Clinical and co-clinical genomics of paediatric solid tumours with a special focus in DIPG

Elisa Pilar Izquierdo Delgado

Submitted to the University of London/The Institute of Cancer Research for the Degree of Doctor of Philosophy

April 2020

Declaration

I hereby declare that this thesis reports on my own original work. Any contribution made by others with whom I have worked is explicitly acknowledged in the thesis.

Elisa Izquierdo Delgado

Abstract

Cure rates for paediatric solid cancers have not improved in the last decade. Delivering precision medicine to stratify patients based on their molecular profiling is a key goal to support existing and upcoming adaptive clinical trials. To this end, I have developed and validated a clinically targeted sequencing panel comprising of 78 genes for a first version and 92 genes for an expanded version of the assay. Using the paediatric panel, a total of 255 patients were sequenced and potentially targetable mutations were found in 50% of the patients. This methodology is currently being used for clinical decision making in every child with a solid tumour in the UK, as part of the National Health Service diagnostic service. Furthermore, I have developed a specific panel for the detection of structural variants in paediatric brain tumours which includes 24 genes. Using the fusion-panel, relevant structural variants were identified in the HERBY clinical trial, including novel internal tandem duplications in *NTRK2*. In addition, plasma and CSF samples from the HERBY clinical trial and other studies were used to identify molecular alterations in circulating tumour DNA in patients with paediatric HGG patients and DIPG. Moreover, in a prospective biopsy-stratified clinical trial (BIOMEDE), I combined detailed molecular profiling linked to drug screening in newly-established patient-derived models of DIPG *in vitro* and *in vivo*. A high degree of *in vitro* sensitivity to the MEK inhibitor trametinib was identified in samples harbouring genetic alterations targeting the MAPK pathway. However, treatment of PDX models and the patient with trametinib at relapse failed to elicit a significant response. Resistant clones in a *BRAF*-G469V model were generated identifying the emergence of acquired mutations in *MEK1/2*. These cells showed the hallmarks of mesenchymal transition revealing up-regulation of invasion and migration biomarkers. Resistant clones were sensitive to tyrosine kinase inhibitor dasatinib and combinations of trametinib with dasatinib and the ERK inhibitor ulixertinib showed synergistic effects *in vitro*. This work demonstrates the value of targeted sequencing for a precise classification and stratification of novel, molecularly based therapies in children with cancer. Furthermore, I show the feasibility in generating patient-specific, testable hypotheses that may be clinically translated in a subset of patients. In addition, I describe the detection of MAPK pathway alterations as a therapeutic target in DIPG and demonstrate the importance of resistance modelling to find rational combinatorial treatments.

Acknowledgements

Throughout the course of this PhD I have been extremely fortunate to have the infinite support of my supervisor, funders, family, colleagues and friends who have contributed in one way or another to this work. For this, I will always be grateful.

Firstly, I would like to thank my supervisor and mentor Chris Jones. Thank you for giving me the opportunity four years ago to work in your lab. Thank you for your constant support and guidance. For always encouraging me to become a better scientist by allowing me to explore and research endless hypotheses. I have been very happy working with you and I am very proud of the work we have carried out.

I would like to thank Christopher's Smile. Kevin and Karen, you have been an inspiration since the first day I met you. I am eternally thankful for your confidence in me. To me, you will always be an example of strength and determination. I believe your kindness and persistence is making a difference to treating children with cancer.

To the Glioma Team. I would like to especially thank Diana Carvalho and Alan Mackay. Both of you have been by my side every step of the way, with your guidance, motivation and care. I could not have asked for better supervisors. Thanks Sara, Matt, Ketty and Diana (again) for listening to me and "just" being there. The Glioma Team will always be with me, wherever I go the "sheepcorn" spirit will remind me of the bright and fun days. I hope there will be many bottles of champagne to open in the forthcoming years and millions of dings! I have learnt so many things from all of you and I always knew I could count on you. You have been my family in the UK and no matter what the situation was I always felt safe and loved .

To my friend Patricia, over the last fifteen years your star has, in a way, guided me and brought me here. To my uncle Tomas, his constant perseverance and brilliant mind has always been very present in my heart. To my brothers, David and Dioni, for always rescuing me when I needed them the most at any stage of my life.

Thank you, Dad and Mum, for always believing in me and teaching me to be a better person every day. Your unconditional love and kindness have enabled me to be where I am. Dad, I know you are looking after us and somehow you have given me the strength to finish and enjoy writing this thesis. I will always try to honour your big heart by seeking to be the person you taught me to be.

Table of contents

List of figures

List of tables

CHAPTER 1 : INTRODUCTION

1.1 Introduction to paediatric solid tumours

Cancer is a disease caused as a result of changes occurring in the DNA of the cells which are positively selected due to their ability to proliferate uncontrollably, enabling tumour growth and metastatic dissemination [1]. Although paediatric cancer is rare, representing 1% of all cancers, is the leading cause of death in children under 19 years of age in developed countries [2]. In the last decade, only a limited improvement in survival has been observed for a number of paediatric solid tumours and long-term side effects due to aggressive treatments including surgery, high-dose chemotherapy and radiation can be devastating [3-5]. Therefore, new treatments are urgently needed to improve survival and reduce side-effects in children with high-risk cancer.

The most common types of childhood cancer are leukaemia, CNS tumours lymphomas, sarcomas, sympathetic nervous system nervous and bone tumours whose distribution varies depending on age and differs substantially to those seen in adults **{Figure 1-1}** [6]**.** Unlike most adult cancers, which are epithelial carcinomas and frequently linked to lifestyle and environmental factors where cells might accumulate genetic alterations, the majority of childhood cancers are fundamentally diseases of dysregulated development and it is thought that most of them arise from stem or progenitor cells [7-9].

Remarkable discoveries in unravelling the genetic repertoire of paediatric cancer have demonstrated that the spectrum of mutations in childhood cancers is different to adult cancer, exhibiting a lower tumour burden and a higher incidence of alterations involved in the epigenetic machinery and oncogenic fusions that activate crucial genes in development [10-12]. Moreover, these studies have revealed that childhood cancer is often driven by a single genetic event, mostly disease-specific, compared to adult cancer which is frequently shaped by multiple driving alterations commonly shared by different tumour types [11, 12]. In addition, numerous studies have described that approximately 50% of paediatric cancers harbour targetable alterations and around 8-10% of patients carry predisposition germline variants [12-16].

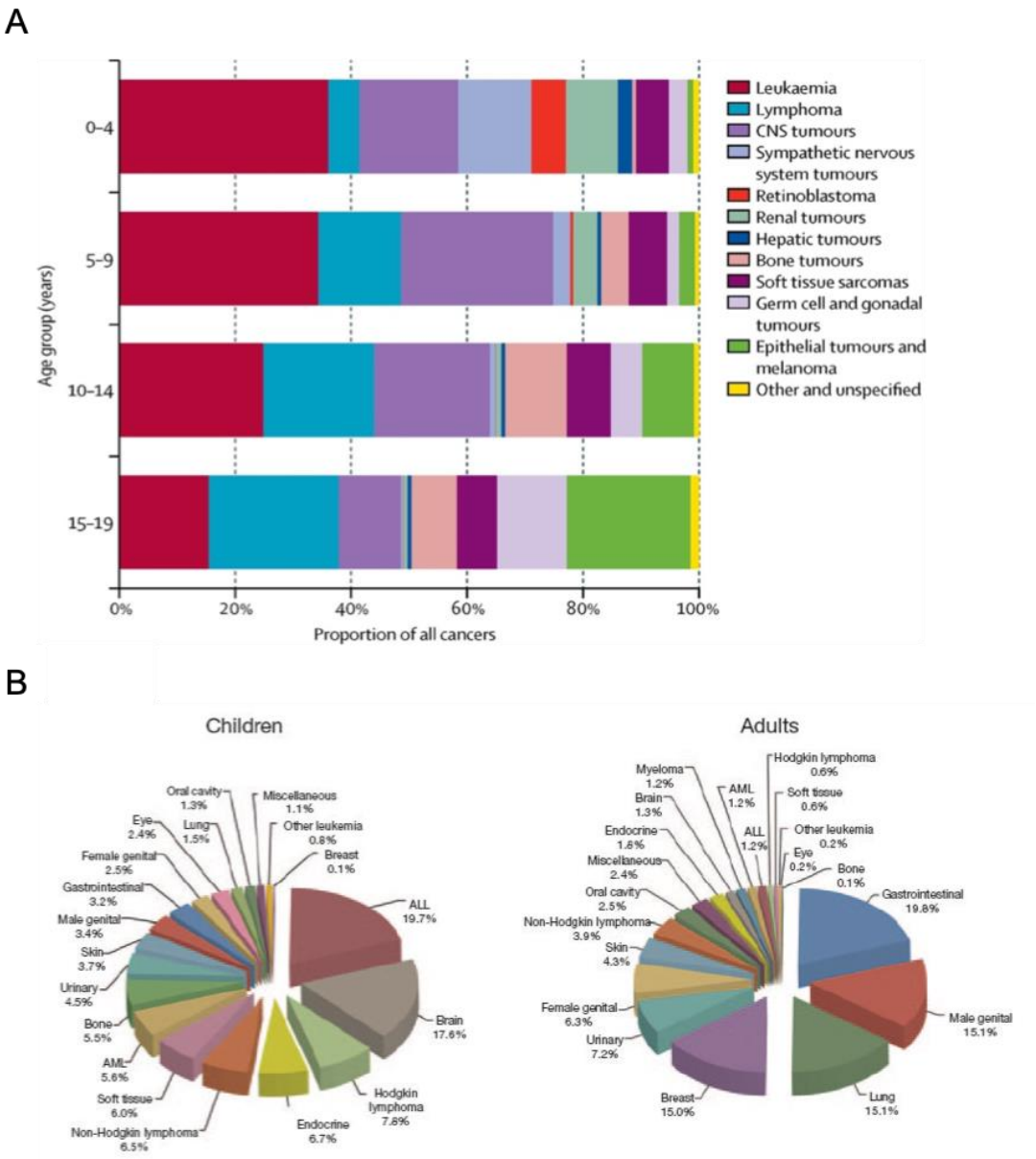


Figure 1-1 Childhood cancer distribution. (A) Distribution of worldwide childhood cancer type by age group, 2001-2010 for children aged 0-19 years, based on data collected from 62 countries, figure taken from [17]. (B) Pie chart of frequency of children and adult cancers diagnosis on the basis of 2012 Surveillance, Epidemiology and End Results (SEER) data, figure taken from [6].

1.2 Intracranial paediatric tumours

Paediatric central nervous system (CNS) tumours are the most common childhood solid tumours, arising at various anatomical locations in the CNS. Brain tumours remain the leading cause of cancer related death in children 0-14 years of age, and long-term side-effects for those who survive can often limit their quality of life as a result of conventional treatments [4, 18, 19].

Paediatric brain tumours comprised a diverse group of tumours of >100 distinct entities, with clinically and biologically distinct features which can make diagnosis challenging. However, remarkable efforts have been made using genome-wide methylation and next generation sequencing (NGS) profiling, which have transformed the classification of paediatric brain tumours, leading to the emergence of new entities and distinct tumour subgroups [20-24].

1.2.1 Glioma

Glioma is the leading cause of cancer-related death in children and they account for approximately 50% of all CNS paediatric tumours [25]. Gliomas are traditionally divided in four grades based on histology and morphology by using light microscopy, with grades I and II considered low-grade gliomas, and grades III and IV high-grade gliomas. More recently, the latest 2016 WHO classification have incorporated molecular features to define specific tumour diagnoses; glioma is notably affected by incorporating IDH mutant, IDH wildtype and diffuse midline glioma (DMG) H3K27M-mutant as distinct entities [26].

1.2.1.1 Paediatric Low-Grade Glioma

Paediatric low-grade glioma (pLGG) is the most common brain tumour in children, accounting for over 30% of all paediatric CNS tumours. Current treatment for pLGG includes total resection correlating with an excellent prognosis, chemotherapy used for unresected and/or recurrent pLGG, and radiation which is reserved for those challenging tumours unresponsive to chemotherapy [27, 28]. Although the majority of pLGG arise sporadically, predisposition to pLGG has been associated with neurofibromatosis type 1 syndrome, characterised by inactivating mutations in *NF1* specially observed in optic pathways gliomas (OPGs), and tuberous sclerosis complex in subependymal giant cell astrocytoma (SEGA) [29, 30].

pLGG can occur anywhere in the brain, however they exhibit a specific anatomical distribution, with different clinical features and molecular alterations suggesting different cells of origin **{Figure 1-2}** [31, 32]. pLGG is usually considered a single alteration-driven disease and is comprised of many molecular subgroups of which a high proportion harbour alterations affecting mitogen-activated protein kinase (MAPK) pathway [33, 34]. Though there is not a 100% correlation between pLGG histology and molecular alterations, there are some associations such as *KIAA1549:BRAF*

tandem duplication being a strong diagnostic biomarker of pilocytic astrocytoma (PA), yet can also occur with low frequency in other rare entities including diffuse leptomeningeal glioneuronal tumour and anaplastic astrocytoma with piloid features [35]. *BRAF*-V600E can be found in a wide number of pLGG, mostly in gangliogliomas and PXAs, but also in pHGG, hence its detection should be used as a predictive biomarker of BRAFi rather than for diagnosis [35]. In addition, *FGFR1* somatic and germline alterations appear as point mutations (N546 and K656), internal tandem duplication (ITD) or as a fusion with a higher incidence in dysembryoplastic neuroepithelial tumours (DNET), and less often in other histologies such as gangliogliomas and PAs [33, 36, 37]. *MYB*/*MYBL1* rearrangements are commonly found in two distinct tumour entities - angiocentric gliomas and isomorphic diffuse gliomas [33, 38, 39]. *MYB:QKI* fusion is the most common rearrangement, with *in vitro* and *in vivo* studies showing that this fusion promotes tumorigenesis via three genetic and epigenetic mechanisms: MYB activation by truncation, aberrant *MYB-QKI* expression and hemizygous loss of *QKI* [40].

Given the high proportion of pLGG tumours with alterations in MAPK pathway, the use of BRAF and MEK1/2 inhibitors has had promising results and currently a phase II trial (NCT03363217) is evaluating the efficacy of trametinib as single agent for treatment of progressing/refractory pLGG tumours with MAK/ERK pathway activation, as well as in combination with dabrafenib for *BRAF* V600 mutation-positive gliomas (NCT02684058) [41-45]. Moreover, the mTOR inhibitor everolimus has been approved by the FDA to treat SEGA based on numerous studies which have shown safety and tumour reduction for long-term treatment (over 5 years) [46, 47].

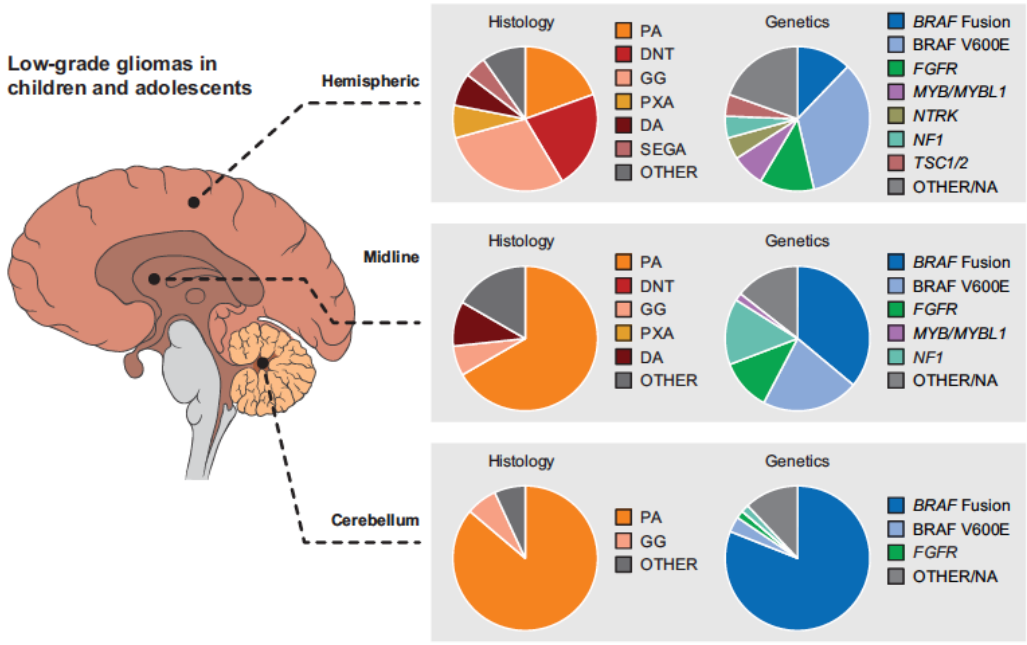


Figure 1-2 Anatomical location of paediatric low grade glioma. For each location (hemispheric, midline or cerebellum) a pie chart coloured by histology and genetic alterations is shown. Figure taken from [32].

1.2.1.2 Paediatric high-grade glioma and DIPG

1.2.1.2.1 Overview

Paediatric high-grade glioma (pHGG) and diffuse intrinsic pontine glioma (DIPG) account for around 20% of all paediatric brain tumours, with a peak incidence in children of 5-9 years of age [25, 48]. Unlike adult GBM, in which tumours are mostly restricted to hemispheric regions of the brain, pHGG and DIPG can occur in different anatomical locations throughout the CNS, correlating with specific age and outcome **{Figure 1-3}** [22]. While pHGG can arise in the cerebral hemispheres, midline structures including the thalamus, and less often cerebellum and spine, DIPG is by definition restricted to the brainstem.

The diagnosis of pHGG is based on histology with WHO grade III and IV features, whereas traditionally DIPG has been diagnosed by radiological and clinical features. There has been no improvement in the survival of pHGG and DIPG in decades, presenting an extremely poor prognosis with a median OS of 9-15 months which represents the greatest cause of cancer-related death in children under 19 years of age [49, 50]. pHGG and DIPG are mainly sporadic cancers, however a fraction of children with these tumours can be linked with three main predisposition syndromes:

Li Fraumeni syndrome (LFS), constitutional mismatch repair deficiency (CMMRD) and neurofibromatosis type 1 [51]. In addition, secondary pHGG are found in children that have been exposed to high-dose ionizing radiation typically for the treatment of medulloblastoma or acute leukaemia [52-54].

