB. Both methylation profiling (left; n = 129) and gene expression profiling (right; n = 103) reveal two SHH-MB subgroups identified by unsupervised *k*-means consensus clustering. Each row represents a methylation probe/expression probeset, each column represents a sample. The level of DNA methylation (b value) is represented with a color scale as depicted. For each sample patient age (blue, infants; yellow, children; and pink, adults) and clustering according to expression data or methylation data (when available) is shown. Grey indicates that no data were available.

(B) GISTIC2 significance plots of amplifications (red) and deletions (blue) observed in SHH-MB infants, children, and adults. Candidate genes mapping significant regions have been indicated.

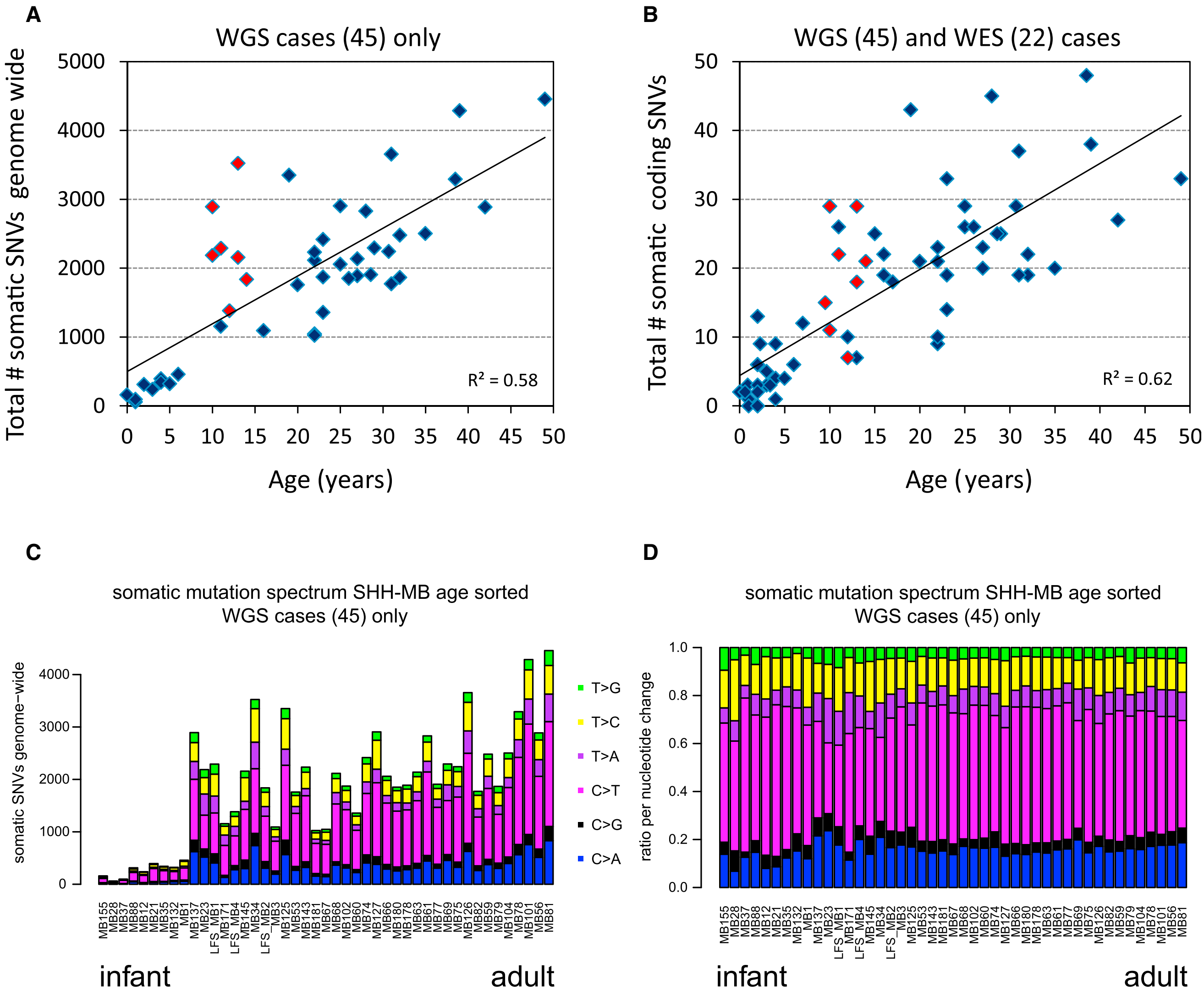


Figure 2. Number and Type of Somatic Mutations in Medulloblastoma Tumors in Relation to the Age of the Patient

(A) Total number of somatic mutations genome wide correlates with age of the patient. Plotted are the total number of somatic SNVs identified genome wide versus age of the patient for all cases for which we performed whole genome sequencing (WGS; n = 45). Red indicates patients harboring a *TP53* mutation.

(B) Same as in (A), but only the total number of coding SNVs is plotted versus age for all cases for which we performed either whole genome or whole exome sequencing (WGS and WES, n = 67).

(C) Mutation signatures. Plotted are the total numbers of somatic mutations genome wide sorted by age of the patient. Coloring of bars represents the ratio of the six possible nucleotide changes (C > A, C > G, C > T, T > A, T > C, and T > G) for each sample.

(D) Normalized mutation signatures sorted by age.

modifiers were more closely associated with *SMO* mutations, like the ones in *BRPF1/3*, while mutations in *CREBBP* or *KDM3B* were more often found in *PTCH1*-mutated tumors.

PI3K/AKT Signaling Activated in Adult SHH-MB Associates with Poor Outcome

As we identified recurrent mutations affecting the PI3K/AKT/mTOR-pathway in SHH-MBs (*PIK3CA*, *PTEN*, and *PIK3C2G* are all mutated in >5% of SHH-MBs; Figures 4A–4C), which could lead to GLI activation independent of SMO (Wang et al., 2012), targeting this pathway could be an option for combination therapies. To investigate which SHH patients would be most suitable for targeting PI3K/AKT/mTOR-signaling, we examined activation of the pathway in a large series of SHH-MBs (n = 155) by immunohistochemistry using antibodies for p-AKT and p-S6. p-AKT and p-S6 positivity were each detected in 17% of cases, with 12% positive for both (Figures 5A–5F). Surprisingly, the vast majority of positive cases were tumors from adult patients, with 31% and 30% of the adult SHH-MBs staining positive for p-AKT or p-S6, respectively. Moreover, survival analysis showed that both p-AKT and p-S6 positivity were strongly associated with a poor outcome in adult patients with SHH-MB (Figure 5G). Other factors shown to be associated with a poor outcome in SHH-MB patients, like

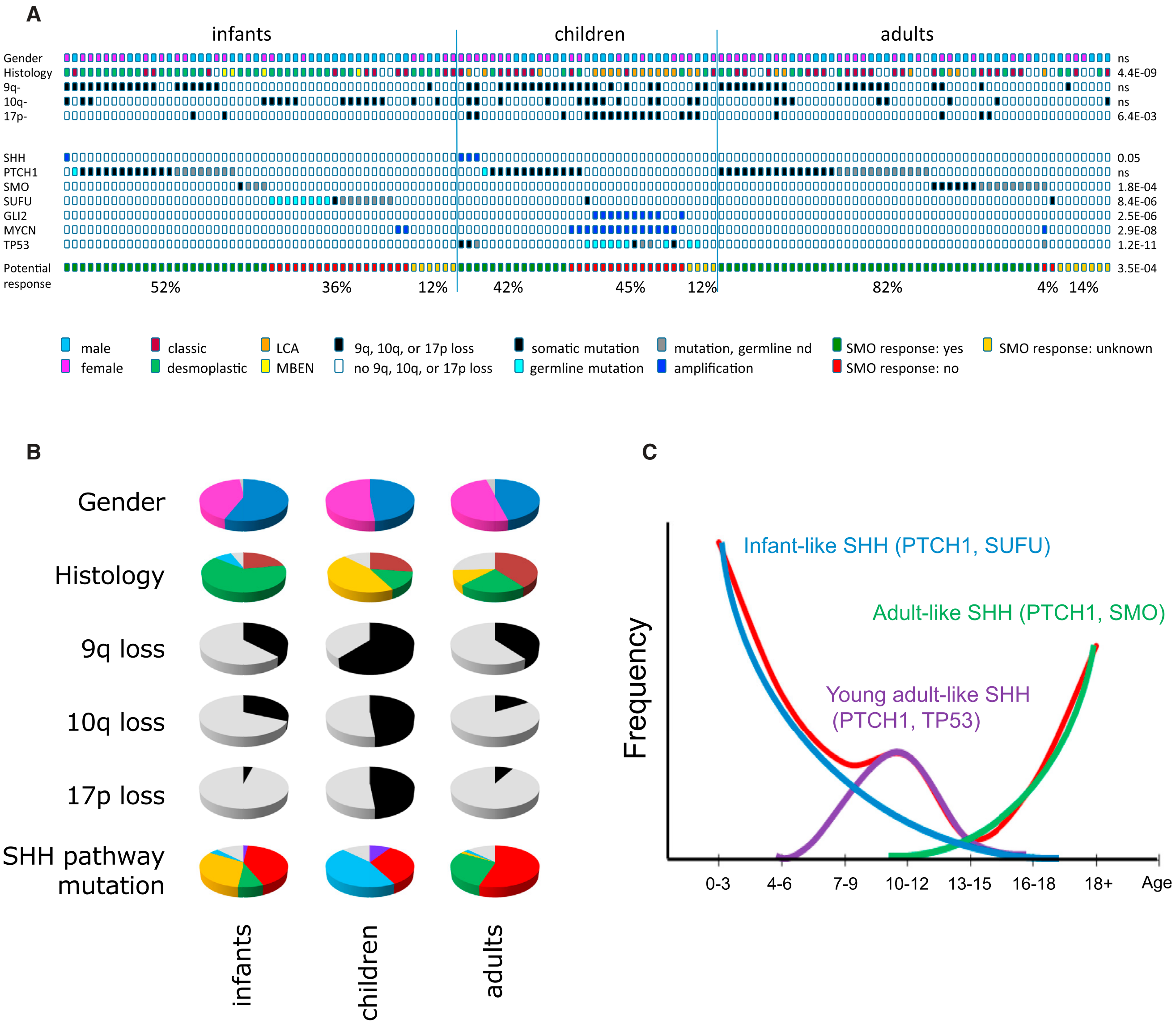


Figure 3. Genetic and Histological Differences between SHH-MBs from Infants, Children, and Adults

(A) SHH pathway mutations, gender, histology and 9q/10q/17p aberrations in all sequenced 133 SHH-MB. Cases have been split up in infants, children and adults, and are sorted based on type of mutation in the SHH-pathway. Potential response to SMO inhibition: cases with *SHH* amplifications, *PTCH1* mutations, or *SMO* mutations will likely respond to SMO inhibition (indicated in green). Cases with *SUFU* mutations or *MYCN* or *GLI2* amplifications will likely not respond to SMO inhibition (indicated in red). In cases for which no mutations in the SHH pathway were detected, it is not clear whether they will respond to SMO inhibitors (indicated in yellow). Percentages indicate fraction of infants, children, or adults, respectively, of each category. p Values indicate whether distributions are significantly different among infants, children, and adults.

(B) Pie charts showing in infants, children, and adults with SHH the distribution of gender (male, blue; female, pink; unknown, gray), histology (classic, dark red; nodular/desmoplastic, green; large cell/anaplastic LCA, orange; MBEN, yellow; and unknown, gray), 9q loss (yes, black; no, gray), 10q loss (yes, black; no, gray), 17p loss (yes, black; no, gray), and type of SHH pathway mutation (*SHH* amp, purple; *PTCH1* mut, red; *SMO* mut, green; *SUFU* mut, orange; *GLI2*/*MYCN* amp, blue; and unknown, gray).

(C) Trimodal age distribution of patients with SHH-MB. Red line indicates age distribution of all patients with SHH-MB. Three subgroups make up this age distribution: young children with *PTCH1* and *SUFU* mutations (blue line), older children with *PTCH1* and *TP53* mutations (purple line), and adults who mostly have *PTCH1* or *SMO* mutations (green line).

See also Figure S1.

MYCN or GLI2 amplification, LCA histology or metastasis at diagnosis, are all exceptionally rare in adult SHH-MB patients (Figures 1C and 3A; Kool et al., 2012), and could therefore not explain the poor outcome of these p-AKT/p-S6-positive subgroup of patients. Our results suggest that adult patients with SHH-MB may be the best group to benefit from combination therapies of SMO inhibitors with PI3K/AKT/mTOR inhibition.

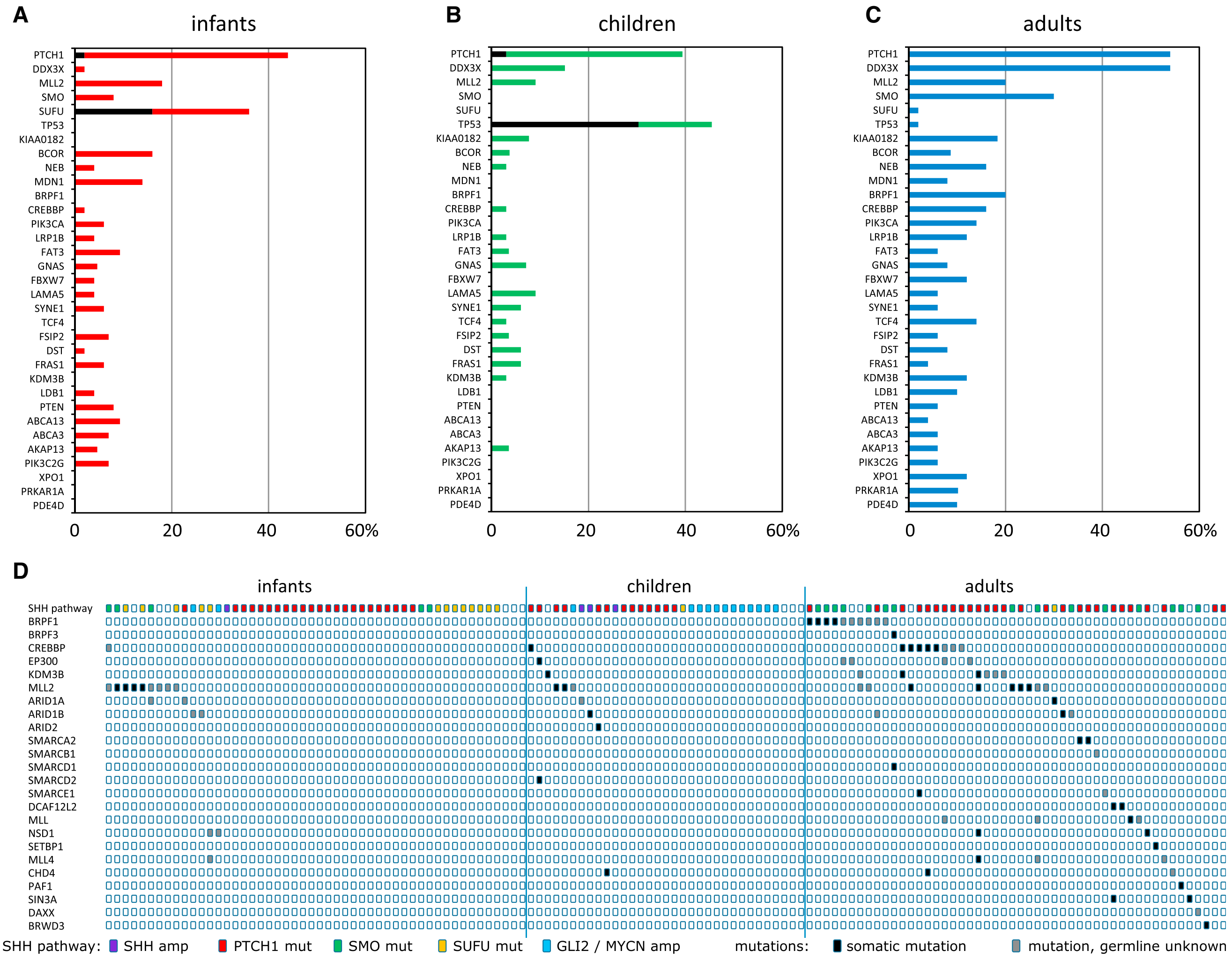


Figure 4. Most Frequently Mutated Genes in SHH-MB and the Mutual Exclusivity of Mutations in Chromatin Modifier Genes

(A–C) Mutation frequencies of 33 genes that are mutated either in $\geq$5% of all SHH-MB cases or in $\geq$10% of SHH-MB cases in one of the age categories. Mutation frequencies for these 33 genes are shown in infants (A), children (B), and adults (C). Black indicates the fraction of mutations that is found in the germline.

(D) Mutations in chromatin modifiers in infants, children, and adults with SHH-MB. The top line shows the mutations in the SHH pathway for each case.

See also Figure S2.

SHH Medulloblastomas with Mutations Downstream of SMO Are Resistant to LDE-225

Assuming a linear pathway, we anticipate that patients with mutations in the SHH pathway downstream of SMO (e.g., *SUFU*, *GLI2*, and *MYCN*) show primary resistance to targeted SMO inhibition. To test this hypothesis, we used xenografts from three SHH-associated MBs (DMB-012, RCMB-018, and RCMB-025; Yeh-Nayre et al., 2012). These xenografts were generated by stereotaxic orthotopic xenotransplantation of cells immediately after surgical resection, maintained by serial intracranial transplantation, and harvested only for use in short-term experiments, allowing them to maintain the characteristics of the original tumors (Shu et al., 2008; Zhao et al., 2012). WES showed that each xenograft harbored a different alteration in the SHH pathway (Figure 6A). Cells from each xenograft line were treated in vitro with NVP-LDE225, an SMO inhibitor that is currently being applied in phase III clinical trials for relapsed childhood and adult SHH-MB (Geoerger et al., 2012). Proliferation was measured based on incorporation of tritiated thymidine. Treatment with LDE225 significantly inhibited the proliferation of DMB-012 cells (*PTCH1* mutant), but did not affect the proliferation of RCMB-018 (*MYCN* amplification) or RCMB-025 cells (*SUFU* deletion; Figures 6B–6D). Preclinical testing in vivo also demonstrated a strong inhibition of tumor growth by LDE225 in DMB-012 (Figure 6E), but not RCMB-018 (Figure 6F and Figure S3), confirming the in vitro data. Survival analyses indeed show that mice with DMB-012 tumors live longer when treated with LDE-225 (Figure 6G), but mice with RCMB-018 tumors do not (Figure 6H). Finally, we have tested whether RCMB-018 cells, resistant to LDE-225, are responsive to arsenic trioxide (ATO) targeting cells at the level of GLI (Beauchamp et al., 2011). Figure 6I illustrates that RCMB-018 cells are responsive to ATO. At concentrations of 5–10 μM, cells are markedly inhibited in growth. Our data

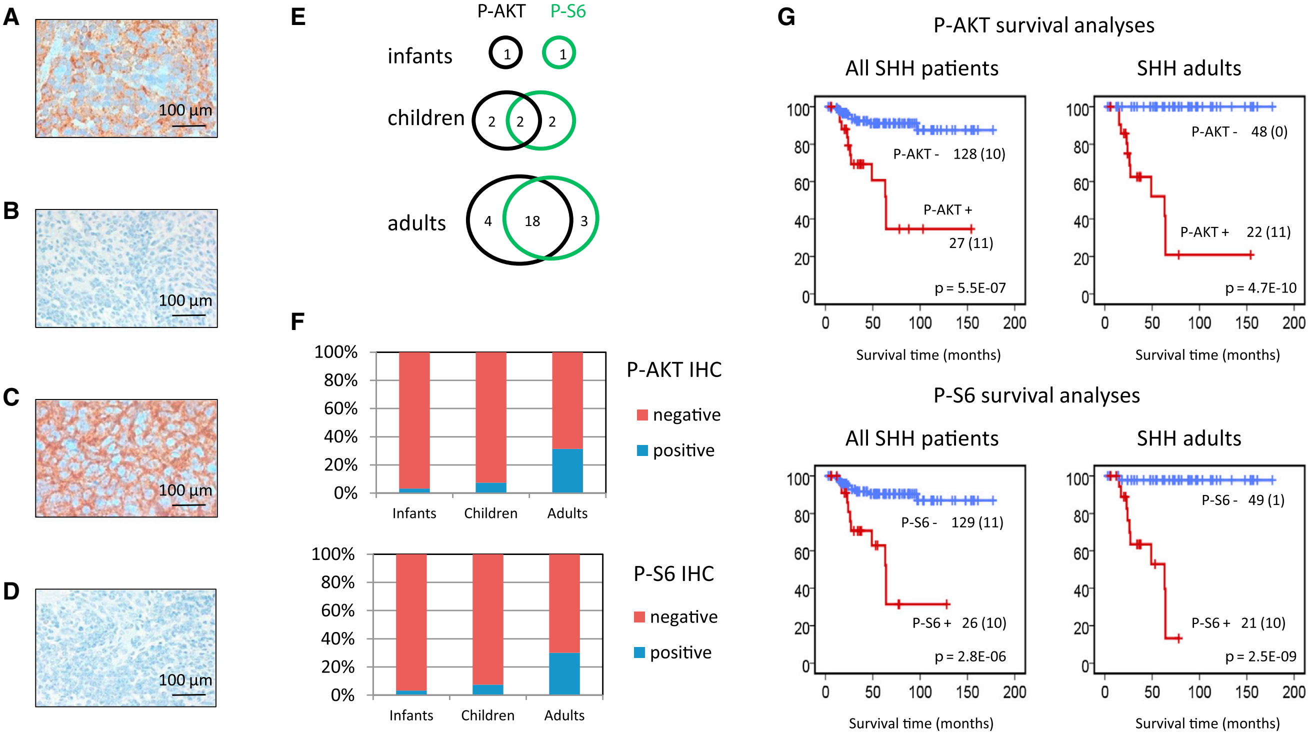


Figure 5. Immunohistochemical Staining of MB Tissue Arrays for p-AKT and p-S6

(A) Example of positive p-AKT MB.

(B) Example of negative p-AKT MB.

(C) Example of positive p-S6 MB.

(D) Example of negative p-S6 MB.

(E) Overlap in staining results between p-AKT and p-S6.

(F) Frequencies of p-AKT and p-S6 staining in infants, children, and adults.

(G) Survival analysis for p-AKT and p-S6 in all SHH patients and in adults only. Numbers on the y-axis indicate the fraction of surviving patients. Numbers on the x-axis indicate the follow-up time in months. The number of patients per group is indicated next to the graphs plus the number of events within that group (between brackets). For infants and children, the number of patients staining positive was too low to draw conclusions from separate survival analyses.

show that classification as an SHH-MB using a five-gene expression signature currently being applied in clinical trials is not sufficient as a predictive biomarker for response to SMO antagonists, because all SHH-MBs are detected by this signature, regardless of their underlying genetic makeup (Amakye et al., 2012).

DISCUSSION

Herein we have shown that genetic hits in SHH-MBs are very heterogeneous. Tumors in infants, children, and adults strongly differ in transcriptome, methylome, and copy-number aberrations as well as in number and type of mutations they contain. Hereditary predisposition syndromes involving germline mutations of *SUFU* (or rarely *PTCH1*; Gorlin syndrome) are highly prevalent in infant (0–3 years old) SHH-MBs, while germline *TP53* mutations (Li-Fraumeni syndrome) are common in older children (>3 years old), especially in children between 8 and 17 years old. Strikingly, almost all adults harbored somatic mutations in the *TERT* promoter, whereas they were much less common in pediatric patients. Our data show that three groups of SHH-MBs should be considered: young children with mostly *PTCH1* or *SUFU* mutations, older children with frequent germline *TP53* mutations and chromothripsis-associated amplifications of SHH pathway genes, and adults harboring mostly *PTCH1* and *SMO* mutations (Figure 3C). Recent data showing that SHH-MBs can arise from different precursor cells in the cerebellum or brainstem (Grammel et al., 2012) suggest that infant SHH-MBs may have a different cellular origin or hit the same progenitor cell at a different stage of differentiation than childhood or adult SHH-MBs (which were more similar at the transcriptome/methylome levels).

Most importantly, our results show that patients with different underlying SHH mutations should be stratified accordingly. We have demonstrated that targeting the SHH pathway in SHH-MB using SMO antagonists will most likely give the best results in adult patients. A vast majority (82%) of adult patients harbor tumors with mutations in either *PTCH1* or *SMO*, rendering them likely responsive to these drugs. In contrast, infant (36%) and childhood (45%) SHH-MBs frequently have mutations downstream of SMO, which makes these tumors intrinsically resistant to drugs targeting SMO. Indeed, SHH-MB xenografts harboring these downstream mutations did not respond to SMO antagonists. The impact of bone developmental toxicity may additionally limit the use of SMO inhibitors in infants (Kimura et al., 2008).

A

case #	MB subgroup	patient age	gender	tumor histology	mutation
DMB-012	SHH	3	female	desmoplastic	*PTCH1* (4 bp del, hemizygous)
RCMB-018	SHH	7	male	anaplastic	*TP53* (R175H, hemizygous) *MYCN* amplification
RCMB-025	SHH	0	female	MBEN	*SUFU* deletion (heterozygous)

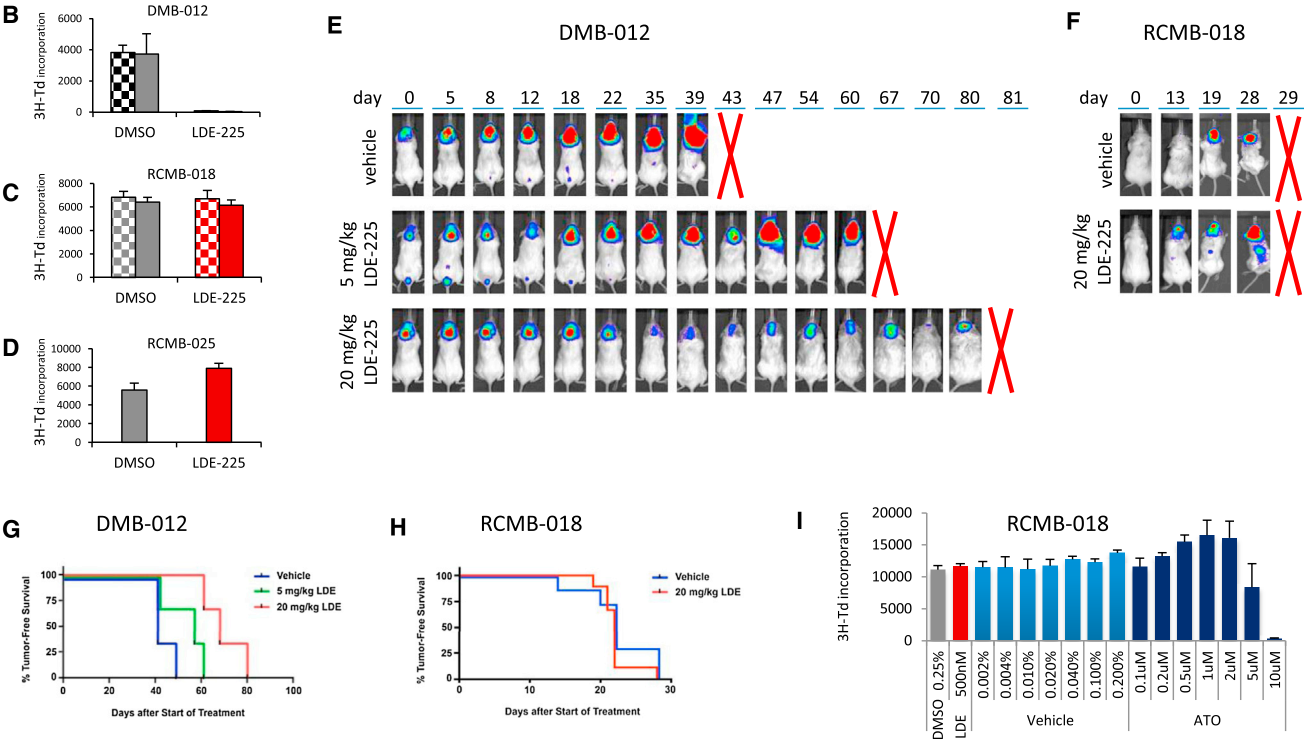


Figure 6. SMO Antagonists Do Not Suppress Proliferation of All SHH-Associated MB Tumors

(A) Characteristics of SHH-MB models treated with LDE225.

(B–D) Cells from patient-derived xenografts of SHH-associated MB were treated with DMSO (0.05% [hatched bars] or 0.25% [solid bars]) or LDE-225 (100 [hatched bars] or 500 nM [solid bars]). Cells were pulsed with [methyl-^{3}H]thymidine (^{3}H-Td) after 48 hr and harvested for analysis of ^{3}H-Td incorporation at 66 hr. In DMB-012 (B), LDE-225 significantly inhibited ^{3}H incorporation compared to DMSO control ($p < 0.01$ based on paired two-tailed t test). In RCMB-018 (C) and RCMB-025 (D), LDE-225 did not significantly inhibit ^{3}H incorporation ($p > 0.5$ and $p > 0.1$, respectively). Data represent means of triplicate samples ± SD.

(E and F) Cells from MB xenograft DMB-012 (E) or RCMB-018 (F) were infected with luciferase virus and transplanted into NSG mice. Bioluminescence images were taken pretreatment (day 0) and at different time points after daily treatment with vehicle or SHH antagonist (LDE-225, 5 or 20 mg/kg/day). Five mice per group were used. Representative examples from each group are shown. Other examples are shown in Figure S3. A red cross indicates when mice were sacrificed.

(G and H) Kaplan-Meier survival plots for the mice harboring DMB-012 tumors (G) or RCMB-018 tumors (H) and treated with vehicle or LDE-225.