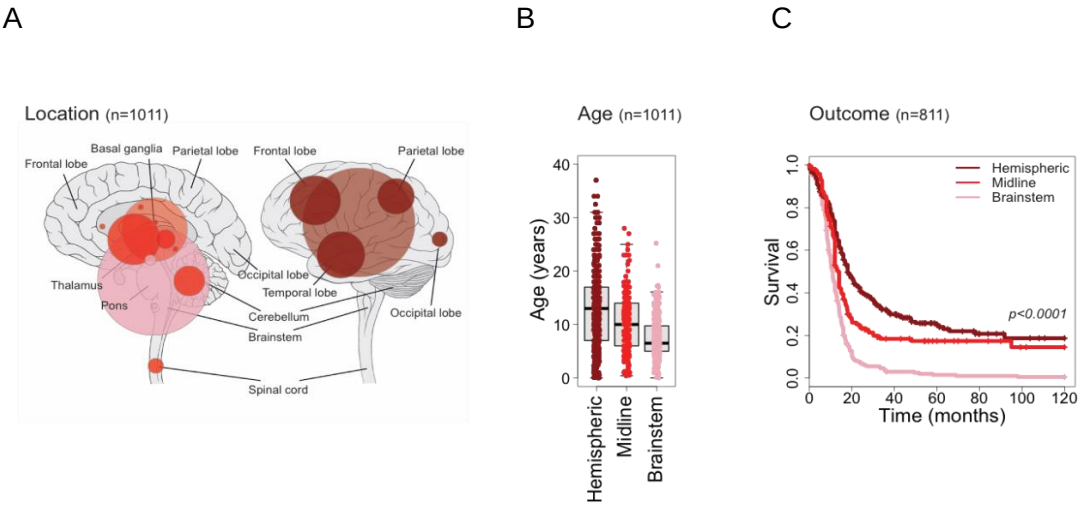


Figure 1-3 Anatomical location, and clinical distribution of pHGG. The different colours represent the anatomical regions of the brain, cerebral hemispheres are shown in dark red, midline structures in red and, the brainstem in pink. (A) Anatomical location of all pHGG (n=1,033). The size of the circle is proportional to the number of cases. The lighter shaded circles correspond to cases that are not located in any of the main three locations described (hemispheric, midline or brainstem) (B). Box-plot showing the age distribution across 1,011 cases per anatomical location. The thick line represents the median, the lower and upper limits represent the first and the third quartiles, and the whiskers the interquartile range (C). Kaplan-Meier plot of overall survival across 811 cases per anatomical location. The p-value was calculated using log rank test. Figure modified from the meta-analysis carried by the Jones lab [22].

1.2.1.2.2 Clinical presentation and current treatment

pHGG and DIPG patients typically present with short history of symptoms (1-3 months) prior to diagnosis. The symptoms are usually a consequence of the increased intracranial pressure and include headaches, behaviour changes, diplopia and emesis [48]. Children can also manifest more specific symptoms such as focal motor deficits, hemiplegia, pyramidal tract findings and dysmetria [48]. Patients with DIPG commonly experienced what is known as “classic triad” which includes cerebral signs (ataxia, dysmetria, dysarthria), long-tract signs (motor deficit, hyperreflexia, clonus, etc) and isolated or multiple cranial nerve palsies (unilateral or bilateral) [55]. At least two of the three clinical features are required for clinical diagnosis of DIPG.

Although some tumours might be diagnosed by using non-contrast computerized tomography (CT) scans, contrast-enhanced magnetic resonance imaging (MRI) is usually the suggested method of choice, and it is considered diagnostic for DIPG and gliomatosis cerebri (GC). GC is a rare growth pattern of diffuse gliomas, characterised as infiltration involving at least three lobes, with diffuse enlargement of anatomic

structures and are often presented bilaterally [56-58]. On MRI scans, pHGG will generally appear as irregularly shaped tumours with heterogeneous enhancement patterns often presenting with areas of cystic necrosis and oedema; diffuse infiltrative margins are however common [59]. DIPG presents ill-defined margins centred in the pons, resulting in diffuse enlargement of this region and they exhibit very little to no enhancement [56, 59, 60]. The involved areas are shown on MRI as hypo-intense on T1 images and hyper-intense signals on T2/flair sequences [59-61]. Extensive infiltration throughout the brain is observed, especially in GC and at disease progression in DIPG [60, 62, 63].

Treatment of pHGG includes maximal safe surgical resection, when feasible, followed by focal radiotherapy (with the exception of infants and younger children), concomitant with the chemotherapeutic drug temozolomide [64]. Gross total resection is a positive prognostic biomarker for pHGG, however due to the infiltrative nature of these tumours complete resection is not always possible. Surgery is not indicated for patients with DIPG as they arise in the ventral pons, an area of the brain that controls many critical nervous system functions. Chemotherapy has not been shown to add any clinical benefit for DIPG patients, leaving radiotherapy (dose of 54-60Gy in daily fractions of 1.5-2Gy over 6 weeks) as the only treatment with proven efficacy in prolonging progression-free survival (PFS) [65, 66].

Clinical trials evaluating the addition of different drug combinations (cetuximab plus irinotecan, bevacizumab with/without vorinostat), as well as targeted therapies (erlotinib, dasatinib, lapatinib) to standard therapy, have not found an improvement in patient outcomes [67-72]. This might be because the patients were not selected according to their tumour biology prior to treatment, underappreciation of the intrinsic tumour heterogeneity of pHGG and DIPG, and/or poor drug penetrance into the CNS.

1.2.1.2.3 Molecular landscape of pHGG and DIPG

The understanding of the biology underpinning pHGG and DIPG has remarkably advanced throughout the use of multi-omics profiling techniques, revealing that gene expression, copy number and the spectrum of mutations are very different from their older counterparts **{Figure 1-4}** and **{Figure 1-5}** [22, 73-78]**.** Adult GBM are genetically characterised by the presence of *IDH1/2* mutations. These can co-segregate with *ATRX, TP53, CDK4* amp and *CDKN2A/B* deletion or co-deletion of chromosomes 1p/19q, commonly found with *TERT* promoter mutations. *IDH1/2* wild-type GBM (classic, mesenchymal, PA-like) often are found with *EGFR* and *PDGFRA*

alterations [79, 80]. *IDH1/2* mutations are seen in a small proportion of tumours (~6%), mostly in hemispheric adolescent pHGG, and are associated with better prognosis compared to other subgroups [22].

By contrast, one of the key distinguishing features of paediatric glioblastoma is the presence of K27M and G34R/V mutations in genes encoding histone H3 variants in pHGG and DIPG (~50%), which are rarely seen in adult tumours (0.2%) [22, 80]. Further description of the histone mutations nature is described in the next section. H3.3 G34R/V mutations found in pHGG usually co-segregate with *ATRX*, *DAXX*, *TP53* and *FBXW7* [22]. Alternative lengthening of the telomeres (ALT) is frequently found in pHGG, especially with H3.3 G34R/V mutations, and represents a distinct telomere maintenance mechanism to the *TERT* promoter hotspot mutations C228T and C250T, commonly seen in adult GBM [73, 81-83]. H3 K27M tumours, now defined by the WHO classification as "diffuse midline gliomas, H3 K27-mutant", are commonly found with *TP53* mutations, *PDGFRA* focal amplifications (predominantly found in the pons) and *FGFR1* mutations (more frequently in thalamic tumours) [22, 84]. In addition, *ACVR1* mutations are commonly present alongside H3.1 K27M and are restricted to DIPG patients, often also harbouring genetic alterations in genes involved in the MAPK/PI3K/mTOR pathway (*PIK3CA, PI3R1, PTEN*) [22, 76]. In addition to the histone mutations, alterations in other chromatin regulators are regularly seen in pHGG and DIPG such as *SETD2, KMT5B/C, KDM6B, BCOR, ASXL1* and *ARID1A/B* [22].

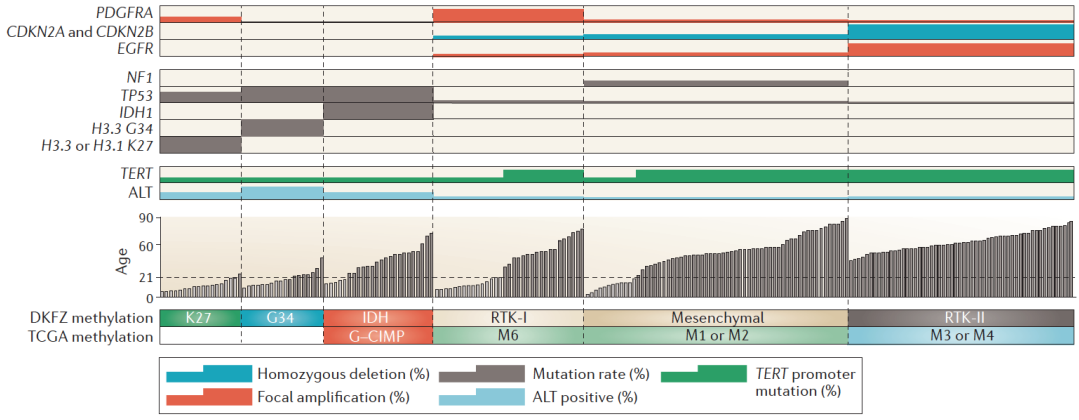


Figure 1-4 Schematic representation of molecular alterations in pHGG and adult glioblastoma. Data taken from the TCGA (The Cancer Genome Atlas) and DKFZ (German Cancer Research). Figure taken from [85].

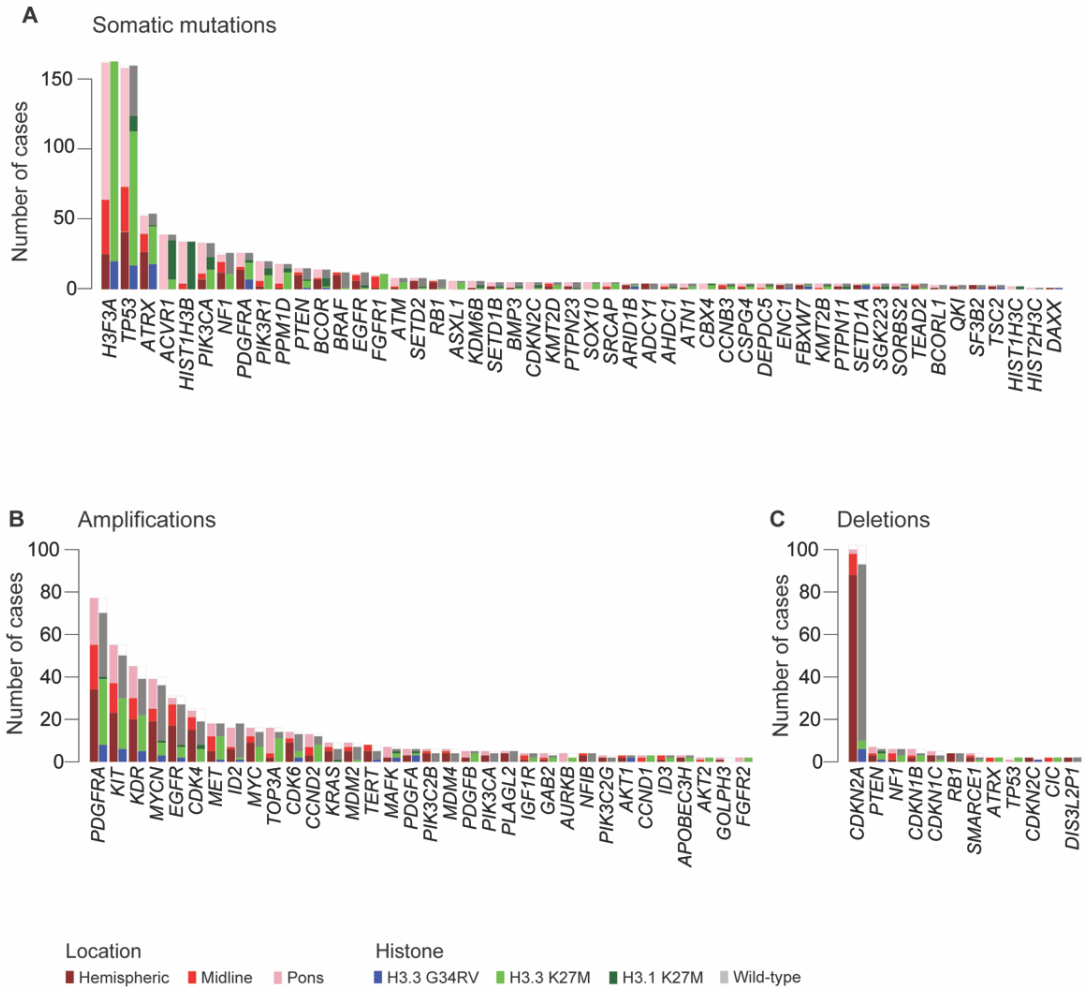


Figure 1-5 Bar-plot of recurrent of recurrent alterations across pHGG and DIPG. Alterations are coloured by location and histone status. (A) Somatic mutations, n=326; (B) amplifications, n=834; and (C) deletions n=834. Figure modified from the meta-analysis carried by the Jones lab [22].

A high mutational burden (>250 mutations/Mb) is observed in ~5% of pHGG associated with CMMRD, typically with the presence of germline mutations in *PMS2, MSH1, MSH2, MSH6, MLH1, POLD1,* and *POLE,* [22, 86, 87]. In addition, a study by our laboratory, identified a high percentage of CD8+ cells in hypermutator cases [87]. The remaining pHGG or DIPG (IDH-WT, histone-WT and non-hypermutator), constitute a group of heterogeneous tumours, with distinct outcomes and driven by different molecular alterations: WT-A (MAPK pathway alterations including *BRAF*-V600E), WT-B (*EGFR/MYCN/CDK6*) and WT-C (*PDGFRA/MET*) [22]. Similarly, genome-wide methylation studies found three molecular sub-types with different clinical outcome in IDH/histone wild-type tumours, which included pedGBM_MYCN, pedGBM_RTK1 and pedGBM_RTK2 with an enrichment of *MYCN*, *PDGFRA* and *EGFR* amplifications respectively [88].

Furthermore, a recent study by our laboratory, identified a new entity of infant HGG with a distinct methylation profile, characterised by the presence of fusion genes (*NTRK1/2/3, ALK, MET and ROS1),* and associated with better outcome than HGG in older children [89].

1.2.1.2.4 Histone mutations

In 2012, two independent studies led to the discovery of unique recurrent mutations in the genes encoding histone H3 variants, found to occur in ~60% cases of pHGG and ~80% cases of DIPG. The mutations occur in two residues in the histone tail: lysine to methionine substitution at position 27 (K27M) in histone 3.3 (*H3F3A),* 3.1 (*HIST1H3B* and *HIST1H3C*) and less often in 3.2 (*HIST2H3C*), and glycine to arginine or valine substitution at position 34 (G34R/V) in histone 3.3 [22, 73, 78, 90]. H3.3 K27M mutations have been occasionally found in posterior fossa ependymoma, and different variants in the histone genes have been reported in other childhood tumours such as H3.3 G34W/L in 92% of giant cell tumour of the bone, and *H3F3A/B* K36M in 95% of chondrosarcoma [91-93].

Later studies have demonstrated that H3 K27M mutations and H3.3 G34R/V are associated with different anatomical location, age and outcome, and may represent distinct entities **{Figure 1-6}** [22]**.** H3.3 K27M mutations are distributed throughout the midline structures (thalamus, brainstem, cerebellum and spine), H3.1 K27M are restricted to the pons and H3.3 G34R/V tumours are found uniquely in the cerebral hemispheres [22]. H3 K27M tumours occur at younger age, and have a worse prognosis compared to G34R/V mutant patients, while H3.1 K27M have a slightly better survival and occur in a younger age group than H3.3 K27M [22]. Moreover, H3.1 and H3.3 K27M tumours appear to show unique transcriptional and epigenetic signatures [94-96].

pHGG and DIPG present a striking correlation of age and molecular pattern of driver alterations, supporting the hypothesis that the cell of origin and developmental time-point in which the initiating mutations occur will determine the molecular sub-type [97, 98]. Several studies have demonstrated that H3 K27M contributes to self-renewal, gliomagenesis and tumour growth when introduced into embryonic or postnatal neural progenitor cells [99-101]. There is evidence that H3 K27M is likely to be the first initiating event in DMG-K27M tumours, however H3 K27M is insufficient to generate tumourigenesis alone. With the addition of *TP53* and *PDGFR* alterations however, tumours were generated recapitulating key features of pHGG and DIPG [102-105]. In

addition, other studies have shown that H3 K27M depleted tumours have expression signatures associated with differentiation along the glial lineage, supporting the idea that H3 K27M maintains the cells in a stem-cell like status [97, 106, 107].

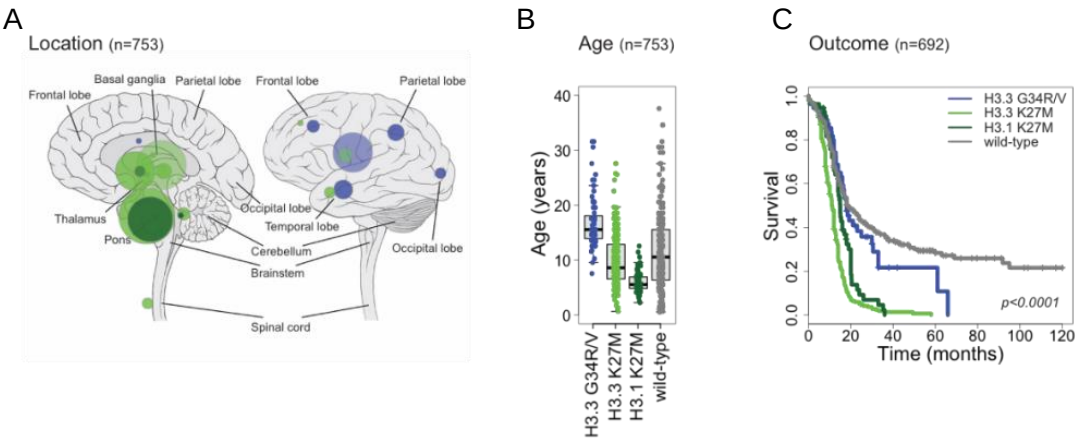


Figure 1-6 Anatomical location, age at diagnosis and clinical outcome of paediatric high grade glioma subgrouped by recurrent hotspot mutations in histone H3.3 and H3.1 genes. The different colours represent the molecular status of the histones: i) H3.3 G34R/V are shown in blue and are found exclusively in the cerebral hemispheres, ii) H3.3 K27M in green present in the midline structures, iii) H3.1 K27M in darker green are mainly restricted to the pons and iv) patients wild-type for the histones genes are displayed in grey. (A) Anatomical location of 755 cases separated by histone mutation, the size of the circles is proportional to the number of cases. The lighter shaded circles correspond to cases that are not located in any of the main three locations described (hemispheric, midline or brainstem). (B) Box-plot showing the age distribution across 753 cases per histone mutation. The thick line represents the median, the lower and upper limits represent the first and the third quartiles, and the whiskers the interquartile range. (C) Kaplan-Meier plot of overall survival across 692 cases per histone mutation. The p-value was calculated using log rank test. Figure modified from the meta-analysis carried by the Jones lab [22].

The mutations are located at or close to key regulatory sites of the histone tail marked by post-translational modifications (PTMs), modulating gene transcription by methylation, acetylation, phosphorylation and ubiquitination **{Figure 1-7}** [108]. H3 K27M mutation results in a global reduction of H3 di- and trimethylation (H3K27me2 and H3K27me3), even when the mutant proteins constitute 4-18% of the total histone pool [109-111]. H3 K27M has a dominant negative effect on EZH2, an enzymatic subunit of polycomb repressive complex 2 (PRC2). PRC2 inhibition prevents H3K27me3 deposition leading to extensive transcriptional reprograming in K27M gliomas [110, 111]. H3.3 G34R/V tumours present reduced H3K36 trimethylation (H3K36me3) as well as lower SETD2 methyltransferase activity [109]. Interestingly a study by the Jones lab led to the discovery that H3.3 G34 mutation upregulates MYCN through redistributed H3K36 binding [112].

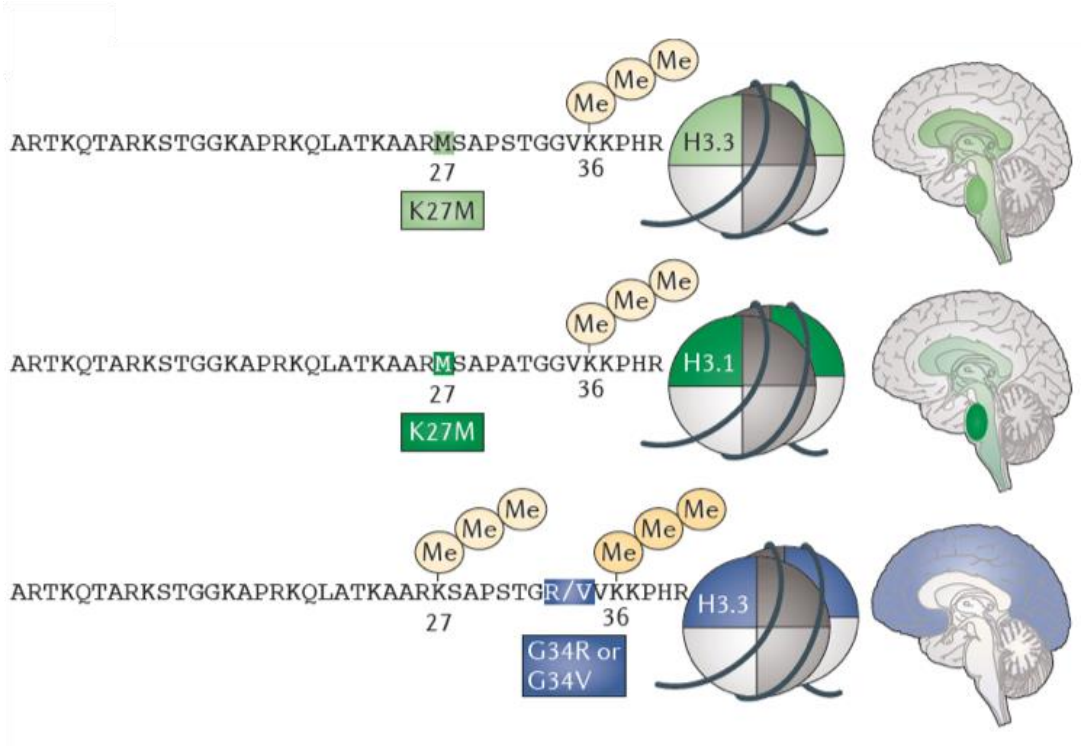


Figure 1-7 Histone mutations in paediatric high-grade glioma and DIPG. Genomic location and anatomical location are shown for K27M in H3.3 (thalamus and pons) and H3.1 (pons) and H3.3 G34R/V (cerebral hemispheres). Figure taken from [108].

1.2.1.2.5 The Blood Brain Barrier

The blood brain barrier (BBB) is a multicellular vascular structure composed of endothelial cells connected by tight junctions surrounded by pericytes and astrocytes, which regulate homeostasis of the CNS **{Figure 1-8}** [113]. Transporter proteins play an important role mediating the uptake/influx rates of certain molecules [114]. Together, these members of the BBB are responsible for the selective transport of molecules, allowing the traffic of essential metabolites (oxygen, glucose, amino acids and electrolytes), but preventing paracellular diffusion and blocking toxic molecules and pathogens from crossing into the brain parenchyma [114]. Whilst the role of the BBB is to protect the CNS from harmful substances, this often implies a restricted penetration of many compounds utilised in the treatment of CNS disease.

Effective treatment of brain tumours requires delivery of the adequate drug concentration to the tumour site. To achieve this, compounds need to be able to cross the BBB. Multiple ABC transporters expressed in the BBB, are involved in drug resistance, in particular, P-glycoprotein (P-gp or ABCB1) and multidrug resistance proteins (MRPs) [113]. Many compounds have been identified to present high affinity

for the efflux pumps P-gp and MRP1, resulting in a decreased uptake rate of therapeutic agents used to treat CNS tumours [115]. In addition, efflux and influx transporters have a different spatial expression across distinct CNS locations [114]. Although the data is limited there is some evidence that BBB in the brainstem is more intact, having poorer permeability compared to other regions of the brain, limiting even further drug penetration in DIPG patients [114, 116, 117].

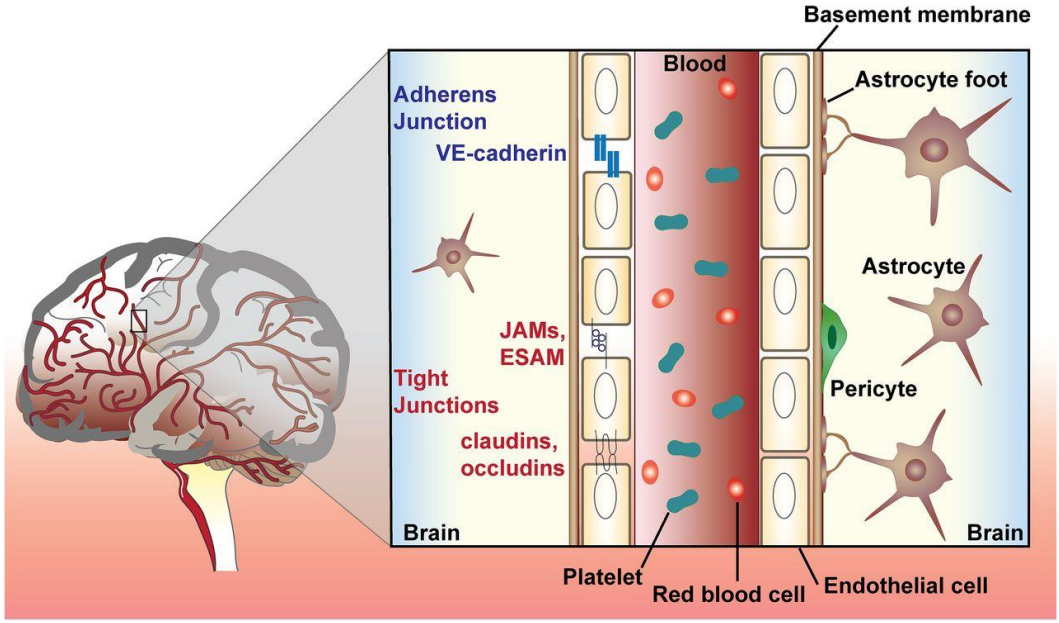


Figure 1-8 Cartoon showing the major components of the brain blood barrier (BBB). The low permeability of the BBB is due to its composition. A monolayer of endothelial cells bound together by very strong of cellular junctions: intracellular adherens junctions (cadherin, actinin and catenin) and tight junctions (occluding, claudin, and junction adhesion molecules). In addition, endothelial cells are surrounded by astrocytes and pericytes. Figure taken from [118].

1.2.1.2.6 New treatment strategies in pHGG and DIPG

New treatment strategies based on the biology underlying pHGG and DIPG have emerged including targeted agents, the use of immunotherapy and oncolytic viruses. In addition, novel methods of drug delivery such as convection-enhanced delivery (CED), which employs catheters to deliver drugs directly into the brain, or super selective intra-arterial cerebral infusion (SIACI), are currently being assessed to bypass the BBB (NCT03086616, NCT03566199, NCT01502917, NCT01884740) [119-123]. Additionally, focused ultrasound (FUD) is a promising non-invasive method to disrupt the BBB and there is there is an ongoing clinical evaluating its safety and feasibility in high-grade glioma (NCT03551249). FUS delivers low frequency ultrasound waves leading to disruption of endothelial cells and causing BBB drug permeability [124].

Pre-clinical work using patient derived *in vitro* and *in vivo* models have identified promising targets against commonly altered genes in pHGG and DIPG such as the HDAC inhibitor, panobinostat, the dual HDAC/PI3K inhibitor, CUDC-902, dabrafenib a BRAF inhibitor, as single agent or in combination with the MEK inhibitor trametinib, EZH2 inhibitors (GSK343, EPZ6438), ALK2 inhibitors (LDN-193189 and LDN-214117), the combination of the PDGFRA inhibitor dasatinib with the mTORi everolimus, BMI-1 inhibitor PTC-596, among others [100, 125-131]. Additionally, the use of immunotherapy by using CAR-T therapy (anti-GD2 CAR-T), peptide vaccine therapy specifically to target K27M-mutated DMG among others, and immune checkpoint inhibition (pidilizumab, nivolumab, durvalumab) are promising therapeutic strategies for pHGG and DIPG tumours [97, 130, 132-137]. In particular, the use of immune checkpoint inhibition in pHGG with biallelic MMRD has been demonstrated to have a clinical benefit [138]. Furthermore, the use of oncolytic viruses is an interesting field that is emerging to tackle pHGG and DIPG, in particular the adenovirus Delta-24-RGD has shown promising pre-clinical results in pHGG and DIPG mouse models [139, 140]. Some of these studies and others have led to phase I/II clinical trials in pHGG and DIPG **{Table 1-1}.**

Clinical Trial	Agent	Therapy type	Indication
NCT02717455	Panobinostat	HDAC inhibitor	DIPG
NCT03893487	Fimepinostat (CUDC-907)	Dual HDAC inhibitor PI3K inhibitor	DIPG and pHGG
NCT03363217	Trametinib	MEK1/2 inhibitor	Glioma with MAPK/ERK activation
NCT02684058	Dabrafenib + Trametinib	BRAF inihibitor/MEK1/2 inhibtior	Glioma BRAF-V600E positive
NCT03352427	Dasatinib + Everolimus	PDGFRA inhibtor/mTOR inhibitor	DIPG and pHGG
NCT03355794	Ribociclib + Everolimus	CDK inhibitor/mTOR inhibitor	DIPG and RB positive
NCT03605550	PTC-596	BMI1 inhibtor	DIPG and pHGG
NCT03749187	BGB-290 + TMZ	PARP inhibtior	IDH1/2 glioma
NCT03581292	Veliparib (ABT-888) + TMZ + Radiation	PARP inhibtior	K27M and BRAF-V600E wild-type glioma
NCT04164901	AG881-C-004 (Vorasidenib)	Pan IDH inhibitor	IDH1/2 glioma
NCT03528642	CB-839 + TMZ + Radiation	Glutaminase Inhibitor CB-839 Hydrochloride	IDH1/2 glioma
NCT03598244	Savolitinib (AZD6094)	MET inhibitor	Recurrent/Refractory CNS tumours with MET activation
NCT03696355	GDC-0084	PI3K/AKT inhibtor	DMG-K27
NCT03416530	ONC201	DRD2 inhibitor	DMG-K27
NCT02644460	Abemaciclib (LY2835219)	CDK 4/6 inhibitor	DIPG or Recurrent/Refractory Solid Tumours
NCT03620032	Nimotuzumab + Vinorelbine + Radiation	EGFR inhibitor	DIPG
NCT01884740	SIACI Erbitux (Cetuximab) + Bevacizumab*	EGFR inhibitor/VEGF inhibitor	Relapsed or refractory glioma
NCT03690869	REGN2810 (Cemiplimab) + Radiation	PD-1 inhibitor	Relapsed or refractory glioma
NCT02992964	Nivolumab	PD-1 inhibitor	Hypermutant cancers
NCT02359565	Pembrolizumab	PD-1 inhibitor	Recurrent/Refractory CNS tumours
NCT02793466	Durvalumab	PD-L1 inhibitor	Recurrent/Refractory paediatric tumours
NCT04049669	Indoximod + TMZ + Radiation	indoleamine 2,3-dioxygenase inhibitor	DIPG and pHGG and recurrent CNS
NCT04212351	NF1 Frameshift Peptides	Vaccine	Neurofibromatosis Type 1 glioma
NCT02722512	HSPPC-96 + surgery + radiation	Vaccine	pHGG and ependymoma

NCT01130077	HLA-A2 restricted glioma antigen peptides vaccine	Vaccine	pHGG
NCT03879512	depletion of regulatory T cells + surgery + vaccine	Vaccine	Resectable pHGG
NCT03334305	anti-tumor T-cells and anti-tumor DC vaccines + TMZ	Vaccine	Resectable pHGG
NCT03916757	V-Boost Immunitor	Vaccine	GBM
NCT04185038	SCRI-CARB7H3(s) CAR-T**	CAR-T	DIPG, DMG, recurrent/refractory CNS
NCT03500991	HER2-specific CAR T	CAR-T	HER2-positive CNS tumor no DIPG
NCT02208362	Autologous IL13(EQ)BBzeta/CD19t +Tcm-enriched T cells****	CAR-T	DIPG and pHGG
NCT03911388	G207 oncolytic herpes simplex virus-1 (HSV)***	Virus therapy	CNS tumor no DIPG
NCT03043391	Polio/Rhinovirus Recombinant (PVSRIPO)*****	Virus therapy	pHGG
NCT03178032	DNX-2401 adenovirus	Virus therapy	DIPG
NCT03086616	CED With Irinotecan Liposome Injection	Chemotherapy	DIPG

*SIACI (Super-selective Intra-arterial Cerebral Infusion)
**Catheter into the tumour resection cavity or ventricular system
***Infused through catheters into region(s) of tumour
****Given via intratumourally, intracavitary or intraventricular catheter
*****Delivered by CED (convection-enhanced delivery)

Table 1-1 Summary table including current clinical trials available for children with paediatric high-grade glioma and DIPG. The table includes clinical trial ID, agent, therapy type and drug indication.

1.2.2 Medulloblastoma

Medulloblastoma (MB) accounts for 20% of all paediatric brain tumours with a peak occurrence of 6-8 years of age and a male predominance overall (male-to-female ratio of 1.8:1) [141]. MBs encompass a group of embryonal tumours in the cerebellum presenting distinct clinical, histological and biological features. There are known syndromes associated with germline mutations involved in MB pathogenesis that increase MB predisposition [142].

In 2012, four molecular subgroups of medulloblastoma were proposed and are currently recognised by the WHO classification of CNS tumours: WNT-MB (*CTNNB1* mutations present in ~80%), SHH-MB (further divided in *TP53* wild-type and *TP53*-mutant), Group 3 MB and Group 4 MB [23]. Each of these four subgroups have subsequently been further subdivided according to unique methylation and expression patterns associated with distinct clinical presentations **{Figure 1-9}** [143-145]. Due to the distinctive epigenetic and transcriptional profiles of the different MB subgroups, they are likely to have different cellular origins [146]. In this line of work two recent studies using single-cell RNA sequencing have demonstrated that MB exhibit putative subgroup-specific origins, with Group 3 being dominated by undifferentiated progenitor-like cells, whereas Group 4 MB is constituted by neuronal-like cells [147, 148]. There appears also to be a subset of intermediate tumours by DNA methylation which exhibit multiple lineages and stages of cell differentiation, demonstrating the challenge associated with confidently classifying MB Group 3 and 4 [147, 148].

Recent studies have identified germline mutations in ~6% of MB, involving *APC, BRCA2, PALB2, PTCH1, SUFU, GPR161* and *TP53* as consensus medulloblastoma predisposition genes with higher incidence in MB-WNT and MB-SHH subgroups [149, 150]. A recent study just published by Waskzak *et a*l., has identified the presence of germline loss-of function variants across *ELP1* in 14% of MB-SHH and found to be associated with protein homeostasis [151]. In addition, a study by Hiromichi Suzuki et al. have identified highly recurrent hotspot mutations (r.3A>G) of U1 spliceosomal small nuclear RNA (snRNAs) in 50% of SHH-MB [152].

Surgical resection followed by risk-adapted cranio-spinal irradiation and adjuvant chemotherapy are used to treat MB. Due to treatment aggressivity, “cured” children can develop multiple complications some of which can be devastating, in addition to

secondary malignancies including radiation-induced glioblastoma [54, 142]. Therefore, there are ongoing trials adjusting radiation intensity especially for WNT-MB which have a better prognosis (NCT02066220, NCT02212574 and NCT02724579) [142, 153]. Moreover, NCT01878617 is a phase 2 trial in which patients are stratified to different treatment approaches (radiation, chemotherapy and SHH inhibitor) based on their clinical risk and molecular sub-type.

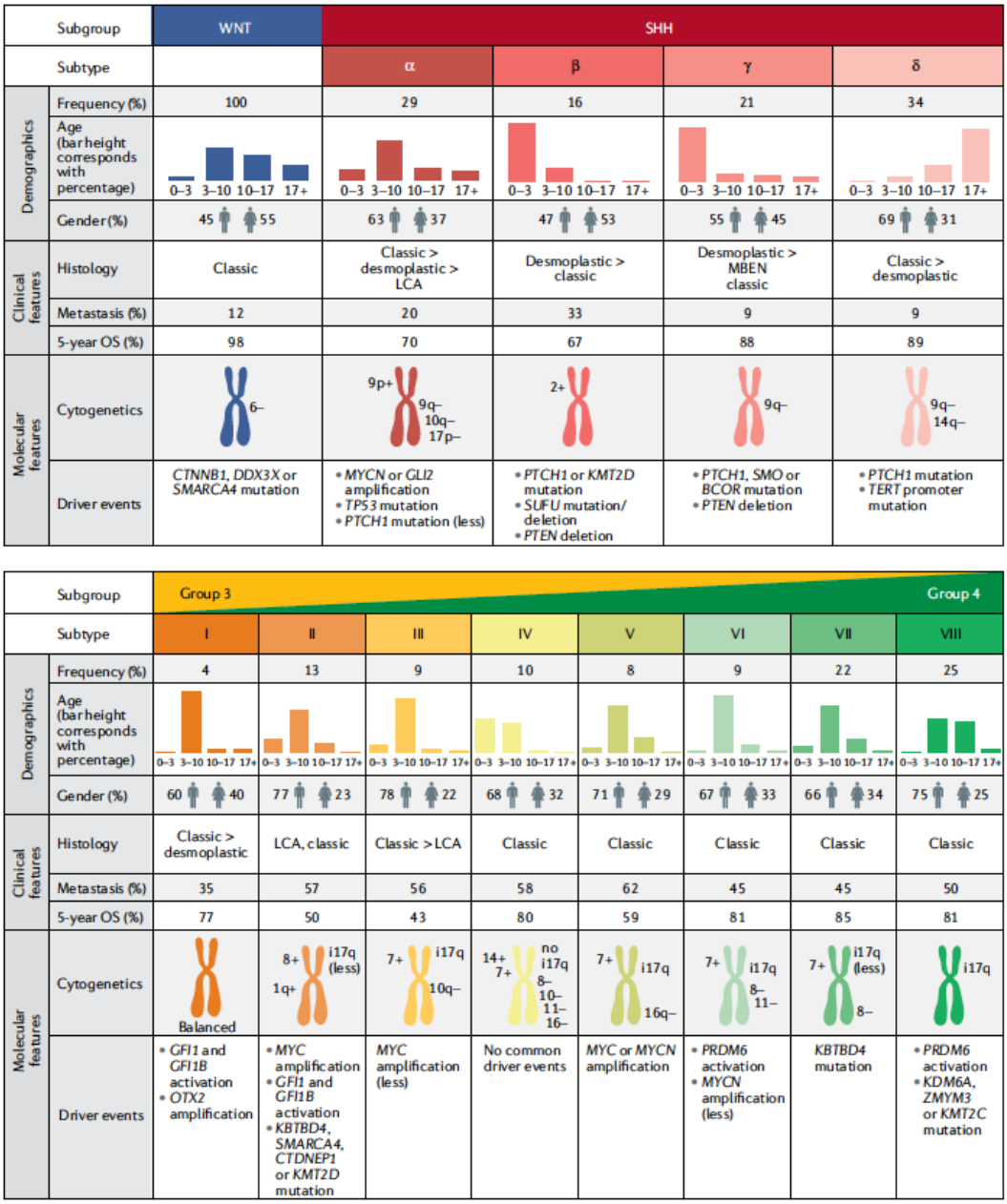


Figure 1-9 Schematic representation of medulloblastoma sub-types. The figure includes demographics, clinical and molecular features derived from the following studies [143, 153-155]. Figure taken from [145].