(I) RCMB-018 cells were treated with DMSO (0.25%; gray bar), LDE-225 (500 nM; red bar), vehicle (PBS + 0.01 N NaOH; light blue bars), or increasing concentrations of ATO (dark blue bars). Cells were pulsed with [methyl-^{3}H]thymidine (^{3}H-Td) after 48 hr and harvested for analysis of ^{3}H-Td incorporation at 66 hr. LDE-225 did not inhibit ^{3}H incorporation compared to DMSO control, but ATO did at 5 and 10 μM concentrations. Data represent means of triplicate samples ± SD.

See also Figure S3.

Furthermore, our results strongly suggest that each patient with a SHH-MB, but especially those between 4 and 17 years of age with LCA histology, should be tested for germline *TP53* mutations. Currently, these patients with Li-Fraumeni syndrome (LFS)-MB are often not recognized and therefore treated with standard protocols, including ionizing radiotherapy and alkylating chemotherapy. Moreover, as almost all patients with germline *TP53* mutations have tumors with LCA histology, they are often stratified as high risk and will therefore get even higher doses of radiotherapy and chemotherapy. It seems that these patients are often cured of their MB, but frequently die of secondary malignancies induced by previous radio-chemotherapy. This may partly explain why *TP53* mutations in SHH-MBs are associated with a particularly poor outcome (Zhukova et al., 2013), and is also in line with the finding that *MYCN* amplification in SHH-MBs is associated with an inferior prognosis (Kool et al., 2012; Korshunov et al., 2012; Ryan et al., 2012). We therefore strongly suggest that separate LFS-MB trials

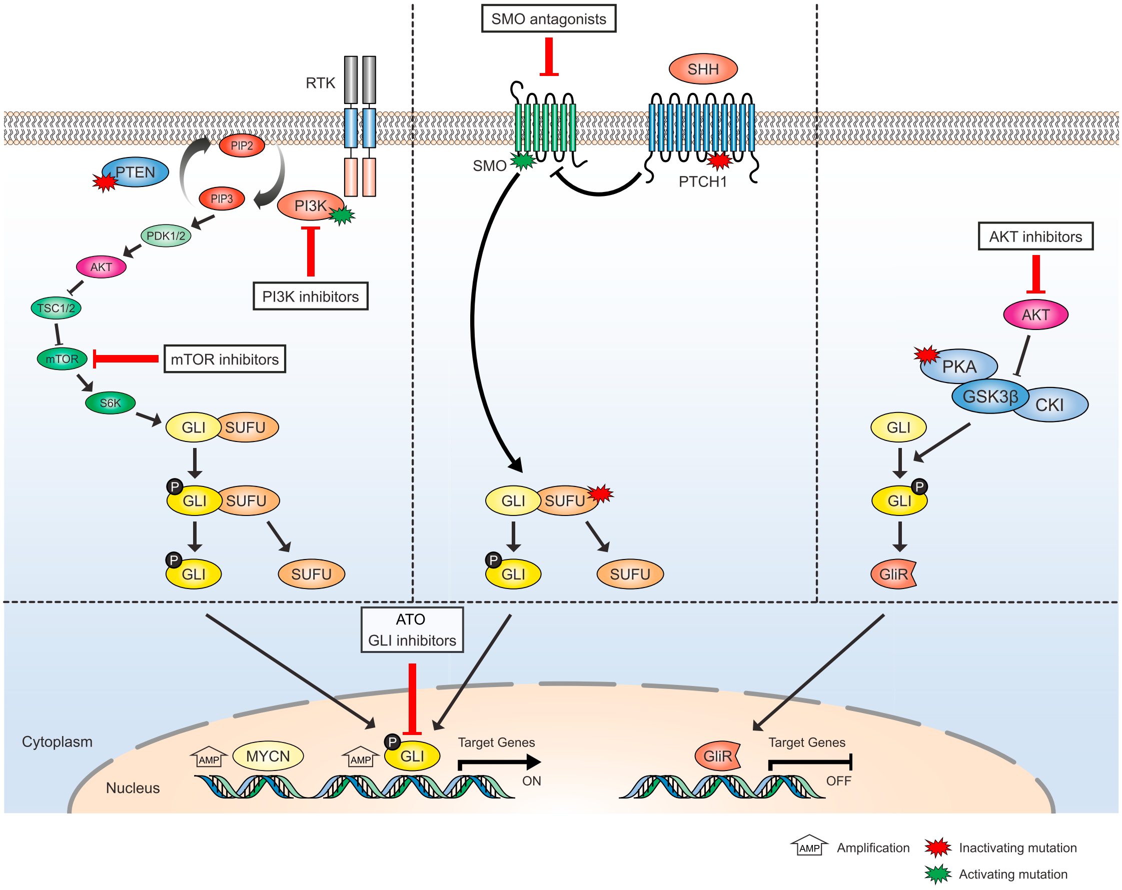


Figure 7. Schematic Overview of SHH-, PI3K/AKT/mTOR-, and PKA Pathways and How They Interact

Genes that were found in the genomic analyses of SHH-MBs to harbor activating mutations (green stars), inactivating mutations (red stars), or were found to be amplified (*MYCN* and *GLI*) are indicated. All these mutations lead to activation of GLI proteins and their downstream pathways. Options for targeted treatment are indicated. Patients harboring mutations in either *PTCH1* or *SMO* should be responsive to SMO inhibitors, whereas patients harboring mutations more downstream in the SHH pathway (*SUFU*, *MYCN*, and *GLI*) or in the PI3K/AKT/mTOR and/or PKA-pathways may be treated using arsenic trioxide (ATO) or other more specific GLI-inhibitors or PI3K/AKT/mTOR inhibitors.

should be developed using chemotherapy-only protocols and excluding alkylating drugs.

We further strongly advocate that the next generation of SMO inhibitor trials should be based on underlying tumor genetics because many patients with SHH-MB will not respond to these inhibitors. Alternative treatment options could include arsenic trioxide (ATO) targeting GLI transcription factors by degrading the protein (Figure 7; Kim et al., 2010, 2013). ATO and the antifungal agent itraconazole (which acts on SMO) have also been suggested in preclinical experiments for use in SHH-MBs that become resistant after treatment with SMO antagonists (Kim et al., 2013) or in combination with SMO inhibitors upfront knowing that *GLI2* amplifications comprise a common mechanism of secondary resistance to SMO inhibition in preclinical models (Buonamici et al., 2010; Dijkgraaf et al., 2011). Other options for combination therapies to avoid or delay the development of resistance include drugs targeting PI3K/AKT/mTOR- or PKA-signaling pathways (Figure 7), both mutated in a subset of patients with SHH and both also leading to GLI activation (Metcalfe et al., 2013; Milenkovic and Scott, 2010; Wang et al., 2012), or epigenetic drugs.

EXPERIMENTAL PROCEDURES

Patient Samples

Patient materials were collected after receiving informed consent according to International Cancer Genome Consortium guidelines (http://www.icgc.org) and as approved by the institutional review board of contributing centers. DNA derived from SHH-MBs and matched normal blood from 45 patients was subjected to whole genome sequencing (WGS) using Illumina technologies. Two additional tumor-normal pairs were sequenced by whole exome sequencing (WES). WGS data for 13/45 and WES data for another 20 pediatric tumor-normal pairs were previously reported (Jones et al., 2012; Pugh et al., 2012). All patients in this discovery cohort (n = 67) were confirmed to have a MB of the SHH subtype by either gene expression profiling, DNA methylation, or immunohistochemistry (SFRP1 Northcott et al., 2011b and GAB1 Ellison et al., 2011). In addition, we used data from 12 pediatric SHH-MB tumor-normal pairs that were sequenced for 2,734 genes as part of a previously reported replication cohort (Jones et al., 2012). Finally, a set of 400 genes, including those identified as recurrently mutated in SHH-MBs in our discovery cohort, was investigated in another independent set of pediatric (31) and adult

(23) SHH-MBs, for which only tumor DNA was available. In total, sequencing data for 133 (83 pediatric and 50 adult) SHH-MBs are presented in this study. Patient details are listed in Table S1.

Animals

Immunocompromised (NOD-scid IL2Rgammanull or NSG) mice used for transplantation were purchased from Jackson Labs. Mice were maintained in the Animal Facility at Sanford-Burnham. All experiments were performed in accordance with national guidelines and regulations, and with the approval of the animal care and use committee at Sanford-Burnham.

The experimental procedures used in this study are described in more detail in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

The Gene Expression Omnibus accession numbers for the complete CpG methylation values are GSE49576 and GSE49377; for the complete gene expression values, the number is GSE49243. The European Genome-phenome Archive accession number for the sequencing data is EGAS00001000607.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.ccr.2014.02.004.

ACKNOWLEDGMENTS

For technical support and expertise, we thank the people from the DKFZ Genomics and Proteomics Core Facility, Malaika Knopf from the NCT Heidelberg, and GATC Biotech AG for sequencing services. This work was principally supported by the PedBrain Tumor Project (http://www.pedbraintumor.org) contributing to the International Cancer Genome Consortium, funded by German Cancer Aid (109252) and by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, grant numbers 01KU1201A, MedSys 0315416C, and NGFNplus 01GS0883). Additional support came from the Dutch Cancer Foundations KWF (2010-4713) and KIKA (to M.K.), and the Brain Tumor Charity (to V.P.C.). G.R. received a research grant from Roche and honoraria for advisory boards from Merck Serono and Roche. R.B. and O.W. are consultants for and received grant funding from Novartis.

Received: August 5, 2013
Revised: November 19, 2013
Accepted: February 13, 2014
Published: March 17, 2014

REFERENCES

Amakye, D.D., Robinson, D., Rose, K., Cho, Y., Ligon, K., Sharp, T., Haider, A., Bandaru, R., Ando, Y., Geoerger, B., et al. (2012). Abstract 4818: The predictive value of a 5-gene signature as a patient pre-selection tool in medulloblastoma for Hedgehog pathway inhibitor therapy. Cancer Res. *72*, http://dx.doi.org/10.1158/1538-7445.AM2012-4818.

Beauchamp, E.M., Ringer, L., Bulut, G., Sajwan, K.P., Hall, M.D., Lee, Y.C., Peaceman, D., Ozdemirli, M., Rodriguez, O., Macdonald, T.J., et al. (2011). Arsenic trioxide inhibits human cancer cell growth and tumor development in mice by blocking Hedgehog/GLI pathway. J. Clin. Invest. *121*, 148–160.

Brandes, A.A., Franceschi, E., Tosoni, A., Reni, M., Gatta, G., Vecht, C., and Kortmann, R.D. (2009). Adult neuroectodermal tumors of posterior fossa (medulloblastoma) and of supratentorial sites (stPNET). Crit. Rev. Oncol. Hematol. *71*, 165–179.

Brugières, L., Remenieras, A., Pierron, G., Varlet, P., Forget, S., Byrde, V., Bombled, J., Puget, S., Caron, O., Dufour, C., et al. (2012). High frequency of germline SUFU mutations in children with desmoplastic/nodular medulloblastoma younger than 3 years of age. J. Clin. Oncol. *30*, 2087–2093.

Buonamici, S., Williams, J., Morrissey, M., Wang, A., Guo, R., Vattay, A., Hsiao, K., Yuan, J., Green, J., Ospina, B., et al. (2010). Interfering with resistance to smoothened antagonists by inhibition of the PI3K pathway in medulloblastoma. Sci. Transl. Med. *2*, 51ra70.

Dijkgraaf, G.J., Alicke, B., Weinmann, L., Januario, T., West, K., Modrusan, Z., Burdick, D., Goldsmith, R., Robarge, K., Sutherlin, D., et al. (2011). Small molecule inhibition of GDC-0449 refractory smoothened mutants and downstream mechanisms of drug resistance. Cancer Res. *71*, 435–444.

Ellison, D.W., Dalton, J., Kocak, M., Nicholson, S.L., Fraga, C., Neale, G., Kenney, A.M., Brat, D.J., Perry, A., Yong, W.H., et al. (2011). Medulloblastoma: clinicopathological correlates of SHH, WNT, and non-SHH/WNT molecular subgroups. Acta Neuropathol. *121*, 381–396.

Geoerger, B., Aerts, I., Casanova, M., Chisholm, J.C., Hargrave, D., Leary, S., Ashley, D.M., Bouffet, E., MacDonald, T.J., Hurh, E., et al. (2012). A phase I/II study of LDE225, a smoothened (Smo) antagonist, in pediatric patients with recurrent medulloblastoma (MB) or other solid tumors. J. Clin. Oncol. *30*, 9519.

Grammel, D., Warmuth-Metz, M., von Bueren, A.O., Kool, M., Pietsch, T., Kretzschmar, H.A., Rowitch, D.H., Rutkowski, S., Pfister, S.M., and Schüller, U. (2012). Sonic hedgehog-associated medulloblastoma arising from the cochlear nuclei of the brainstem. Acta Neuropathol. *123*, 601–614.

Jones, D.T., Jäger, N., Kool, M., Zichner, T., Hutter, B., Sultan, M., Cho, Y.J., Pugh, T.J., Hovestadt, V., Stütz, A.M., et al. (2012). Dissecting the genomic complexity underlying medulloblastoma. Nature *488*, 100–105.

Killela, P.J., Reitman, Z.J., Jiao, Y., Bettegowda, C., Agrawal, N., Diaz, L.A., Jr., Friedman, A.H., Friedman, H., Gallia, G.L., Giovanella, B.C., et al. (2013). TERT promoter mutations occur frequently in gliomas and a subset of tumors derived from cells with low rates of self-renewal. Proc. Natl. Acad. Sci. USA *110*, 6021–6026.

Kim, J., Tang, J.Y., Gong, R., Kim, J., Lee, J.J., Clemons, K.V., Chong, C.R., Chang, K.S., Fereshteh, M., Gardner, D., et al. (2010). Itraconazole, a commonly used antifungal that inhibits Hedgehog pathway activity and cancer growth. Cancer Cell *17*, 388–399.

Kim, J., Aftab, B.T., Tang, J.Y., Kim, D., Lee, A.H., Rezaee, M., Kim, J., Chen, B., King, E.M., Borodovsky, A., et al. (2013). Itraconazole and arsenic trioxide inhibit Hedgehog pathway activation and tumor growth associated with acquired resistance to smoothened antagonists. Cancer Cell *23*, 23–34.

Kimura, H., Ng, J.M., and Curran, T. (2008). Transient inhibition of the Hedgehog pathway in young mice causes permanent defects in bone structure. Cancer Cell *13*, 249–260.

Koelsche, C., Sahm, F., Capper, D., Reuss, D., Sturm, D., Jones, D.T., Kool, M., Northcott, P.A., Wiestler, B., Böhmer, K., et al. (2013). Distribution of TERT promoter mutations in pediatric and adult tumors of the nervous system. Acta Neuropathol. *126*, 907–915.

Kool, M., Korshunov, A., Remke, M., Jones, D.T., Schlanstein, M., Northcott, P.A., Cho, Y.J., Koster, J., Schouten-van Meeteren, A., van Vuurden, D., et al. (2012). Molecular subgroups of medulloblastoma: an international meta-analysis of transcriptome, genetic aberrations, and clinical data of WNT, SHH, Group 3, and Group 4 medulloblastomas. Acta Neuropathol. *123*, 473–484.

Korshunov, A., Remke, M., Kool, M., Hielscher, T., Northcott, P.A., Williamson, D., Pfaff, E., Witt, H., Jones, D.T., Ryzhova, M., et al. (2012). Biological and clinical heterogeneity of MYCN-amplified medulloblastoma. Acta Neuropathol. *123*, 515–527.

Lee, Y., Kawagoe, R., Sasai, K., Li, Y., Russell, H.R., Curran, T., and McKinnon, P.J. (2007). Loss of suppressor-of-fused function promotes tumorigenesis. Oncogene *26*, 6442–6447.

Lin, T.L., and Matsui, W. (2012). Hedgehog pathway as a drug target: Smoothened inhibitors in development. OncoTargets and therapy *5*, 47–58.

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

Low, J.A., and de Sauvage, F.J. (2010). Clinical experience with Hedgehog pathway inhibitors. J. Clin. Oncol. *28*, 5321–5326.

Metcalfe, C., Alicke, B., Crow, A., Lamoureux, M., Dijkgraaf, G.J., Peale, F., Gould, S.E., and de Sauvage, F.J. (2013). PTEN loss mitigates the response of medulloblastoma to Hedgehog pathway inhibition. Cancer Res. *73*, 7034–7042.

Milenkovic, L., and Scott, M.P. (2010). Not lost in space: trafficking in the hedgehog signaling pathway. Sci. Signal. *3*, pe14.

Ng, J.M., and Curran, T. (2011). The Hedgehog's tale: developing strategies for targeting cancer. Nat. Rev. Cancer *11*, 493–501.

Northcott, P.A., Hielscher, T., Dubuc, A., Mack, S., Shih, D., Remke, M., Al-Halabi, H., Albrecht, S., Jabado, N., Eberhart, C.G., et al. (2011a). Pediatric and adult sonic hedgehog medulloblastomas are clinically and molecularly distinct. Acta Neuropathol. *122*, 231–240.

Northcott, P.A., Korshunov, A., Witt, H., Hielscher, T., Eberhart, C.G., Mack, S., Bouffet, E., Clifford, S.C., Hawkins, C.E., French, P., et al. (2011b). Medulloblastoma comprises four distinct molecular variants. J. Clin. Oncol. *29*, 1408–1414.

Northcott, P.A., Jones, D.T., Kool, M., Robinson, G.W., Gilbertson, R.J., Cho, Y.J., Pomeroy, S.L., Korshunov, A., Lichter, P., Taylor, M.D., and Pfister, S.M. (2012a). Medulloblastomics: the end of the beginning. Nat. Rev. Cancer *12*, 818–834.

Northcott, P.A., Shih, D.J., Peacock, J., Garzia, L., Morrissy, A.S., Zichner, T., Stütz, A.M., Korshunov, A., Reimand, J., Schumacher, S.E., et al. (2012b). Subgroup-specific structural variation across 1,000 medulloblastoma genomes. Nature *488*, 49–56.

Padovani, L., Sunyach, M.P., Perol, D., Mercier, C., Alapetite, C., Haie-Meder, C., Hoffstetter, S., Muracciole, X., Kerr, C., Wagner, J.P., et al. (2007). Common strategy for adult and pediatric medulloblastoma: a multicenter series of 253 adults. Int. J. Radiat. Oncol. Biol. Phys. *68*, 433–440.

Parsons, D.W., Li, M., Zhang, X., Jones, S., Leary, R.J., Lin, J.C., Boca, S.M., Carter, H., Samayoa, J., Bettegowda, C., et al. (2011). The genetic landscape of the childhood cancer medulloblastoma. Science *331*, 435–439.

Pugh, T.J., Weeraratne, S.D., Archer, T.C., Pomeranz Krummel, D.A., Auclair, D., Bochicchio, J., Carneiro, M.O., Carter, S.L., Cibulskis, K., Erlich, R.L., et al. (2012). Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. Nature *488*, 106–110.

Rausch, T., Jones, D.T.W., Zapatka, M., Stütz, A.M., Zichner, T., Weischenfeldt, J., Jäger, N., Remke, M., Shih, D., Northcott, P.A., et al. (2012). Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with TP53 mutations. Cell *148*, 59–71.

Remke, M., Ramaswamy, V., Peacock, J., Shih, D.J., Koelsche, C., Northcott, P.A., Hill, N., Cavalli, F.M., Kool, M., Wang, X., et al. (2013). TERT promoter mutations are highly recurrent in SHH subgroup medulloblastoma. Acta Neuropathol. *126*, 917–929.

Robinson, G., Parker, M., Kranenburg, T.A., Lu, C., Chen, X., Ding, L., Phoenix, T.N., Hedlund, E., Wei, L., Zhu, X., et al. (2012). Novel mutations target distinct subgroups of medulloblastoma. Nature *488*, 43–48.

Rudin, C.M., Hann, C.L., Laterra, J., Yauch, R.L., Callahan, C.A., Fu, L., Holcomb, T., Stinson, J., Gould, S.E., Coleman, B., et al. (2009). Treatment of medulloblastoma with hedgehog pathway inhibitor GDC-0449. N. Engl. J. Med. *361*, 1173–1178.

Ryan, S.L., Schwalbe, E.C., Cole, M., Lu, Y., Lusher, M.E., Megahed, H., O'Toole, K., Nicholson, S.L., Bognar, L., Garami, M., et al. (2012). MYC family amplification and clinical risk-factors interact to predict an extremely poor prognosis in childhood medulloblastoma. Acta Neuropathol. *123*, 501–513.

Shu, Q., Wong, K.K., Su, J.M., Adesina, A.M., Yu, L.T., Tsang, Y.T., Antalffy, B.C., Baxter, P., Perlaky, L., Yang, J., et al. (2008). Direct orthotopic transplantation of fresh surgical specimen preserves CD133+ tumor cells in clinically relevant mouse models of medulloblastoma and glioma. Stem Cells *26*, 1414–1424.

Spreafico, F., Massimino, M., Gandola, L., Cefalo, G., Mazza, E., Landonio, G., Pignoli, E., Poggi, G., Terenziani, M., Pedrazzoli, P., et al. (2005). Survival of adults treated for medulloblastoma using paediatric protocols. Eur. J. Cancer *41*, 1304–1310.

Taylor, M.D., Northcott, P.A., Korshunov, A., Remke, M., Cho, Y.J., Clifford, S.C., Eberhart, C.G., Parsons, D.W., Rutkowski, S., Gajjar, A., et al. (2012). Molecular subgroups of medulloblastoma: the current consensus. Acta Neuropathol. *123*, 465–472.

Wang, Y., Ding, Q., Yen, C.J., Xia, W., Izzo, J.G., Lang, J.Y., Li, C.W., Hsu, J.L., Miller, S.A., Wang, X., et al. (2012). The crosstalk of mTOR/S6K1 and Hedgehog pathways. Cancer Cell *21*, 374–387.

Welch, J.S., Ley, T.J., Link, D.C., Miller, C.A., Larson, D.E., Koboldt, D.C., Wartman, L.D., Lamprecht, T.L., Liu, F., Xia, J., et al. (2012). The origin and evolution of mutations in acute myeloid leukemia. Cell *150*, 264–278.

Yauch, R.L., Dijkgraaf, G.J., Alicke, B., Januario, T., Ahn, C.P., Holcomb, T., Pujara, K., Stinson, J., Callahan, C.A., Tang, T., et al. (2009). Smoothened mutation confers resistance to a Hedgehog pathway inhibitor in medulloblastoma. Science *326*, 572–574.

Yeh-Nayre, L.A., Malicki, D.M., Vinocur, D.N., and Crawford, J.R. (2012). Medulloblastoma with excessive nodularity: radiographic features and pathologic correlate. Case reports in radiology *2012*, 310359.

Zhao, X., Liu, Z., Yu, L., Zhang, Y., Baxter, P., Voicu, H., Gurusiddappa, S., Luan, J., Su, J.M., Leung, H.C., and Li, X.N. (2012). Global gene expression profiling confirms the molecular fidelity of primary tumor-based orthotopic xenograft mouse models of medulloblastoma. Neuro-oncol. *14*, 574–583.

Zhukova, N., Ramaswamy, V., Remke, M., Pfaff, E., Shih, D.J., Martin, D.C., Castelo-Branco, P., Baskin, B., Ray, P.N., Bouffet, E., et al. (2013). Subgroup-specific prognostic implications of TP53 mutation in medulloblastoma. J. Clin. Oncol. *31*, 2927–2935.

4) Long-term outcome of 4040 children diagnosed with pediatric low-grade gliomas: An analysis of the Surveillance Epidemiology and End Results (SEERs) database

Pratiti Bandopadhayay, Guillaume Bergthold, Wendy B. London, Liliana C. Goumnerova, Andres Morales La Madrid, Karen J Marcus, Dongjing Guo, Nicole J. Ullrich, Nathan Robison, Susan N. Chi, Rameen Beroukhim, Mark W. Kieran, Peter E. Manley;

Pediatr Blood Cancer. 2014 Jul;61(7):1173-9

Pediatr Blood Cancer 2014;61:1173–1179

Long-Term Outcome of 4,040 Children Diagnosed With Pediatric Low-Grade Gliomas: An Analysis of the Surveillance Epidemiology and End Results (SEER) Database

Pratiti Bandopadhayay, MBBS, PhD,[1,2] Guillaume Bergthold, MD,[2] Wendy B. London, PhD,[3] Liliana C. Goumnerova, MD,[1,4] Andres Morales La Madrid, MD,[1] Karen J. Marcus, MD,[5] Dongjing Guo, MPH,[3] Nicole J. Ullrich, MD, PhD,[1,6] Nathan J. Robison, MD,[8] Susan N. Chi, MD,[1] Rameen Beroukhim, MD, PhD,[2,7] Mark W. Kieran, MD, PhD,[1] and Peter E. Manley, MD[1]*

Background. Children with pediatric low-grade gliomas (PLGG) are known to have excellent 10-year survival rates; however the outcomes of adult survivors of PLGG are unknown. We identified patients diagnosed with PLGG diagnosed between 1973 and 2008 through the Surveillance Epidemiology and End Results (SEER) database to examine outcomes of adult survivors of PLGG. ***Procedure.*** Four thousand and forty patients with either WHO grade I or II PLGG were identified and outcome data retrieved. Two analyses were performed to assess survival and risk of death from tumor. Competing risks analysis was conducted and cumulative incidence curves of death due to disease were generated. Cox proportional hazards regression was performed, with adjustment for non-disease death. Kaplan–Meier curves for overall cancer specific survival (OS) were also generated. ***Results.*** The 20-year OS was 87% ± 0.8% and the 20-year cumulative incidence of death due to glioma was 12% ± 0.8%. The incidence of death after transition to adulthood (age greater than 22 years) was slightly lower, with 20-year cumulative incidence of disease death of 7% ± 1.8%. Year of diagnosis, age of diagnosis, histology, WHO grade, primary site, radiation, and degree of initial resection were prognostic in univariate analysis, while the administration of radiation was the greatest risk of death in multivariate analysis of OS (hazard ratio = 3.9). ***Conclusions.*** PLGGs are associated with an excellent long-term survival, with a low likelihood of PLGG related death in adult survivors. Treatment strategies for pediatric tumors should therefore aim for disease control during childhood and adolescence with an emphasis on minimizing long-term treatment induced toxicities. Pediatr Blood Cancer 2014;61:1173–1179.

Key words: outcome; pediatric low-grade glioma; SEER

INTRODUCTION

Children with PLGGs have been reported to have superior 10-year overall survival (OS) rates compared to adults diagnosed with low-grade gliomas during adulthood (ALGG) [1]. While children have an excellent 10-year OS, adult patients diagnosed with low-grade gliomas have a more aggressive clinical course and poor long-term survival rates, with a high incidence of malignant transformation and death [2,3]. PLGGs have been reported to spontaneously stabilize or regress [4,5], particularly in children with neurofibromatosis type 1 [6,7], a phenomenon which is rare in ALGG. Although small case series and anecdotal evidence have suggested that PLGGs rarely undergo malignant transformation [8–10], long-term outcomes of PLGG remain unknown. In particular, it remains unclear whether these tumors undergo malignant transformation to cause death in adulthood. We performed a comparative retrospective analysis of survival of patients with PLGG diagnosed between 1973 and 2008 as identified in the National Institute Surveillance, Epidemiology and End Results (SEER) database of the National Cancer Institute (NCI), to analyze the survival rates of adult survivors of pediatric low-grade gliomas.

MATERIALS AND METHODS

Approval to access the SEER data was granted by the NCI to use all datasets up until 2011 [11]. Pediatric patients (0–19 years of age) diagnosed with a grade I or II glioma between 1973 and 2008 were identified. Variables recorded for each patient included age, year of diagnosis, race, sex, location of tumor, histological diagnosis and grade of tumor, extent of surgery, and whether the child received radiation therapy. Histological groupings used to identify patients included pilocytic astrocytoma, diffuse astrocytoma, astrocytoma not otherwise specified (NOS), glioma NOS, mixed glioma, or unique astrocytoma variants. Only tumors classified as grade I or II were included. Outcomes were recorded for status (alive or dead), time from diagnosis to last follow-up or death, and cause of death. Deaths were classified as death due to glioma, or death not due to glioma. Secondary malignancies, such as radiation induced tumors, were not classified as deaths due to glioma.

---

[1]Division of Pediatric Hematology/Oncology, Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston Children's Hospital, Boston, Massachusetts; [2]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, Massachusetts; [3]Department of Medicine, Boston Children's Hospital, Boston, Massachusetts; [4]Department of Pediatric Neurosurgery, Boston Children's Hospital, Boston, Massachusetts; [5]Department of Radiation Oncology, Boston Children's Hospital, Dana-Farber Cancer Institute and Brigham and Women's Hospital, Boston, Massachusetts; [6]Department of Neurology, Boston Children's Hospital, Boston, Massachusetts; [7]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts; [8]Children's Center for Cancer and Blood Diseases, Children's Hospital Los Angeles, University of Southern California Keck School of Medicine, Los Angeles, California

Grant sponsor: Stop&Shop Pediatric Brain Tumor Program; Grant sponsor: Andrysiak Fund for LGG, Pediatric Low-Grade Astrocytoma Foundation; Grant sponsor: Friends of DFCI; Grant sponsor: Nuovo-Soldati Foundation; Grant sponsor: Philippe Foundation; Grant sponsor: St Baldrick's Foundation

Conflict of interest: Nothing to declare.