1.2.3 Atypical teratoid rhabdoid tumour

Atypical teratoid/rhabdoid tumours (ATRT) are very aggressive tumours of the CNS that occur in younger children, and are one of the most common CNS tumours in children below 1 year of age [25, 156]. ATRT is an embryonal tumour arising in diverse locations including posterior fossa, diencephalon, cerebrum and midbrain [157]. Overall survival of ATRT is poor, and patients usually succumb to their disease between 6 months and 1 year from diagnosis [158]. ATRT treatment varies depending on the location, age and disease stage at diagnosis [158]. Surgery, craniospinal radiation and intensive chemotherapy are used to treat ATRTs, however these therapies are often not an option, especially in patients of young age.

The hallmark genetic event of ATRT is the inactivation of *SMARCB1* and less often *SMARCA4,* both members of the SWI/SNF chromatin-remodelling complex. The loss of SMARCB1 has shown to affect the epigenome by a global loss of H3K27ac and H3K27me3 [159]. A study by Grotzer *et al.* has demonstrated that ATRT is not a homogeneous disease, and although most ATRT tumours have *SMARCB1* inactivation they show distinct epigenetic and expression patterns determining that ATRT is comprised of three epigenetic subgroups: *ATRT-TYR, ATRT-SHH and ATRT-MYC. ATRT-SHH* has been further subdivided in two groups *ATRT-SHH-1* with predominant supratentorial location and *ATRT-SHH-2* with infratentorial location showing a distinct methylation profile **{Figure 1-10}** [160]. Additionally, an integrative meta-analysis has identified that *ATRT-MYC* exhibit increased infiltration of CD8+ cytotoxic cells which has been associated with response to immune checkpoint inhibition in other tumour types [161].

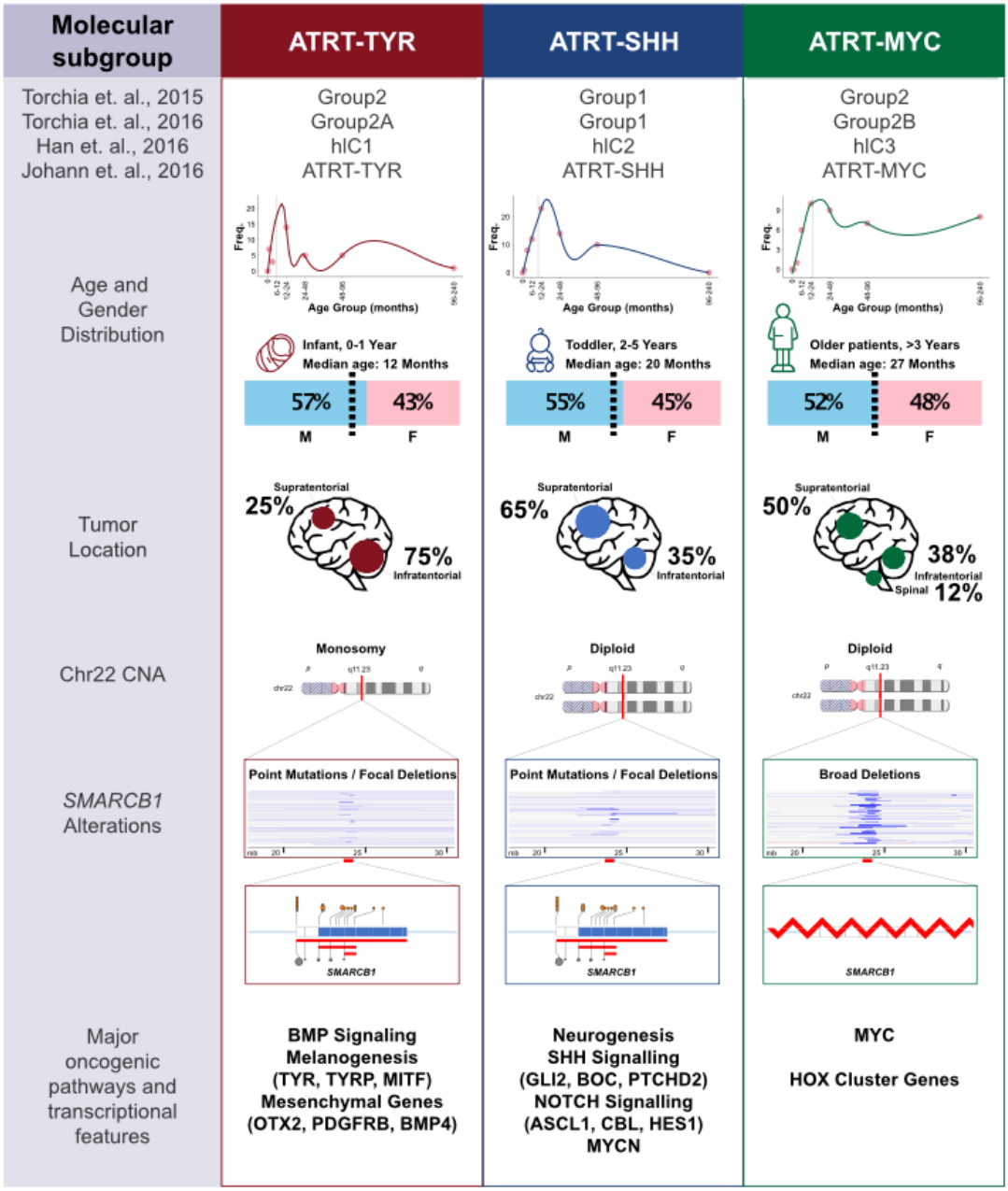


Figure 1-10 Schematic representation of ATRT subgroups. The figure includes demographics, tumour location, chromosome 22 copy number alteration, SMARCB1 type of alterations and genetic/transcriptional features associated to the three different ATRT subgroups (ATRT-TYR, ATRT-SHH and ATRT-MYCN). Figure taken from [160].

1.2.4 Ependymoma

Ependymomas are neuroepithelial tumours accounting for 6-10% of all paediatric brain tumours with nearly 90% occurring intracranially (2/3 located in the posterior fossa, PF and 1/3 are supratentorial, SP), and the remaining 10% in the spine [25, 162, 163]. Ependymomas can arise at any time during childhood, though they present

a peak at 0-4 years of age with a male-to-female ratio of 1.77:1 [164]. The familial syndrome Neurofibromatosis Type 2 (NF2) has been associated with an increased risk of developing ependymoma, especially in the spine [165]. Maximal safe resection followed by adjuvant therapy (radiotherapy with or without chemotherapy) is used for the treatment of ependymoma, except for infants [166, 167]. Paediatric ependymoma patients exhibit inferior survival rates compared to adults and have dismal outcomes, with a 10-year OS rate of ~50% and PFS of ~29%, with younger age patients presenting a worse outcome [168].

An important study carried out in 2015 by Pajtler *et al.* using genome-wide DNA methylation on 500 ependymoma-naïve patients revolutionised the original classification of ependymoma, which was previously based solely on histological features. This study identified nine distinct ependymoma entities associated with clinical and molecular characteristics **{Figure 1-11}** [24]. Supratentorial tumours are subclassified as *EPN-ST-RELA, EPN-ST-YAP1*, both subgroups defined by the presence of highly recurrent gene fusions in the NF-kB subunit gene *RELA* or the transcription factor *YAP1*, as well as a third less common subgroup *EPN-ST-SE* (subependymoma) with an unknown driver event. Posterior fossa tumours are subclassified as *EPN-PF-A* with a very poor prognosis, *EPN-PF-B* and *PF-PF-SE* as slower growing groups of tumours. A subset of *EPN-PF-A* present low H3K27me3 and express high levels of EZHIP (cXorf67), and less often (4.2%) *H3.3*-K27M mutation [169, 170]. These two mechanisms are mutually exclusive features of EZH2 inhibition, the enzymatic component of the Polycomb Repressive Complex 2 (PRC2) modulating H3K27me3 in *EPN-PF-A* [171, 172]. In addition, subgroup-specific super-enhancer-driven genes have been identified as potential therapeutic targets including *HDAC7, EPHA2, FGFR1* and *CACNA1H* [173]. Spinal ependymal tumours are comprised of three subgroups which show better prognosis and have good concordance with histological grading. These include MPE (grade I, subependymomas and myxopapillary) and EPN (grade II, classic ependymomas) [24], as well as a new molecular subgroup, *SP-EPN-MYCN*, defined as a very aggressive tumour with the presence of *MYCN* amplification and similar OS and PFS to *ST-EPN-RELA* and *PF-EPN-A* [174].

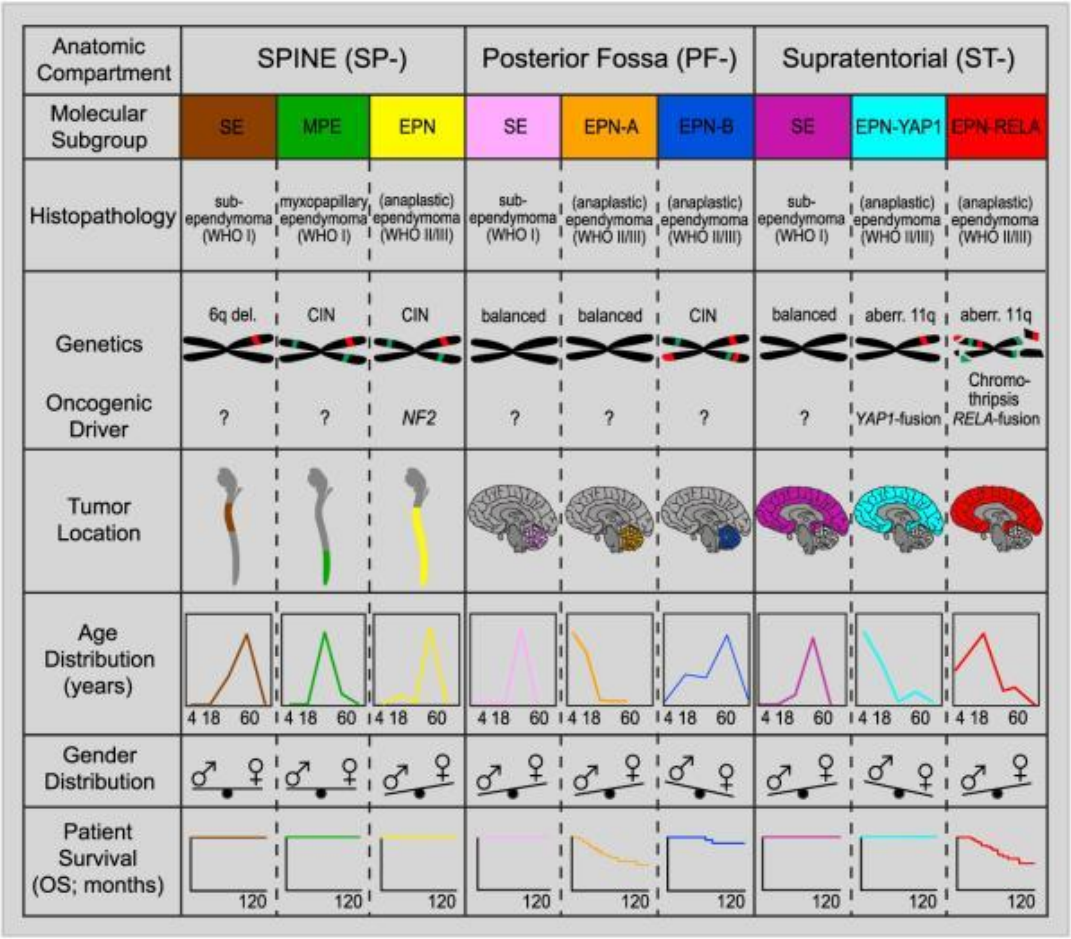


Figure 1-11 Schematic representation of the nine ependymoma identified by methylation profiling by Patjtler and colleagues. The figure shows the following associated features: histopathology, genetic alterations, tumour location, age and gender distribution as well as patient survival. Figure taken from [24].

1.2.5 Embryonal tumours with multi-layered rosettes

ETMR (embryonal tumours with multi-layered rosettes) is a very aggressive embryonal tumour which occurs almost exclusively in infants and young children under 4 years of age. ETMR tumours are histologically heterogeneous, variously described as ependymoblastoma, medulloepithelioma or embryonal tumour with abundant neuropil and true rosettes [26]. ETMR is genetically characterised by the presence of amplification and fusion of a miRNA cluster on chromosome 19 (*C19MC*) with *TTYH1 (or others, MIRLET7BHG)* as well as the overexpression of LIN28A protein [175-178]. Sin-Chan *et al.* discovered a C19MC-LIN28A-MYCN super-enhancer-dependent oncogenic circuit in ETMR arising in embryonic neural progenitors, and identified *in vitro* sensitivity to the BET inhibitor JQ1 [179]. Furthermore, a study by Lambo *et al.* have identified germline mutations in *DICER1*

or somatic amplifications of *mir17-92* (*MIR17HG*) in ETMRs which are *C19MC* non-amplified [180]. Additionally, by using WGS they identified a high frequency of structural variants and R-loops many of which were found surrounding *C19MC*, and showed that topoisomerase and PARP inhibitors might be an effective treatment for ETMR [180]. A schematic representation of ETMR with clinical and molecular features is shown in **{Figure 1-12}**.

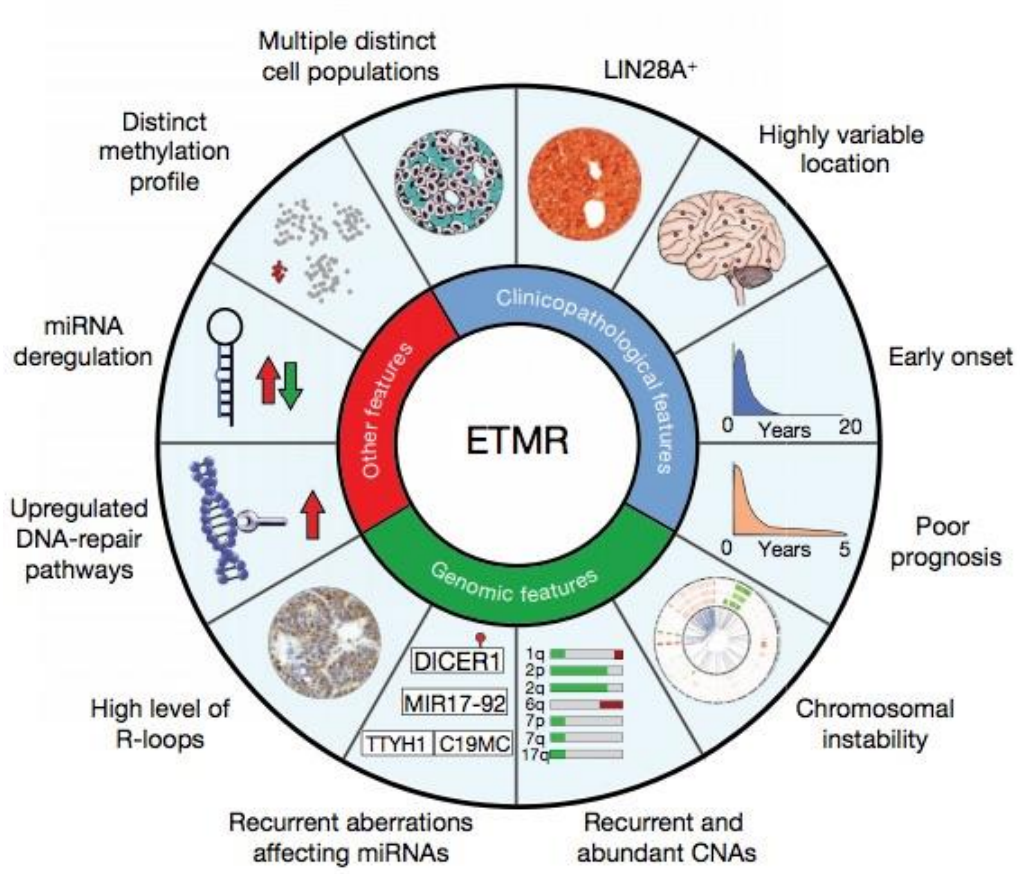


Figure 1-12 Schematic representation of ETMR features. Figure taken from [180].

1.2.6 CNS-PNET

CNS-PNET (primitive neuroectodermal tumours) originally constituted a molecularly heterogenous group of embryonal tumours characterised by an aggressive behaviour (5-year overall survival of ~50%), undifferentiated or poorly differentiated with the presence of variable type of cells - neuronal, ependymal or glial [181, 182]. However, in 2016 a study by Sturm *et al.* revealed that many of the designated CNS-PNET tumours correspond to well-defined CNS tumours (ETMRs, gliomas, ependymomas, medulloblastomas amongst others). In addition, they identified four new entities, each

of them associated with specific genetic alterations, clinical and histopathological features [21]. The recurrent molecular alterations included *FOXR2* rearrangements with increased gene expression in *CNS NB-FOXR2* (CNS neuroblastoma with FOXR2 activation), genetic alterations involving the capicua transcriptional repressor *CIC* in *CNS EFT-CIC* (CNS ewing sarcoma family tumour with *CIC* alterations), interchromosomal gene fusion implicating the *MN1* gene in CNS *HGNET-MN1* (CNS high-grade neuroepithelial tumour with *MN1* alterations) and tandem duplication as well as frameshift mutations in *BCOR* characteristic of CNS *HGNET-BCOR* (CNS high-grade neuroepithelial tumour with *BCOR* alteration). In addition, the authors showed that the four new entities exhibit distinct gene expression signatures including various transcription factors and potential drug targets.

1.3 Extracranial paediatric solid tumours

1.3.7 Neuroblastoma

Neuroblastoma is an embryonal tumour that arises from neural-crest derived progenitor cells and is the most common extracranial paediatric solid tumour. Neuroblastoma has a median age diagnosis of 18 months and occurs almost exclusively in children, with 90% of patients diagnosed under 10 years of age [183]. Neuroblastoma is characterised by heterogenous clinical manifestations, from spontaneous regression, often in infant *MYCN* non-amplified cases, to treatment-refractory progression, mostly in older children who can succumb to the disease despite months of intensive therapy [184-186]. A recent study from Ackermann *et al.* have proposed three different neuroblastoma subgroups defined by the absence (low risk) or presence of telomerase activation (high risk) with dramatically inferior survival when accompanied by RAS/p53 pathway alterations (very high risk) **{Figure 1-13}** [187]. Common genetic events driving neuroblastoma are *MYCN* amplifications, *ALK* germline and spontaneous alterations including point mutations and amplifications, *ATRX* inactivating alterations as well as *TERT* rearrangements [188-195]. In addition, extrachromosomal circular DNA (ecDNA) has recently been found to be a frequent cause of somatic rearrangements contributing to genome remodelling in neuroblastoma [196].

Extensive pre-clinical efforts have been made to translate findings from the underlying biological features of neuroblastoma to the clinic, however to-date, these have mostly been early-phase clinical trials for patients with recurrent high-risk tumours [197].

Clinical trials using the first generation *ALK* inhibitor crizotinib have shown differential sensitivity identifying a group of patients harbouring mutations in *ALK* (F1174L and F1245C) who were resistant to the drug [198]. Therefore, clinical trials using second generation inhibitors such as ceritinib (NCT01742286) and entrectinib (NCT02650401, NCT02097810) are currently under evaluation [199]. A phase I trial led by Yael Mosse at the Children's Hospital of Philadelphia using the AURKA inhibitor alisertib (MLN8237) with irinotecan and temozolomide has shown promising response and progression-free survival rates, however this was not associated with *MYCN* status [200].

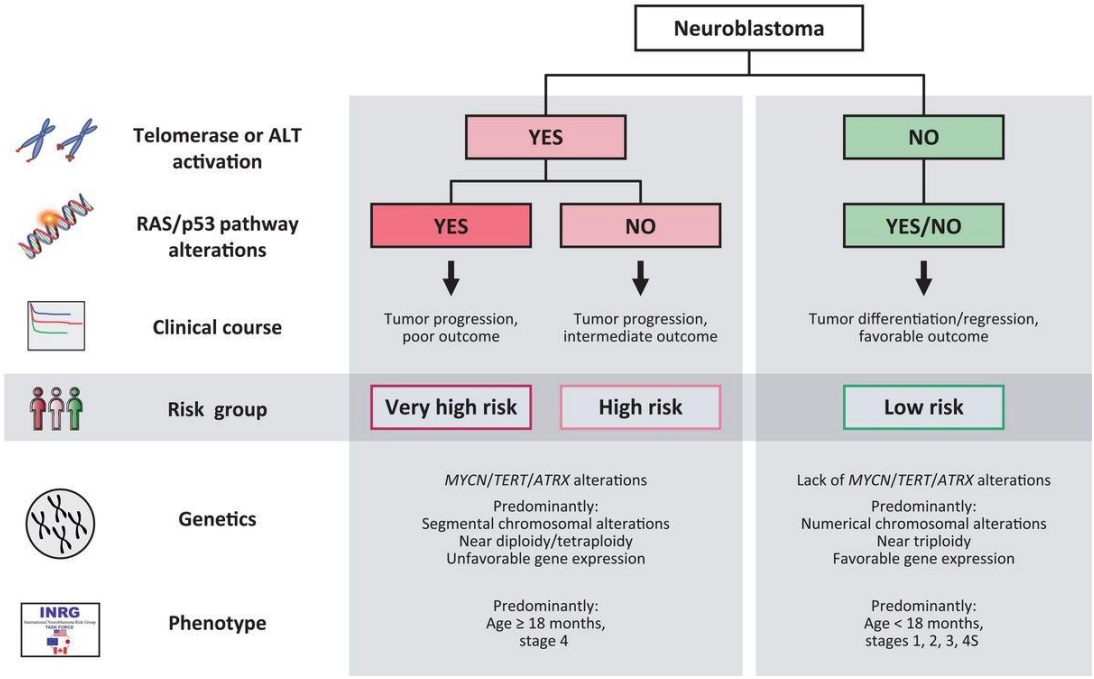


Figure 1-13 Schematic representation of the proposed neuroblastoma subgroups by S. Ackermann and colleagues. Neuroblastoma samples are proposed to be classified on the basis of telomerase or ALT activation as well as RAS and TP53 pathway alterations associated with specific genetic and phenotypic features. Figure taken from [187].

1.3.8 Sarcomas

Sarcomas are a heterogeneous group of tumours with distinct clinical and pathological features, representing 10-15% of childhood cancers compared to 1% in the adult population [201, 202]. They are comprised of bone and soft tissue tumours and are generally classified according to the type of tissue that they arise from. Despite improvements in survival, the outcome for patients with metastatic or recurrent sarcomas remain dismal [203].

Accurate diagnosis involves the detection of characteristic gene fusions highly associated with specific sarcoma sub-types, which can be challenging as there are over 140 different fusions reported **{Table 1-2 }** [204, 205]. Genome-wide methylation array can be a powerful tool to produce unequivocal diagnosis, supported by a recent study by Koelsche *et al.* who have used DNA-methylation status to precisely classified undifferentiated tumours with small blue round cell histology into well-defined sarcoma entities [206].

Histology	Chromosomal aberration	Involved genes
Alveolar soft parts sarcoma	t (x;17)(p11.2;q25)	ASPL/TFE3
Angiomatoid fibrous histiocytoma	t (12;16)(q13;p11), t (2;22)(q33;q12), t (12;22)(q13;q12)	FUS/ATF1, EWSR1/CREB1, EWSR1/ATF1
BCOR-rearranged sarcoma	inv (X)(p11.4;p11.2)	BCOR/CCNB3
CIC-rearranged sarcoma	t (4;19)(q35;q13), t (10;19)(q26;q13)	CIC-DUX4
Clear cell sarcoma	t (12;22)(q13;q12), t (2;22)(q33;q12)	ATF1/EWSR1, EWSR1/CREB1
Congenital (infantile) fibrosarcoma/ mesoblastic nephroma	t (12;15)(p13;q25)	ETC-NTRK3
Dermatofibrosarcoma protuberans	t (17;22)(q22;q13)	COL1A1/PDGFB
Desmoid fibromatosis	Trisomy 8 or 20, loss of 5q21	CTNNB1 or APC mutations
Desmoplastic small round cell tumor	t (11;22)(p13;q12)	EWSR1/WT1
Epithelioid hemangioendothelioma	t (1;3)(p36;q25)	WWTR1/CAMTA1
Epithelioid sarcoma	Inactivation SMARCB1	SMARCB1
Extraskeletal myxoid chondrosarcoma	t (9;22)(q22;q12), t (9;17)(q22;q12), t (9;15)(q22;q21), t (3;9)(q11;q22)	EWSR1/NR4A3, TAF2N/NR4A3, TCF12/NR4A3, TGF/NR4A3
Hemangiopericytoma	t (12;19)(q13;q13.3) and t (13;22)(q22;q13.3)	LMNA-NTRK3
Infantile fibrosarcoma	t (12;15)(p13;q25)	ETV6/NTRK3
Inflammatory myofibroblastic tumor	t (1;2)(q23;q23), t (2;19)(q23;q13), t (2;17)(q23;q23), t (2;2)(p23;q13), t (2;11)(p23;q15)	TPM3/ALK, TPM4/ALK, CLTC/ALK, RANBP2/ALK, RAS
Infantile myofibromatosis		PDGFRB
Low-grade fibromyxoid sarcoma	t (7;16)(q33;p11), t (11;16)(p11;p11)	FUS/CREB3L2, FUS/CREB3L1
Malignant peripheral nerve sheath tumor	17q11.2, loss or rearrangement 10p, 11q, 17q, 22q	NF1
Mesenchymal chondrosarcoma	Del[8](q13.3;q21.1)	HEY1/NCOA2
Myoepithelioma	t (19;22)(q13;q12), t (1;22)(q23;q12), t (6;22)(p21;q12)	EWSR1/ZNF44, EWSR1/PBX1, EWSR1/POU5F1
Myxoid/round cell liposarcoma	t (12;16)(q13;p11), t (12;22)(q13;q12)	FUS/DD1T3, EWSR1/DD1T3
Primitive myxoid mesenchymal tumor of infancy		BCOR internal tandem duplications
Rhabdoid tumor	Inactivation SMARCB1	SMARCB1
Sclerosing epithelioid fibrosarcoma		EWSR1/CREB3L2
Solitary fibrous tumor	Inv[12](q13;q13)	NAB2/STAT6
Synovial sarcoma	t (x;18)(p11.2;q11.2)	SYT/SSX
Tenosynovial giant cell tumor	t (1;2)(p13;q35)	COL6A3/CSF1

Table 1-2 List of frequent chromosomal aberration and genes involved by histology in paediatric soft tissue sarcomas. Table taken from [207]

1.3.8.3 Soft tissue sarcomas

Rhabdomyosarcoma (RMS) is the most common soft-tissue sarcoma in children accounting for ~7% of paediatric malignancies [208]. RMS is a mesenchymal tumour arising in many distinct anatomical sites normally associated with the musculature

[209]. Although the cell of origin has not yet been well characterised, it is thought that RMS may arise from cells programmed to express a complement of skeletal myocyte genes [210]. RMS tumours are classified histologically in two major groups, alveolar (ARMS) and embryonal (ERMS). The most common pathway and gene alterations found in RMS are shown in **{Figure 1-14}**.The prognosis for ERMS when localised is better than ARMS [211]. ARMS is mostly driven by the presence of a fusion between *PAX3* or *PAX7* and *FOXO1* genes which occurs in 80% of this group, often harbouring amplification of *MYCN* or *CDK4* [212-214]. ERMS is characterised by combination of copy number alterations (uniparental disomy of chromosome 11p, loss of chromosome 10 and 15, polysomy of chromosome 8 and gains of 11q and 7q regions) and RAS pathway activation (mutations in *NRAS, KRAS, HRAS, NF1* and *PIK3CA*) as well as mutations in *CTNNB1*, *FGFR4* and *BCOR* [214, 215]. Additionally, a rare subset of ARMS presenting spindle and sclerosing cells, defines an aggressive sub-type associated with poor outcome and is frequently driven by mutations in the myogenic transcription factor *MYOD1,* often with concomitant mutations in the PI3K-AKT pathway [216, 217]

Other non-RMS sarcomas include a diverse group of tumours of over 35 histologies often presenting distinct molecular features [218]. The most common non-RMS sarcomas are synovial sarcoma (*SS18:SSX1/2*) [219], infantile fibrosarcoma (*ETV6:NTRK3*)[220], inflammatory myofibrastic tumour (*TPM3/4/CLCTC/CARS:ALK*) [221, 222], clear cell sarcoma (*EWSR1:ATF1/CREB1*) [223], desmoplastic small round cell tumour (*EWSR1:WT1/DDIT3*) [224] and myxoid liposarcoma (*TLS:CHOP*) [225].

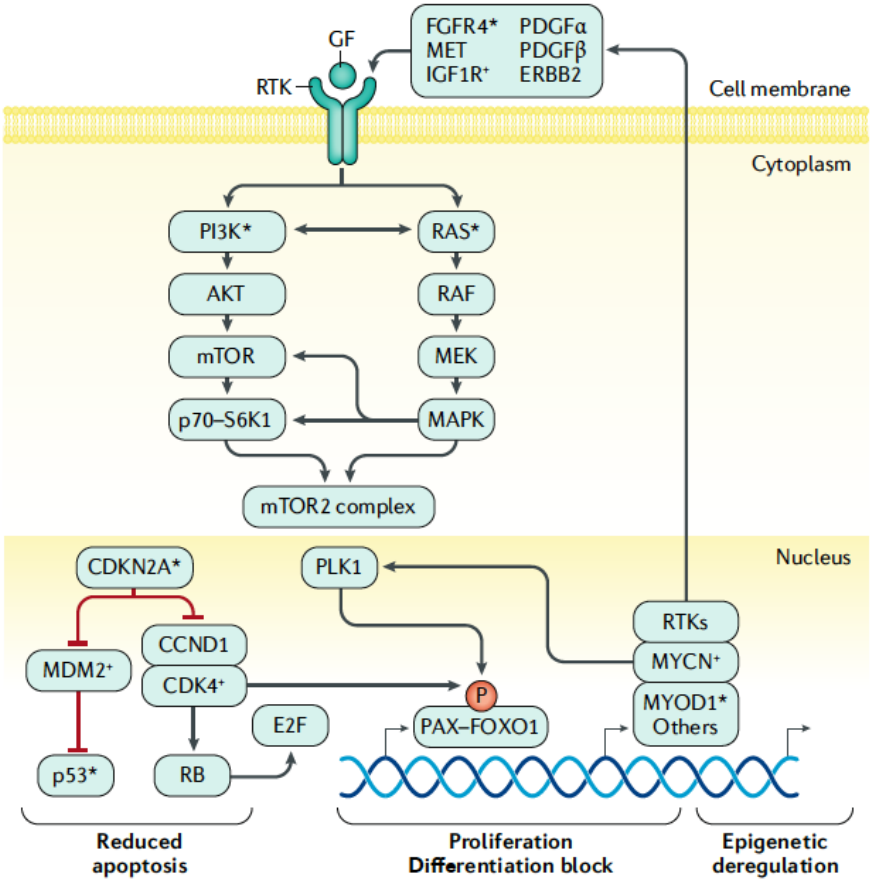


Figure 1-14 Schematic representation of the main pathways altered in rhabdomyosarcoma. Most common mutated genes are denoted by *. Figure taken from [210].

1.3.8.4 Bone sarcomas

Osteosarcoma is the most common bone cancer in children, with a poor prognosis in patients presenting metastatic disease at diagnosis [226]. Osteosarcoma is characterised by the presence of a large number of structural variants, called kataegis, with relatively low rate of recurrent point mutations. Of these the most common mutated genes are *TP53*, *RB1* and *ATRX* [227, 228]. Interestingly, structural variants in exon 1, is one of the most common mechanisms of inactivation of *TP53* [228, 229].

Ewing sarcoma is a small round cell bone tumour and is the second most frequent bone cancer in children, with a peak incidence age of 10-5 years of age. Ewing sarcoma results from a fusion between genes of the TET/FET family (*EWSR1*, *FUS* and *TAF15)* and various members of the ETS family of transcription factors (*FL1*,

ERG, FEV, ETV1, ETV4) [230]. Additional mutations in *STAG2* and *TP53* are present in 15-20% and 4-7% of ES at diagnosis [231, 232].

A rare group of tumours called Ewing-like sarcomas (ELSs) have recently been described, and although they share clinical similarities with Ewing sarcoma there is some evidence that these are distinct molecular entities [233]. ELSs are *FET-ETS* gene fusion wild-type but harbour other structural variants [234]. ELSs include *CIC*-fused sarcomas [235, 236], *BCOR*-rearranged sarcomas [237, 238] and *NFATC2* sarcomas [238, 239].

1.3.9 Kidney tumours

Wilms tumour or nephroblastoma is the most common childhood kidney cancer accounting for 90% of all malignant kidney tumours in children and 7% of childhood malignancies overall [240]. Cure rates (90%) have improved remarkably with the use of chemotherapy, surgery and less often, radiotherapy. Wilms tumours commonly harbour more than one genetic alteration and have a heterogenous gene expression as well as methylation pattern based on different genetic alterations [241]. Recurrent alterations involve *WT1, TP53, CTNNB1, MYCN, DROSA, SIX1/2, WTX, BCOR, BCORL1* and *ARID1A* amongst others [241-243]. Recent studies have found a high rate of germline mutations (*TP53, WT1, PALB2, CHEK2 DIS3L2*, *DICER1, ARID1A* and *EP300)* that may demonstrate an inherited component to Wilms formation tumours [241].

Other childhood renal tumours include renal cell carcinoma typically seen in Von Hippel-Lindau disease, with ~40% of cases presenting *TFE3* gene fusions [244, 245] and clear cell sarcoma of the kidney which is characterised by in-frame internal tandem duplication of *BCOR* [246, 247].

1.3.10 Retinoblastoma

Retinoblastoma is rare tumour of the eye arising in the retina accounting for 2-3% of all paediatric cancers, usually presenting in younger children <2 years of age. Dismal outcome is associated with late diagnosis, usually in low-income countries, presenting an overall survival of 30% compared to high-income countries >95% [248]. It is caused in most cases by germline or somatic biallelic loss of the tumour suppressor *RB1,* which lies on chromosome 13q14 and encodes for retinoblastoma protein [249].

Other alterations found in retinoblastoma include mainly copy number changes in *DDX2, KIF14, MDM4, OTX2* and loss of *CHD11* and *RBL2*, additionally a very low number of point mutations involving *BCOR* and *CREBBP* [250, 251]. Non-RB1 retinoblastoma has been reported in 2.7% cases and is probably initiated by amplification of *MYCN,* representing highly aggressive tumours harbouring specific histological characteristics with multiple nucleoli and blast cells [252].

1.4 Childhood cancer predisposition syndromes

Recent studies have suggested that about 7-9% of childhood cancer carry a known or likely pathogenic germline variant in cancer-predisposition genes, frequently in DNA repair genes such as *TP53, BRCA2, CHEK2, MSH2, MSH6* and *PMS2*, [12, 13]. However, due to the low number of studies, the prevalence of cancer predisposition syndromes in children might be underestimated and as large scale genetic testing is performed, it is possible that the percentage of cases forming part of hereditary cancer syndromes might increase [14, 253, 254]. The most common syndrome associated with cancer predisposition is Li-Fraumeni syndrome, characterised by *TP53* mutations and correlated with a high risk of developing bone and soft tissue sarcomas, brain tumours, and adrenal cortical carcinomas, among others [255]. APC-associated polyposis conditions predispose to CNS tumours (SHH-medulloblastoma), osteosarcoma and neuroblastoma [256]. Germline mutations in neurofibromatosis 1 (*NF1*) is associated with paediatric glioma (LGG and HGG) as well as malignant peripheral nerve sheath tumours [256]. Gorlin syndrome, characterised by mutations in *SUFU* and *PTHC1* is linked to young children with SHH-MB, rhabdomyosarcoma, ependymoma and foetal rhabdomyoma [256]. Rhabdoid tumour predisposition 1/2 involving *SMARCB1* and *SMARCA4* mutations are associated with renal and extrarenal rhabdoid tumours, meningioma (*SMARCB1*) and schwannomatosis (*SMARCA4*) [256]. *RB1* germline mutations predispose to retinoblastoma and high risk of secondary malignancies such as osteosarcoma, while *DICER1* mutations predispose to rhabdomyosarcoma, pituitary blastoma, PNET, and pineoblastoma among others [256]. A summary of cancer predisposition genes and syndromes is shown in **{Table 1-3}** and **{Figure 1-15}**.