*Correspondence to: Peter Manley, Pediatric Neuro-Oncology Program, D3148, Dana-Farber Cancer Institute, 450 Brookline Ave, Boston, MA 02215. E-mail: Peter_Manley@dfci.harvard.edu

Received 1 November 2013; Accepted 3 January 2014

DOI 10.1002/pbc.24958

A competing risks analysis was conducted according to the methods of Pepe and Mori [12]. With adjustment for non-disease deaths, cumulative incidence curves of death due to disease were plotted for the overall cohort and for subgroups of each potential prognostic factor: age, sex, year of diagnosis, histology, WHO grade of differentiation, primary site, extent/type of radiation therapy, and degree of initial resection. The prognostic ability of the factors was tested using a Cox proportional hazards regression model (univariate and multivariate) for competing risks, according to Rosthoj et al. [13], with adjustment for non-disease death. This method allows for separate estimation/quantification of (i) risk for death due to disease; and (ii) risk for death from non-disease causes. Kaplan–Meier curves of OS were generated to determine overall survival including all causes of death. Cancer specific survival was performed by generating Kaplan–Meier curves including glioma related deaths. Non-disease deaths were censored [3].

To compare the risk of death from disease prior to transition to adulthood (age 22 years) versus after transition to adulthood, a time-dependent covariate for the occurrence of a patient's 22nd birthday was included in the Cox model. The age of 22 was chosen to allow for the examination of the natural history of PLGG after the age at which the majority of brain maturation during the transition from adolescence to adulthood has been completed [14].

RESULTS

A total of 4,040 patients aged $\leq$19 years of age (median age at diagnosis: 9 years, range 0–19 years) at time of diagnosis of a PLGG between 1973 and 2008 were identified (Table I). The 30-year OS of the 4,040 patients was 74.8% ± 2%, including all causes of death. We were unable to determine progression-free survival from the available data. Of the 4,040 patients, 347 patients died as a result of PLGG and 78 patients died of other causes. Cause of death was unknown in 17 patients. Median time of follow-up was 6.9 (0–36.5) years. Importantly, the cohort included 875 patients with at least 15 years of follow-up from diagnosis.

The Kaplan–Meier estimate of overall cancer specific OS at 20 years was 87% ± 0.8% (Fig. 1A), and the 20-year overall survival of the 875 patients for which there was at least 15 years follow-up was 92% ± 3% (Fig. 1B). Pepe–Mori 20-year cumulative incidence of death due to disease was 12% ± 0.8% (Fig. 1C). We examined the overall survival of adult survivors of pediatric low-grade gliomas by calculating OS after transition to adulthood. This analysis included children diagnosed with all subtypes of PLGG. After reaching adulthood (defined as age 22), survivors of PLGG (including all histological subtypes) had a 30-year OS of 93% ± 1.7 (Fig. 1D). These patients had a 20-year cumulative incidence of death due to disease of 7% ± 1.8% (Fig. 1E). For patients prior to reaching their 22nd birthday, the risk of death due to disease is slightly higher ($HR = 1.37$, $P = 0.24$) than the risk of death from any cause ($HR = 1.20$, P-value $= 0.40$) in comparison to patients after age 22 years, though not significantly higher in either case.

Univariate analysis of risk of death identified age <2 years at diagnosis ($P < 0.0001$, Fig. 2A), year of diagnosis between 1970 and 1989 ($P < 0.0001$, Fig. 2B), histology other than pilocytic astrocytoma ($P < 0.0001$, Fig. 2C), tumor site other than cerebellum ($P < 0.0001$, Fig. 2D) and WHO grade II ($P < 0.0001$, Fig. 2E) as prognostic of increased risk of death. There was no significant difference in overall survival between children who had a gross total surgical resection and children with residual disease ($P = 0.4$, Fig. 2F).

We performed a multivariate Cox model analysis (Table II) and identified an increased risk of cancer-specific death in children who were treated with radiation therapy ($HR = 3.9$, $P < 0.0001$). Children who did not have cerebellar disease ($HR = 2.3$, $P < 0.0001$) or had non-pilocytic astrocytoma histology ($HR = 2.2$, $P < 0.0001$) also had increased risk of disease-related death. Despite this, children with all histological subtypes of pediatric low-grade gliomas still had an excellent long-term overall survival (Fig. 3). Children who were less than 2 years of age had greater risk of disease-related death compared to older children ($HR = 2$, $P < 0.0001$). Although not significant on univariate analysis, patients with subtotal or no resection had a small but statistically

TABLE I. Demographics of Pediatric Patients in Cohort

Category	Number (%)
Gender	
Male	2,059 (51)
Female	1,981 (49)
Age (median)	9 years
Histology	
Pilocytic astrocytoma	2,648 (65)
Astrocytoma NOS	841 (21)
NOS grade I	227
NOS grade II	614
Diffuse astrocytoma	260 (6)
Glioma NOS (total)	189 (5)
Grade I	68
Grade II	121
Mixed glioma (total)	75 (2)
Grade I	16
Grade II	59
Unique astrocytoma variant	27 (1)
Grade I	12
Grade II	15
Grade (all histologies)	
Grade I	2,971 (74)
Grade II	1,069 (26)
Location	
Supratentorial	1,264 (31)
Ventricular	189 (5)
Cerebellum	1,170 (29)
Brainstem	504 (12)
Overlapping or NOS	718 (18)
Spinal cord	195 (5)
Extent of resection	
No resection	373 (9)
Biopsy	9 (0.2)
Subtotal resection	1,769 (44)
Gross total	1,094 (27)
Not otherwise specified	593 (15)
Unknown	202 (5)
Radiation	
No radiation therapy	3,235 (80)
Radiation therapy	736 (18)
Unknown	69 (2)
Number of deaths	442 (11)
PLGG	347
Non-PLGG	78
Unknown	17

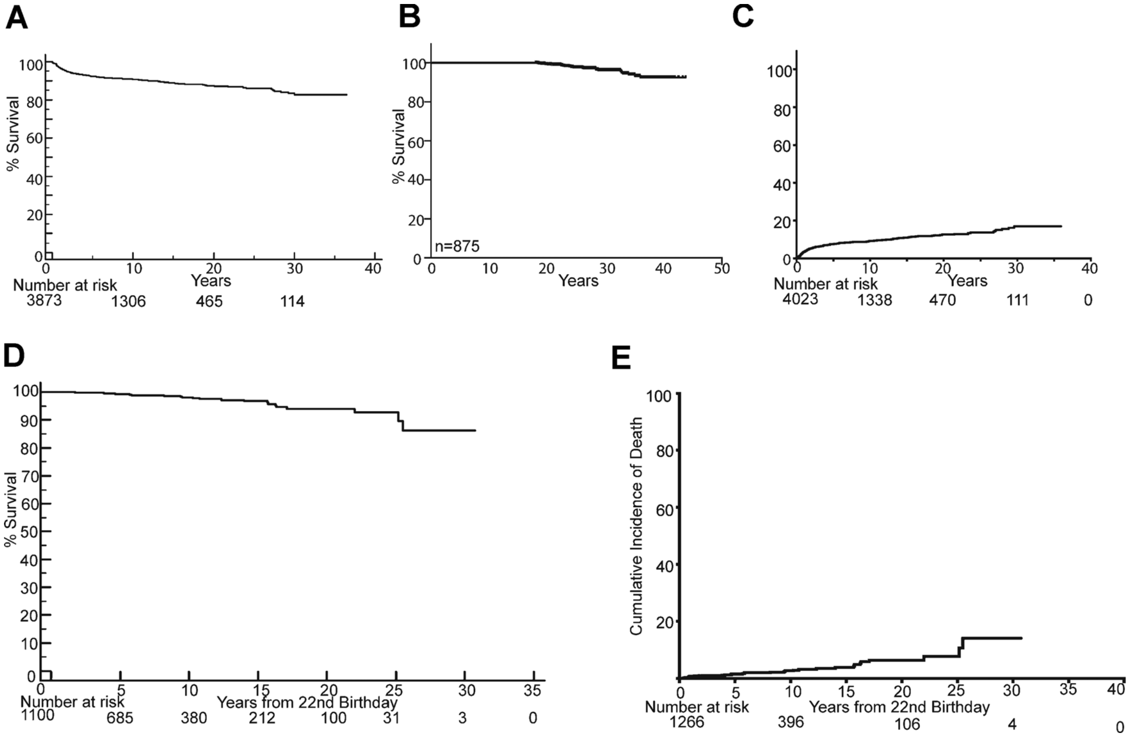


Fig. 1. Adult survivors of pediatric low-grade gliomas have excellent overall survival with low rates of mortality after patients transition into adulthood. **A**: Kaplan–Meier overall survival curve of patients with PLGG including only tumor related deaths. **B**: Kaplan–Meier overall survival curve of patients for which there is at least 15 years of follow-up. **C**: Pepe–Mori cumulative incidence of tumor specific death curve of patients diagnosed with PLGG. **D**, **A**: Kaplan–Meier overall survival curve of patients with PLGG showing survival starting from the patient's 22nd birthday. **E**: Pepe–Mori cumulative incidence of tumor specific death curves of patients starting from patient's 22nd birthday.

significant increase in risk of disease-related death compared to those with total resection on multivariate analysis (HR = 1.5, 95% CI: 1.01, 2.1, $P = 0.04$).

Treatment with radiation therapy was an adverse prognostic factor. Regardless of the extent of surgical resection, children who were treated with radiation therapy had a worse OS and greater risk of disease death compared to those that were not treated with radiation therapy (HR = 3.9, $P < 0.0001$; Fig. 4, respectively). Within each treatment group (radiation or no radiation), there was no difference in OS or cumulative incidence of disease death in children who had received a gross total resection in comparison to children who had residual tumor. Further, children in whom only a subtotal resection was achieved but did not undergo radiation had a significantly superior outcome compared to children who had a gross total resection but treated with radiation. The hazard ratio of death due to glioma in patients who were treated with radiation was 3.9 (P-value < 0.0001). Data were not available in the SEER database to permit a determination of the incidence of radiation-induced secondary malignancies; however the hazard ratio of death due to non-glioma related causes of 2.4 (P-value $= 0.0006$) in patients treated with radiation suggests that radiation-induced mortality may have accounted for some of the deaths that were not directly due to glioma.

The risk of death due to disease was slightly higher for patients before age 22 compared to after age 22 (hazard ratio = 1.37), although not statistically significant ($P = 0.24$). The risk of death from any cause was slightly higher for patients before age 22 compared to after age 22 (hazard ratio = 1.20), although not statistically significant ($P = 0.40$). Within the group that had received radiation, the hazard ratio of dying from causes all causes compared to patients less than 22 years of age was 1.59 (P-value $= 0.1$) and from tumor 1.74 (P-value $= 0.1$).

DISCUSSION

Our outcome analysis is the largest ever reported for very long-term survivors of PLGGs, and confirms the excellent prognosis of PLGG. While it is known that patients with PLGGs have superior 5- and 10-year OS [15,16] the data presented here confirm for the first time that adult survivors of pediatric disease have a low incidence of glioma related death. Importantly this highlights that PLGGs are very unlikely to undergo malignant transformation resulting in tumor-related death, which contrasts with the natural history of ALGGs [1,3]. Further, the data show that children who are treated with radiation therapy, regardless of their surgical status, have significantly inferior outcomes compared to those children who are not treated with radiation therapy.

We have found that children who were treated with radiation had inferior outcomes. It is however important to acknowledge that it is difficult to determine from the SEER data how much of this is due to radiation itself and how much is a reflection of selection bias. While some degree of selection bias is likely, the finding that children in whom a gross total resection was achieved and were irradiated had an inferior outcome compared to those children with a subtotal resection who were not irradiated suggests that radiation induced mortality (rather than PLGG related mortality) also accounts for a

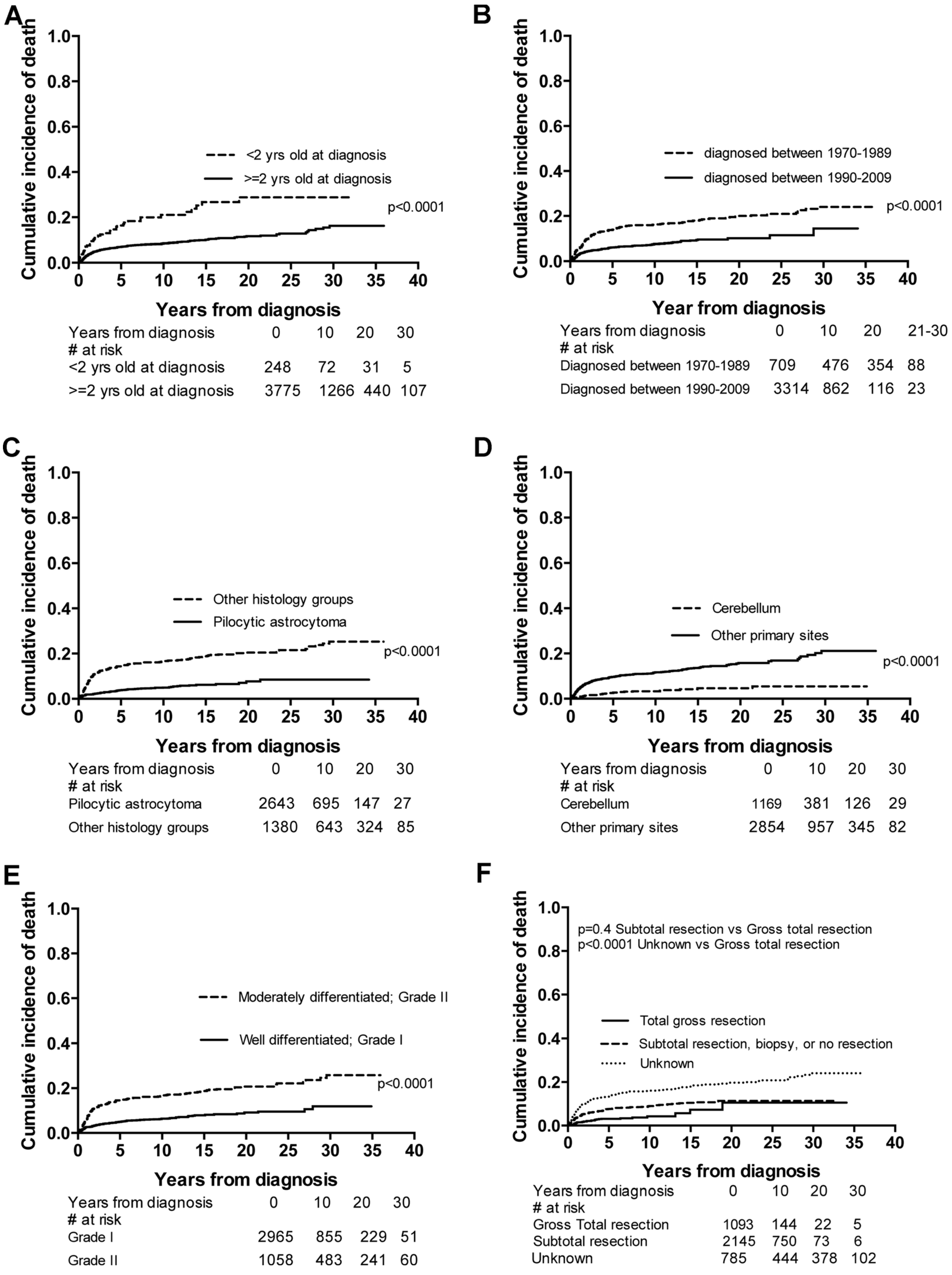


Fig. 2. Pepe–Mori cumulative incidence of death curves depicting univariate analysis including **A**, age at diagnosis **B**, year of diagnosis **C**, histology **D**, location of primary tumor **E**, grade and **F**, extent of resection. Number of patients at each time point are shown below.

number of these deaths. The HR of dying from disease after radiation was 3.9 and the HR of dying from non-glioma related disease after radiation was 2.4. These data suggest that the observed increased in mortality in children treated with radiation therapy is likely to be a combination of both selection bias (i.e., children with harder to control disease are more likely to have received radiation) and radiation induced mortality, such as the increased risk for secondary malignancies and vasculopathy. Due to the inherent

TABLE II. Multivariate Analysis of Risk Factors in Patients With Pediatric Low-Grade Glioma

	Death due to disease		Death from non-disease causes	
Factor (reference level)	HR (95% CI)	*P*-value	HR (95% CI)	*P*-value
Beam radiation, with or without implants or isotopes (no radiation)	3.9 (3.0, 4.9)	<0.0001	2.4 (1.4, 3.9)	0.0006
Degree of radiation unknown (no radiation)	3.1 (1.7, 5.8)	0.0003	NA	1.0
Primary site (cerebellum)	2.3 (1.6, 3.2)	<0.0001	NA	1.0
Histology group (pilocytic astrocytoma)	2.2 (1.7, 2.8)	<0.0001	NA	0.05
Age of diagnosis (≥2 years old)	2.0 (1.5, 2.8)	<0.0001	2.2 (1.1, 4.5)	0.03
Subtotal resection, biopsy or no resection (total resection)	1.5 (1.01, 2.1)	0.04	3.0 (1.1, 8.6)	0.04

HR, hazard ratio (increased risk of death due to disease for the Factor in comparison to the reference level); CI, confidence interval.

limitations associated with SEER database, we are unable to determine the proportion of children who were treated with radiation after failing chemotherapeutic approaches.

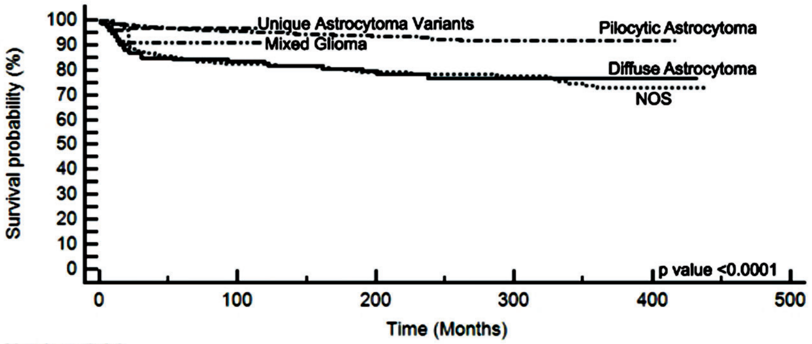
GTR of PLGGs have been associated with superior event free and overall survival [16]. In the SEER cohort of PLGG we have found that children who do not have a GTR also have an excellent overall survival. This analysis included all histological subtypes of pediatric low-grade gliomas. Thus, the aim of surgery for these children should be to resect the maximal amount of tumor that can be safely resected while minimizing the risk of long-term neurological sequalae.

The overall long-term survival of patients diagnosed with PLGGs including all causes of death is superior to the outcomes of patients diagnosed with glioma as adults [3]. Our analysis, using two methodologies to examine disease-specific deaths of patients with PLGG, confirms that PLGG infrequently causes death in adult survivors. In fact, the risk of death decreases once patients reach adulthood.

Our study reports outcomes of grade II low-grade gliomas in a large SEER-based cohort of pediatric patients, and our finding of excellent overall survival in this cohort has important clinical implications. Many physicians have been concerned that residual grade II PLGGs may undergo malignant transformation later in life, and have thus advocated potentially morbid therapies including radical surgical resection and/or radiation therapy. However, in our series, the majority of children with grade II PLGGs also had excellent overall survival, and we did not observe increased rates of death with time as have been reported in low-grade gliomas that arise in adults, and which are prone to undergo malignant transformation.

Unfortunately, data to perform progression-free survival analysis are not available from the SEER database; however, the observation of a very low mortality in adult survivors supports the hypothesis that PLGG become stable, possibly quiescent as children transition into their early adult years. It is possible that loss of follow-up of adult survivors of PLGG may have affected outcome measures in this study and may result in under-estimation of overall survival as adult survivors who are well and not requiring medical assistance may not be captured in follow-up data while adult survivors who have disease recurrence and subsequent death are more likely to have been reported.

The confirmation that the majority of adult survivors of PLGG do not eventually succumb to their disease is an important



Number at risk
Group: Diffuse astrocytoma (protoplasma, fibrillary)
239 109 72 20 4 0
Group: Mixed glioma
75 9 0 0 0 0
Group: Not otherwise specified
971 579 361 164 13 0
Group: Pilocytic astrocytoma
2562 922 257 67 3 0
Group: Unique astrocytoma variants
26 1 0 0 0 0

Fig. 3. Patients with all histological subtypes of PLGG have excellent very long-term overall survival. Kaplan–Meyer curves of outcomes of patients diagnosed with different histological subtypes of PLGG.

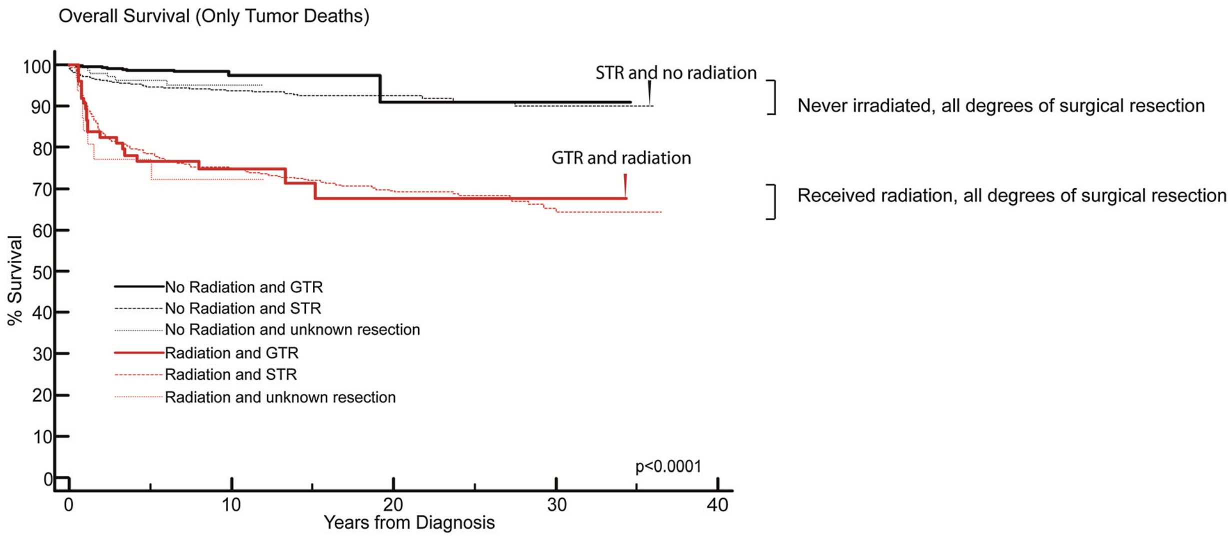


Fig. 4. Patients with PLGG who received radiation therapy have inferior overall survival compared to patients that did not receive radiation therapy. Kaplan–Meier overall survival curve of patients with PLGG including only tumor related deaths showing outcomes of children who received radiation therapy and extent of surgical resection. Patients with a STR who did not receive radiation (curve shown by black arrow) had a superior outcome than patients with a GTR who did receive radiation (curve shown by red arrow).

observation that must be taken into account when considering therapeutic options for children with low-grade gliomas. Given these children are expected to survive into adulthood, the aim of treatment should be to control disease and minimize long-term morbidity from both the tumor itself and treatment. This can be achieved with surgery if the tumor can be safely resected without resulting in neurological sequelae, chemotherapy or novel targeted therapeutic agents. Treatment strategies should aim to minimize treatment induced long-term morbidity such as those associated with radiation therapy, including neuro-cognitive deficits [17–20], hormone deficiencies and secondary malignancies [20,21]. Although some groups are currently using radiation therapy as first-line therapy [22], the data we present suggest that treatment with radiation therapy should be reserved for the small number of children in whom tumor control cannot be achieved with surgery, chemotherapy or targeted agents.

These data confirm that the natural history of PLGGs is distinct from their adult counterparts [1,3], and it is likely that this is a reflection of differences in the underlying biology of the tumors. Although histologically these tumors have the same characteristics in children and adults, the higher rate of genetic alterations observed in ALGGs compared to pediatric low-grade gliomas suggest a difference in genomic stability. While mutations of *IDH* [23] and *p53* [24] are reported in ALGGs and predict malignant transformation, these are rare in their pediatric counterparts, which more commonly harbor alterations of *BRAF* [25–28]. While the propensity of ALGGs to slowly progress and undergo malignant transformation is secondary to the acquisition of driver mutations, it remains to be determined why PLGGs do not also acquire these alterations with time to cause transformation. It can be hypothesized that normal processes that guide development and maturation of the brain, for example, epigenetic processes, may cause PLGGs to become quiescent as children transition into adulthood. The factors that govern the stability of PLGG remain unknown.

There are limitations to the data obtainable from the SEERs database. In particular, the quality of the data extracted is dependent on how the data are entered into the database, histology and radiology results have not been centrally reviewed, and interpretation of the grading of low-grade gliomas according to the current WHO classification can vary. The limitations of the WHO classification are highlighted by the finding that a significant proportion of PLGG in the SEER database are designated as "not otherwise specified," rather than given a distinct classification according to the WHO criteria. This suggests a need for better molecular markers to diagnose and stratify these gliomas. Further, it is possible that some of the tumors that have been designated as pediatric low-grade gliomas not otherwise specified may have represented higher-grade gliomas, thus accounting for the inferior survival of those children whose tumors could not be classified according to the WHO classification. This may have resulted in an over-estimation of the actual death rate of children with true low-grade gliomas.

In addition, the SEER database lacks information about chemotherapy regimens that were used, and how many courses of chemotherapy children were treated with prior to radiation therapy. This information would have been important for those patients in whom a subtotal resection was achieved, in determining whether radiation therapy was used more frequently in tumors in which control had not been achieved with radiation therapy.

In conclusion, we show that PLGGs are not the cause of death in the majority of adult survivors and that patients diagnosed with PLGGs have excellent long-term survival. This holds true for both grade (I or II) and degree of resection (GTR or STR) in pediatric low-grade gliomas. Therapeutic strategies should be designed to provide tumor control while avoiding those that cause irreversible long-term toxicity should be avoided as patients can be expected to survive long into adulthood.

ACKNOWLEDGMENTS

Support for this study included the Stop&Shop Pediatric Brain Tumor Program (PB, NR, SNC, MWK, PEM), Andrysiak Fund for

LGG (PB, MWK, PEM), Pediatric Low-Grade Astrocytoma Foundation (PB, GB, RB, MWK), Friends of DFCI (PB), Nuovo-Soldati Foundation (GB), Philippe Foundation (GB), and St Baldrick's Foundation (AMLM). This work has been presented in part at the 13th International Conference of Long-Term Complications of Treatment of Children and Adolescents for Cancer, Memphis, TN, 2013.

REFERENCES

1. Rees J, Watt H, Jager HR, et al. Volumes and growth rates of untreated adult low-grade gliomas indicate risk of early malignant transformation. Eur J Radiol 2009;72:54–64.
2. Youland RS, Brown PD, Giannini C, et al. Adult low-grade glioma: 19-year experience at a single institution. Am J Clin Oncol 2013;36:612–619.
3. Johnson DR, Brown PD, Galanis E, et al. Pilocytic astrocytoma survival in adults: Analysis of the surveillance, epidemiology, and end results program of the National Cancer Institute. J Neurooncol 2012;108:187–193.
4. Gunny RS, Haywood RD, Phipps KP, et al. Spontaneous regression of residual low-grade cerebellar pilocytic astrocytomas in children. Pediatr Radiol 2005;35:1086–1091.
5. Rozen WM, Joseph S, Lo PA. Spontaneous regression of low-grade gliomas in pediatric patients without neurofibromatosis. Pediatr Neurosurg 2008;44:324–328.
6. Perilongo G, Moras P, Carollo C, et al. Spontaneous partial regression of low-grade glioma in children with neurofibromatosis-1: A real possibility. J Child Neurol 1999;14:352–356.
7. Schmandt SM, Packer RJ, Vezina LG, et al. Spontaneous regression of low-grade astrocytomas in childhood. Pediatr Neurosurg 2000;32:132–136.
8. Broniscer A, Baker SJ, West AN, et al. Clinical and molecular characteristics of malignant transformation of low-grade glioma in children. J Clin Oncol 2007;25:682–689.
9. Afra D, Osztie E, Sipos L, et al. Preoperative history and postoperative survival of supratentorial low-grade astrocytomas. Br J Neurosurg 1999;13:299–305.
10. Parsa CF, Givrad S. Juvenile pilocytic astrocytomas do not undergo spontaneous malignant transformation: Grounds for designation as hamartomas. Br J Ophthalmol 2008;92:40–46.
11. National Cancer Institute, Surveillance, Epidemiology and End Results (SEER) program SEER*Stat database:incidence: SEER 17 Regs research data + hurricane katrina impacted Louisiana cases, Nov 2010 Sub (1973–2008 varying): Linked to County Attributes, DCCPS, Surveillance Research Program, Cancer Statistics Branch. 2011.
12. Pepe MS, Mori M. Kaplan-Meier, marginal or conditional probability curves in summarizing competing risks failure time data? Stat Med 1993;12:737–751.
13. Rosthoj S, Andersen PK, Abildstrom SZ. SAS macros for estimation of the cumulative incidence functions based on a Cox regression model for competing risks survival data. Comput Methods Programs Biomed 2004;74:69–75.
14. Lebel C, Walker L, Leemans A, et al., Microstructural maturation of the human brain from childhood to adulthood. NeuroImage 2008;40:1044–1055.
15. Ater JL, Zhou T, Holmes E, et al. Randomized study of two chemotherapy regimens for treatment of low-grade glioma in young children: A report from the Children's Oncology Group. J Clin Oncol 2012;30:2641–2647.
16. Gnekow AK, Falkenstein F, von Hornstein S, et al. Long-term follow-up of the multicenter, multidisciplinary treatment study HIT-LGG-1996 for low-grade glioma in children and adolescents of the German Speaking Society of Pediatric Oncology and Hematology. Neuro Oncol 2012;14: 1265–1284.
17. Keene DL, Johnston DL, Grimard L, et al. Vascular complications of cranial radiation. Childs Nerv Syst 2006;22:547–555.
18. Campen CJ, Kranick SM, Kasner SE, et al. Cranial irradiation increases risk of stroke in pediatric brain tumor survivors. Stroke 2012;43:3035–3040.
19. Ellenberg L, Liu Q, Gioia G, et al. Neurocognitive status in long-term survivors of childhood CNS malignancies: A report from the Childhood Cancer Survivor Study. Neuropsychology 2009;23:705–717.
20. Armstrong GT. Long-term survivors of childhood central nervous system malignancies: The experience of the Childhood Cancer Survivor Study. Eur J Paediatr Neurol 2010;14:298–303.
21. Galloway TJ, Indelicato DJ, Amdur RJ, et al. Second tumors in pediatric patients treated with radiotherapy to the central nervous system. Am J Clin Oncol 2012;35:279–283.
22. Müller K, Gnekow A, Falkenstein F, et al. Radiotherapy in pediatric pilocytic astrocytomas. A subgroup analysis within the prospective multicenter study HIT-LGG 1996 by the German Society of Pediatric Oncology and Hematology (GPOH). Strahlenther Onkol 2013;189:647–655.
23. Juratli TA, Kirsch M, Robel K, et al. IDH mutations as an early and consistent marker in low-grade astrocytomas WHO grade II and their consecutive secondary high-grade gliomas. J Neurooncol 2012;108:403–410.
24. Kim YH, Nobusawa S, Mittelbronn M, et al. Molecular classification of low-grade diffuse gliomas. Am J Pathol 2010;177:2708–2714.
25. MacConaill LE, Campbell CD, Kehoe SM, et al. Profiling critical cancer gene mutations in clinical tumor samples. PLoS ONE 2009;4:e7887.
26. Lin A, Rodriguez FJ, Karajannis MA, et al. BRAF alterations in primary glial and glioneuronal neoplasms of the central nervous system with identification of 2 novel KIAA1549: BRAF fusion variants. J Neuropathol Exp Neurol 2012;71:66–72.
27. Tian Y, Rich BE, Vena N, et al. Detection of KIAA 1549-BRAF fusion transcripts in formalin-fixed paraffin-embedded pediatric low-grade gliomas. J Mol Diagn 2011;13:669–677.
28. Zhang J, Wu G, Miller CP, et al. Whole-genome sequencing identifies genetic alterations in pediatric low-grade gliomas. Nat Genet 2013;45:602–612.

5) Genomic analysis of diffuse pediatric low-grade gliomas identifies novel oncogenic MYBL1-truncating rearrangements

Lori A. Ramkissoon, Peleg M. Horowitz, Justin M. Craig, Shakti H. Ramkissoon, Benjamin E. Rich, Steve E. Schumacher, Aaron McKenna, Mike Lawrence, Guillaume Bergthold, Priscilla K. Brastianos, Barbara Tabak, Matthew D. Ducar, Paul van Hummelen, Laura E. MacConaill, Yoon-Jae Cho, Hala Taha, Madeha Mahmoud, Daniel C. Bowers, Linda Margraf, Uri Tabori, Cynthia Hawkins, Roger J. Packer, D. Ashley Hill, Scott L. Pomeroy, Charles G. Eberhart, Ian F. Dunn, Liliana Goumnerova, Gad Getz, Jennifer A. Chan, Sandro Santagata, William C. Hahn, Charles D. Stiles, Azra H. Ligon, Mark W. Kieran, Rameen Beroukhim, Keith L. Ligon;

Proc Natl Acad Sci USA. 2013; 110(20): 8188-93

Genomic analysis of diffuse pediatric low-grade gliomas identifies recurrent oncogenic truncating rearrangements in the transcription factor *MYBL1*

Lori A. Ramkissoon[a,1], Peleg M. Horowitz[a,b,c,d,1], Justin M. Craig[a], Shakti H. Ramkissoon[a,e,f], Benjamin E. Rich[a], Steven E. Schumacher[a,c], Aaron McKenna[c], Michael S. Lawrence[c], Guillaume Bergthold[a,c], Priscilla K. Brastianos[a], Barbara Tabak[a,c], Matthew D. Ducar[g], Paul Van Hummelen[g], Laura E. MacConaill[f,g], Tina Pouissant-Young[h], Yoon-Jae Cho[i], Hala Taha[j], Madeha Mahmoud[j], Daniel C. Bowers[k], Linda Margraf[l], Uri Tabori[m], Cynthia Hawkins[n], Roger J. Packer[o,p], D. Ashley Hill[q], Scott L. Pomeroy[c,r], Charles G. Eberhart[s], Ian F. Dunn[b], Liliana Goumnerova[d], Gad Getz[c], Jennifer A. Chan[t], Sandro Santagata[e,f,u], William C. Hahn[a,c,g,v], Charles D. Stiles[w], Azra H. Ligon[f,u,x], Mark W. Kieran[y,z], Rameen Beroukhim[a,c,g,v,2], and Keith L. Ligon[a,e,f,u,x,2]

Departments of [a]Medical Oncology, [w]Cancer Biology, and [y]Pediatric Oncology, [g]Center for Cancer Genome Discovery, and [x]Center for Molecular Oncologic Pathology, Dana-Farber Cancer Institute, Boston, MA 02115; Departments of [b]Neurosurgery and [f]Pathology, Brigham and Women's Hospital, Boston, MA 02115; [c]Broad Institute, Massachusetts Institute of Technology and Harvard University, Cambridge, MA 02142; Departments of [d]Neurosurgery, [e]Pathology, [h]Radiology, [r]Neurology, and [z]Hematology–Oncology, Boston Children's Hospital, Boston, MA 02115; [i]Departments of Neurology and Neurosurgery, Stanford University School of Medicine, Stanford, CA 94305; [j]Department of Pediatric Oncology and Pathology, Children's Cancer Hospital, 57357 Cairo, Egypt; Departments of [k]Pediatrics and [l]Pathology, University of Texas Southwestern Medical Center, Dallas, TX 75390; Departments of [m]Hematology and Oncology and [n]Pathology, Hospital for Sick Children, Toronto, ON, Canada M5G 1X8; [o]Center for Neuroscience and Behavioral Medicine, [p]Brain Tumor Institute, [q]Department of Pathology, Children's National Medical Center, Washington, DC 20010; [s]Departments of Pathology, Ophthalmology, and Oncology, The Johns Hopkins University School of Medicine, Baltimore, MD 21287; [t]Department of Pathology and Laboratory Medicine, University of Calgary, Calgary, AB, Canada T2N 1N4; and Departments of [u]Pathology and [v]Medicine, Harvard Medical School, Boston, MA 02115

Edited by Webster K. Cavenee, Ludwig Institute for Cancer Research, University of California San Diego, La Jolla, CA, and approved March 28, 2013 (received for review January 7, 2013)

Pediatric low-grade gliomas (PLGGs) are among the most common solid tumors in children but, apart from *BRAF* kinase mutations or duplications in specific subclasses, few genetic driver events are known. Diffuse PLGGs comprise a set of uncommon subtypes that exhibit invasive growth and are therefore especially challenging clinically. We performed high-resolution copy-number analysis on 44 formalin-fixed, paraffin-embedded diffuse PLGGs to identify recurrent alterations. Diffuse PLGGs exhibited fewer such alterations than adult low-grade gliomas, but we identified several significantly recurrent events. The most significant event, 8q13.1 gain, was observed in 28% of diffuse astrocytoma grade IIs and resulted in partial duplication of the transcription factor *MYBL1* with truncation of its C-terminal negative-regulatory domain. A similar recurrent deletion-truncation breakpoint was identified in two angiocentric gliomas in the related gene *v-myb avian myeloblastosis viral oncogene homolog* (*MYB*) on 6q23.3. Whole-genome sequencing of a *MYBL1*-rearranged diffuse astrocytoma grade II demonstrated *MYBL1* tandem duplication and few other events. Truncated *MYBL1* transcripts identified in this tumor induced anchorage-independent growth in 3T3 cells and tumor formation in nude mice. Truncated transcripts were also expressed in two additional tumors with *MYBL1* partial duplication. Our results define clinically relevant molecular subclasses of diffuse PLGGs and highlight a potential role for the MYB family in the biology of low-grade gliomas.

cancer | aCGH | A-myb

Pediatric low-grade gliomas (PLGGs) are the most common brain tumors in children and, collectively with other CNS tumors, have surpassed leukemias as the leading cause of cancer-related deaths in children and young adults (1). PLGGs are generally categorized as "nondiffuse" or "diffuse" based on their extent of brain infiltration. Nondiffuse tumors exhibit minimal infiltration and are predominantly benign World Health Organization (WHO) grade I pilocytic astrocytomas (PAs), which are most often cured by surgery alone. In contrast, diffuse gliomas are associated with less favorable clinical outcomes, including recurrence after initial resection, by virtue of their extensive infiltration and invasion into the brain. These tumors are also more likely to progress to glioblastoma. PLGGs with diffuse growth patterns are further subclassified histologically as diffuse astrocytoma grade IIs (DA2s), gangliogliomas (GGs), angiocentric gliomas (AGs), pleomorphic xanthoastrocytomas (PXAs), and several other rare glioma types (2). However, these subclasses exhibit extensive heterogeneity and histologic overlap, often precluding categorical diagnosis. PLGGs that cannot be categorized are often referred to as low-grade gliomas, not otherwise specified (LGG-NOS) and represent nearly one-third of all PLGGs. Moreover, these histologic categories do not reliably predict biologic behavior and risk of malignant transformation.

Unifying genetic events have been identified in some PLGG subtypes, including *v-raf murine sarcoma viral oncogene homolog B1* (*BRAF*) fusions in PAs and *BRAF* V600E mutations in PXAs and GGs, with substantial diagnostic, prognostic, and therapeutic implications (3–9). Identification of genetic alterations in diffuse PLGGs would increase biologic understanding of tumor behavior as well as define diagnostic molecular subclasses. However, unlike pilocytic astrocytomas, the rarity and diversity of diffuse PLGGs combined with the scarcity of frozen tissue available for genomic analyses has historically impeded identification of genetic alterations specific to these tumors. Prior studies have

---

Author contributions: L.A.R., P.M.H., B.E.R., R.B., and K.L.L. designed research; L.A.R., P.M.H., J.M.C., S.H.R., and G.B. performed research; S.E.S., A.M., M.S.L., Y.-J.C., H.T., M.M., D.C.B., L.M., U.T., C.H., R.J.P., D.A.H., S.L.P., C.G.E., I.F.D., L.G., G.G., W.C.H., C.D.S., and M.W.K. contributed new reagents/analytic tools; L.A.R., P.M.H., S.H.R., B.E.R., S.E.S., A.M., M.S.L., P.K.B., B.T., M.D.D., P.V.H., L.E.M., T.P.-Y., G.G., J.A.C., S.S., A.H.L., R.B., and K.L.L. analyzed data; and L.A.R., P.M.H., R.B., and K.L.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The copy number data reported in this paper has been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE45843). The sequence reported in this paper has been deposited in the database of Genotypes and Phenotypes (dbGaP), www.ncbi.nlm.nih.gov/gap (accession no. phs000614.v1.p1).

[1] L.A.R and P.M.H. contributed equally to this work.

[2] To whom correspondence may be addressed. E-mail: keith_ligon@dfci.harvard.edu or rameen_beroukhim@dfci.harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1300252110/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1300252110

found that *BRAF-KIAA1549* fusions are rare to nonexistent in diffuse PLGGs, particularly in DA2s (10). Diffuse PLGGs included in large cohorts of low- and high-grade gliomas were suggested to have increased expression of the proto-oncogene *v-myb avian myeloblastosis viral oncogene homolog* (*MYB*), including rare cases with genomic aberrations involving the gene (11), but no unifying recurrent genetic events have been identified.

Here we describe high-resolution copy-number profiles of 44 diffuse PLGGs, the largest collection ever to have been analyzed, and whole-genome sequencing of a diffuse PLGG. These studies reveal a recurrent rearrangement of the transcription factor *v-myb avian myeloblastosis viral oncogene homolog-like 1* (*MYBL1*) that induces anchorage-independent growth of 3T3 cells as well as tumor growth in vivo. These findings indicate oncogenic events that define subclasses of diffuse PLGGs.

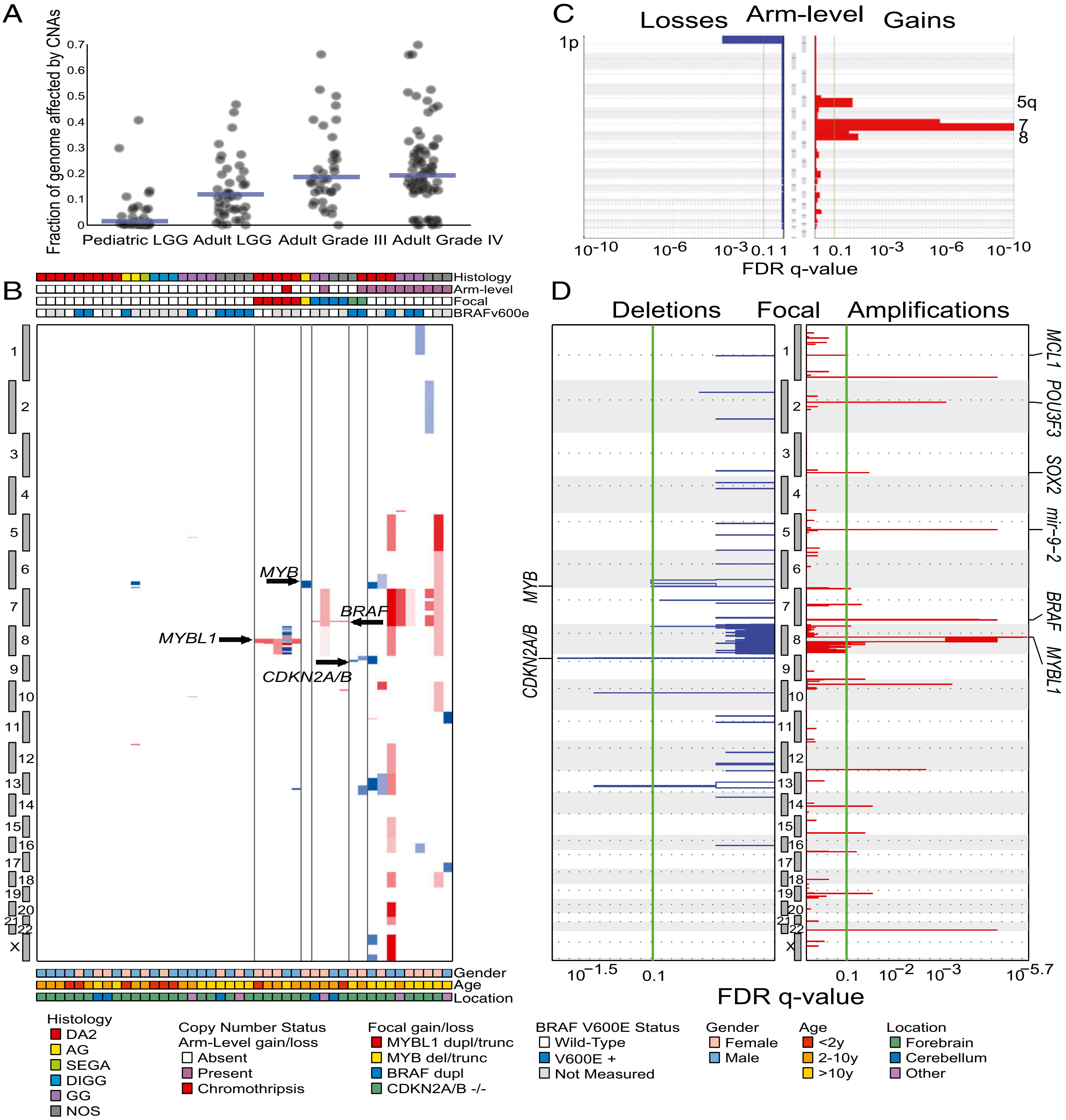


Fig. 1. CNAs among diffuse PLGGs. (*A*) Fraction of the genome altered by CNAs is lower among diffuse PLGGs compared with adult LGGs and high-grade gliomas ($P < 10^{-6}$). Blue bars indicate medians. (*B*) Amplifications (red) and deletions (blue) among 44 diffuse PLGGs (*x* axis) across the genome (*y* axis) ordered by copy-number status. Significance (*x* axis) of (*C*) arm-level and (*D*) focal deletions (*Left*, blue) and amplifications (*Right*, red) across the genome (*y* axes). Putative gene targets within the peak regions are indicated where known.

Results

Characteristics of the Diffuse PLGG Cohort. To focus our studies on diffuse PLGGs, we collected a diverse set of carefully screened tumors through an international consortium of seven institutions (Table S1 and Fig. S1). Our cohort specifically excluded the more common, nondiffuse pilocytic astrocytomas, which are known to be driven primarily by *BRAF* alterations and are the focus of separate ongoing international collaborative sequencing efforts [International Cancer Genome Consortium (Germany)]. Our cohort included 18 DA2, three AG, three desmoplastic infantile ganglioglioma, nine GG, one subependymal giant cell astrocytoma, and 10 LGG-NOS tumors. Given the infiltrative growth and rarity of certain categories of diffuse PLGGs, the samples that we acquired were mostly archival formalin-fixed, paraffin-embedded (FFPE) tissue. We recently developed a method for reliable performance of array comparative genomic hybridization (aCGH) on FFPE archival samples (12) and used this technique to determine copy-number status at 1 million loci genome-wide. We also performed deep whole-genome or whole-exome sequencing to define and validate recurrent genetic events that drive tumorigenesis in these rare pediatric cancers.

Genomic Identification of Significant Targets in Cancer Analysis Identifies Significant Recurrent Events in Specific PLGG Subtypes. The percentage of the genome altered by copy-number alterations (CNAs) in diffuse PLGGs was significantly lower than among previously profiled adult low- and high-grade gliomas ($P < 10^{-6}$, Mann–Whitney test, Fig. 1*A*) (13, 14). Few (12/44; 27%) of these tumors harbored alterations affecting more than 90% of the length of a chromosome arm (Fig. 1*B*), compared with an 83–97% rate among adult low- and high-grade tumors (15). One of the PLGG samples exhibited chromothripsis on chromosome 8 (chr8) (highlighted in Fig. 2*A*, PLGG27). The most significantly recurrent arm-level CNAs were gains of chromosomes 7 (11% of tumors), 8 (7%), and 5q (5%) and loss of 1p (2%) (Fig. 1*C*). These events have all been described in pediatric high-grade gliomas and adult gliomas with varying frequencies (15, 16).

We found 6 significantly recurrent regions of focal deletion and 17 significantly recurrent regions of focal amplification (Fig. 1*D* and Table S2). One deleted region on 9p21.3 contained *cyclin-dependent kinase inhibitor 2A and 2B* (*CDKN2A* and *CDKN2B*), known tumor suppressors that had previously been reported in diffuse PLGGs (17); a second region was immediately adjacent to this one. A third region (6q26) contained 252 genes, including the proto-oncogene *MYB*. Two regions (10q21.3 and 8p22) contained single genes with no known relation to cancer or neural development, *catenin (cadherin-associated protein), alpha 3* (*CTNNA3*) and *zeta sarcoglycan* (*SGCZ*), respectively. The sixth region (13q31.3) contained 48 genes and was adjacent to the known tumor suppressor *RB1*. We did not identify any focal deletions of other known tumor suppressors involved in adult or pediatric brain tumors such as *neurofibromin 1* (*NF1*), *phosphatase and tensin homolog* (*PTEN*), or *cyclin-dependent kinase inhibitor 1C* (*CDKN1C*).

One of the 17 focally gained regions contained *BRAF*. However, the canonical *BRAF-KIAA1549* duplication-fusion was detected in only four samples: two GGs and two LGG-NOS. This is in contrast to pilocytic astrocytomas, among which >80% of tumors harbor a *BRAF* duplication (18) ($P < 0.0001$, Fisher's exact test). We also determined *BRAF* V600E mutation status in 24 tumors with sufficient DNA for sequencing. We found mutations in 54% of the diffuse PLGGs (Fig. 1*A* and Table S1), consistent with previously published rates for diffuse PLGGs (8).

A second focally gained region (3q26.33) contained the stem cell and glial transcription factor *sex determining region Y-box 2* (*SOX2*), which is amplified in adult glioblastomas (19). Two additional regions (2q12.1 and 5q14.3) contained factors that control telencephalic neural progenitor proliferation and differentiation: *POU class 3 homeobox 3* (*POU3F3*) (also known as *BRN1*) and *microRNA 9-2* (20, 21). A fifth region (1q21.3) contained *myeloid cell leukemia sequence 1* (*MCL1*), a known oncogene amplified in several cancer types (22). Twelve regions either contained over 150 genes or did not contain genes with known roles in cancer or neural development. We did not observe any high-level amplification of receptor tyrosine kinases (e.g., *EGFR*, *PDGFRA*), which are observed frequently in both adult and pediatric high-grade gliomas (19, 23).

The most statistically significant recurrent focal aberration ($q = 3.37 \times 10^{-6}$) was a gain on chromosome 8q involving the transcription factor *MYBL1*. Although *MYBL1* is not a known oncogene, it is closely related to the proto-oncogene *MYB*. In contrast to prior reports (11), no amplifications or gains of the proto-oncogene *MYB* were identified in our study set. All of the focal 8q gains occurred in DA2s ($P = 0.0057$, Fisher's exact test), comprising 28% (5/18) of this histologic subtype. In contrast, *MYBL1* was not in a significant amplification peak across 3,131 cancers comprising multiple other cancer types that we had previously analyzed (22) or, specifically, among adult low- or high-grade gliomas (15).

All five DA2 samples with 8q focal gains exhibited a common centromeric breakpoint within *MYBL1* after exon 9, including the sample with chromothripsis of chr8 (Fig. 2*A*). To confirm the *MYBL1* centromeric breakpoint, we performed fluorescence in situ hybridization (FISH) using a probe slightly telomeric to the breakpoint on all eight DA2 samples with sufficient tissue available (Fig. 2*B* and Fig. S2). All DA2 samples with 8q gain (3/3) demonstrated duplication of one allele in more than 60% of the nuclei in each tumor whereas none of the other DA2 samples showed duplication (0/5) ($P = 0.018$, Fisher's exact test).

The tight clustering of these breakpoint sites, and particularly their location immediately preceding the C-terminal negative regulatory domains of *MYBL1* and *MYB*, suggested a mechanism for rearrangement and creation of functional, truncated genes reminiscent of the viral oncogene *v-MYB* (Fig. 2*C*). Indeed, we also identified a homologous breakpoint between exons 10 and 11 of *MYB* on 6q in one angiocentric glioma with a focal 6q deletion (Fig. S3) similar to that previously reported in a single angiocentric glioma with a 6q deletion (11). An additional angiocentric glioma (PLGG45), not included in our initial diffuse PLGG cohort and genomic identification of significant targets in cancer (GISTIC) analysis, also exhibited a deletion in 6q at the same location in *MYB* as seen in PLGG29 (Fig. S3).

Whole-Genome Sequencing of a DA2 with 8q Focal Gain Defines a Tandem Duplication–Truncation of *MYBL1*. To further characterize the *MYBL1* amplicon and its genetic context, we performed 90× whole-genome sequencing of a DA2 sample with *MYBL1* gain but no other CNAs (PLGG24, Table S3). Whole-genome sequencing of PLGG24 determined the centromeric breakpoint of the 8q amplicon to single-base resolution between exons 9 and 10 of *MYBL1* (Fig. 3*A*). The telomeric sequence was located in an intergenic region 38 kb from *matrix metallopeptidase 16* (*MMP16*). We validated the breakpoint locations in this sample using PCR on native genomic DNA from the same tumor (Fig. S2 *C* and *D*). Taken together, our data define a tandem duplication–truncation of *MYBL1*.

Apart from this event, whole-genome sequencing of PLGG24 revealed a sparsely altered genome. No other CNAs or fusion events were identified, and the *BRAF* V600E mutation was not present. Three nonsynonymous mutations in exons were identified (Table S4); none of these have been reported in association with cancer. The genome-wide mutation rate (1.48/Mb) and the number of nonsynonymous mutations in exons (three per genome) were low compared with pediatric and adult high-grade astrocytomas (mutation count means: 15 and 47.3 per genome,

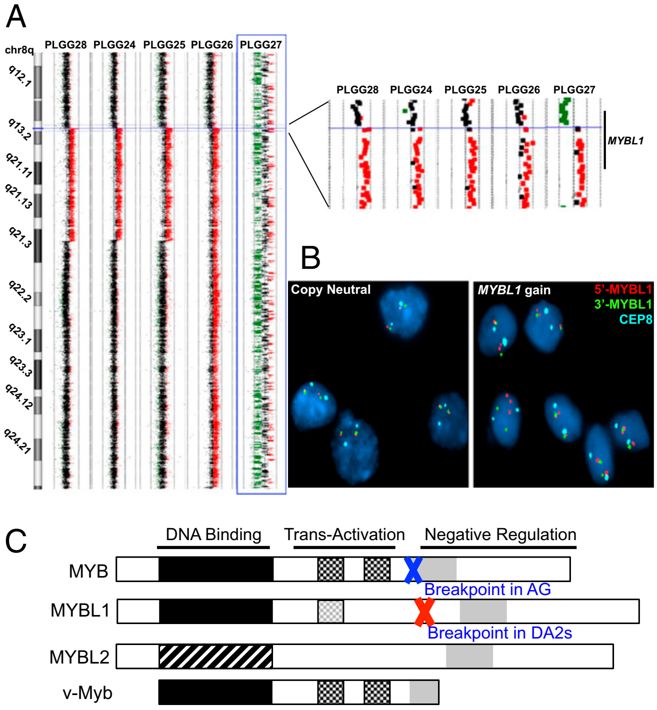


Fig. 2. DA2 samples with focal 8q gains identified by aCGH share a common centromeric breakpoint within *MYBL1*. (*A*) aCGH data for the five DA2 samples with 8q gain, magnified on the right. Red: copy-number gain; green: loss. The highlighted sample (PLGG27, blue box) exhibits chromothripsis of chr8 but shows no other copy-number changes on other chromosomes. (*B*) FISH probes corresponding to sequences immediately distal (red) and proximal (green) of *MYBL1* confirm that the single-copy gain identified by aCGH is a duplication involving one *MYBL1* allele. The D8Z2 centromere enumeration probe (aqua) was used as a control. (*C*) Schematic representation of the proto-oncogene *MYB* family and breakpoints observed in our cohort in relation to the viral oncogene *v-MYB*.

respectively) (16). The low rate of CNAs, mutations, and translocations in this sample highlight the potential biological impact of *MYBL1* duplication–truncation.

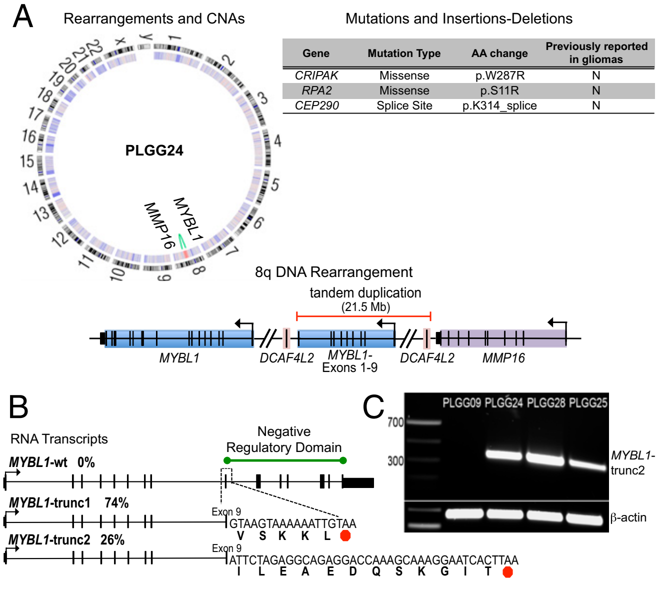
Tandem Duplication–Truncation of *MYBL1* Results in Expression of Oncogenic Transcripts. To determine whether the *MYBL1* duplication–truncation resulted in expression of fused transcripts, we performed 3′-rapid amplification of cDNA ends (3′-RACE) on cDNA generated from RNA of PLGG24. We identified two populations of *MYBL1* transcripts, both of which contained *MYBL1* exons 1–9 but also acquired short noncanonical sequences fused to the 3′ ends, leading to premature translation stops (Fig. 3*B*). The first transcript (*MYBL1*-trunc1) contained an additional 15 bp of intronic sequence, and the second transcript (*MYBL1*-trunc2) acquired 36 bp from the intergenic region near *MMP16*. *MYBL1*-trunc1 was retrieved more frequently (74%) than *MYBL1*-trunc2 (26%); the wild-type transcript was not observed. We then performed RT-PCR to detect *MYBL1*-trunc1 and *MYBL1*-trunc2 in the two other PLGGs with 8q focal amplification for which we had sufficient RNA (PLGG25 and PLGG28). We detected *MYBL1*-trunc2 in both of these samples (Fig. 3*C*), indicating that this transcript is recurrently expressed in DA2s.

To assess the oncogenic potential of these transcripts, we transduced 3T3 cells with lentiviruses containing *MYBL1*-trunc1, *MYBL1*-trunc2, full-length *MYBL1*-wt, or GFP control constructs and plated them in soft agar. Both aberrantly truncated *MYBL1* sequences produced soft agar colony growth indicative of transformation (Fig. 4*A*). No evidence of colony formation was noted with full-length *MYBL1*-wt or *GFP* control. Both *MYBL1*-trunc1– and *MYBL1*-trunc2–transformed 3T3 cells were also able to form tumors with malignant histology in Nude mice, whereas cells transduced with full-length *MYBL1*-wt constructs showed no evidence of tumor formation (Fig. 4 *B*–*D*). These results suggest that the truncation of *MYBL1* is oncogenic.