APC-associated polyposis conditions (MIM 175100) (Jasperson & Burt, 1998)
Mutation of *APC*; features may include multiple mainly colonic polyps, dental anomalies, and congenital hypertrophy of the retinal pigment epithelium; benign and malignant neoplasia including colon cancer, pancreatic cancer, papillary thyroid carcinoma, medulloblastoma, hepatoblastoma, desmoid tumors, and osteomas
Beckwith–Wiedemann syndrome (MIM 130650) (Shuman, Beckwith, & Weksberg, 2000)
Complex spectrum of epigenetic and genetic alterations of chromosomal region 11p15; features may include neonatal hypoglycemia, macrosomia, macroglossia, hemihyperplasia, omphalocele, visceromegaly, adrenocortical cytomegaly, renal abnormalities, and ear anomalies; benign and malignant neoplasia including Wilms tumor, hepatoblastoma, neuroblastoma, and rhabdomyosarcoma
Down syndrome (MIM 190685) (Roberts & Izraeli, 2014)
Trisomy 21, mainly isolated cases, be aware of translocation trisomies including Robertsonian translocations involving chromosome 21 in the parents; typical phenotype with short stature, mental retardation, congenital heart defect, muscular hypotonia, single transverse palmar crease, and characteristic facial appearance including epicanthal folds, upslanting palpebral fissures, and protruding tongue; predisposes to childhood ALL, transient myeloid neoplasms, and acute megakaryocytic leukemia (AMKL), the latter two being associated with somatic *GATA1s* mutations
DICER1-related conditions (MIM 606241) (Doros, Schultz et al., 2014)
Mutation of *DICER1*; no obvious syndromic features except in rare patients with GLOW (Global developmental delay, lung cysts, overgrowth, and Wilms tumor) syndrome (*DICER1* mosaic) (Klein et al., 2014); benign and malignant neoplasia including pleuropulmonary blastoma, ovarian sex cord-stromal tumors, cystic nephroma, multinodular goiter and thyroid cancer, ciliary body medulloepithelioma, botryoid-type embryonal rhabdomyosarcoma, nasal chondromesenchymal hamartoma, pituitary blastoma, primitive neuroectodermal tumor (PNET), and pineoblastoma
Fanconi anemia (e.g., MIM 227650) (Alter, 2014)
Clinically and genetically heterogenous mostly autosomal-recessive disease caused by mutations in genes encoding proteins involved in DNA repair; homozygous or compound heterozygous mutations causing genomic instability; pre- and postnatal short stature, microcephaly, hypoplastic/aplastic thumb, small or absent radius, hyperpigmented skin macules, bone marrow failure, and developmental abnormalities in major organ systems; predisposition for acute leukemia; medulloblastoma, hepatocellular carcinoma, nephroblastoma, and also solid tumors in adulthood, e.g., squamous cell carcinoma and breast cancer
GATA2 deficiency (MIM 137295) (Wlodarski et al., 2016)
Mutation of *GATA2*, a crucial hematopoietic transcription factor; germline mutations account for 15% of primary pediatric advanced MDS and 70% of adolescents with MDS and monosomy 7; clinically heterogenous, also associated with MonoMAC/DCML syndrome (MIM 614172), Emberger syndrome (MIM 614038), might be associated with higher skin and breast cancer risk
Gorlin syndrome (MIM 109400) (Evans & Farndon, 2002)
Mutation of *SUFU* or *PTCH1*; features may include multiple jaw keratocysts, macrocephaly, frontal bossing, coarse face, facial milia, skeletal anomalies, and ectopic calcification of the falx; benign and malignant neoplasia including medulloblastomas, basal cell carcinomas, cardiac and ovarian fibromas, rhabdomyosarcomas, and ependymoma or rhabdomyomas
Li–Fraumeni syndrome (MIM 151623) (Schneider, Zelley, Nichols, & Garber, 1999)
Mutation of *TP53*; no obvious syndromic features; benign and malignant neoplasia including soft tissue sarcoma, osteosarcoma, pre-menopausal breast cancer, choroid plexus carcinoma, medulloblastoma, adrenocortical carcinoma, neuroblastoma, leukemia, and others; multiple neoplasms
Neurofibromatosis 1 (MIM 162200) (Friedman, 1998)
Mutation of *NF1*; features may include multiple café-au-lait spots, axillary and inguinal freckling, multiple cutaneous neurofibromas, Lisch nodules, learning difficulties, scoliosis, tibial dysplasia, and vasculopathy; benign and malignant neoplasia including plexiform neurofibromas, low- and high-grade gliomas, malignant peripheral nerve sheath tumors, juvenile myelomonocytic leukemia, and others
Retinoblastoma (MIM 180200) (Lohmann & Gallie, 2000)
Mutation of *RB1*, prototypic tumor suppressor gene – two-hit hypothesis described by Alfred G. Knudson (Croce, 2016); no obvious syndromic features; development of unilateral, bilateral and trilateral retinoblastomas (bilateral, or rarely unilateral retinoblastoma plus pineoblastoma); high risk of secondary primary malignancies, e.g., osteosarcoma, especially following external radiation, or melanoma
Rhabdoid tumor predisposition 1/2 (MIM 609322/613325) (Sevenet et al., 1999)
Mutation of *SMARCB1* or, more rarely, of *SMARCA4*; autosomal dominant; predisposition to renal and extrarenal rhabdoid tumors, meningioma (*SMARCB1*), and schwannomatosis (*SMARCB1*). Malignant peripheral nerve sheath tumors can also occur (*SMARCB1*). *SMARCA4* germline mutations can lead to small-cell carcinoma of the ovary, hypercalcemic type.

Table 1-3 Summary table of the most common cancer predisposition syndromes in children with cancer. Table taken from [256].

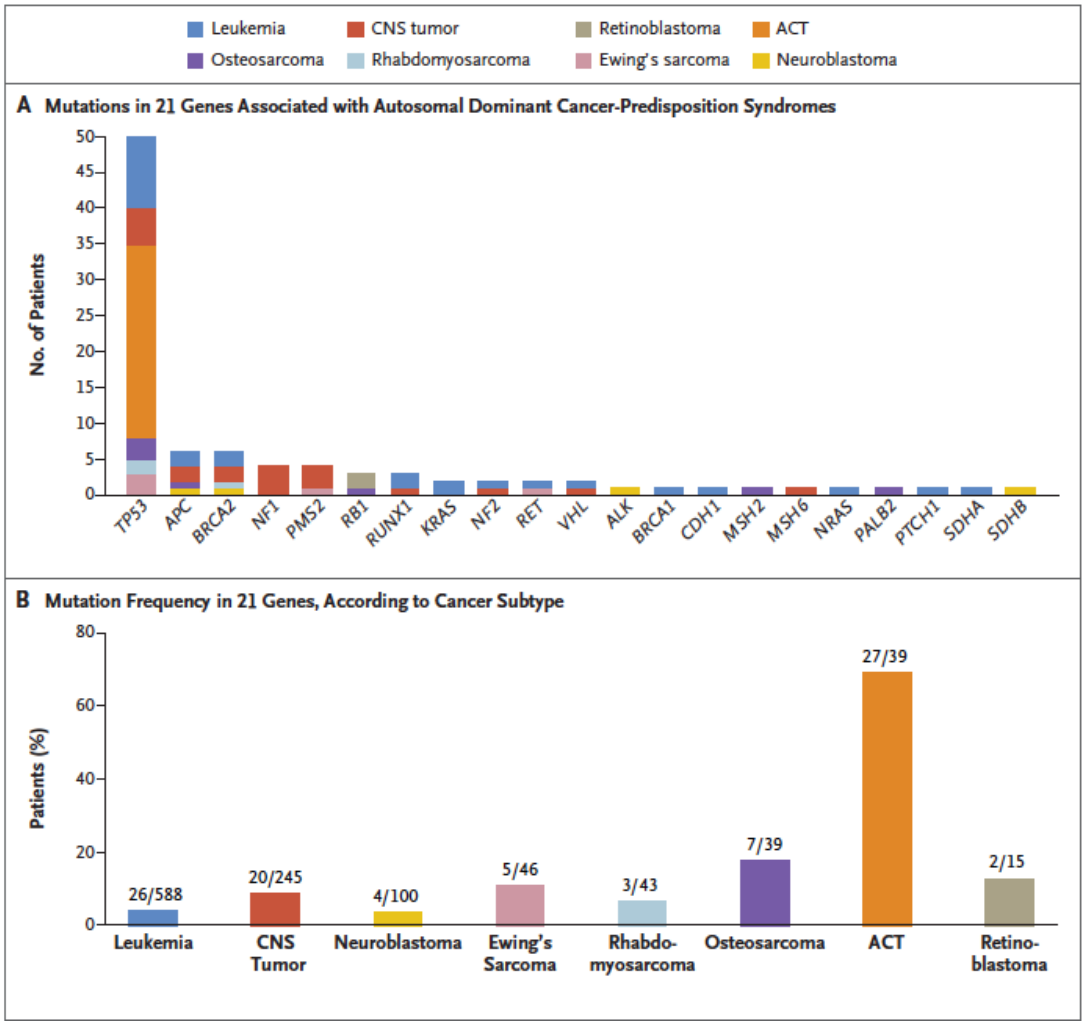


Figure 1-15 Germline mutations in paediatric cancer. (A) Bar-plot showing mutations in 21 genes associated with autosomal dominant cancer predisposition syndromes coloured by tumour type. (B) Bar-plot showing germline mutation frequency in 21 genes by tumour type. Figure taken from [13].

1.5 Next generation sequencing

In the last decade, advances in next generation sequencing (NGS) have facilitated the understanding of childhood cancer genomics which have revolutionised the way translational and clinical studies are conducted.

1.5.11 NGS technology review

NGS methodology varies depending on the strategy and platform chosen **{Figure 1-15}** [257-259]. The first step is the **library preparation**, which is accomplished by random fragmentation of the genomic DNA and can be conducted mechanically by sonication or by enzymatic digestion of the DNA. The fragment size needs to be compatible with the sequencer, and can vary from small fragments for short-read

sequencing (200-300bp) to longer for long-read sequencing (20,000bp) **{Figure 1-17A}**. After fragmentation, the DNA ends are repaired and adaptors are ligated followed by DNA amplification. In addition, the inclusion of unique molecular identifiers (UMI) can be utilised to identify sampling bias and correct for the effects of PCR duplicates. From this step, libraries can be used for WGS. An enrichment step is needed for WES or targeted-sequencing where capture probes that bind to the DNA sequence of interest are employed. Amplicon-based approaches involve PCR and use primers flanking the region of interest**.** The next step is the **library sequencing** generating clonal clusters from any given library. The details of the sequencing chemistry are different according to the platform chosen **{Figure 1-17B}**. After sequencing, FastQC is run to check quality on raw data and reads are aligned to the reference genome. This can be done by utilising different software (BWA, Bowtie, Novoaling) generating sequence alignment/map (SAM) and binary alignment map (BAM) format files. These are then used to perform variant calling and for visualisation using genome browsers such as the Integrative genomics viewer IGV. The most common variant calling tools include Samtools, GATK Unified Genotyper and MuTect. After the variants are detected, annotation is performed to include gene and transcript identifiers, as well as clinical significance and pathogenic prediction by using numerous cancer databases (ENSEMBL, COSMIC) and tools (Polyphen2, Oncotator, SnpEff or Alamut) [257].

The implementation of NGS-based assays in molecular diagnostics require a complex process of validation prior to clinical use, which require overall test performance and the assessment of crucial specifications such as limit of detection, accuracy, sensitivity and specificity amongst others [260]. The use of the appropriate reference standards is essential to account for specific challenges and ensure the accuracy and robustness required in CLIA (Clinical Laboratory Improvement Amendments) accredited laboratories which are routinely monitored with proficiency testing [261]. To assist with the validation of NGS testing assays, numerous guidelines and recommendations for standardised NGS framework have been published including technical and bioinformatic quality control steps [260, 262-264].

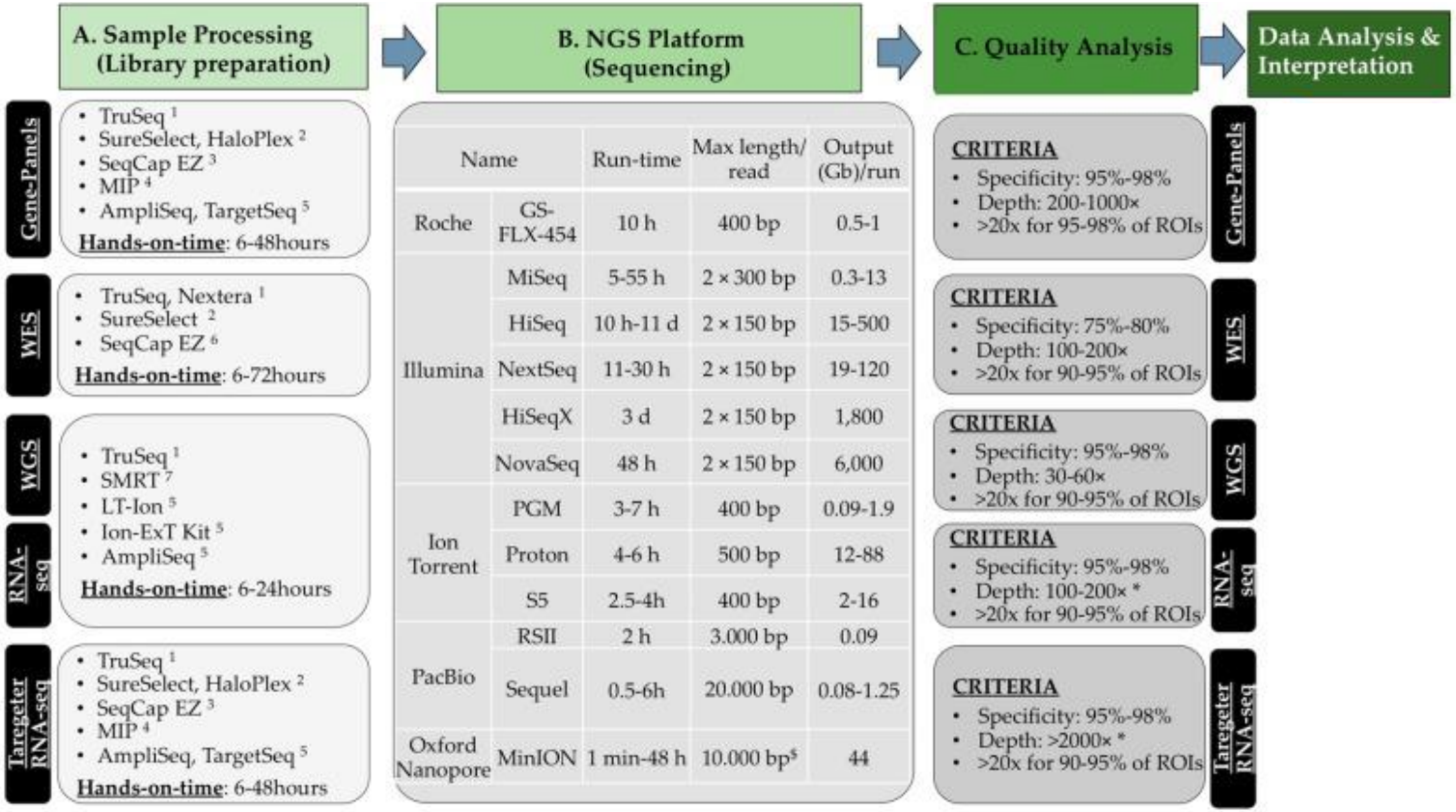


Figure 1-16 Schema showing the most common used NGS platforms. 1 Illumina, 2 Agilent, 3 Nimblegen, 4 Molecular Inversion Probe, 5 Thermo Fisher, 6 Roche and 7 PacBio. (A) Library preparation is divided by NGS approach and hand-on-time. (B) Sequencing platforms are shown indicating run time and sequencing read length and (C) Quality Control criteria is shown per NGS approach. Figure taken from [259].

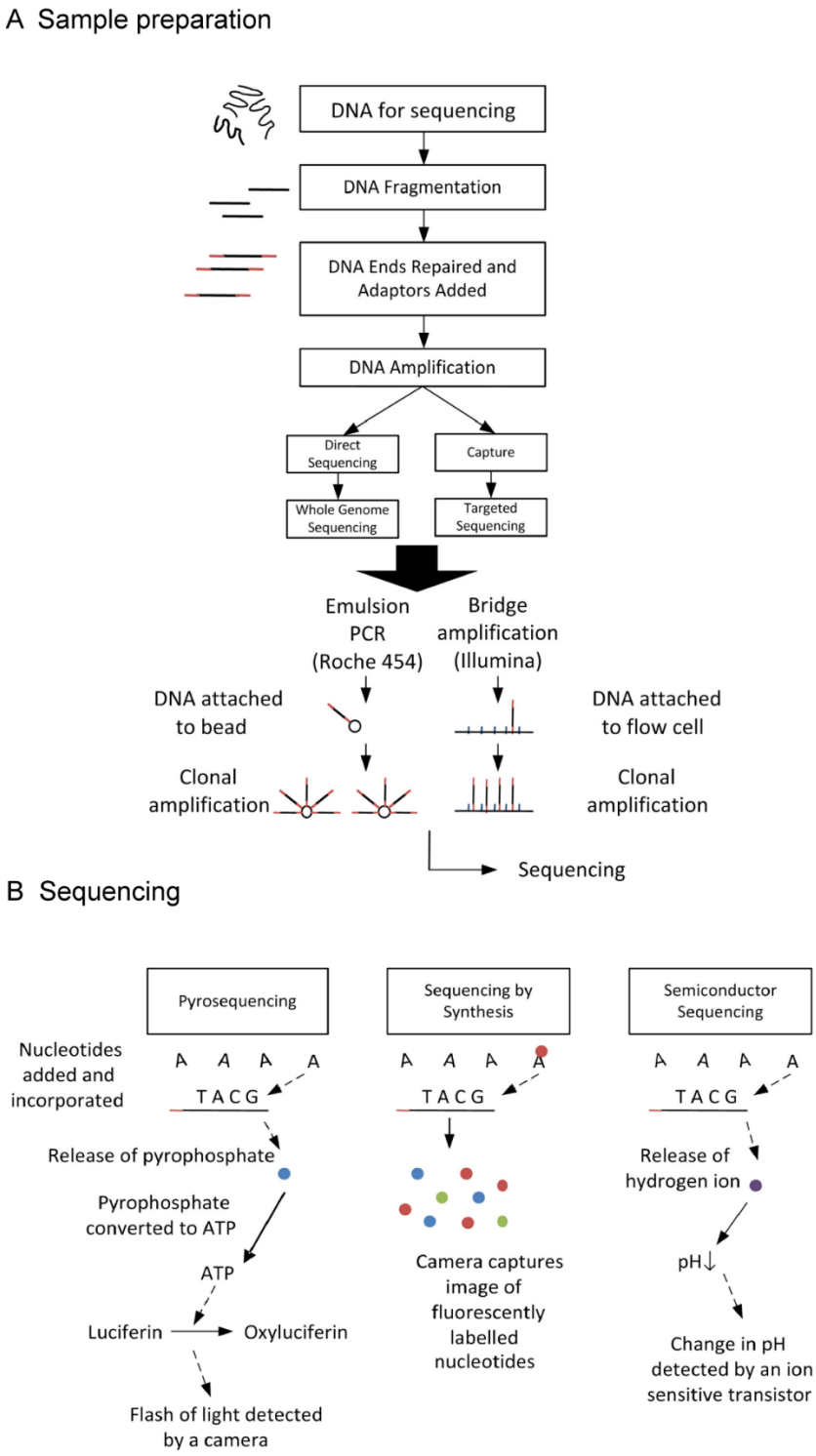


Figure 1-17 Overview of next generation sequencing. (A) Library preparation starts with DNA fragmentation followed by end-repair and adaptor ligation. PCR amplification is required to create enough copies of the DNA which can be carried out by emulsion-PCR (Roche and Life Technologies) or bridge-PCR (Illumina) depending on the platform used. (B) Sequencing platforms. There are three main platforms to detect the signal produced when the nucleotide is added i) Roche 454 uses pyrosequencing, whereby each cycle with the addition of a nucleotide a pyrophosphate is realised and converted to ATP, which transform luciferin to oxyluciferin producing light; ii) Sequencing by synthesis, whereby each cycle includes the addition of four fluorescently-labelled nucleotides, captured by a camera with four channels (Illumina); and iii) Semiconductor sequencing, whereby in each cycle a different nucleotide is applied resulting in release of a proton, altering the pH which is recorded as a change of voltage (Life Technologies). Figure taken from [257].