Discussion

Our data identify several recurrent somatic genetic events in PLGG, some of which track with specific histologic types. The most significant of these was recurrent focal amplification of *MYBL1*, found exclusively in DA2 samples, none of which had other PLGG-associated lesions such as a *BRAF* duplication. Focal deletion of 6q involving the *MYB* locus was also observed in two angiocentric gliomas in our dataset, similar to deletions in two individual cases of angiocentric glioma noted in prior studies. Such aberrations may therefore be useful in identifying this tumor type (11, 24). In other cases, genetic events span histologic subtypes and tumor grades, such as duplications of *BRAF* in GG and NOS samples in our diffuse PLGG cohort as well as deletions involving *CDKN2A/B* (Fig. 5). Although previous work had identified high-level amplification of *MYB* and 6q deletions involving *MYB* in several DA2 tumors, we detected no such events in our DA2 cohort (11).

These findings provide genetic and functional evidence of a role for *MYBL1* and the MYB proto-oncogene family of transcription factors in low-grade gliomagenesis. MYB transcription factors are known to regulate cell-cycle progression in multiple cellular contexts, and *MYB* has been identified as an oncogene in T-ALL and adenoid cystic carcinomas (25). The MYB family members share extensive homology in their DNA-binding domains; however, they differ in their C-terminal domains. In vitro and in vivo studies have



Gene	Mutation Type	AA change	Previously reported in gliomas
CRIPAK	Missense	p.W287R	N
RPA2	Missense	p.S11R	N
CEP290	Splice Site	p.K314_splice	N

GTAAGTAAAAAATTGTAA
V S K K L

ATTCTAGAGGCAGAGGACCAAAGCAAAGGAATCACTTAA
I L E A E D Q S K K G I T

Fig. 3. Characterization of the *MYBL1*-truncating rearrangement in a DA2 sample (PLGG24). (*A*) Circos plot (*Left*) of 90× whole-genome sequence data showing a single-copy-number gain on 8q (red, inner heatmap ring) and corresponding intrachromosomal rearrangement (green) involving *MYBL1* and an intergenic region outside *MMP16*. A schematic diagram represents the 8q tandem duplication (*Lower Right*). All nonsynonymous mutations and insertions and deletions are also listed (*Upper Right*). (*B*) Truncated *MYBL1* transcripts (*MYBL1*-trunc1 and *MYBL1*-trunc2) identified by 3′-RACE. (*C*) RT-PCR demonstrating *MYBL1*-trunc2 expression in two additional PLGG samples with *MYBL1* gain.

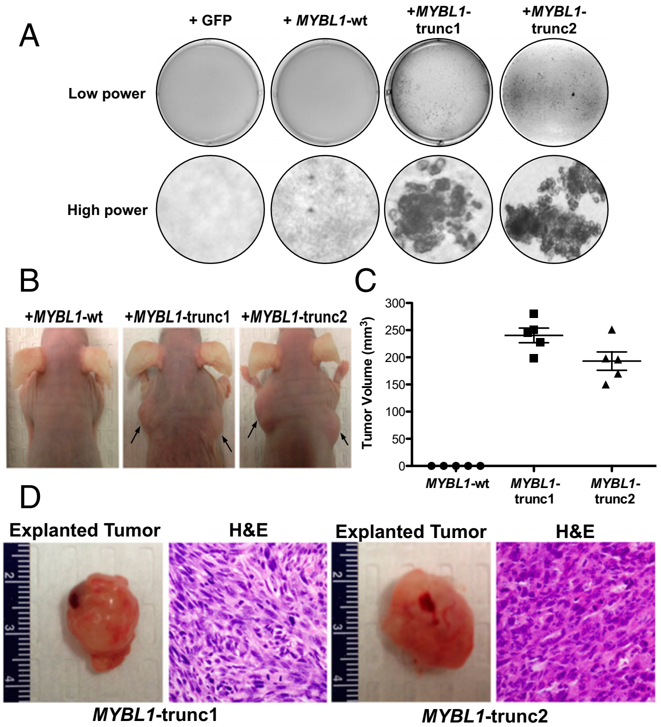


Fig. 4. (*A*) Anchorage-independent growth of 3T3 cells transduced with *MYBL1*-wt, *MYBL1*-trunc1, or *MYBL1*-trunc2 retroviruses, relative to *GFP* controls. (*B*) In vivo tumor formation after injection of 3T3 cells transduced with *MYBL1*-wt (n = 5), *MYBL1*-trunc1 (n = 5), or *MYBL1*-trunc2 (n = 5) (arrows) into the flanks of Nude mice. Tumor volume was calculated 6 wk post injection (*C*), and tumors were subjected to histologic analysis by H&E staining (*D*).

demonstrated roles for each transcription factor in cell-cycle regulation mediated by interactions with and modifications of the C terminus (26). Truncating deletions may affect the transcriptional activity of MYB transcription factors by loss of a conserved negative regulatory domain or may result in increased gene expression due to loss of a 3′ UTR targeted by microRNA 15a/16 and microRNA 150 (25). Our results suggest that truncations affecting the C terminus of MYB family transcription factors may be sufficient to drive oncogenesis in a discrete molecular subclass of diffuse PLGGs by acting as gain-of-function mutations.

Methods

Patients and Samples. Institutional review board approval from all institutions (Boston Children's Hospital, Dana-Farber Cancer Institute, The University of Texas School of Medicine Southwestern, Children's Cancer Hospital-Egypt, The Johns Hopkins University School of Medicine, Children's National Medical Center, Hospital for Sick Children, and the Mayo Clinic) was obtained, and all samples were from patients who provided informed consent or were studied with waiver of the requirement for informed consent by the Dana-Farber Cancer Institutional review board. Samples of various histologic subtypes were identified, collected at multiple institutions, and central histopathologic review was performed by at least three board-certified neuropathologists using WHO criteria (K.L.L., S.S., S.H.R., or J.A.C.).

aCGH and Data Processing. DNA extraction from archival FFPE samples and aCGH were performed as previously described (12). GC-normalized copy-number data for the samples were then cleaned of known germ-line copy-number variations. Circular Binary Segmentation was used to segment the copy-number data, using parameters ($\alpha = 0.001$, undo.splits = sdundo, undo.SD = 1.5, minimum width = 5). Segmented data were analyzed with GISTIC 2.0 to determine statistically significant recurrent broad and focal CNAs using the following parameters: minimum segment size = 8, lesion amplitude threshold = 0.2, focal/broad cutoff = 0.9× chromosome arm length, q-value threshold = 0.10, and gene confidence level = 0.95. For comparison of diffuse PLGG data to previously published adult low-grade glioma (LGG) and high-grade glioma (HGG) data (13, 14), previously segmented copy-number data were subjected to the same GISTIC analysis parameters as above.

Whole-Genome Sequencing and Data Processing. DNA from fresh-frozen tissue from PLGG24 and paired blood was extracted using the QIAGEN DNA Blood and Tissue kit. A target depth of 90× in blood and tumor was set for Illumina sequencing, using two different insert-size libraries (500 and 800 bp) to maximize detection of rearrangements (27). Sequencing quality control (QC) metrics are shown in Table S3.

Sequence data were aligned to the hg19 (b37) reference genome with the Burrows-Wheeler Aligner (28) with parameters [-q 5 -l 32 -k 2 -t 4 -o 1]. Aligned data were sorted, normalized, mate-fixed, duplicate-marked, and indexed with Samtools and Picard tools (29). Base-quality score recalibration and local realignment around insertions and deletions was achieved with the Genome Analysis Toolkit (30, 31).

Somatic mutations and small insertions-deletions were called with MuTect and Indelocator, filtered against a panel of normals, and annotated to genes with Oncotator (29, 32, 33). CNAs were called with SegSeq and standard parameters (34). Somatic rearrangements were identified with dRanger and BreakPointer algorithms (32, 33) with the cutoff SN parameter increased to 1,000 bp (reflecting the larger-than-normal insert sizes used for sequencing). Results are reported with high confidence if the dRanger score was ≥8 and the BreakPointer algorithm identified the exact breakpoints on both ends (equivalent to 8× high-confidence coverage of read pairs spanning the breakpoint).

Whole-Exome Sequencing. DNA was extracted from tumors as above, and 250-bp libraries were prepared by Covaris sonication, followed by double-size selection (Agencourt AMPure XP beads) and ligation to specific barcoded

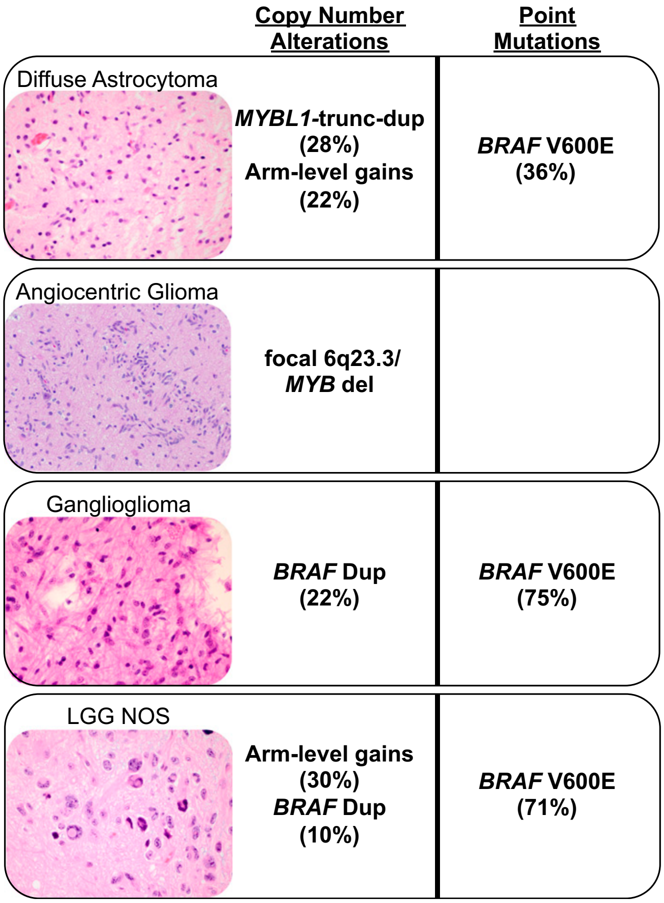


Fig. 5. Schematic summarizing the highly characteristic genomic aberrations detected in our diffuse PLGG cohort.

adaptors (Illumina TruSeq) for multiplexed analysis. Exome hybrid capture was performed with the Agilent Human All Exon v2 (44 Mb) or Illumina TruSeq bait sets, and samples were sequenced as above. Tumors were manually reviewed for the presence of the *BRAF* V600E mutation.

FISH and PCR. FISH was performed on 4-μm tissue sections using methods described previously (35) and Homebrew probes RP11-110J18 (5′ to *MYBL1*; directly labeled in SpectrumOrange) and RP11-707M3 (3′ to *MYBL1*; directly labeled in SpectrumGreen) that map to 8q13.1. *MYBL1* status was assessed in 50 tumor nuclei per sample. PCR was performed on genomic DNA from PLGG0024 and control samples to confirm breakpoint sequences identified by dRanger. Primer sequences (5′-AATGCTATCCCTCCCCACTC-3′ and 5′-GAGGGAGCTTGGAAATTTGA-3′) targeting *MYBL1* and intergenic sequences, respectively, amplified a 450-bp fragment. The band was gel-purified, cloned (TOPO TA Cloning; Invitrogen, and sequenced by Sanger sequencing to validate the fusion.

RT-PCR and 3′-RACE. *MYBL1*-trunc1 and *MYBL1*-trunc2 were cloned from PLGG24 frozen tumor tissue using a 3′ RACE kit (Invitrogen) per manufacturer's instructions. To amplify all *MYBL1* specific transcripts, primers targeting the 5′ end of *MYBL1* (5′-AAAACCCTGCAGGAGACTG-3′) were used in conjunction with a universal amplification primer (UAP). A second PCR was performed using a nested *MYBL1* primer (5′-TGCGGTACTTGAAGGATGG-3′) along with the UAP. PCR products were subjected to Sanger sequencing, and results were aligned to the hg19 reference genome. For RT-PCR reactions, RNA was extracted from FFPE samples of PLGG09, PLGG25, and PLGG28 using RNeasy FFPE kit (Qiagen). cDNA was generated from 500 ng RNA using the iScript cDNA synthesis kit (Bio-Rad). PCR to detect the presence of *MYBL1*-trunc2 was performed using primers targeting the *MYBL1* exon 7–8 junction (5′-ATTGTATAGAACATGTTCAGCCT-3′) and the *MYBL1*-trunc2 sequence (5′-GGTCCTCTGCCTCTAGAATAGATTC-3′).

Expression Constructs and Lentiviral Production. Full-length *MYBL1-wt* (Open Biosystems), *MYBL1*-trunc1, and *MYBL1*-trunc2 cDNA sequences were subcloned into the pLenti7.3/V5 vector using the Gateway system (Invitrogen). For retroviral production, 293FT packaging cells were cotransfected with pBabe expression clones, Gag-pol, and vs.v-g. Viral supernatant was harvested 48 h after transfection, filtered through a 45-μm filter, and concentrated by ultracentrifugation.

Anchorage-Independent Growth Assay. Wild-type 3T3 mouse embryonic fibroblasts (ATCC) were transduced by addition of viral supernatant to the growth medium. Twenty-four hours later, infection efficiency was evaluated based on GFP expression and determined to be >80%. Cells were harvested and mixed with growth medium containing 0.33% bactoagar, and 1×10^4 cells were plated in triplicate onto a bottom layer of medium with 0.5% agar in a six-well plate. Soft agar colonies were counted 2 wk later. Images were acquired using the AlphaInnotch FluorChem HD2 Imager.

Flank Tumor Growth Assay. The Institutional Animal Care and Use Committee at Dana-Farber Cancer Institute preapproved all animal experiments. 3T3-*MYBL1*-wt, 3T3-*MYBL1*-trunc1, or 3T3-*MYBL1*-trunc2 cells (10^6) were suspended in 150 μL PBS, mixed with 150 μL Matrigel (BD Biosciences), and then injected into the flank of 6-wk-old male Nude (NU-Foxn1, Charles River) mice. Mice were then monitored for signs of distress or tumor growth. Six weeks post injection the mice were euthanized and analyzed for tumor growth. Tumors were subjected to standard histologic analysis.

ACKNOWLEDGMENTS. The authors thank Denise Garcia and Heather Homer for technical help with FISH experiments; Yotam Drier and Cheng-Zhong Zhang for assistance with analytical algorithms; and Anna Pisarek-Horowitz and Sophie Horowitz for critical reading of the manuscript. This work was supported by the Pediatric Low Grade Astrocytoma Foundation; The Sontag Foundation (K.L.L., R.B.); Goldhirsh Foundation (K.L.L.); Boston Children's Hospital Team Path to Cure (K.L.L.); American Brain Tumor Association Basic Research Fellowship in Honor of Joel A. Gingras, Jr. (to L.A.R.); the Eleanor and Miles Shore 50th Anniversary Fellowship Program for Scholars in Medicine (to A.H.L.); Nuovo-Soldati Foundation and Philippe Foundation, Inc. (to G.B.); and National Institutes of Health Grants R01 CA109467 and P30 HD018655 (to S.L.P.), U54CA143789 and K08CA122833 (to R.B.), and P01CA142536 (to K.L.L., W.C.H., S.S., C.D.S., L.E.M., and M.W.K.).

1. Horner MJRL, et al. (2009) *SEER Cancer Statistics Review 1975–2006* (National Cancer Institute, Bethesda, MD).
2. Louis DN, et al. (2007) The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol* 114(2):97–109.
3. Pfister S, et al. (2008) BRAF gene duplication constitutes a mechanism of MAPK pathway activation in low-grade astrocytomas. *J Clin Invest* 118(5):1739–1749.
4. Dias-Santagata D, et al. (2011) BRAF V600E mutations are common in pleomorphic xanthoastrocytoma: Diagnostic and therapeutic implications. *PLoS ONE* 6(3):e17948.
5. Sievert AJ, et al. (2009) Duplication of 7q34 in pediatric low-grade astrocytomas detected by high-density single-nucleotide polymorphism-based genotype arrays results in a novel BRAF fusion gene. *Brain Pathol* 19(3):449–458.
6. Jones DT, et al. (2008) Tandem duplication producing a novel oncogenic BRAF fusion gene defines the majority of pilocytic astrocytomas. *Cancer Res* 68(21):8673–8677.
7. Bar EE, Lin A, Tihan T, Burger PC, Eberhart CG (2008) Frequent gains at chromosome 7q34 involving BRAF in pilocytic astrocytoma. *J Neuropathol Exp Neurol* 67(9): 878–887.
8. Dougherty MJ, et al. (2010) Activating mutations in BRAF characterize a spectrum of pediatric low-grade gliomas. *Neuro Oncol* 12(7):621–630.
9. Schindler G, et al. (2011) Analysis of BRAF V600E mutation in 1,320 nervous system tumors reveals high mutation frequencies in pleomorphic xanthoastrocytoma, ganglioglioma and extra-cerebellar pilocytic astrocytoma. *Acta Neuropathol* 121(3): 397–405.
10. Lin A, et al. (2012) BRAF alterations in primary glial and glioneuronal neoplasms of the central nervous system with identification of 2 novel KIAA1549:BRAF fusion variants. *J Neuropathol Exp Neurol* 71(1):66–72.
11. Tatevossian RG, et al. (2010) MYB upregulation and genetic aberrations in a subset of pediatric low-grade gliomas. *Acta Neuropathol* 120(6):731–743.
12. Craig JM, et al. (2012) DNA fragmentation simulation method (FSM) and fragment size matching improve aCGH performance of FFPE tissues. *PLoS ONE* 7(6):e38881.
13. Kotliarov Y, et al. (2006) High-resolution global genomic survey of 178 gliomas reveals novel regions of copy number alteration and allelic imbalances. *Cancer Res* 66(19):9428–9436.
14. Anonymous (2005) *REMBRANDT home page* (National Cancer Institute, Bethesda, MD).
15. Beroukhim R, et al. (2007) Assessing the significance of chromosomal aberrations in cancer: Methodology and application to glioma. *Proc Natl Acad Sci USA* 104(50): 20007–20012.
16. Schwartzentruber J, et al. (2012) Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* 482(7384):226–231.
17. Schiffman JD, et al. (2010) Oncogenic BRAF mutation with CDKN2A inactivation is characteristic of a subset of pediatric malignant astrocytomas. *Cancer Res* 70(2): 512–519.
18. Tatevossian RG, et al. (2010) MAPK pathway activation and the origins of pediatric low-grade astrocytomas. *J Cell Physiol* 222(3):509–514.
19. Atlas TCG; Cancer Genome Atlas Research Network (2008) Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 455(7216):1061–1068.
20. McEvilly RJ, de Diaz MO, Schonemann MD, Hooshmand F, Rosenfeld MG (2002) Transcriptional regulation of cortical neuron migration by POU domain factors. *Science* 295(5559):1528–1532.
21. Shibata M, Nakao H, Kiyonari H, Abe T, Aizawa S (2011) MicroRNA-9 regulates neurogenesis in mouse telencephalon by targeting multiple transcription factors. *J Neurosci* 31(9):3407–3422.
22. Beroukhim R, et al. (2010) The landscape of somatic copy-number alteration across human cancers. *Nature* 463(7283):899–905.
23. Paugh BS, et al. (2010) Integrated molecular genetic profiling of pediatric high-grade gliomas reveals key differences with the adult disease. *J Clin Oncol* 28(18):3061–3068.
24. Preusser M, et al. (2007) Angiocentric glioma: Report of clinico-pathologic and genetic findings in 8 cases. *Am J Surg Pathol* 31(11):1709–1718.
25. Persson M, et al. (2009) Recurrent fusion of MYB and NFIB transcription factor genes in carcinomas of the breast and head and neck. *Proc Natl Acad Sci USA* 106(44): 18740–18744.
26. Lei W, Rushton JJ, Davis LM, Liu F, Ness SA (2004) Positive and negative determinants of target gene specificity in myb transcription factors. *J Biol Chem* 279(28): 29519–29527.
27. Brastianos PK, et al. (2013) Genomic sequencing of meningiomas identifies oncogenic SMO and AKT1 mutations. *Nat Genet* 45(3):285–289.
28. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14):1754–1760.
29. Network TCGA; Cancer Genome Atlas Research Network (2011) Integrated genomic analyses of ovarian carcinoma. *Nature* 474(7353):609–615.
30. McKenna A, et al. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20(9):1297–1303.
31. DePristo MA, et al. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43(5):491–498.
32. Berger MF, et al. (2011) The genomic complexity of primary human prostate cancer. *Nature* 470(7333):214–220.
33. Chapman MA, et al. (2011) Initial genome sequencing and analysis of multiple myeloma. *Nature* 471(7339):467–472.
34. Chiang DY, et al. (2009) High-resolution mapping of copy-number alterations with massively parallel sequencing. *Nat Methods* 6(1):99–103.
35. Firestein R, et al. (2008) CDK8 is a colorectal cancer oncogene that regulates beta-catenin activity. *Nature* 455(7212):547–551.

6) Predicting clinical response to anticancer drugs using an ex vivo platform that captures tumor heterogeneity

Majumder B, Baraneedharan U, Thiyagarajan S, Radhakrishnan P, Narasimhan H, Dhandapani M, Brijwani N, Pinto DD, Prasath A, Shanthappa BU, Thayakumar A, Surendran R, Babu GK, Shenoy AM, Kuriakose MA, Bergthold G, Horowitz P, Loda M, Beroukhim R, Agarwal S, Sengupta S, Sundaram M, Majumder PK.

Nat Commun. 2015 Feb 27;6:6169.

nature COMMUNICATIONS

ARTICLE

Received 3 Oct 2014 | Accepted 22 Dec 2014 | Published 27 Feb 2015

DOI: 10.1038/ncomms7169

OPEN

Predicting clinical response to anticancer drugs using an *ex vivo* platform that captures tumour heterogeneity

Biswanath Majumder[1], Ulaganathan Baraneedharan[1,*], Saravanan Thiyagarajan[1,*], Padhma Radhakrishnan[1], Harikrishna Narasimhan[2], Muthu Dhandapani[1], Nilesh Brijwani[1], Dency D. Pinto[1], Arun Prasath[1], Basavaraja U. Shanthappa[1], Allen Thayakumar[1], Rajagopalan Surendran[3], Govind K. Babu[4], Ashok M. Shenoy[4], Moni A. Kuriakose[5], Guillaume Bergthold[6], Peleg Horowitz[6,7,8], Massimo Loda[6,7], Rameen Beroukhim[7,8], Shivani Agarwal[2], Shiladitya Sengupta[7,9,10,*], Mallikarjun Sundaram[1,*] & Pradip K. Majumder[1,9,*]

Predicting clinical response to anticancer drugs remains a major challenge in cancer treatment. Emerging reports indicate that the tumour microenvironment and heterogeneity can limit the predictive power of current biomarker-guided strategies for chemotherapy. Here we report the engineering of personalized tumour ecosystems that contextually conserve the tumour heterogeneity, and phenocopy the tumour microenvironment using tumour explants maintained in defined tumour grade-matched matrix support and autologous patient serum. The functional response of tumour ecosystems, engineered from 109 patients, to anticancer drugs, together with the corresponding clinical outcomes, is used to train a machine learning algorithm; the learned model is then applied to predict the clinical response in an independent validation group of 55 patients, where we achieve 100% sensitivity in predictions while keeping specificity in a desired high range. The tumour ecosystem and algorithm, together termed the CANScript technology, can emerge as a powerful platform for enabling personalized medicine.

---

[1] Mitra Biotech, Bangalore 560099, India. [2] Indian Institute of Science, Bangalore 560012, India. [3] Government Stanley Medical College, Chennai 600001, India. [4] Kidwai Memorial Institute of Oncology, Bangalore 560030, India. [5] Mazumdar-Shaw Cancer Center, Bangalore 560099, India. [6] The Broad Institute of The Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts 02142, USA. [7] Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. [8] Children's Hospital, Boston, Massachusetts 02115, USA. [9] India Innovation Research Center, New Delhi 110092, India. [10] Harvard-MIT Division of Health Sciences and Technology, Cambridge, Massachusetts 02139, USA. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to P.K.M. (email: pradip@mitrabiotech.com).

The ability to predict patient tumour response to cytotoxic or target defined therapeutic agents remains a holy grail. While molecular and genetic profiling is driving the evolution of subtype-specific personalized therapy[1,2], the presence of a biomarker often does not translate into a successful clinical outcome[3–5]. For example, epidermal growth factor receptor (EGFR) inhibitors, cetuximab and panitumumab, are approved for metastatic colorectal carcinoma with wild-type *KRAS,* but provide clinical benefit in only 10–20% of selected patients[1,6,7]. A technology that can identify drug sensitivity and predict clinical benefit can significantly advance the clinical management of cancer.

Emerging evidence implicates intratumoral heterogeneity, both hierarchical and stochastic, in the variability of response to chemotherapy, which is not captured by the existing cancer cell biomarker-based approaches. Genetic and epigenetic distinctions within clonal populations could critically determine whether a particular drug combination will benefit a patient or result in resistance[8–13]. In addition, the contribution of the tumour microenvironment to these phenotypes is increasingly being appreciated[9,10,14,15]. Indeed, the spatial distribution of cancer and stromal cells within the tumour microenvironment can affect how they interact with each other and their microenvironment, which in turn can impact proliferation, differentiation, morphology and a range of cellular functions[16–18]. We rationalized that to predict the clinical outcome of chemotherapy with high accuracy, it is therefore important to conserve this clinical ‘global’ heterogeneity with high fidelity in terms of cancer and stromal cells, tumour microenvironment and architecture. Unfortunately, current gold-standard *in vitro* and *ex vivo* preclinical approaches that employ cell lines and spheroids[3,12,19] or *ex vivo* organotypic tumour models are all limited by their inability to capture the full biological approximation of the native tumour, resulting in poor mapping to clinical outcomes[19–22].

To create a clinically relevant predictive model, here we engineered an *ex vivo* tumour ecosystem, where thin tumour sections with conserved cellular and microenvironmental heterogeneity and architecture were cultured in tissue culture wells coated with grade-matched tumour matrix support in the presence of autologous serum (AS) containing endogenous ligands. The integration of the tumour ecosystems with a novel machine learning algorithm formed the CANScript platform, which reliably predicted the therapeutic efficacy of targeted and cytotoxic drugs in patients with head and neck squamous cell carcinoma (HNSCC) and colorectal cancer (CRC). The robustness of this platform in predicting clinical response could potentially be useful for personalizing cancer treatment.

Results

Role of matched tumour matrix proteins in CANScript platform. We depict the schematic for the development and validation of the CANScript platform in Fig. 1. A detailed patient demography and tumour subtypes used in this study are provided in Supplementary Table 1. As a first step towards mimicking the patient tumour ecosystem, we studied the contribution of cancer and grade-specific human tumour-stromal matrix proteins (TMPs) in preserving tumour morphology of HNSCC and CRC explants in an *ex vivo* setting. Indeed, three-dimensional (3D) matrix support is emerging as a critical factor that dynamically determines the fate of tumours in terms of integrity, survival, metastasis and response to chemotherapy[23–25]. We isolated and characterized the matrix components from clinical HNSCC and CRC tumours using processes described in detail in Supplementary Methods and Supplementary Fig. 1. The overall relative abundance of different TMP in tumour (both HNSCC and CRC) biopsies was analysed by liquid chromatography–mass spectrometry (LCMS/MS; Fig. 2a). Interestingly, a systematic analysis of the major TMP components not only revealed distinct compositions between the two tumour types and between high- and low-grade tumours of the same type (Fig. 2b,c), but also heterogeneity within the patient population as demonstrated using heat maps (Supplementary Figs 2a,d and 3a,d). Venn diagrams reveal unique matrix proteins that were conserved across the patient cohort within each tumour type and grade (Supplementary Figs 2b,e and 3b,e), which together with their abundance (median) (Supplementary Figs 2c,f and 3c,f) formed the basis for selection of the proteins to create the tumour- and grade-matched cocktails (listed in Supplementary Figs 2,3). We coated tissue culture microwells with these defined cancer- and grade-specific TMPs, which was confirmed using scanning electron microscopy and matrix proteins-specific immunofluorescence (Fig. 2d). Thin section tumour explants were then cultured in these TMP-coated wells. As compared with uncoated control, type- and grade-matched TMP showed a dose-dependent improvement in the maintenance of tissue morphology, proliferation and cell viability of the tumour explants (Fig. 2e,f). Furthermore, scanning electron microscopy analysis of native tumour extracellular matrix structure post culture indicated that integrity was better preserved in tumour explant tissues that were provided with TMP support (Fig. 2g). To further understand the role of grade-matched TMP cocktail, we did a cross-comparison analysis where high- and low-grade tumours were cultured in matched and unmatched TMP-coated plates. As shown in Fig. 2h, explants cultured on matched TMPs better retained native (T_0) proliferation (Ki-67) state compared with the corresponding unmatched counterparts and no matrix controls. As expected, high-grade tumours did exhibit a greater capacity to preserve the proliferation profile even in low-grade TMP. Low-grade tumours in high-grade matrix performed poorly (Fig. 2h and Supplementary Fig. 4a). Next we compared the effects of different commercially available matrix proteins with TMP coating in maintaining the proliferation, viability and signalling activation of the explants to the native state (T_0 baseline). As shown in Fig. 2i–j and Supplementary Fig. 4b, explants cultured in non-coated wells lost tumour architecture and exhibited decreased viability, proliferation and activation of oncogenic pathways compared with T_0 baseline. While gelatin coating was no better than non-coated condition, collagen partially supported tumour proliferation, tumour area and phosphorylation of ERK1/2 but not cell viability. Interestingly, Matrigel, a widely used murine tumour-derived matrix, resulted in increased cell viability, tumour area and phospho-ERK but not in proliferation (Fig. 2j and Supplementary Fig. 4b). In contrast, explants cultured in matched TMPs retained tumour morphology, viability, proliferation and phospho-ERK1/2 status similar to the T_0 baseline parameters. This observation is consistent with recent reports that highlight context-dependent stromal-epithelial interaction as a critical requirement of tumour cell survival and maintenance[10].

Autologous ligands maintain the signalling and phenotypes. A heterogenous tumour microenvironment represents a diverse network of oncogenic signalling pathways, which are activated in both ligand-dependent and -independent manner and can spatiotemporally and dynamically cross-talk[26–30]. Indeed, a reverse phase phosphoprotein array (RPPA)-based profiling of key receptor tyrosine kinases (RTKs) and their nodal proteins in the tumour biopsies revealed a heterogeneity in the baseline activation levels of these receptors and downstream signals

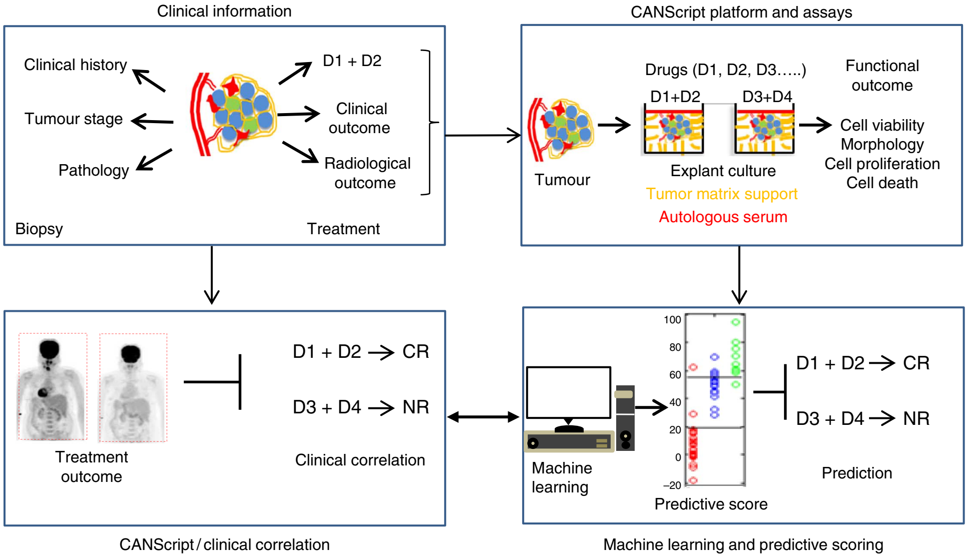


Figure 1 | Schematic showing the development and validation of the CANScript technology. Four critical modules were integrated in generating and validating the CANScript platform. The first module involved collecting tumour core or surgical biopsy with tumour staging and pathology information besides clinical/treatment history. In the second module, tumour biopsy was rapidly processed into thin explants. Tumour biopsies were also used to generate either *in vivo* implants in mice, or processed for isolation and analysis of tumour matrix, which was used to develop the TMP cocktail. The explants were cultured in tumour- and grade-matched TMP and AS and incubated with selected drug regimens. While multiple drug regimens can be used, the one used by the oncologist for the patient was always included in the tumour explant culture. Functional outcome of treatment in terms of cell viability, pathological and morphological analysis, cell proliferation and cell death was quantified. In module three, these quantitative scores from the explants were aggregated using a machine learning algorithm to assign a final score, which helped rank the outcomes as CR,PR or NR. The learning algorithm was trained on data from 109 patients. In the final module, these predictions were tested against clinical outcomes from 55 new patients to validate the approach. D1, D2, D3 and D4 indicate different drug regimens.

(Fig. 3a and Supplementary Fig. 5a,b). This led us to hypothesize that a balanced induction of these receptors using their original ligands in an individualized setting is critical to mimic the baseline networks of the parent tumour *ex vivo*.

Autocrine–paracrine loops of growth factors enriched in patient sera contribute to the activation of signalling networks and survival cascades in cancer cells[10,31,32]. As the second step towards fabricating the CANScript platform, we therefore studied the functional attributes of AS. As shown in Fig. 3b a number of growth factors (represented by EGF, hepatocyte growth factor (HGF), vascular endothelial growth factor (VEGF) and macrophage colony-stimulating factor (MCSF)) were found to be within clinically detectable ranges in patient sera. The variability in the levels of these growth factors that exists between individuals further underlined the importance of using the complete AS for a balanced induction of signal transduction pathways as opposed to an artificial combination of growth factors. We first performed a dose–response analysis, where increasing concentrations of AS was used in combination with a reducing percentage of fetal bovine serum (FBS) in the culture for 72 h. A concentration dependent increase in cell proliferation in the explants was observed while supplementing the system with AS that attained the peak at 2% (Fig. 3c). Concomitantly, 2% AS also mimicked the native state (morphology and proliferation) of tumours at T_0 baseline (Fig. 3d). The decline above this concentration is consistent with earlier observations with growth factor ligands and possibly arises due to the downregulation of targets[33]. In addition, 2% AS (+8% FBS) resulted in significant increase in ATP utilization and cell proliferation compared with 10% FBS or recombinant EGF alone (Fig. 3e,f). Furthermore, compared with exogenous EGF controls, the addition of AS significantly preserved the major signalling networks as measured by phosphorylation of EGFR, Met and downstream target, ERK1/2. It is interesting to note that $1\,\mathrm{ng\,ml^{-1}}$ concentration of EGF predominantly activates EGFR pathway alone. In contrast, 2% AS showed the capacity to activate both EGFR and HGFR/Met pathways along with downstream ERK1/2 comparable to the T_0 baseline, consistent with the balanced effect of patient-derived ligands in its natural milieu. The enhanced response to AS was reduced to $T_{72\,\mathrm{h}}$ baseline (that is, no AS control) using neutralizing antibodies to EGFR, which is consistent with the aberrant activation of EGFR pathway in a majority of HNSCC and other cancers of epithelial origin[27,29]. However, the neutralizing antibody failed to fully abrogate the proliferation below the level of $T_{72\,\mathrm{h}}$ control, suggesting that despite the predominant role of EGFR in some individual tumours additional constitutive mechanisms exist that might contribute to minimal maintenance of these tumours (Fig. 3g–i). It is obvious that the survival of tumour is not a consequence of dependency on single pathway lineage or network.

To further validate the contribution of autologous sera in personalizing the explant culture, we compared the individual effects of heterologous/allogenic sera (HS) obtained from treatment naïve patients (age, sex and cancer-type matched) with AS and recombinant EGF. As shown in Fig. 3j,k, while EGF resulted in the maximum effect in inducing EGFR phosphorylation, 2% AS efficiently maintained both EGFR and Met phosphorylation. In contrast, 2% HS, while exerting a greater effect than no ligand ($T_{72\,\mathrm{h}}$) control, was significantly inferior to AS. Similar pattern was observed for Ki-67 (Supplementary Fig. 6a). Taken together, these results indicate that presenting the entire repertoire of growth-promoting ligands by using AS is critical to fully capture the parental activation status of important receptors in the personalized explant setting. Indeed, RPPA array-based analysis of the parent HNSCC tumours (T_0 baseline) showed that a bulk of the proteins in RTK cascades that were upregulated are largely conserved in the tumour explants cultured in 2% AS (Fig. 3l and Supplementary Fig. 6b).

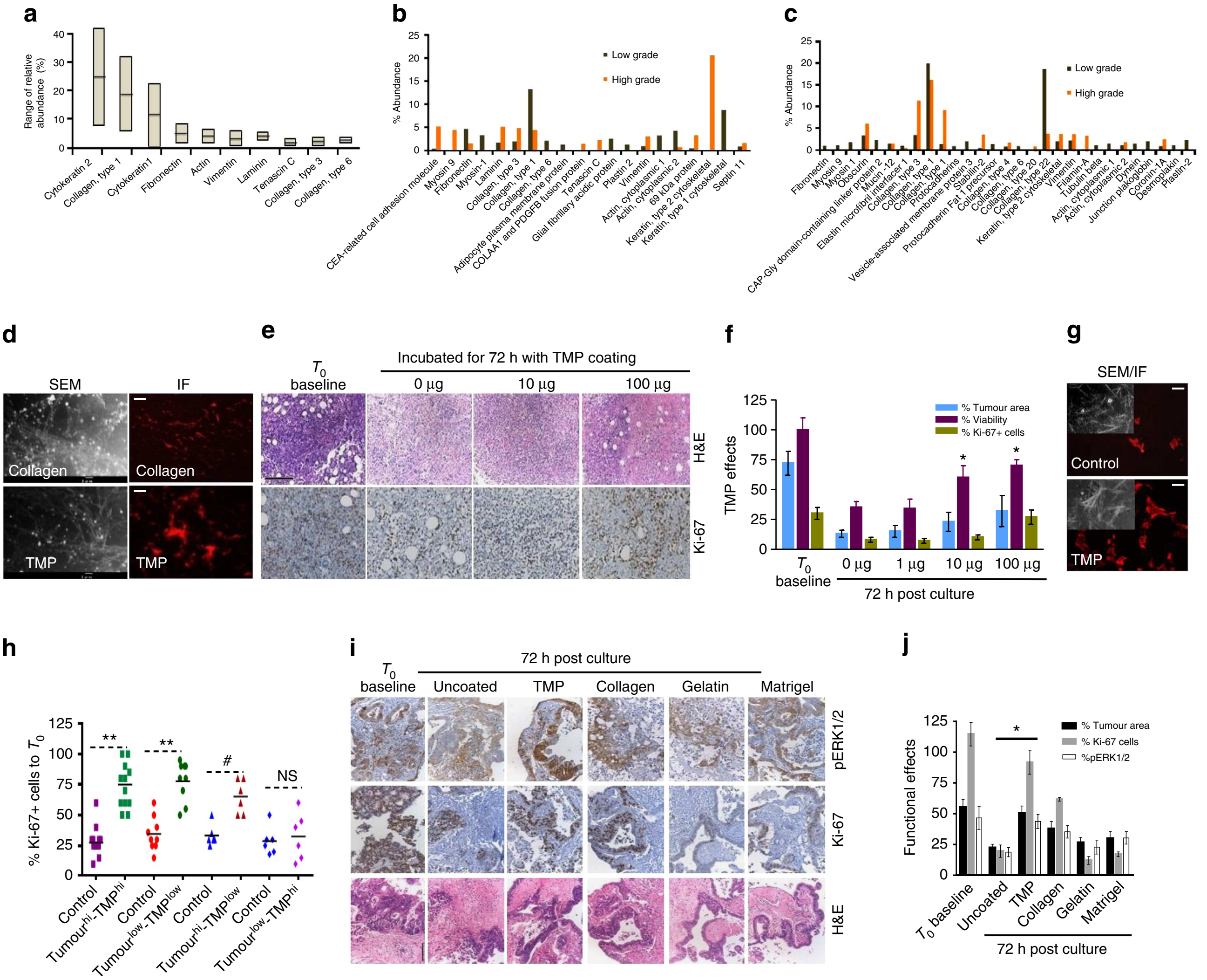


Figure 2 | TMPs are critical for conserving primary tumour characteristics in explants. (**a**) The average composition and abundance range of key components of TMP. Abundance range was measured based on area under the peak using Pearson's correlation for clustering of protein features. The line within each notch box represents the median, and the lower and upper boundaries of the box indicate first and third quartiles, respectively ($n = 24$), of some of the key TMPs isolated from (**b**) HNSCC ($n = 12$) and (**c**) CRC tumours ($n = 12$) (**d**) scanning electron microscopy (SEM) images of plastic surface precoated with Collagen-I (top) or TMP cocktails (bottom). Scale bars, 1 µm. Adherence of the component proteins to the surface and their ability to form networks is shown following immunofluorescence (IF) staining using human Collagen-I antibody. Adherence was measured by detecting specific fluorescence signal in coated area contrasting to uncoated area of the same surface. Scale bars, 200 µm (right). (**e**) HNSCC explants were cultured for 72 h in plates coated with different concentrations of TMP as indicated. Maintenance of overall intratumoral heterogeneity and integrity was determined by hematoxylin and eosin staining (H&E; top) and tumour cell proliferation by Ki-67 staining (bottom). Scale bar, 100 µm. (**f**) Tumours from HNSCC patients were sliced. Explants were cultured for 72 h in plates coated with different concentrations of TMP as indicated. Percent tumour area, cell viability and Ki-67 + cells per field was measured (mean ± s.d.). $^{*}P \leq 0.05$ compared with uncoated control using paired *t*-test. Data represent one of the five independent experiments performed in triplicates. (**g**) HNSCC tumour slices cultured for 72 h with or without TMP were subjected to extraction of native extracellular matrix (ECM). Preservation of ECM 72 h post culture was determined by IF staining of extracted ECM parallel to SEM imaging (inset). (**h**) HNSCC tumours of high and low grades were sectioned cultured for 72 h in plates coated with matched and unmatched TMP (high and low grade) Scatter plot indicates the effects of grade-matched and unmatched TMP on retaining the proliferation profile. Percent Ki-67–positive cells from HNSCC explants were calculated at the end of 72 h based on T_0 score. $^{**}P < 0.0002$, $^{\#}P < 0.05$ for the high-grade tumours cultured in presence of low-grade TMP by paired Student's *t*-test. NS, not significant ($n = 12$). (**i**) Representative images show the effects of CRC-specific TMP and other coating materials on pERK status (top), proliferation (middle) and morphology (bottom) of tumour explants. Scale bar, 100 µm. (**j**) Quantitative analysis of TMP on proliferation, tumour area and pERK status in CRC explants. $^{**}P < 0.01$ compared with T72 control (analysis of variance, $n = 8$).

Reconstructing a tumour ecosystem. As the final step towards constructing the CANScript tumour ecosystem, both conditions (that is, TMP and AS) were contextually integrated in the explant system. Immunohistochemistry (IHC) labelling was used to evaluate a number of static and dynamic phenotypic markers associated with functional heterogeneity of tumour microenvironment. Profiling for CD68 (marker for immune component)[34], VEGFR (marker for angiogenesis), CD34 (marker for angiogenesis and progenitors)[35], E-Cadherin and Vimentin (markers for epithelial mesenchymal transition (EMT)) revealed that the combination of AS and TMP conserved the parental (T_0) phenotypes better than $T_{72\,h}$ control or EGF + TMP(Fig. 4a,b). Similar effects of AS and TMP were also observed for EMT-specific markers (Fig. 4c). Furthermore, cell viability, proliferation

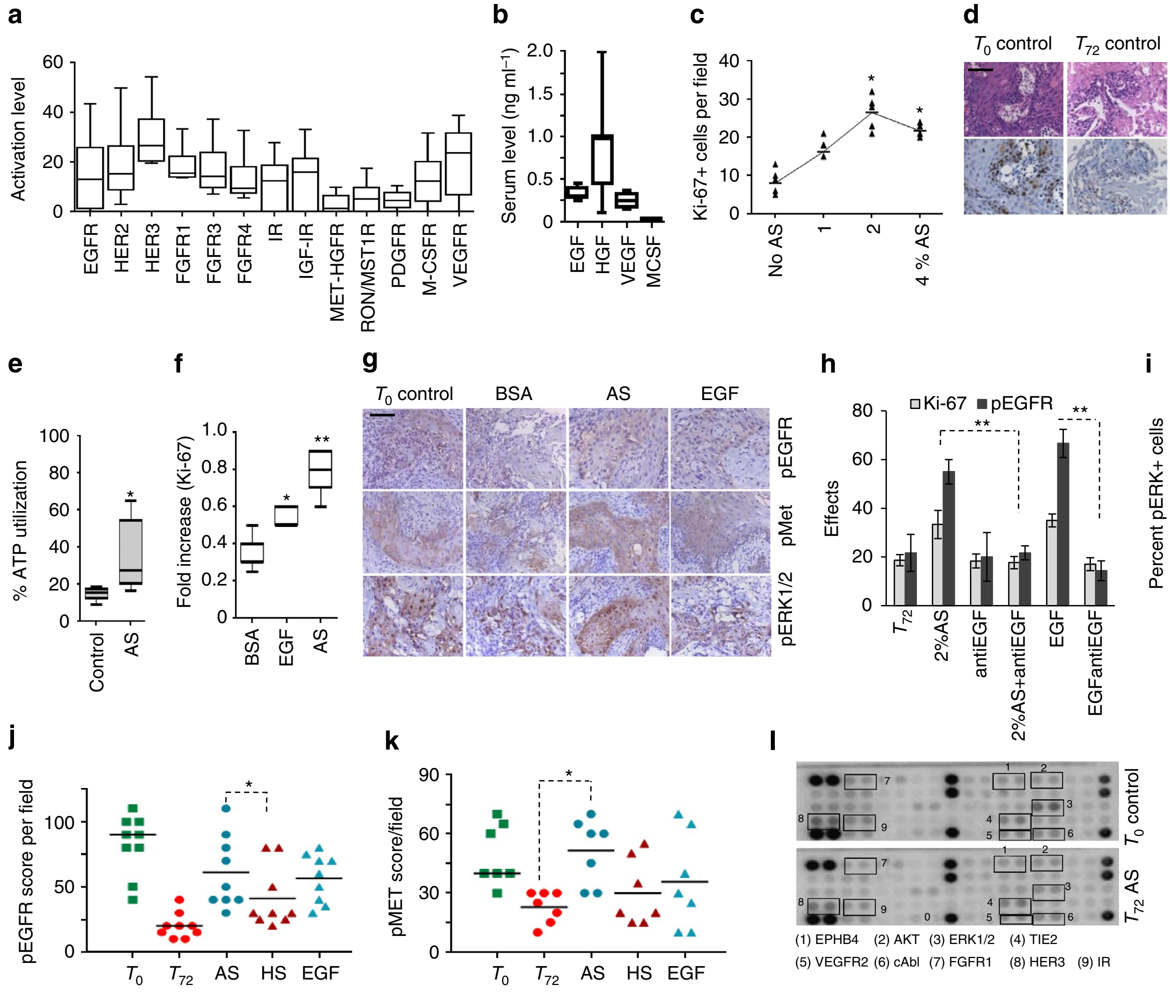


Figure 3 | Autologous serum conserves the integrity of tumour explants. (**a**) Activation levels of major RTKs by RPPA profiling of patient tumours ($n=5$). Quantification of RTK activation was performed by measuring the signal intensity of individual analytes normalized to negative control. The line within each notch box represents the median, and the lower and upper boundaries of the box indicate first and third quartiles, respectively. Error bars (whiskers) represent the interquartile range. (**b**) Serum growth factor (EGF, HGF, VEGF and MCSF) profiles of HNSCC patients by ELISA ($n=8$). Horizontal line represents median and error bars indicate the interquartile range. (**c**) The dose-dependent effect of AS in HNSCC was measured by Ki-67. $^*P<0.001$ by one way analysis of variance (ANOVA; $n=9$) compared with no AS control. (**d**) Tumour slices cultured in the presence or absence of autologous ligands for 72 h and stained with hematoxylin and eosin stain (top) and Ki-67 (bottom). Scale bar, 50 µm. (**e**) Box plot shows ATP utilization ($^*P<0.05$ by t-test, $n=6$) at 72 h in the presence of AS. (**f**) Box plot shows fold increase in Ki-67-positive cells cultured with AS and EGF ($^*P<0.05$ and $^{**}P<0.01$, t-test, $n=8$). (**g**) Impact of AS on the balanced activation of different signalling receptors close to T_0 baseline. Tumour explants were treated with 2% AS, 1 ng ml^{-1} per h EGF or 8% FBS + 2% BSA (BSA Control) for 72 h. Tumours were stained for pEGFR (top), pMet (middle) and pERK1/2 (bottom). Scale bar, 100 µm. (**h**) Graph shows quantification of effects of different treatments on the proliferation and phosphorylated EGFR status in the explants. HNSCC samples were cultured in the presence of 2% AS or EGF up to 6 h for pEGFR and 72 h for detecting proliferation. Appropriate controls (no serum, no antibody and antibody alone) were included. Anti-EGF was added 1 h before stimulation. The effect was assessed by pEGFR and Ki-67 staining. All data ($n=8$) are represented as mean ± s.d. $^{**}P<0.01$ by t-test. (**i**) Box plot shows percent pERK positivity ($n=8$) $^{**}P<0.01$ (by analysis of variance). Horizontal line represents median and error bars indicate the interquartile range. Graphs shows comparison of the capacity of AS, HS and EGF in activating (**j**) EGFR ($^*P<0.02$ by t-test, $n=9$) and (**k**) in maintaining phospho-Met expression ($^*P<0.0001$ by t-test, $n=7$). (**l**) Global RTK profiles of cultured HNSCC tumour explants and corresponding T_0 baseline was compared following stimulation with 2% AS for 72 h. Total cell lysates were applied to array slides precoated with different antibodies against RTK pathways. Signal was detected by chemiluminiscence method.

index, and metabolic state of the explants in the CANScript tumour ecosystem was similar to native (parent/T_0 baseline) tumour and significantly enhanced compared with control explants cultured without AS and TMP, or with either AS or TMP alone. The pattern of augmentation of Ki-67 upon AS + TMP was found to be consistent and significant (Fig. 4d,e). Together these results indicate that the native tumour-stromal micro-architecture and phenotypic features were largely conserved in the CANScript tumour ecosystem compared with the culture conditions with only TMP or AS or EGF-supplemented TMP. Next, we used microarray profiling to compare the transcriptome of primary tumours at baseline (T_0) and serially sectioned tumour explants cultured under different conditions. Indeed, a high degree of conserved global transcriptomic profile consistent with the primary tumour was observed only in the case of the CANScript platform that integrated both TMP and AS, while supplementing the explant cultures with either AS or TMP(+ EGF) alone resulted in distinct transcriptomic signatures (Fig. 4f,g). Concurrent to the phenotypic expression as shown in Fig. 4a,b, conservation of

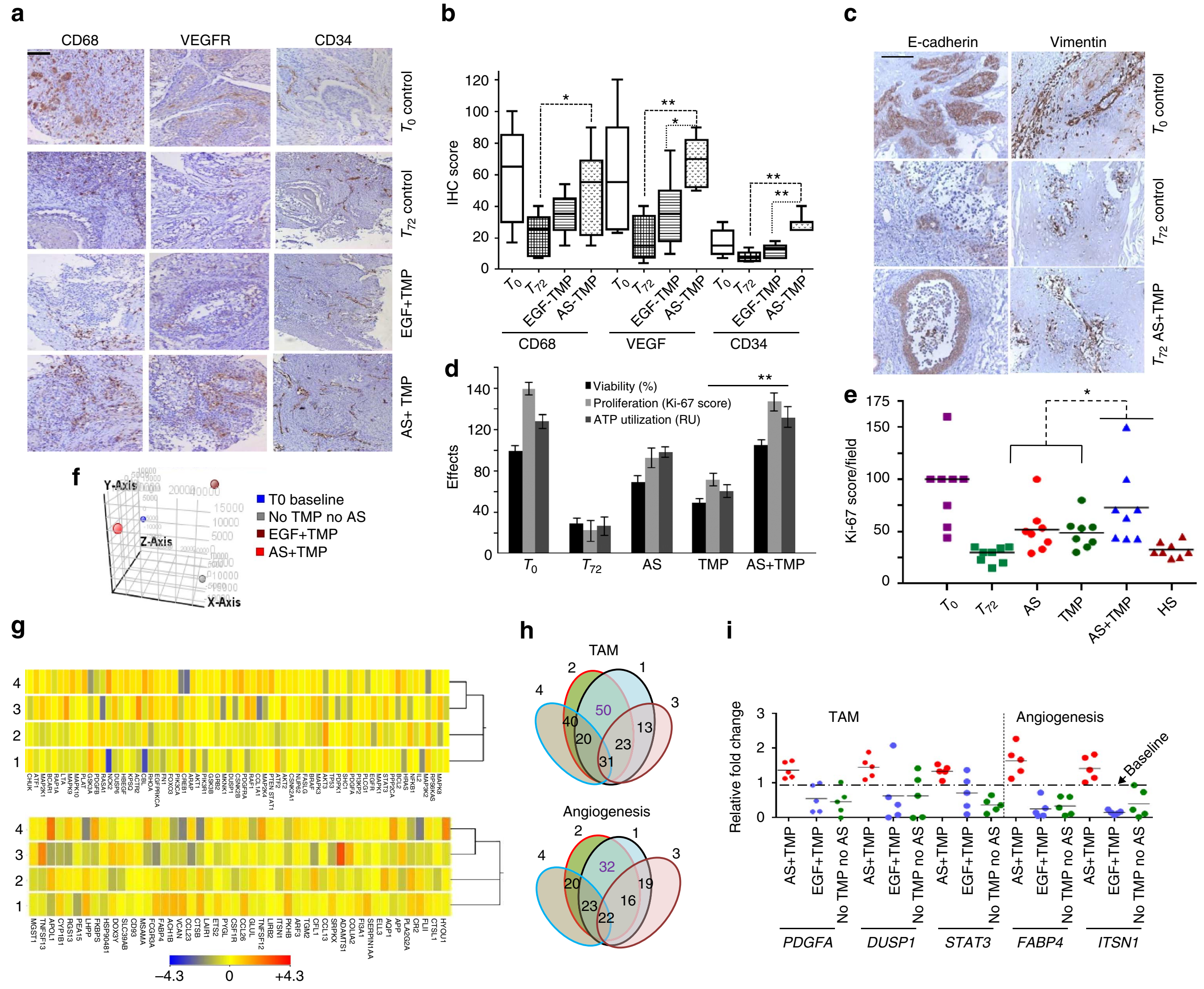
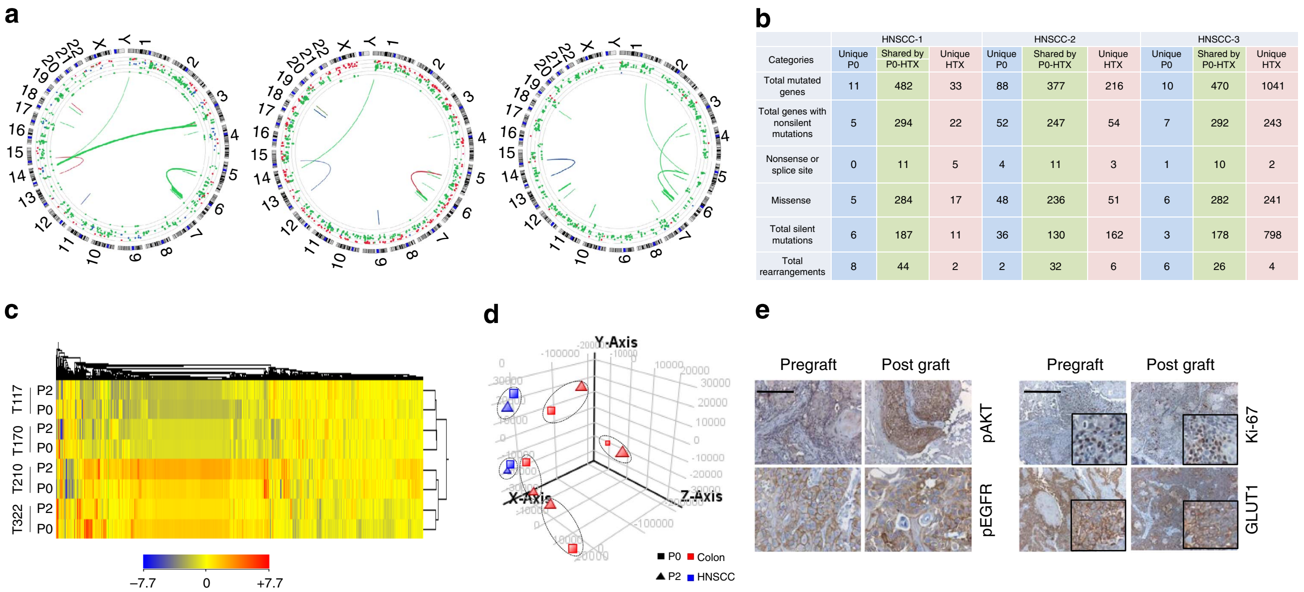


Figure 4 | Integration of both TMP and AS in the CANScripts maintains the tumour ecosystem. (**a**) Representative IHC images show the effect of AS and matchedTMP on the phenotypic stability of tumour explants 72 h post culture. Tumour sections were stained for CD68, VEGFR and CD34. Scale bar, 100 µm. (**b**) Quantitative IHC based box plot indicates CD68, VEGFR and CD34-positive cells in the explants maintained under different conditions. Horizontal line represents median and error bars indicate the interquartile range. $^{*}P<0.05$ and $^{**}P<0.001$, respectively (paired *t*-test, $n=8$). (**c**) Representative IHC images show EMT related markers of tumour microenvironment in the CANScript explants. Scale bar, 100 µm. (**d**) Graph shows the combined effects of AS and TMP on the functional integrity of the explants. Tumour sections were cultured for 72 h. Number of Ki-67-positive cells were counted and plotted along with percent viability and ATP utilization per section in triplicates (mean ± s.d.). $^{**}P<0.01$ (by analysis of variance). (**e**) The combined effects of AS and TMP on the functional integrity of explants are represented as scatter plot ($n=8$). Number of Ki-67-positive cells were counted and plotted. HS was run as a control. $^{*}P<0.05$ (by paired *t*-test). (**f**) 3D-PCA plot showing global gene expression patterns between different culture conditions (that is, no AS and no TMP, EGF + TMP, AS + TMP and T_0 baseline) obtained from HNSCC tumour explants after 12 h. After initial normalization of data analysis was performed compared with baseline. (**g**) Heat map analysis of the microarray data showing the genes related to TAM (top) and angiogenesis (bottom). Tumours explants were cultured in TMP-coated plates with AS (AS + TMP, lane 2) or EGF (EGF + TMP, lane 3) or in uncoated plates without AS (No TMP and no AS Control, lane 4) and transcriptomic pattern was compared with base line tumour (lane 1). Heat map scale indicates the expression range. Clustering of genes was performed by k- means algorithm. Distance was measured by Euclidean distance metric. (**h**) Venn diagram showing number of overlapped genes related to TAM and angiogenesis between the three culture conditions. (**i**) Validation of microarray gene signature by qRT–PCR for TAM (left) and angiogenesis (right); selected genes from each signature was run in triplicates (technical replicates) normalized to baseline expression (biological replicates) and compared between conditions as indicated in the scatter plot ($n=5$).

stromal gene expression signatures, specifically linked to tumour-associated macrophages and angiogenesis were also observed (Fig. 4g). To further confirm these results, a selected panel of genes relevant to TAM (that is, *PDGFA*, *DUSP1* and *STAT3*) and angiogenesis (that is, *FABP4* and *ITSN1*) signatures (Supplementary Table 2) was analysed under different conditions using qRT–PCR. As shown in Fig. 4i, the expression of these markers were conserved only under AS + TMP condition but not when either is absent. In addition, expression of tumour-associated key cytokine/chemokines, such as interleukin-6, interleukin- 8 and CXCR-4, matrix degrading enzyme matrix metallopeptidase 9 (MMP-9) and cancer stem cell markers like CD44 and ALDH1 observed in the parent HNSCC tumours were also preserved in the CANScript tumour ecosystem (Supplementary Fig. 7a–c). It is important to note that unlike common synthetic organotypic inserts, the CANScript platform exhibited enhanced preservation of native tumour morphology and proliferation status (Supplementary Fig. 7d). Taken together, these results suggest that a number of phenotypic markers characteristic of EMT, immune cells and cytokines as well as cancer stem cell phenotypes are more consistently and collectively better conserved in this platform compared with culture conditions with either TMP or AS or EGF-supplemented TMP.