1.5.12 NGS paediatric cancer platforms

In the last four years, numerous studies have shown that comprehensive sequencing approaches from targeted sequencing, RNA-sequencing, methylation, WES to WGS can be applied for clinical decision-making in children with refractory/relapsed cancer [254, 265-269]. These studies have revealed a similar success rate for obtaining high quality sequencing data (81-89%) and for discovering molecular alterations that could help in the clinical management of these children (34-61%)**.** Such strategies are capable of supporting innovative clinical trials for children with cancer based-upon specific molecular alterations. **Paediatric MATCH** (Molecular Analysis for Therapy Choice, NCT03155620) is a single agent basket trial for children with refractory or recurrent tumours in the United States with seven treatments arms [270]. **ESMART** (European Proof-of-Concept Therapeutic Stratification Trial of Molecular Anomalies in Relapsed or Refractory Tumours, NCT02813135) is a basket trial using combination of targeted therapies plus chemotherapy in Europe with seven treatment arms, each of them as an independent trial with phase 1 dose escalation and phase two as expansion phase. **INFORM2** (Individualized therapy For Relapsed Malignancies in childhood) is a German study including sites in several European countries and Australia that uses a combination of targeted/immune therapies with established chemotherapy treatments based on the INFORM Registry findings [267].

Additionally, the sequencing platforms support disease-specific molecular-driven trials such as NEPENTHE (Next Generation Personalised Neuroblastoma Therapy) and PNOC008 (Pacific Pediatric Neuro-Oncology Consortium) initiatives. NEPENTHE (NCT02780128) is a phase I trial led by Yoel Mosse where genetic alterations are matched to targeted agents in children with neuroblastoma at time of relapse. PNOC008 (NCT03739372) is a two-strata pilot trial using WES, targeted panel sequencing and RNA-seq to test efficacy of personalised treatment in children and young adults, with high-grade glioma.

Besides the advances in clinical sequencing trials, the use of biomarkers predictive of response to an FDA-approved drug for the management of children with cancer remains very low compare to the adult population [271].

1.6 Liquid biopsy in cancer

Obtaining tumour tissue by surgical procedure at diagnosis is a routine practise for most paediatric solid tumours, however this is not always feasible for some inaccessible brain tumours, such as DIPG, resulting in very invasive procedures with associated complications and risks [272, 273]. In addition, tissue biopsy often fails to capture tumour heterogeneity and potential differences in metastatic tumours [274, 275]. Liquid biopsy represents a powerful tool to overcome sampling limitations including sample accessibility for longitudinal studies, enabling the monitoring of predictive biomarkers of response to treatment. In addition, liquid biopsy may offer an advantage to evaluate tumour heterogeneity and detection of acquired resistance alterations in comparison with a single core biopsy [276]. In adult cancers, circulating tumour DNA (ctDNA) has been broadly applied for early cancer detection, to monitor treatment, predict patient relapse as well as for the discovery of drug-resistance mutations [277-284].

Liquid biopsies refer to biological fluids derived from cancer tissue which can represent a source of cancer biomarkers **{Figure 1-18}.** These include circulating tumour cells (CTCs), circulating tumour nucleic acids (ctDNA and ctRNA), miRNAs, extracellular vesicles (EVs), tumour educated platelets (TEPs) and proteins. Although blood is the most typical source, nearly all body fluids can be utilised as liquid biopsy. This will depend on the anatomical location of the primary or metastatic tumour, and can include cerebrospinal fluid (CSF) for tumours affecting the CNS, urine for urogenital cancers, saliva for head and neck squamous cell carcinoma, pleural effusions for lung cancer and mesothelioma [285].

Cell-free DNA (cfDNA) is shed into the bloodstream from dying cells and constitutes short fragments of double-strand DNA. The half-life of cfDNA in the circulation is <2.5h [286]. ctDNA is the proportion of cfDNA that derives from cancerous cells that undergo cell death (apoptosis or necrosis), lysis of CTCs and secretion from the tumour. The average fragment size of ctDNA varies from 40-200bp with a peak size of ~166bp, corresponding to the length wrapped around a nucleosome plus ~20bp linker bound to histone H1 [287]. The amount of ctDNA can vary depending on tumour size, stage of disease, extent of metastatic spread and disease burden [287]. Furthermore, detectable levels of ctDNA in plasma can differ depending on the tumour type, for example ovarian or colorectal cancer are found to have a median of 100-1000 copies of DNA per 5 mL of plasma whereas brain tumours have been shown to have less

than 10 copies per 5 mL of plasma **{Figure 1-19}** [288]. Importantly, the proportion of ctDNA in primary brain tumours or solid tumours that metastasizes in the brain is significantly higher in CSF than in plasma [289-292].

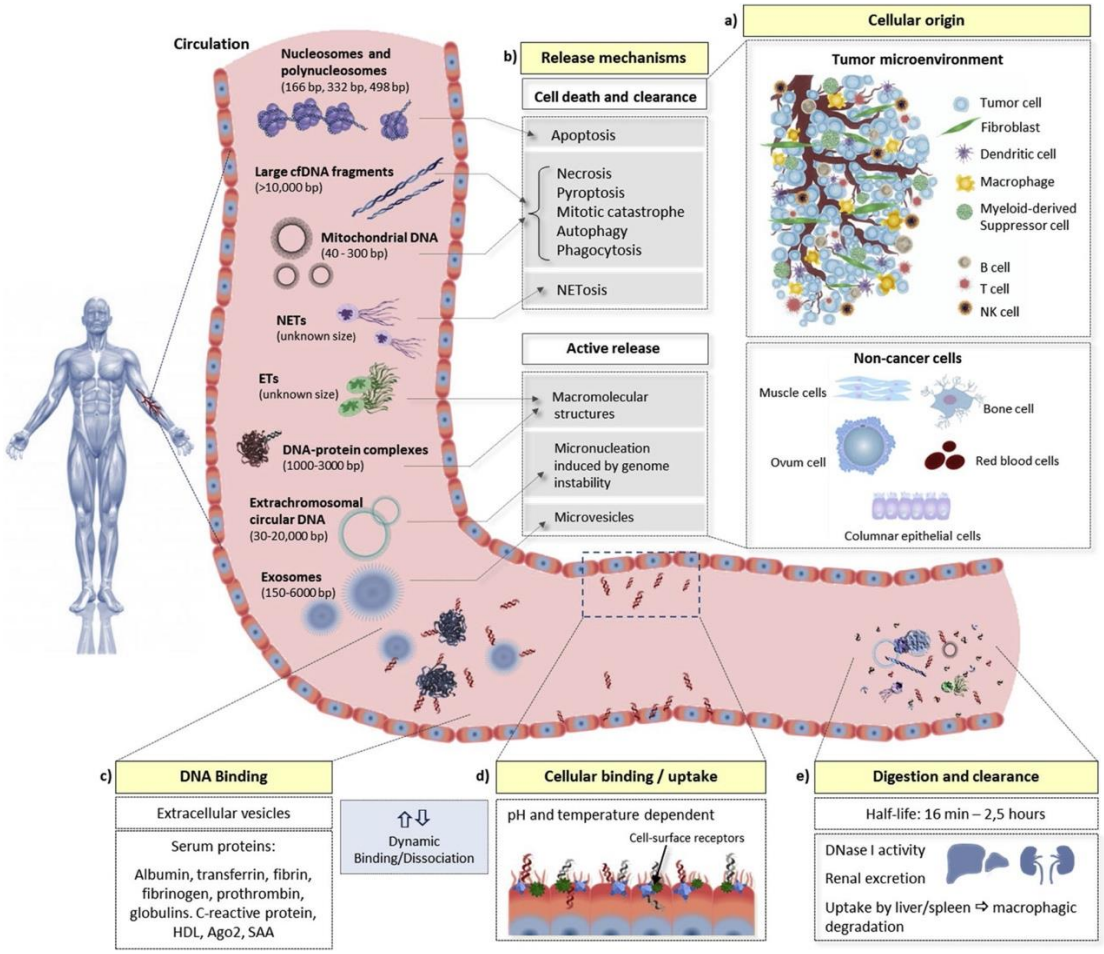


Figure 1-18 Cell-free DNA (cfDNA) origin and characteristics. (A) cfDNA can derived from different cells including tumour cells (ctDNA). (B) DNA is released to the circulation by different mechanisms most commonly by apoptosis, necrosis and active secretion which will determine DNA fragment size. (C) cfDNA levels are influenced by association and disassociation with extracellular vesicles and serum proteins (D) with different rate of binding depending on pH, temperature, and can be inhibited by certain substances such as heparin (E). cfDNA has a half-life of 16 min to 2.5hours. Figure taken from [293].

Given that the fraction of ctDNA in relation to the total cfDNA can be very low, often <0.01%, highly sensitive methods are required to identify ctDNA [294]. These include allele-specific PCR such as quantitative PCR (qPCR) COBAS kits for the detection of *EGFR* and *KRAS* mutations, droplet digital PCR (ddPCR) and BEAMing digital PCR [294-296]. In addition, deep-sequencing NGS-based assays allow the detection of multiple gene alterations. Amplicon-based NGS methods include TAM-seq (tagged-amplicon sequencing) [297], InVisionFirst™ [298], Safe-SeqS (Safe Sequencing System). Hybrid capture-based NGS assays were developed to improve the detection

of mutations and avoid potential PCR amplification bias, and include CAPP-seq (Cancer Personalised Profiling by Deep Sequencing) [299, 300], Guardant360® (Guardant Health, Inc.) [301, 302], and FoundationOne® Liquid (Foundation Medicine Inc). Furthermore, CancerSEEK uses a combination approach of multiplex PCR-based NGS assay and protein biomarker assessment from ctDNA [303].

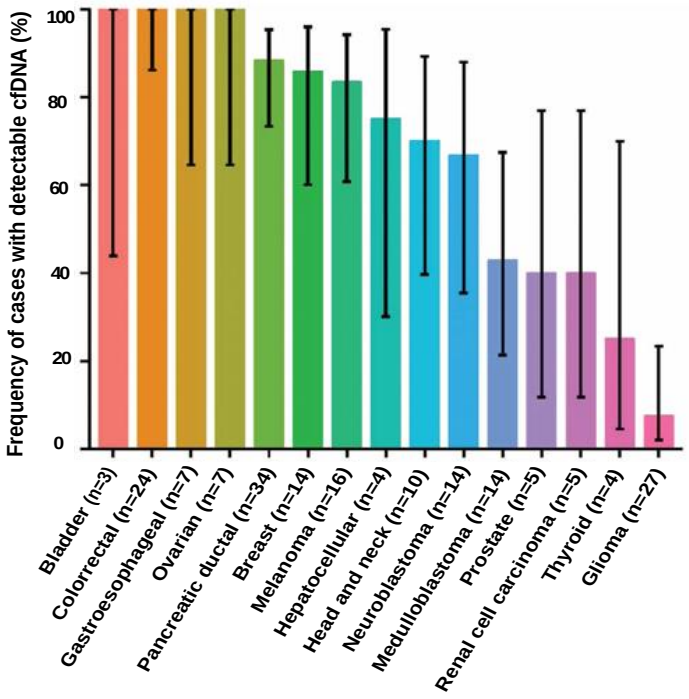


Figure 1-19 Bar-plot showing the frequency of patients with detectable circulating tumour DNA per tumour type. Figure modified from [288].

1.6.13 Liquid biopsy in childhood cancer

Although the study of liquid biopsies in paediatric cancer has increased over the last few years, it has not advanced to the same degree as in adult oncology. Recurrent alterations have however been detected in neuroblastoma (*MYCN, ALK* and *ATRX*) [304-306], Ewing sarcoma (*EWSR1* fusion, *TP53* and *STAG2*) [307, 308], rhabdomyosarcoma (*PAX3:FOXO1*) [309], retinoblastoma (RB1 mutations) [310, 311], Wilms tumours (*TP53*) [312], CCSK (ITD-*BCOR*) [313], HGNET-BCOR (ITD-*BCOR*) [314], DMG (H3-K27M) [315-318] and *BRAF*-V600E in Langerhans cell histiocytosis [319].

While few studies have compared the levels of plasma-ctDNA amongst paediatric tumours, it appears that children with neuroblastoma have higher levels of ctDNA compared to other non-CNS solid tumours (Ewing sarcoma, osteosarcoma, Wilms tumours and ARMS) [320]. Less is known about ctDNA in paediatric brain tumours and the few studies published are mainly in DMG-K27M tumours in which they have used ddPCR to detect K27M mutations [315-318]. In these studies, similarly to adult CNS tumours, it has been observed that ctDNA levels are higher in CSF than in plasma, with a mean mutant allele frequency of >10% and <0.1% respectively [317].

There are currently a small number of ongoing clinical trials evaluating the use of liquid biopsies. NGSKids (NCT02546453) and MICCHADO (NCT03496402), are two clinical trials open in France and led by Gudrun Schlieirmacher, utilising NGS-based technology to detect genetic alterations in ctDNA derived from blood, CSF, bone marrow and urine in metastatic, high-risk and low-risk paediatric tumours. Additionally, PRISM (PReciSion Medicine for children with cancer, NCT03336931) is a multicentre prospective study carried in Australia for children with high-risk cancer. In this study liquid biopsy is incorporated in their molecular profiling pipeline.

1.7 Aims

Aim 1. Establishment and clinical validation of a paediatric solid tumour NGS panels.

1a Development and validation, within a clinical pathology accredited laboratory (CPA, UK), of a paediatric solid tumour sequencing assay for use with either routine FFPE or fresh frozen (FF) samples.

1b Generation of a comprehensive genomic classification of paediatric solid tumours using targeted sequencing to identify recurrent, prognostic and predictive genomic alterations in clinical samples linked to outcome data.

Aim 2. Using circulating biomarker analysis to track tumour evolution and response

2a. Investigation of the value of liquid biopsies (cfDNA and CSF) as circulating biomarkers correlating DNA concentrations from plasma and CSF with tumour burden and clinical parameters.

2b. Molecular analysis through droplet digital PCR (ddPCR) will be used as a tool for minimal residual disease monitoring of therapy, prediction of early relapse and aid in non-invasive tumour stratification.

Aim 2. Identifying alternative treatments based upon genetic dependencies in a co-clinical DIPG trial (BIOMEDE)

3a. To identify rational therapeutic options with individualised preclinical evidence as to their efficacy in a co-clinical trial model for DIPG patients.

3b. To study specific evolutionary trajectories from DIPG derived cultures under drug exposure, generating resistance culture populations, and explore the resistance mechanism to identify more effective treatments.

CHAPTER 2 : Material and methods

2.1 Cases and clinical trials

2.1.1 Local and UK samples

A representative selection of common paediatric tumours and cell lines were used for the validation of a paediatric NGS panel, comprising a total of 132 samples. Furthermore, plasma and CSF samples from local hospitals were consented under the **CCR-4873** study. Local institutional review board approval was obtained in addition to the separate approvals from the contributing tumour banks (The Children's Cancer and Leukaemia Group Tumour Bank and the Queensland Children's Tumour Bank).

2.1.2 CCR-4294

A Royal Marsden Hospital (RMH) pilot study **CCR-4294** for patients aged ≤24 years with solid tumours treated on the Children and Young People's' Unit commenced in March 2016 and was subsequently expanded nationally for children aged ≤16 years. Ethical approval was obtained from the National Research Ethics Service (reference: 15/LO/07) and the Biological Studies Steering Group of the Children's Cancer and Leukaemia Group (reference: 2015 BS 09) by Dr Sally George. Participants and/or guardians gave informed consent. Patients were eligible to enrol at any time including diagnosis and relapse/progression. Blood was taken for germline DNA analysis, plasma was isolated, and archival tissue retrieved from the most recent surgery. If indicated, a repeat biopsy could be requested at the treating clinician's discretion. Samples from this pilot study were used to assess the implementation of targeted sequencing into clinical practice in order to identify diagnostic, prognostic and predictive in paediatric solid tumours linked to outcome data.

2.1.3 HERBY

The **HERBY** trial (NCT01390948, 2011-2015) was a multicentre, randomised phase II study of the addition of bevacizumab (BEV, a monoclonal antibody to vascular endothelial growth factor) to radiotherapy and temozolomide (RT/TMZ, standard of care for children with HGG, ≥3 years) in non-brainstem pHGG [69, 87, 321]. All patient samples were collected after signed consent to the HERBY translational research program, under full Research Ethics Committee approval at each participating centre.

As part of the correlative translational research programme, tumour tissue was taken at diagnosis. In addition, serial plasma samples were taken at five different time-points during the course of treatment**.** A subset of samples from this cohort was used to investigate the presence of recurrent structural variants (SV) as well as plasma samples utilised to interrogate their use in the detection of circulating biomarkers to track tumour evolution and treatment response.

2.1.4 BIOMEDE

The Biological Medicine for DIPG Eradication trial, **BIOMEDE** (NCT02233049) is an ongoing multicentre, randomised phase II study with molecular stratification of DIPG after upfront biopsy. Patients are randomised and stratified into one of three different arms according to the expression of PDGFRA, EGFR and PTEN, for treatment with dasatinib, erlotinib or everolimus. All patient samples were collected after signed consent to the BIOMEDE translational research program, under full Research Ethics Committee approval at each participating centre. Within the UK, snap-frozen tissue, whole-blood, live tissue, and optionally plasma, taken at diagnosis were collected and sent to our laboratory. The samples were used to characterise the molecular alterations driving the tumours as well to establish *in vitro* and *in vivo* models to identify alternative therapeutic targets. Once the *in vitro* models were established, a drug screen was performed based upon the molecular alterations detected in the original patient tumour sample. When possible, we selected a compound identified in the drug screen to perform efficacy studies in the serial patient derived xenograft (PDX) or cell derived xenograft (CDX). These results were fed back to the treating clinician in order to guide patient treatment at relapse.