CANScript predicts response to cytotoxic and targeted drugs. The conservation of patient tumour heterogeneity in the CANScript tumour ecosystem prompted us to explore the possibilities of using this as a preclinical tool to predict anticancer drug response. To assess this, we first compared drug response in human tumour-derived xenotransplants (HTX) and in matched CANScripts explants (constructed from passage 2, that is, P2-HTX). Primary HNSCC tissues were propagated in severe combined immunodeficiency mice up to second passage (P2-HTX). Since response and resistance to a particular drug combination can be intrinsically controlled by deregulation at the genetic and epigenetic levels,[11,36–38] we first mapped the degree to which a xenotransplanted tumour (at P2) conserves the descriptors of the primary tumour. Interestingly, exome data from three different primary samples, HNSCC-1, HNSCC-2 and HNSCC-3, and their matched P2-HTX, showed that while the overall events of mutation and translocation of primary tumours were largely preserved when passaged in immunocompromised mice, there were mutations that were unique to original parental P0 and P2-HTX, respectively (Fig. 5a,b and Supplementary Table 3). However, global transcriptome pattern showed a good association between P0 and matched HTXs (Fig. 5c,d). Furthermore, histopathological characterization of P2-HTX revealed that the HTX successfully conserved key morphological and molecular characteristics of original parental (P0) tumours, including the expression of proliferation marker (Ki-67), glucose transport (GLUT1), phospho-EGFR and phospho-AKT (Fig. 5e). Subsequently, these extensively characterized P2-HTX were used as surrogates for initial functional validation of the CANScripts. HTX-derived CANScripts were concurrently treated with the clinically approved cytotoxic drug regimen of docetaxel, cisplatin and 5-fluorouracil (TPF), segregated into two groups of responders and non-responders based on viability, ATP utilization, proliferation status and loss of tumour area/nuclear fragmentation (Fig. 6a–c and Supplementary Fig. 8a–c). Interestingly, we noticed an excellent correlation between the outcomes in the CANScript platform and the response to chemotherapy in the HTX studies. For example, cases predicted as responders using the CANScript tumour ecosystem mapped to a significant inhibition of tumour growth when the animals were treated at maximum tolerated dose daily for up to 21 days (Fig. 6d). The results were further validated at the molecular level by determining the end point changes in mean tumour area/nuclear size in sections, Ki-67 and concomitant drug-induced increase in apoptotic cells by staining with TUNEL method (Fig. 6e,f). Similarly, cases predicted as non-responders using the CANScript tumour ecosystem did not show any effect in HTX system, as defined by the lack of any distinctions in Ki-67 and active Caspase-3 expression between the treated and untreated groups (Supplementary Fig. 8a–e).

The *ex vivo* to *in vivo* correlation in response to a general cytotoxic drug combination that we observed in HNSCC samples



	HNSCC-1			HNSCC-2			HNSCC-3		
Categories	Unique P0	Shared by P0-HTX	Unique HTX	Unique P0	Shared by P0-HTX	Unique HTX	Unique P0	Shared by P0-HTX	Unique HTX
Total mutated genes	11	482	33	88	377	216	10	470	1041
Total genes with nonsilent mutations	5	294	22	52	247	54	7	292	243
Nonsense or splice site	0	11	5	4	11	3	1	10	2
Missense	5	284	17	48	236	51	6	282	241
Total silent mutations	6	187	11	36	130	162	3	178	798
Total rearrangements	8	44	2	2	32	6	6	26	4

Figure 5 | Comparative profiling of patient tumours and their corresponding xenografts. (**a**) Mutational and translocation spectrum obtained from the whole-exome sequence analysis (Agilent 44 Mb, × 50 coverage) of HNSCC patient tumours and also their corresponding xenograft tumour (passage no.2) tissues. In representative Circos plots each dot represents a mutation and line represents translocation. Blue colour stands for original tumours, green colour stands for overlapped original and xenografted tumours and red colour denotes the events in xenograft only. (**b**) Exome data table illustrates that the HNSCC tumours when passaged in mice (P2) retain majority of genomic characteristics of the baseline tumour. (**c**) Unsupervised 2D hierarchical clustering performed on colon samples shows that expressed genes in primary tumour (P0) are associated with HTX and stably expressed when passaged in mice (P2). Scale represents expression ranges (fold normalized changes, bottom). (**d**) 3D-PCA plots generated by GeneSpring GX software to show the clustering of samples of same origin and serial passage. The plot shows six distinct clusters comprising of four pairs of colon carcinoma and two pairs of HNSCC samples. (**e**) Representative IHC images of early passages of HTXs and matched primary tumours. Primary HNSCC tissues (pregrafts) were propagated up to passage 2 (post grafts) in SCID mice. Tumours from both pregrafts and post grafts were stained with anti-Ki-67 (right, top), antibodies against pAKT (left, top) and pEGFR (left, bottom). Scale bars indicate 100 μm for Ki-67 and pAKT and 50 μm for pEGFR. Tumours from both pre- and post grafts were also stained for expression of GLUT1 (right, bottom) using specific antibodies. Scale bars, 100 and 50 μm (inset).

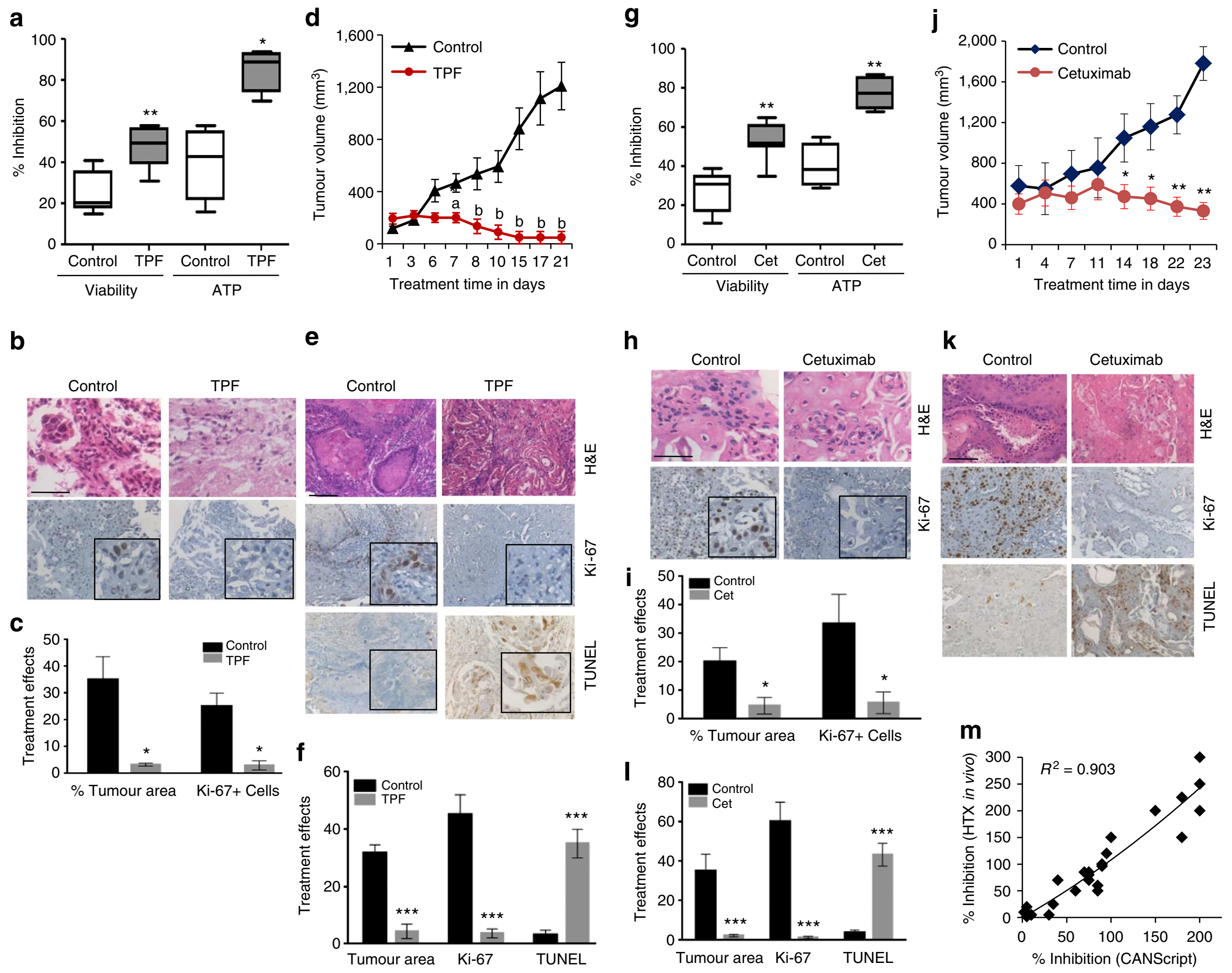


Figure 6 | Application of CANScripts in functional *ex vivo* and *in vivo* correlation. (**a**) Antitumour effects of TPF chemotherapy regimen on HNSCC tumour explant culture. Box plots show inhibition of viability and ATP utilization in TPF-treated tumours as calculated using T_0 values as baseline. Horizontal line represents median and error bars indicate the interquartile ranges. *$P<0.01$ and **$P<0.001$ compared with vehicle-treated control ($n=20$) by *t*-test. (**b**) Representative images show corresponding IHC profile. Tumour sections were stained with H&E (hematoxylin and eosin stain; top) and Ki-67 (bottom). Scale bar, 50 µm. (**c**) Graph shows percent tumour area and Ki-67-positive cells from vehicle- and TPF- treated explants. Data shown are mean ± s.d. *$P<0.05$ by *t*-test ($n=3$). (**d**) Graph shows *in vivo* tumour growth inhibition in xenografts following TPF treatment till 21 days of treatment. Data shown are mean tumour volume ± s.d. with six mice per group ($^{a}P<0.01$ and $^{b}P<0.001$ by analysis of variance, ANOVA) compared with vehicle control. (**e**) Representative IHC images of pre- and post-treatment tumours stained with H&E stain (top), Ki-67 antibodies (middle) and TdT-mediated dUTP nick end labeling (TUNEL; bottom). Scale bar, 100 µm. (**f**) Graph shows quantitative analysis of IHC for tumour area, Ki-67+ cells and TUNEL+ cells from control and TPF treatment. Data are mean ± s.d. of six mice per group. ***$P<0.001$ compared with vehicle control (paired t-test). (**g**) Cetuximab efficacy in HSNCC CANScripts. Explants were treated with DMSO or cetuximab. Box plots represent percent inhibition of cell viability and ATP utilization ($n=20$). Horizontal line represents median and error bars indicate the interquartile range. **$P<0.001$ by *t*-test. (**h**) Representative images of tumour sections labelled with H&E for morphology (top panel) and anti-Ki-67 antibodies for proliferation (bottom panel). Scale bar, 50 µm. (**i**) Graph shows quantification of effects of treatment in the CANScripts in terms of percent tumour area and Ki-67-positive cells from control and cetuximab treatment. Data represented as mean ± s.d. ($n=3$). *$P<0.001$ by *t*-test. (**j**) Graph shows tumour growth inhibition in cetuximab-treated mice. Data are mean tumour volume ± s.d. $n=10$. *$P<0.02$ and **$P<0.001$ (ANOVA) versus vehicle control. (**k**) Representative IHC reveals changes in H&E (top), Ki-67 (middle) and TUNEL (bottom). Scale bar, 100 µm. (**l**) Quantitative analysis of tumour area, Ki-67 and TUNEL from control and cetuximab-treated mice. All data indicate mean ± s.d. ***$P<0.001$ (*t*-test). $n=6$. (**m**) The correlation observed between the efficacy data from TE explants and *in vivo* studies. R^2 was calculated using Spearman's correlation coefficient method.

encouraged us to further validate the predictive ability of CANScript for targeted therapeutics. For this purpose we used HTXs generated from HNSCCs harbouring wild-type or mutant *KRAS*. Consistent with the results observed earlier with cytotoxics, a positive response in the CANScript explants with cetuximab (Fig. 6g–i) was mirrored by tumour inhibition *in vivo* (Fig. 6j–l).The functional outcome was correlated with a decrease in Ki-67 positivity, increased TUNEL and a reduction in phospho-EGFR levels in both the CANScript explants and *in vivo* (Fig. 6h,i,k,l and Supplementary Fig. 9a,b). We next tested the effect of cetuximab in HTX and CANScript explants generated from CRCs. As shown in Supplementary Fig. 10a–g, an inhibitory outcome in the CANScript explants correlated with a significant tumour growth inhibition *in vivo*, while in the absence of an inhibitory effect in the tumour ecosystem (TE), minimal tumour growth inhibition was evident *in vivo* (Supplementary

Fig. 10h–n). In the cetuximab-treated groups, responders showed a decrease in Ki-67 and phospho-ERK levels and increase in cleaved caspase-3 expression (Supplementary Fig. 10b,c,e–g). This was not evident in the non-responders (Supplementary Fig. 10i,j,l–n). Collectively, we observed a linear correlation ($R^2 = 0.903$, $n = 26$, by Spearman's correlation coefficient) between CANScript explants outcomes and *in vivo* HTX responses (Fig. 6m).

CANScript as a tool to predict treatment outcome in patients. The concordance in outcome between HTX *in vivo* and corresponding CANScript studies suggested the possibility of using the latter for predicting the treatment outcome in patients. The CANScript explants were generated from biopsies of CRC and HNSCC tumours from 109 patients and were incubated with the same drug combination as that administered to the patient, that is, docetaxel, cisplatin and 5-fluoro uracil (5-Fu) for the 70 HNSCC patients and cetuximab + FOLFIRI for the 39 CRC patients. The functional read-outs from these CANScripts, quantified in terms of viability, histopathology, proliferation and apoptosis, together with the observed clinical response in the matched patients, classified as progressive disease/non-response (NR), partial response (PR) or complete response (CR) based on PERCIST guidelines (Fig. 7a), were then used as the training set for a novel machine learning algorithm. In this algorithm, as the first step, we classified patients as simply responders or non-responders, with a focus on ensuring high sensitivity (true positive rate). This was formulated by maximizing the partial area under the receiver operating characteristic (ROC) curve (partial area under the curve (AUC)) up to an acceptable false positive range (Fig. 7b). To this end, PR and CR were grouped together into a responder (R) category and a linear prediction model was learned using SVMpAUC, a recently proposed structural support vector machine algorithm for optimizing partial AUC. The learned model was designed to maximize partial AUC while achieving at least 75% specificity (that is, at most 25% false positive rate) on the training set, and assigned coefficients of 0.2977, 0.5562, 0.0073 and 0.1388 to the viability, histology, proliferation and apoptosis read-outs, respectively, together with a threshold of 19.1 (that is, cases assigned a weighted score > 19.1 by the learned model were predicted to be responders).The model achieved 96.77% sensitivity on the training set (Fig. 7c). We then tested the learned algorithm on a new test group of 55 patients, consisting of 42 HNSCC and 13 CRC patients treated with the same drugs as above, where the model achieved 91.67% specificity and 100% sensitivity (Fig. 7d). In particular, no potential responders (PR or CR patients) in the test set were predicted as NR (Fig. 7d).

In the next step, the learned model was refined to classify the predicted responders into partial and complete responders (PR and CR), by selecting a threshold that maximized PR versus CR prediction accuracy on the training set. Following this, scores between 19.1 and 55.14 were classified as PR, and those > 55.14 as CR. As can be seen in Fig. 7e,f, the coefficients assigned to the four read-outs by the SVMpAUC-learned model, together with the above thresholds, resulted in predictions that were significantly better than what could be achieved by predicting using any one of the functional read-outs alone. Confusion matrices summarizing predictions in each category on both the training and test sets are shown in Fig. 7g,h; break-ups among HNSCC and CRC cases are shown in Fig. 7i–l. The resulting predictions had 87.27% accuracy on the test set (Fig. 7h). In particular, among the 55 test cases, there were only seven prediction errors: four PRs were predicted as CR; one CR was predicted as PR; one NR was predicted as PR; and one NR was predicted as CR (Fig. 7h). This is the benefit of using the SVMpAUC machine learning algorithm, which explicitly encourages high sensitivity in the learned model (indeed, a standard support vector ordinal regression algorithm which directly classified the patients into one of the three categories yielded a lower accuracy of 81.82% on the test set, making a total of 10 prediction errors on the 55 cases, which included 1 PR case predicted as NR). Again, it is worth emphasizing that these errors using the SVMpAUC machine learning algorithm were all 'benign', in that no potential responder (PR or CR) was predicted as a NR. While such 'benign' errors do mean unwarranted drug use that can result in potential side effects, it also means that no patient who would respond to chemotherapy is denied a drug based on a false prediction. Indeed, current clinical practice also assumes this principle, where the error rate is significantly higher as seen in our study. For example, as shown in Fig. 7m, biomarker analysis selected all 13 CRC patients in the test set, all of whom were positive for wild-type *KRAS*, to receive cetuximab. However, as can be seen, only 3 of these 13 wild-type *KRAS* patients actually responded to the drug (1 exhibited CR and 2 exhibited PR), while the remaining 10 presented with progressive disease. Interestingly, the CANScript platform predicted two CRs, two PRs and nine NRs, with only one actual NR case being wrongly predicted as CR. As shown in Fig. 7n, based on standard practice, all 42 HNSCC patients in the test set received TPF. However, 14 of these patients did not respond to the drug combination. The CANScript platform could identify 13 of these as NRs. Again, importantly, all patients predicted by the platform as NRs were indeed NRs. It should be noted that 13 and 42 are small sample sizes, and that larger-scale studies are needed in the future to establish similar results on larger sample sizes; however based on the observed improvements over the standard/biomarker-based approach, we anticipate that the CANScript platform can emerge as a powerful strategy for predicting chemotherapy outcomes.

Discussion

While biomarker driven personalized cancer therapy has emerged as a powerful concept, the mere presence of a biomarker in a cancer cell may not translate into clinical efficacy[1,6,7,39]. This arises from heterogeneity, where multiple genetic, epigenetic and phenotypic alterations along with immune and metabolic changes represent a complex state of the neoplastic transformation[40]. Indeed, in the current study, of the 52 patients who received cetuximab based on wild-type *KRAS* status, only 1 exhibited CR and 12 exhibited PR, and the remaining 39 presented with progressive disease. While the use of more than one biomarkers, for example, the use of wild-type *KRAS* and *BRAF* to select patients eligible for cetuximab[41] is the emerging trend, the ability to predict chemotherapy outcomes accurately at an early time point still remains a holy grail in the management of cancer. Here we have demonstrated the development of a novel technology platform that integrates a comprehensive explant culture with a machine learning algorithm to better predict chemotherapy outcomes. As we have demonstrated in this study, the CANScript platform is versatile in its ability to predict the outcomes of both cytotoxic chemotherapy regimens and targeted therapeutics.

A key attribute of the CANScript platform is its ability to capture the intratumoral heterogeneity to a greater degree than achieved by biomarker-based selection of cancer cells. Cancer stem cells, stromal cells such as intra and peritumoral immune cells, and vascular components can further add to the heterogeneity and contribute towards tumour survival, progression and metastasis[15,17,34,42], suggesting that an explant culture that globally conserves these distinct cellular components in their

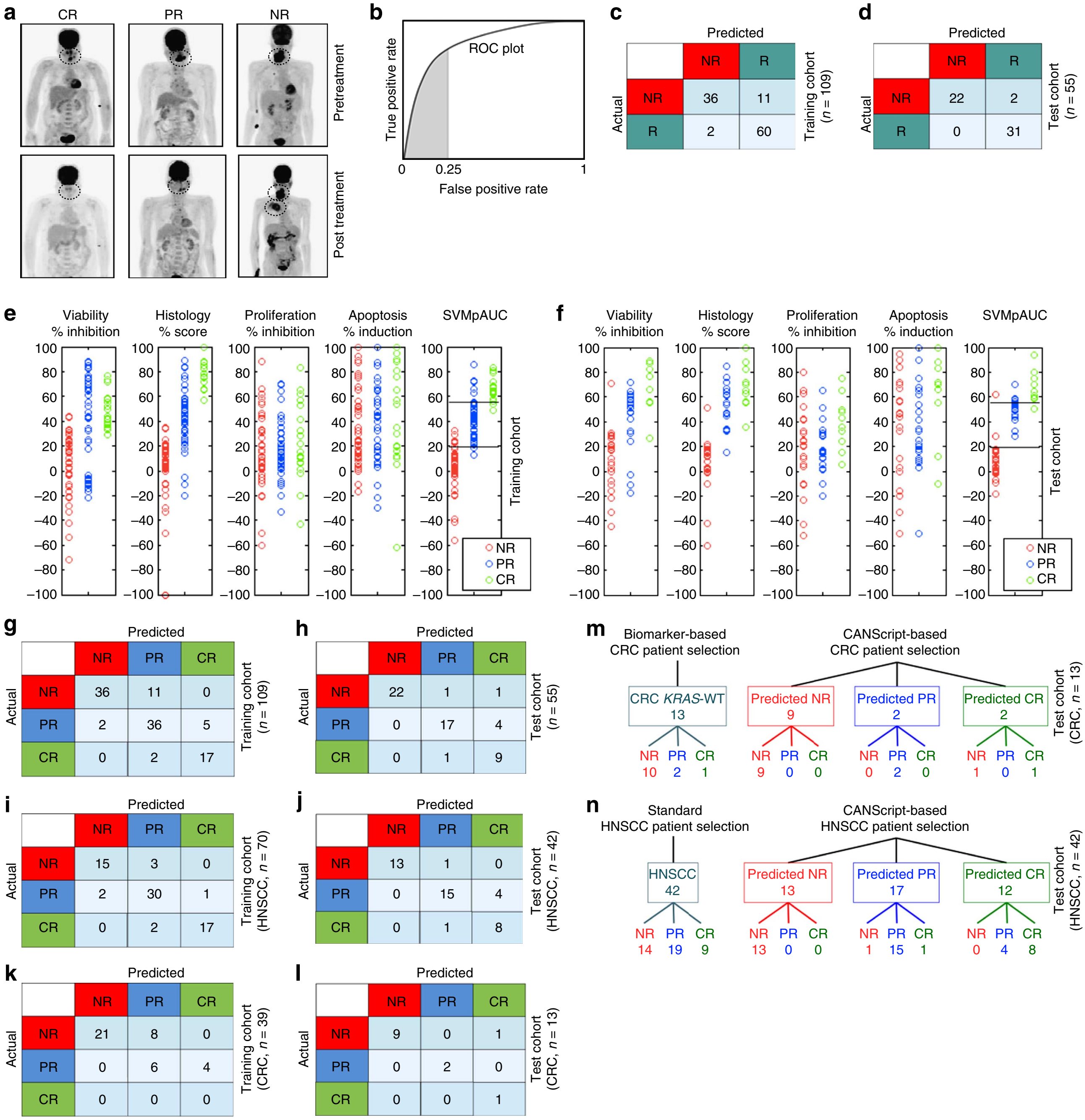


Figure 7 | Validation of CANScript platform using clinical data. (**a**) Positron emission tomography–computed tomography (PET–CT) for representative cases of CR (left), PR (middle) and NR (right) as determined by PERCIST. Primary treatment-naïve HNSCC patients underwent FDG-PET-CT scan examination before (predose) and after three cycles of TPF treatment (post dose). Clinical response to the drugs for individual patients was evaluated based on PERCIST data. (**b**) ROC plot showing true positive rate (sensitivity) and false positive rate (one minus specificity); the shaded area represents the partial area under the ROC curve up to false positive rate 0.25. The SVMpAUC algorithm used to learn a NR/R model to distinguish the non-responders from responders maximized the partial area under the ROC curve up to false positive rate of 0.25 on the training set. This encourages learning a model with high sensitivity, minimizing the number of potential responders (PR or CR patients) that are predicted to be NR while keeping specificity at least 75%. (**c**) Performance of learned NR/R model on the training set. Confusion matrix displays the number of patients with various actual and predicted responses to TPF for HNSCC and cetuximab + FOLFIRI for CRC in the training set ($n = 109$). (**d**) Performance of learned NR/R model on the test set. Confusion matrix displays the number of patients with various actual and predicted responses in the test set ($n = 55$). (**e**) Plots showing values of the functional read-outs from the CANScripts (that is, viability, histology, proliferation and apoptosis), as well as scores assigned by the SVMpAUC-learned model to patients in the training set, and (**f**) in the test set. (**g**) Performance of final refined NR/PR/CR prediction model on the training set. Confusion matrix displays the number of patients with various actual and predicted responses to TPF for HNSCC and cetuximab + FOLFIRI for CRC in the training set. (**h**) Performance of final refined NR/PR/CR prediction model on the test set. Confusion matrix displays the number of patients with various actual and predicted responses in the test set. (**i**,**j**,**k** and **l**) Performance of final refined NR/PR/CR prediction model on HNSCC cases alone in the training and test sets, and on CRC cases alone in the training and test sets, respectively.(**m**) CANScript-based model is a better tool than biomarker- *(KRAS)* based prediction of response to cetuximab and FOLFIRI in CRC. (**n**) CANScript-based model is a better tool than standard patient selection for response to TPF in HNSCC.

original architecture, as evident in the CANScript platform, is important for increasing the probability of predicting a chemotherapy outcome. Indeed, anticancer drugs have been reported to exert their effects by altering both cancer cells and tumour microenvironment[10,11,34,40].

Interestingly, short-term cultivation of primary explants of human tumours had been explored previously, for example, growing the specimens in plasma clots[43] or using engineered 3D explant cultures[20,25,44–48]. While these explant models did capture the heterogenous cancer and stromal cell population to certain degree, and were used to study tumour heterogeneity, invasiveness and response to treatment[15–17,21,25,49], these attempts did not show full viability at diverse functional levels. While extracellular matrix support was shown to help preserve and recreate many important morphologic and phenotypic properties in these 3D spheroids and organotypic cultures, these studies did not elucidate the importance of conserving tumour type- and grade-matched matrix factors in maintaining functional organization and dynamics. Indeed, our results indicate that the composition of TMPs is distinct between tumour types and also between grades. Nor did these studies recreate the oncogenic signalling networks encompassing the activation of diverse RTK signalling with extensive heterogeneity and cross-talks[20,27,30,50–52]. Importantly, our results with mismatched matrix or HS controls indicate the criticality of a matched tumour microenvironment together with AS in preserving the phenotypic and molecular features of the native tumour. It is evident from the RPPA profiling of key signalling pathways and physiologically relevant growth factors detectable in patient serum that extensive heterogeneity exists between patients and that a truly personalized milieu with an active balance of multiple parallel signalling cascades can therefore be successfully created by AS[53–57]. It should, however, be noted that while AS and TMP independently and collectively improved explant culture quality, not all aspects are necessarily dependent on dual presence of AS and TMP. For example, growth factor dependent features are better sustained in presence of AS, whereas TMP plays a dominant role in tumour heterogeneity and at the phenotypic level in maintaining survival and proliferation.

The ability to predict outcomes is not only attractive from a clinical perspective, but also has major implications on preclinical cancer research, where the focus has been to develop assays that can bridge the translational gap. While animal models have been used as the front line in predicting efficacy, the predictive value of these models is debatable, a consequence of using cell lines cultured over years that are no longer representative of the original tumour. Furthermore, transgenic murine models may recapitulate a specific cancer pathway, but fail to capture the true heterogeneity that is characteristic of human tumours. For example, we observed in our study that while EGFR generally plays a critical role in HNSCC, additional driver mechanisms such EphB4, AKT, ERK1/2, Tie2, VEGFR2, cAbl, FGFR1, HER3 and IR are activated. Indeed, such stochastic heterogeneity has been implicated in the induction of adaptive resistance. There is therefore a resurgence in the use of early-passage patient-derived xenografts for predicting clinical responses. Consistent with these recent studies, we observed a good concordance in terms of histopathology and gene expression between the tumour biopsy (P0) and the 2ndpassage xenografts (P2-HTX). However, we did observe unique mutations between the P0 biopsy and the P2-HTX xenografts. It is possible that these differences between P0 and P2-HTX arise due to intratumour heterogeneity at the time of implantation[50]. In our study, the presence of fewer unique P2-HTX mutations in HNSCC-1, a clinical responder, versus the high number of unique P2-HTX mutations in HNSCC-2 and HNSCC-3 tumours, clinically classified as partial and non-responders, respectively, could indicate a propensity for the acquisition of new mutations and/or rearrangements during tumour propagation, consistent with the genetic instability. Furthermore, the 'take rate' in the current studies was $<50\%$, consistent with published reports, which together with the long time required to establish a graft has been a limiting factor for translation of xenotransplant of primary models for predictive studies[58]. The *ex vivo* to *in vivo* functional correlation data clearly show the benefit of using CANScript technology as a surrogate of animal modelling. In addition, the minimal amount of tissue required to establish the CANScripts means multiple explants per tumour biopsy, which allows us to better capture the impact of intratumoral heterogeneity on outcome.

A powerful feature of the CANScript platform is its use of a novel machine learning approach that is tailored to make accurate predictions particularly for potential responders. Specifically, the algorithm operates in two stages: it first employs the recently proposed SVMpAUC-based learning algorithm to distinguish between responders and non-responders in a way that maximizes sensitivity (fraction of responders predicted as responders). Indeed, the learned model in our case achieved 100% sensitivity on the test set while keeping specificity in an acceptably high range. In the second stage, the algorithm learns an additional threshold to separate responder predictions into complete responder and partial responder predictions. This approach was found to be superior to the performance of a standard, widely used support vector ordinal regression algorithm that directly aims to make predictions in the three categories and does not explicitly incorporate the need for high sensitivity. Interestingly, studies have correlated complete pathological response to the long-term progression-free survival[59], while recent ongoing clinical trials like adjuvant dynamic marker adjusted personalized therapy trial (ADAPT) are using short-term dynamic response prediction biomarkers like decrease in Ki-67 in clinical settings as surrogates for clinical outcome for tailoring personalized treatment[60], indicating that integrating multiple end points into a single score as adopted in this model could make response prediction more comprehensive. Combined together, the comprehensive tumour ecosystem and the SVMpAUC-based algorithm makes the CANScript platform a powerful predictive tool that can be used across different tumour types and treatment regimens, as is evident from the overall response rates observed in HNSCC and CRC tumours to targeted and chemotherapy regimens that was similar to clinical outcome observed in previous studies[6,61]. Moreover, while for this study we have focused on predicting the patient response to a single drug regimen at a time, in the future, the approach can be extended to predicting a rank order among different drug regimens based on their likely outcomes, which could help in prioritizing different treatments. Furthermore, the CANScript platform can afford nearly high-throughput testing while capturing the patient intratumoral heterogeneity at a global level with higher fidelity, allowing predictions to be made within 7 days for truly personalizing chemotherapy.

Methods

Collection of tumour samples and patient sera. Tumours samples were collected by core biopsy at the beginning of treatment and at the time of surgical removal for deserving patients (for patient detail see Supplementary Methods). For each patient 5–10 ml of non-heparinized blood was collected at the time of first biopsy in BD-Vacutainer tubes and serum was separated at 1,000*g* for 15 min. All sera samples were aliquoted and stored at −80 °C for further use.

Isolation of extracellular matrix proteins. Surgically removed fresh tumour tissues were dissected into small sections (~ 1–$2\,\text{mm}^3$) and digested with dispase (Stem cell Technologies Inc.), and subsequently the cells were separated using a sieve[62]. Decellularization process was verified by phase contrast microscopy, and further confirmed by 4,6-diamidino-2-phenylindole staining and DNA quantification. Tissue slices suspended in dispase solution was incubated for 15 min at 48 °C. The tissues were homogenized in a high salt buffer solution containing 0.05 M Tris pH 7.4, 3.4 M sodium chloride, 4 mM of EDTA, 2 mM of N-ethylmaleimide and protease (Roche.11836153001) and phosphatase inhibitors (Sigma-aldrich, P0044 and P5726) using tissue homogenizer (Cole Parmer). The homogenized mixture was centrifuged repeatedly three times at 7,000*g* for 15 min and the supernatant was discarded to retain the pellet. The pellet was incubated in 2 M Urea buffer (0.15 M sodium chloride and 0.05 M Tris pH 7.4) and stirred for 1 h at 50 °C. The complex extracted proteins were solubilised in Urea buffer[63]. The mixture was then finally centrifuged at 14,000*g* for 20 min and resuspended in the 2 M Urea buffer, aliquoted and stored at − 80 °C. In addition, extracted protein samples were run at denaturing conditions in the presence of standard molecular weight ladder. When the run was complete, the gel was transferred into a suitable staining tray and fixed in a solution containing formaldehyde in a shaker for 2 h. The gel was washed three times with 1 × wash solution once in every 5 min. The gel was incubated with sensitizing solution containing sodium-thiosulphate for 2 min with gentle shaking and visualized using silver staining.

Identification of TMP components by nano LCMS/MS. The protein mixture was dissolved at the concentration of 1 μg ml^{-1} in 50 mM ammonium bicarbonate buffer. The pH of protein samples was adjusted to ~8.5. The samples (50 µl) were reduced with 10 mM DTT at 56 °C for 45 min, incubated at 95 °C for 5 min and then allowed to cool. Alkylation was carried out by using 55 mM final concentration of iodoacetamide in the dark. Trypsin (13 ng μl^{-1}) was added at a ratio of 1:30 and enzyme/sample mixture was mixed well. Tubes with sample were placed into thermostat and incubated at 55 °C for 2 h and then 37 °C overnight. Digested samples were subjected for SpeedVac at 30 °C for 2–3 h. 5% formic acid was added for adjusting the pH to 3. The samples were either subjected directly to MS analysis or stored at − 20 °C.

Sample was vacuum dried and reconstituted in 12 µl of 0.1% formic acid containing 12.5 fmol μl^{-1} bovine serum albumin (BSA) or β-gal. One micro liter of this was injected on column. Hence, the area of BSA/ β-gal was used for normalization. A separate Mascot run was performed with carboxymethylation as a dynamic modification to obtain area of BSA or β-gal protein. Area of the respective protein is normalized to the control area of respective sample. Reserpine (Sigma-aldrich) was used as a working standard. Digested peptides were subjected to analysis by injecting into nano LCMS/MS[63]. The instrument (STAR Elite, Q-TOF LCMS, Applied Biosystems) was externally calibrated with standard compounds. In brief, peptide mixtures were dissolved in 25 µl of sample preparation solution and injected (10 µl pick up) into nano-LC through an auto-sampler system. Peptides were eluted using nano-reverse phase column (Michrom C18 5 µm particle, 300 Å pore size, 75 µm ID, 150 mm length) which was further connected to the Nano Spray ESI-Q-TOF system (Qstar Elite, Applied Biosystems). A gradient of water and Acetonitrile was set up for 60 min with a flow rate of 400 nl min^{-1}. Eluted peptides from the column were ionized using ESI source with ion spray voltage 2250 V and temperature 120 °C. Ionized peptides were analysed by one full MS scan and four consecutive product ion scans of the four most intense peaks, using rolling collision energy. An Information Dependant Acquisition (IDA) experiment was used to specify the criteria for selecting each parent ion for fragmentation, which included selection of ions in m/z range: >400 and <1600, of charge state of $+2$ to $+5$, exclusion of former target ions for 30 s, accumulation time of 1 s for a full scan and 2 s for MS/MS. The data generated by the Analyst software were stored in a.wiff format. The machine generated data files were analysed using ProteinPilot version 4.0 software with a combined NCBI Human Database (release 45, containing 39125 non-redundant protein entries, 18.8 Mb), Paragon Algorithm and Proteome Discoverer1.3 software. All searches were performed with tryptic specificity allowing two missed cleavages. Trypsin and keratin entries were retained in the list generated. During the analysis, in the search parameters modification of cysteine by idoacetamide and biological modifications programmed in algorithm were allowed. Mass tolerance for precursor ion and fragment ions were set to 100 p.p.m. and 0.2 Da, respectively. In Paragon Algorithm, protein score was calculated on the basis of percentage confidence level of the peptides identified. Protein score of minimum 0.47 (fit incorrect rate is 0%) corresponding to a confidence level $>66\%$ were used. To rule out false discoveries, we carried out a False Discovery Rate (FDR) analysis[64] using ProteinPilot 4.0 with Paragon algorithm for data analysis. As part of the Paragon analysis method, a FDR analysis of the results was carried out by the Proteomics Performance Evaluation Pipeline Software (PSPEP). Finally, proteins were selected on the basis of their critical FDR value, that is, 1%. To avoid identifications based on redundant peptides in our proteome, we did not include proteins that have no unique peptide identifications. Protein grouping function was disabled for generation of protein list. Proteins that share some peptides as well as have unique peptide identifications were grouped accordingly. Deeper annotations were done by accessing specific published information.

For peptide and protein identification, peak lists were correlated with the human protein database[65–68]. The rationale for spectral counting derived protein abundance is that proteins in higher abundance result in more proteolytic peptides detected by tandem MS and subsequently identified by database searching. Following the matching of peptide peaks, peptide abundances in each of the analysed gradient fractions were calculated from the area under the peak. All data processing steps were manually inspected to ensure correct peak detection and matching; overlapping peaks were discarded. Proteins were considered quantifiable if they were represented in at least 75% of the clinical samples matching the cancer type and grade. There are many inherent variables, like ionization efficiency, sensitivity to digestion and interference at the time of elution might influence in the determination of the relative abundance for a protein. In general the prediction falls within a ratio of twofold compared with the actual one. Both sample distance and protein feature distance were calculated using Pearson's correlation and average linkage was used for the clustering of both samples and protein features.

Preparation, coating and detection of TMP mix. TMP cocktails were prepared, based on the relative abundance of key components obtained from LCMS/MS analysis of HNSCC and CRC patient tumour tissues using human proteins as shown in Supplementary Figs 2 and 3. Sterile culture wells were freshly coated with TMP cocktails (100 μg ml^{-1}) unless mentioned otherwise. To visualize the coat, the matrix was incubated with anti-Collagen1 antibody at a dilution of 1:50 (rabbit polyclonal, Abcam. ab34710) for 1 h at room temperature. After four washes in PBS, slides were incubated with Alexa Fluor 555 (anti-mouse, Cell Signaling Technology.4409) for additional 45 min at room temperature in dark. Slides were washed with PBS and finally mounted with Vecta-Shield DAPI (Vector laboratories. H-1200) to confirm the absence of nuclear contamination in premixed TMP cocktail. Images were visualized under immunofluorescence microscopy setting using red and blue filters (DM4000, Leica Microsystems) and images were captured with DFC 425C (Leica) camera.

Surface scanning electron microscopy. Electron microscope compatible cover slips (Thermanox, Ted Pella Inc.) were coated with freshly prepared TMP cocktails (100 μg ml^{1}) for 4 h, washed twice and were fixed in 10% buffered formalin for 10 min (to resist metal coating and high electron beam), washed in PBS and dehydrated with 70 and 100% ethanol for 5 min each. Immediately before imaging, the slides were coated with gold and the images were captured using a Cambridge scanning electron microscope with EDAX attachment.

Human tumour explant culture. Tumour tissues were sectioned into ~300 µm slices using McIlwain tissue chopper (TedPella). These tumour sections were randomized and cultured in 48-well flat bottom plates coated with stage and grade-matched TMP with RPMI medium supplemented with 2% AS, 8% FBS (Life Technologies. 10270-106), 1 × Insulin-Transferrin-Selenium (ITS, Life Technologies. 41400-045), 1 × GlutaMAX (Life Technologies. 35050-061) and 1 × penicillin, streptomycin and amphotericin B (Life Technologies. 15140-122). Tumour slices ($n=3$) were treated with either anti-EGF neutralizing antibody (rabbit monoclonal, clone D8A1, Cell Signaling Technology. 12157) or with TPF (for HNSCC) or with cetuximab + FOLFIRI (for CRC) or with dimethylsulphoxide (DMSO; vehicle control) for 72 h. The final concentration of DMSO was kept $\geq 0.01\%$. Media with drugs were changed every 24-h interval. A portion of each tumour slice was used for cell viability (assessed by WST) and remaining tumours were fixed in 10% buffered formalin and embedded in paraffin. The paraffin-embedded tumours were used for histological (hematoxylin and eosin stain) and IHC analysis including proliferation and cell death.

Human tumour xenografts. Freshly isolated primary human tumours were washed in normal saline and cut into a small pieces (~5 mm^3) and implanted subcutaneously onto the flanks (both sides) in immune compromised 5–6 weeks old, female severe combined immunodeficiency (C.B-17/IcrHsd-PrkdcscidLystbg, Harlan) mice. Tumour bearing mice (at the time of commencement of treatment maximum tumour size was restricted to 100–150 mm^3) were treated with vehicle (0.9% normal saline;) or concurrent regimen of TPF (cisplatin 2.5 mg kg^{-1} body weight, docetaxel 20 mg kg^{-1} and 5Fu 50 mg kg^{-1}) or single agent cetuximab (4 mg kg^{-1}) for 3–4 weeks. Tumour volume was calculated using the following formula, Tumour volume (mm^3) = (π/6) LWH; where L = length (mm), W = width (mm) and H = height (mm). All mice studies and experimental protocols were approved by the institutional animal ethics committee.

Gene expression, exome and mutational analysis. See Supplementary Methods for details. The data are publically available at Gene Expression Omnibus through GEO series accession number GSE63544 and GSE63545; Biosample accession numbers, SAMN03271711, SAMN03271712 and SAMN03271713.

Machine learning algorithm. We learned a model for predicting patient responses as NR/PR/CR in two stages. At the first stage, PR and CR labels were grouped together into a single responder (R) category, and the recently proposed SVMpAUC algorithm[69] was trained on the training set of 109 patients to learn a model to assign the scores and predict NR/R for new test cases. Specifically, given a training set containing n examples (x_i, y_i), $i = 1,\ldots,n$ (here $n = 109$), where x_i is a feature vector containing the four functional read-outs for the i-th patient and y_i is 1 if the i-th patient is a responder and -1 otherwise, the SVMpAUC algorithm learns a weight vector **w** maximizing (a concave lower bound on) the partial area under the ROC curve (partial AUC) up to a specified false positive rate β (here $\beta = 0.25$), defined as follows[70]

$$pAUC\ (w) = \sum_{i:y_i=1} \sum_{j:y_j=-1} 1(w \cdot x_i > w \cdot x_j) \cdot 1(j \in S_\beta)$$

Where S_β contains indices j of the top β fraction of non-responders in the training set, ranked according to scores $w.x_j$. This produced a weight vector w assigning coefficients of 0.2977, 0.5562, 0.0073 and 0.1388 to the viability, histology, proliferation and apoptosis read-outs, respectively. Together with a threshold of 19.1 corresponding to (approximately) $\beta = 0.25$ false positive rate on the training set, this yielded an initial NR/R prediction model. In the second stage, the above model was further refined to classify the predicted responders as PR and CR; this was done by selecting a threshold (55.14) that maximized PR/CR classification accuracy on the training set.

Statistical analysis. One way analysis of variance and Student's *t*-test, linear regression and Spearman coefficient of correlation was analysed using GraphPad Prism version 5 for Windows, GraphPad Software.

References

1. Siena, S., Sartore-Bianchi, A., Di Nicolantonio, F., Balfour, J. & Bardelli, A. Biomarkers predicting clinical outcome of epidermal growth factor receptor-targeted therapy in metastatic colorectal cancer. *J. Natl Cancer Inst.* **101,** 1308–1324 (2009).
2. van't Veer, L. J. & Bernards, R. Enabling personalized cancer medicine through analysis of gene-expression patterns. *Nature* **452,** 564–570 (2008).
3. Samson, D. J., Seidenfeld, J., Ziegler, K. & Aronson, N. Chemotherapy sensitivity and resistance assays: a systematic review. *J. Clin. Oncol.* **22,** 3618–3630 (2004).
4. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483,** 603–607 (2012).
5. Foekens, J. A. *et al.* Multicenter validation of a gene expression-based prognostic signature in lymph node-negative primary breast cancer. *J. Clin. Oncol.* **24,** 1665–1671 (2006).
6. Cunningham, D. *et al.* Cetuximab monotherapy and cetuximab plus irinotecan in irinotecan-refractory metastatic colorectal cancer. *N. Engl. J. Med.* **351,** 337–345 (2004).
7. Van Cutsem, E. *et al.* Open-label phase III trial of panitumumab plus best supportive care compared with best supportive care alone in patients with chemotherapy-refractory metastatic colorectal cancer. *J. Clin. Oncol.* **25,** 1658–1664 (2007).
8. Nakasone, E. S. *et al.* Imaging tumor-stroma interactions during chemotherapy reveals contributions of the microenvironment to resistance. *Cancer Cell* **21,** 488–503 (2012).
9. McMillin, D. W. *et al.* Tumor cell-specific bioluminescence platform to identify stroma-induced changes to anticancer drug activity. *Nat. Med.* **16,** 483–489 (2010).
10. Straussman, R. *et al.* Tumour micro-environment elicits innate resistance to RAF inhibitors through HGF secretion. *Nature* **487,** 500–504 (2012).
11. Sharma, S. V. *et al.* A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell* **141,** 69–80 (2010).
12. Garnett, M. J. *et al.* Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature* **483,** 570–575 (2012).
13. Kreso, A. *et al.* Variable clonal repopulation dynamics influence chemotherapy response in colorectal cancer. *Science* **339,** 543–548 (2013).
14. Pontiggia, O. *et al.* The tumor microenvironment modulates tamoxifen resistance in breast cancer: a role for soluble stromal factors and fibronectin through beta1 integrin. *Breast Cancer Res. Treat.* **133,** 459–471 (2012).
15. Sethi, T. *et al.* Extracellular matrix proteins protect small cell lung cancer cells against apoptosis: a mechanism for small cell lung cancer growth and drug resistance *in vivo*. *Nat. Med.* **5,** 662–668 (1999).
16. Yamada, K. M. & Cukierman, E. Modeling tissue morphogenesis and cancer in 3D. *Cell* **130,** 601–610 (2007).
17. Weaver, V. M. *et al.* beta4 integrin-dependent formation of polarized three-dimensional architecture confers resistance to apoptosis in normal and malignant mammary epithelium. *Cancer Cell* **2,** 205–216 (2002).
18. Wang, F. *et al.* Phenotypic reversion or death of cancer cells by altering signaling pathways in three-dimensional contexts. *J. Natl Cancer Inst.* **94,** 1494–1503 (2002).
19. Sharma, S. V., Haber, D. A. & Settleman, J. Cell line-based platforms to evaluate the therapeutic efficacy of candidate anticancer agents. *Nat. Rev. Cancer* **10,** 241–253 (2010).
20. Vaira, V. *et al.* Preclinical model of organotypic culture for pharmacodynamic profiling of human tumors. *Proc. Natl Acad. Sci. USA* **107,** 8352–8356 (2010).
21. Ridky, T. W., Chow, J. M., Wong, D. J. & Khavari, P. A. Invasive three-dimensional organotypic neoplasia from multiple normal human epithelia. *Nat. Med.* **16,** 1450–1455 (2010).
22. Jung, P. S. *et al.* Progression-free survival is accurately predicted in patients treated with chemotherapy for epithelial ovarian cancer by the histoculture drug response assay in a prospective correlative clinical trial at a single institution. *Anticancer Res.* **33,** 1029–1034 (2013).
23. Karagiannis, G. S. *et al.* Proteomic signatures of the desmoplastic invasion front reveal collagen type XII as a marker of myofibroblastic differentiation during colorectal cancer metastasis. *Oncotarget* **3,** 267–285 (2012).
24. Genovese, L. *et al.* Cellular localization, invasion, and turnover are differently influenced by healthy and tumor-derived extracellular matrix. *Tissue Eng. Part A* **20,** 2005–2018 (2014).
25. Mazzoleni, G., Di Lorenzo, D. & Steimberg, N. Modelling tissues in 3D: the next future of pharmaco-toxicology and food research? *Genes Nutr.* **4,** 13–22 (2009).
26. Sharma, S. V. & Settleman, J. Oncogene addiction: setting the stage for molecularly targeted cancer therapy. *Genes Dev.* **21,** 3214–3231 (2007).
27. Benedettini, E. *et al.* Met activation in non-small cell lung cancer is associated with *de novo* resistance to EGFR inhibitors and the development of brain metastasis. *Am. J. Pathol.* **177,** 415–423 (2010).
28. Frederick, M. J. *et al.* Phosphoproteomic analysis of signaling pathways in head and neck squamous cell carcinoma patient samples. *Am. J. Pathol.* **178,** 548–571 (2011).
29. Puliyappadamba, V. T. *et al.* Opposing effect of EGFRWT on EGFRvIII-mediated NF-kappaB activation with RIP1 as a cell death switch. *Cell Rep.* **4,** 764–775 (2013).
30. Baudhuin, L. M. *et al.* S1P3-mediated Akt activation and cross-talk with platelet-derived growth factor receptor (PDGFR). *FASEB J.* **18,** 341–343 (2004).
31. Sheng, Q. *et al.* An activated ErbB3/NRG1 autocrine loop supports *in vivo* proliferation in ovarian cancer cells. *Cancer Cell* **17,** 298–310 (2010).
32. Yamada, T. *et al.* Paracrine receptor activation by microenvironment triggers bypass survival signals and ALK inhibitor resistance in EML4-ALK lung cancer cells. *Clin. Cancer Res.* **18,** 3592–3602 (2012).
33. Goh, L. K. & Sorkin, A. Endocytosis of receptor tyrosine kinases. *Cold Spring Harb. Perspect. Biol.* **5,** a017459 (2013).
34. Steidl, C. *et al.* Tumor-associated macrophages and survival in classic Hodgkin's lymphoma. *N. Engl. J. Med.* **362,** 875–885 (2010).
35. Young, M. R. *et al.* Human squamous cell carcinomas of the head and neck chemoattract immune suppressive CD34(+) progenitor cells. *Hum. Immunol.* **62,** 332–341 (2001).
36. Akervall, J. *et al.* Genetic and expression profiles of squamous cell carcinoma of the head and neck correlate with cisplatin sensitivity and resistance in cell lines and patients. *Clin. Cancer Res.* **10,** 8204–8213 (2004).
37. Chang, J. C. *et al.* Patterns of resistance and incomplete response to docetaxel by gene expression profiling in breast cancer patients. *J. Clin. Oncol.* **23,** 1169–1177 (2005).
38. Misale, S. *et al.* Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer. *Nature* **486,** 532–536 (2012).
39. Siu, L. L. *et al.* Phase III randomized, placebo-controlled study of cetuximab plus brivanib alaninate versus cetuximab plus placebo in patients with metastatic, chemotherapy-refractory, wild-type K-RAS colorectal carcinoma: the NCIC Clinical Trials Group and AGITG CO.20 Trial. *J. Clin. Oncol.* **31,** 2477–2484 (2013).
40. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144,** 646–674 (2011).
41. Di Nicolantonio, F. *et al.* Wild-type BRAF is required for response to panitumumab or cetuximab in metastatic colorectal cancer. *J. Clin. Oncol.* **26,** 5705–5712 (2008).
42. Kim, J. *et al.* Chemokine receptor CXCR4 expression in colorectal cancer patients increases the risk for recurrence and for poor survival. *J. Clin. Oncol.* **23,** 2744–2753 (2005).
43. Burnouf, T. *et al.* Human blood-derived fibrin releasates: composition and use for the culture of cell lines and human primary cells. *Biologicals.* **40,** 21–30 (2012).
44. Gerlach, M. M. *et al.* Slice cultures from head and neck squamous cell carcinoma: a novel test system for drug susceptibility and mechanisms of resistance. *Br. J. Cancer.* **110,** 479–488 (2014).
45. Ashley, N., Jones, M., Ouaret, D., Wilding, J. & Bodmer, W. F. Rapidly derived colorectal cancer cultures recapitulate parental cancer characteristics and enable personalized therapeutic assays. *J. Pathol.* **234,** 34–45 (2014).
46. Sekine, H. *et al. In vitro* fabrication of functional three-dimensional tissues with perfusable blood vessels. *Nat. Commun.* **4,** 1399 (2013).
47. Collins, K. B., Patterson, B. K., Naus, G. J., Landers, D. V. & Gupta, P. Development of an *in vitro* organ culture model to study transmission of HIV-1 in the female genital tract. *Nat. Med.* **6,** 475–479 (2000).
48. Hirt, C. *et al.* "*In vitro*" 3D models of tumor-immune system interaction. *Adv. Drug. Deliv. Rev.* **79-80C,** 145–154 (2014).

49. Suchy, S. L., Hancher, L. M., Wang, D., Ervin, Jr. P. R. & Brower, S. L. Chemoresponse assay for evaluating response to sunitinib in primary cultures of breast cancer. *Cancer Biol. Ther.* **11,** 1059–1064 (2011).
50. Radhakrishnan, P. *et al.* Inhibition of rapamycin-induced AKT activation elicits differential antitumor response in head and neck cancers. *Cancer Res.* **73,** 1118–1127 (2013).
51. Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366,** 883–892 (2012).
52. Majumder, P. K. *et al.* mTOR inhibition reverses Akt-dependent prostate intraepithelial neoplasia through regulation of apoptotic and HIF-1-dependent pathways. *Nat. Med.* **10,** 594–601 (2004).
53. Scheel, C. *et al.* Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast. *Cell* **145,** 926–940 (2011).
54. Shumiya, T. *et al.* Evidence for the therapeutic potential of *ex vivo* expanded human endothelial progenitor cells using autologous serum. *Circ. J.* **74,** 1006–1013 (2010).
55. Gohji, K. *et al.* Independent prognostic value of serum hepatocyte growth factor in bladder cancer. *J. Clin. Oncol.* **18,** 2963–2971 (2000).
56. Leitzel, K. *et al.* Elevated plasma platelet-derived growth factor B-chain levels in cancer patients. *Cancer Res.* **51,** 4149–4154 (1991).
57. Zhang, Z., Dong, Z., Lauxen, I. S., Filho, M. S. & Nor, J. E. Endothelial cell-secreted EGF induces epithelial to mesenchymal transition and endows head and neck cancer cells with stem-like phenotype. *Cancer Res.* **74,** 2869–2881 (2014).
58. Seshadri, M. *et al.* Establishment and characterization of patient tumor-derived head and neck squamous cell carcinoma xenografts. *Cancer Biol. Ther.* **8,** 2275–2283 (2009).
59. Liedtke, C. *et al.* Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J. Clin. Oncol.* **26,** 1275–1281 (2008).
60. Hofmann, D. *et al.* WSG ADAPT—adjuvant dynamic marker-adjusted personalized therapy trial optimizing risk assessment and therapy response prediction in early breast cancer: study protocol for a prospective, multi-center, controlled, non-blinded, randomized, investigator initiated phase II/III trial. *Trials* **14,** 261 (2013).
61. Posner, M. R. Paradigm shift in the treatment of head and neck cancer: the role of neoadjuvant chemotherapy. *Oncologist* **10**(Suppl 3): 11–19 (2005).
62. Uriel, S. *et al.* Extraction and assembly of tissue-derived gels for cell culture and tissue engineering. *Tissue Eng. Part C Methods* **15,** 309–321 (2009).
63. Ngoka, L. C. Sample prep for proteomics of breast cancer: proteomics and gene ontology reveal dramatic differences in protein solubilization preferences of radioimmunoprecipitation assay and urea lysis buffers. *Proteome Sci.* **6,** 30 (2008).
64. Naba, A. *et al.* Extracellular matrix signatures of human primary metastatic colon cancers and their metastases to liver. *BMC Cancer* **14,** 518 (2014).
65. Zeng, X. *et al.* Lung cancer serum biomarker discovery using glycoprotein capture and liquid chromatography mass spectrometry. *J. Proteome Res.* **9,** 6440–6449 (2010).
66. Zhou, W., Liotta, L. A. & Petricoin, E. F. The spectra count label-free quantitation in cancer proteomics. *Cancer Genomics Proteomics* **9,** 135–142 (2012).
67. Zhang, G. *et al.* Protein quantitation using mass spectrometry. *Methods Mol. Biol.* **673,** 211–222 (2010).
68. Matzke, M. M. *et al.* A comparative analysis of computational approaches to relative protein quantification using peptide peak intensities in label-free LC-MS proteomics experiments. *Proteomics* **13,** 493–503 (2013).
69. Narasimhan, N. & Agarwal, S. $\mathrm{SVM}_{\mathrm{pAUC}}^{\mathrm{tight}}$. A new support vector method for optimizing partial AUC based on a tight convex upper bound. In *Proceedings of the 19th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* 167–175 (2013).
70. Chu, W. & Keerthi, S. S. Support vector ordinal regression. *Neural Comput.* **19,** 792–815 (2007).

Acknowledgements

We thank Utpal S. Tatu of the Indian Institute of Science, Bangalore, for LCMS-MS analysis, Jayanta Haldar of JNCASR, Bangalore, for electron microscopy, Vikram D. Kekatpure of MSCC for helpful discussion and Pragnashree Mukhopadhyay of Mitra Biotech, Bangalore, for her suggestions on IHC and histology data. We also thank A Kamal, Ayyappan Velu and Rohini S. Nair of Mitra Biotech for their technical assistance.

Author contributions

B.M., P.R., S.A., S.S. and P.K.M. were involved in the design of experiments. B.M., B.U, S.T., M.D., N.B., D.D.P. B.U.S., A.T. conducted *ex vivo*, *in vivo*, microarray, IHC and LCMS experiments and analysed the data. A.P. analysed the assay data. N.H. and S.A. were involved in algorithm development. B.M., S.T., P.R., M.S., S.A., S.S., P.M. wrote the manuscript. R.S., G.K.B., A.M.S., M.A.K. contributed to the clinical aspects of the study. G.B., P.H. conducted exome experiments and data analyses. R.B. and M.L. provided critical inputs for exome and pathology data and results. P.K.M. supervised the study.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: Except H.N., A.M.S., M.A.K., G.B., P.H., R.B. and S.A. all authors are either employees of Mitra Biotech and/or hold equity of Mitra Biotech. The remaining authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Majumder, B. *et al.* Predicting clinical response to anticancer drugs using an *ex vivo* platform that captures tumour heterogeneity. *Nat. Commun.* 6:6169 doi: 10.1038/ncomms7169 (2014).