2.2 Capture sequencing design and sample preparation

2.2.1 Panel design

Four different panels were designed: Paeds-v1 (v1=78 genes, 311 kb), Paeds-v2 (v2=91 genes, 473 kb), CNS fusion detection panel (24 genes, 2.2 Mb) and pHGG-panel (330 genes, 1.2 Mb). The panel design workflow is described in **{Figure 2-1}.** The targets were selected following literature review as either recurrently altered in paediatric cancers (Paeds-v1/v2) or in CNS-tumours (Fusion-panel). The genes in the pHGG-panel included recurrent alterations in genes previously identified from a meta-analysis published from our laboratory [22]. After gene selection, a bed file containing the exact genome location of the region of interest (ROI) was created comprising of

exons for hotspot mutations (*BRAF* exon 12 and 15), all coding exons for tumour suppressor genes (*TP53*), intronic regions for structural variants (SV) (*ALK* intron 19-20) and all coding exons. In addition, common heterozygous SNPs in the population, alongside the gene, were included for copy number variation (CNV) (*MYCN*). The bed file was sent to Nimblegen (Roche) where they first generated an *in-silico* design of the library of customised biotinylated DNA probes complementary to the genes of interest. The *in-silico* design contained a list of the ROI with an estimated-predicted coverage. The design was then reviewed, for the estimated non-predictive coverage regions, if feasible, other ROI were selected (only possible for CNV detection). We also try to obtain a more relaxed probe design, allowing for a maximum of matches in the genome from 20 to 50 in difficult regions, as well as adding extra probes in the areas limiting the poorly captured ROI. This process was reviewed with Nimblegen until we were satisfied with the ROI estimated-predicted coverage and only then was the library of biotinylated probes for the different panels manufactured.

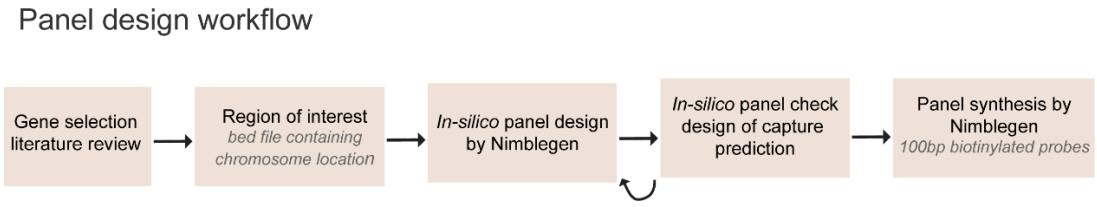


Figure 2-1 Panel design workflow from gene selection to probe synthesis.

2.2.2 Sample preparation and sequencing

For v1/v2, FFPE sample assessment from haematoxylin and eosin (H&E) stained slides was performed by an experienced pathologist (Dr Khin Thway) and a trained biologist (Professor David González de Castro). This was done to mark the region of the section containing tumour, and to estimate neoplastic cell content, defined as the percentage of neoplastic cells out of total nucleated cells in the marked area. Macro-dissection of the marked area was conducted when a distinct area of neoplastic cells from normal cells was observed in a large area and the overall tumour content without macro-dissection would have been <60%. 24 out of the 83 FFPE samples underwent macro-dissection to enrich the tumour content.

The sample workflow is described in **{Figure 2-2}**. DNA from blood and cell lines, FF and FFPE samples was extracted using the QIAamp DNA blood mini kit (Qiagen, 51106), the QIAamp DNA mini kit (Qiagen, 51306) and the QIAamp DNA FFPE tissue

kit (Qiagen, 56404), respectively. For specimens where DNA was extracted at local centres, their own protocols were followed. DNA was quantified with the Qubit 2.0 fluorometer (Invitrogen). To determine the degree of fragmentation of genomic DNA prior to library preparation, FFPE samples were analysed by TapeStation 2200 using the genomic DNA ScreenTape assay (Agilent Technologies, 5067-5366). Based on optimization studies, samples yielding DNA with median fragment length >1000bp were processed using 50-200ng DNA. Samples with DNA <1000 bp were processed using 400ng if there was sufficient DNA.

Library preparation was performed using the KAPA Hyper and HyperPlus Kit (Kapa, Roche KK8502 and KK8514) and SeqCap EZ adapters (Roche, NimbleGen, KK8701 and KK8702), following the manufacturer's protocol, including dual-SPRI size selection of the libraries (150-350bp). For the first 39 samples, I used the KAPA-Hyper Kit and sheared the DNA using Covaris M220 (Covaris, Woburn, MA). For the rest of the samples I used an updated kit, KAPA HyperPlus, which employs enzymatic fragmentation. Optimisation of the process indicated that the change to enzymatic fragmentation resulted in a substantial improvement in library complexity and unique coverage depth compared to sonication [322]. Following fragmentation DNA was end-repaired, A-tailed and indexed adapters ligated. To optimise enrichment and reduce off-target capture, pooled, multiplexed, amplified pre-capture libraries (6 to 10 cycles according to the DNA input) were hybridized twice overnight using 1 µg of the pooled library DNA to a custom design of DNA baits complementary to the genomic regions of interest (NimbleGen SeqCap EZ library, Roche). A 5 cycle PCR was performed between hybridizations to enrich the captured product. After hybridisation, unbound capture baits were washed away and the remaining hybridised DNA was PCR amplified (11 cycles). PCR products were purified using AMPure XP beads (Beckman Coulter) and quantified using the KAPA Quantification q-PCR Kit (KAPA, KK4828) or Qubit dsDNA High Sensitivity Assay Kit (ThermoFisher Scientific, Q32854). For v2 panel, updated CNS fusion-panel and pHGG-panel, as the probes became more specific, a single overnight hybridisation was performed. Samples were sequenced on a MiSeq or NextSeq (Illumina) with 75bp paired-end reads or 150bp paired-end reads chemistry (Illumina, MS-102-3001 and 20024907) according to the manufacturer's instructions. For clinical reported patients where germline matched control was available pools from tumour and control DNA, libraries were multiplexed separately for hybridization and combined prior to sequencing at a ratio of 4:1, increasing the relative number of reads derived from tumour DNA.

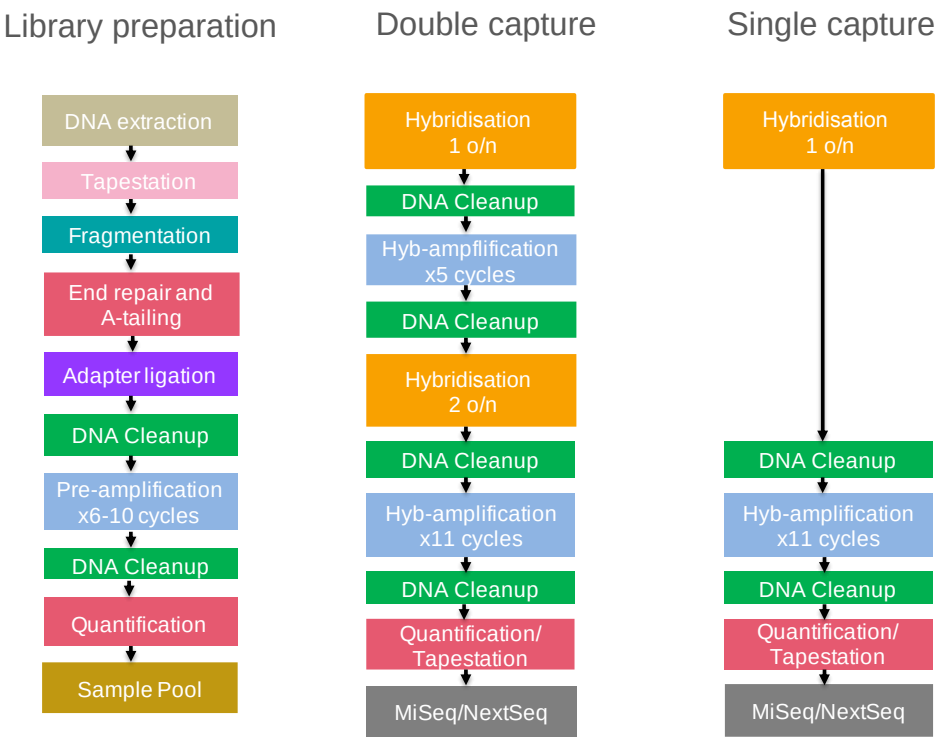


Figure 2-2 Sample workflow description from DNA extraction to library preparation, capture and sequencing

2.2.3 Paeds-v1 and Paeds-v2 data analysis

Primary analysis was performed using MiSeq Reporter Software (v2.5.1; Illumina, (http://emea.support.illumina.com/sequencing/sequencing_software/miseq_reporter/downloads.html), generating nucleotide sequences and base quality scores in Fastq format. Further analysis was performed by bioinformaticians Dr Lina Yuan and Sabri Jamal from the Clinical Genomics Research team (RMH, NHS, Sutton). Resulting sequences were aligned against the human reference sequence build GRCh37/Hg19 to generate binary alignment (BAM) and variant call files (vcf). Secondary analysis was performed in-house using Molecular Diagnostics Information Management System to generate QC, variant annotation, data visualisation and a clinical report. In the Molecular Diagnostics pipeline, reads were deduplicated using Picard (http://broadinstitute.github.io/picard/), and metrics generated for each panel region. Oncotator (v1.5.3.0) (https://portals.broadinstitute.org/oncotator) was used to annotate point mutations and indels using a minimum variant allele frequency (VAF) of 5% and a minimum number of 10 variant reads. Manta (https://github.com/Illumina/manta) was used for the detection of structural variants. Variants were annotated for gene names, nature of variant (e.g. missense), PolyPhen-2 predictions, and cancer-specific annotations from the variant databases including COSMIC, Tumorscape, and published MutSig results. CNV was assessed

using the ratio of GC-normalized depth of ROI in tumour against GC-normalized read depth of ROI in either matched germline DNA (when available) or the male cell line G147A (Promega). Any ratio below 0.65-fold was defined as a potential deletion whereas a ratio above 2.4 was flagged as a potential amplification. All potential mutations, structural variants and CNVs were visualised using Integrative Genomics Viewer (IGV) and two individuals were required to review the mutation report independently. Variant calls from samples with previously known SNVs and indels were checked manually on IGV. For v2, analysis was later executed using an update of in-house developed pipeline, Molecular Diagnostic Information Management System version 3.0 (MDIMSv3) using the following bioinformatics software and versions: demultiplexing was performed using bcl2fastq 2.17.1.14, reads were aligned using BWA 0.7.12, structural variants were identified using Manta 0.29.6, SNVs and indels were called with GATK 3.5.0, and variants were annotated with Oncotator version 1.5.1.0. CNV's were assessed as previously per Paeds-v1 modifying the ratio established for deletions from 0.65-fold to 0.5-fold.

2.2.4 Fusion-panel, pHGG-panel data analysis

Bioinformatic analyses describe in this section were carried out by the senior bioinformatician and my associate supervisor in our laboratory, Alan Mackay.

Capture reads were aligned to the GRCh37/Hg19 build of the human genome using bwa bwa v0.7.12 (bio-bwa.sourceforge.net), PCR duplicates were removed with PicardTools 1.94 (pcard.sourceforge.net) and BEDTools was used for quality control (QC) and generation of metrics for each sample Single nucleotide variants were called using the Genome Analysis Tool Kit v3.4-46 based upon current Best Practices using local re-alignment around indels, downsampling and base recalibration with variants called by the Unified Genotyper (broadinstitute.org/gatk/). Variants were annotated using the Ensembl Variant Effect Predictor v74 (ensembl.org/info/docs/variation/vep) incorporating SIFT (sift.jcvi.org) and PolyPhen (genetics.bwh.harvard.edu/pph2) predictions, COSMIC v64 (sanger.ac.uk/ genetics/CGP/cosmic/), dbSNP build 137 (ncbi.nlm.nih.gov/sites/SNP), ExAc and ANNOVAR annotations. Copy number was obtained by calculating $\log_2$ ratios of tumor/normal coverage binned into exons of known Ensembl genes, smoothed using circular binary segmentation (DNAcopy, www.bioconductor.org) and processed using in-house scripts in R v3.6.0.Manta (https://github.com/Illumina/manta) and Breakdancer (breakdancer.sourceforge.net) were used for the detection of structural variants. The

raw list of candidates provided by Manta were filtered for more than 2 reads covering both genes, common false positive base pairs (bp) positions/fusions outside of the capture set at both ends, common breakpoint/false positives within 10 bp, common false positive gene pairs, fusions within the same gene and homologous sequences greater than 10bp. Breakdancer was used to confirm all the breakpoints in all samples. Sequences either side of the break points were annotated to look for repetitive elements. A BLAT score was obtained to remove loci which were not uniquely mapped. Integrative Genomics Viewer (IGV) was used to view the fusions.

2.3 Panel sequencing validation

2.3.1 Paeds-v1 and v2

2.3.1.1 Validation samples

To validate v1, I used a total of 132 samples, including: i) Four cell blends with validated variants (Tru-Q1-4 HorizonDiscovery, Cambridge, UK), ii) 15 paediatric cell lines iii) 83 paediatric FFPE samples and iv) 30 FF paediatric samples.

The four cell blends contained 163 SNVs and 34 indels common to all four blends (background variants). Additionally, there were 61 SNVs and 17 indels, cancer variants, which were unique between blends, at known VAF, and verified by ddPCR. The four cell blends were used to assess overall performance, repeatability, intermediate precision, sensitivity and limit of detection. Specificity was determined using 87 true negative SNV sites (wild type) where another blend harboured a mutation at the corresponding position. The cell blends were processed and sequenced in two different runs by two independent users.

v2 was validated using the previous described four cell-blends (Tru-Q1-4) as well as 10 paediatric FFPE samples with known variants previously run using v1 (SNVs=554, indels=79). Quality and coverage metrics were calculated across all the samples including i) total reads, ii) percentage of reads mapped to the reference sequence, iii) percentage of duplicates, iv) percentage of bases from unique reads de-duplicated on target and v) mean depth. Sensitivity, specificity and accuracy were determined by comparing the cell-blends and FFPE samples with known variants and known true negatives.

2.3.1.2 Validation metrics

The validation was based according to the guidelines based by Mattocks and others to implement a test in the context of clinical human molecular genetic testing in order to ensure these assays perform to defined standards [260].

2.3.1.2.1 Overall performance

Four cell blends and five FFPE samples were used to measure performance across the capture design. Log mean depth was compared across the panel to the log depth of each region captured for each gene. Regions were classified as underperforming if the depth was lower than 2 x SD of the mean based on $\log_2$ [$\log_2$(ROI)>mean($\log_2$(ROI))-2xSD($\log_2$(ROI))]. GC content and mappability scores were compared against each region captured by the panel. Quality and coverage metrics were calculated across all the samples including i) total reads, ii) percentage of reads mapped to the reference sequence, iii) percentage of duplicates, iv) percentage of bases from unique reads de-duplicated on target, v) mean depth of targeted positions.

2.3.1.2.2 Limit of detection

To assess the limit of detection and determine a reliable cut off for the analysis a unique cancer-specific set of variants was used from the four cell blends introduced at range of VAFs from 4% to 30%, defined by ddPCR.

2.3.1.2.3 Precision

Repeatability (or within-run precision) was determined by comparing the cell blend background variant data across the 4 different samples in the same run for variant detection and VAF. Intra-run pairwise correlation was calculated for two runs where the cell blends were prepared and sequenced by different users generating two sets of repeatability data.

Intermediate precision (or between-run precision) was determined by comparing the cell blend background variant data between two runs for variant detection and VAF. Between-run pairwise correlation was calculated from two different runs prepared by different users and sequenced on different MiSeq instruments.

2.3.1.2.4 Sensitivity and specificity

The sensitivity of the panel was determined by separately comparing the cell blend background variants and the cancer-specific variants introduced at known VAF and specificity by using the cell blend cancer-specific set of data with known variants and known true negative sites. Variants were classified according to the different ranges of frequencies of the variants present in the DNA blends. The Positive-Predictive Value (PPV) and Negative-Predictive Value (NPV) were also determined.

2.3.1.3 Correlation between NGS targeted panel and other technologies

For v1 panel, 13 paediatric cancer cell lines were tested, harbouring a total of 30 known SNVs, deletions and amplifications previously identified by the Cancer Cell Line Encyclopaedia using Target Enrichment Sequencing (Agilent Technologies, Santa Clara, CA) and other published data [191, 323-327]. Furthermore, 33 samples (FF=14, FFPE=19) were used which harboured a total of 65 known genetic alterations including i) SNVs detected by Sanger Sequencing (*H3F3A, TP53, CTNNB1, HIST1H3B, ALK, BRAF*) [112, 328, 329] and RNA-Seq ii) copy number changes by FISH (*MYCN*) [330] and 450k array and iii) rearrangements by Real-Time Quantitative PCR involving *ESWR1* as previously described [331, 332].

2.3.1.4 Fresh frozen vs FFPE samples

For v1 panel, 15 paired FF and FFPE paediatric samples were compared for quality control metrics, coverage and the distribution of library inserts sizes between FFPE and FF paired samples. In addition, the VAF of the total variants found in the paired samples were correlated.

2.3.2 Glioma fusion-panel validation

To test whether the same technology used for Paeds-v1/v2 was suitable for the design of a CNS fusion-specific panel, we first obtained a pilot fusion-panel, containing a small library of the probes complementary to the genes of interest, designed by Nimblegen. To validate the assay, 25 samples were used comprising nine fusion positive cases involving *MN1, NTRK*, and *ALK* genes; four cases with suspected fusions by methylation and twelve WT or unknown fusion status. Quality metrics were generated and the presence of the fusions was assessed by MANTA. To evaluate the uniformity of sequencing coverage the Inter-Quartile Range, (IQR, the difference between the 75^{th} and 25^{th} percentiles of the histogram) was calculated.

This value is a measure of statistical variability, reflecting the uniformity of coverage across the data set. A high IQR indicates high variation across the ROI, while a low IQR reflects more uniform sequence coverage (https://www.illumina.com/science/technology/next-generation-sequencing/plan-experiments/coverage.html).

A series of steps were performed to optimise library preparation suitable for SV detection. To achieve longer DNA-library fragment length (from 150-350bp to 250-500bp), the fragmentation time as well as dual-SPRI size selection was optimised. Read length can also impact the detection of SV, shorter read length can result in less confident mappings [333]. This was done by modifying the number of base-pairs sequence, from 2x75bp to 2x150bp paired-end reads by choosing a different sequencing kit (Illumina, 20024907).

Once the library preparation and sequencing protocol was optimised a new version of the fusion-panel was designed to include *QK1*, *BEND2* and the non-tyrosine kinase domain of *ALK* (exons 1-20). Furthermore, as part of the panel optimisation, an increased number of probes within and limiting the poorly capture ROI were additionally added by Nimblegen when manufacturing the updated panel. To test the updated version, four samples with known SV (*FGFR1*-tandem duplication, *MN1:BEND2*, *ALK:SPECC1L* and *BRAF:KIAA1549*) were prepared, sequenced and the SV presence assessed.

2.4 Multi-omics analysis in the BIOMEDE co-clinical trial

2.4.1.1 Nucleic acid extraction

Nucleic acid extractions, from snap-frozen tissue, whole-blood, cells and patient/cell derived xenografts, were performed with the help of Sara Temelso, Higher Scientific Officer in our lab.

DNA and RNA were isolated from the same piece of tissue or cell pellet by using ZR-Duet DNA/RNA Miniprep Plus (Zymo Research, D7001). If only DNA was extracted, DNeasy Blood & Tissue Kit (Qiagen, 69581) was used and if only RNA was extracted, RNeasy Mini Kit protocol (Qiagen, 74104) was used. DNA and RNA quality were measured using Nanodrop spectrophotometer (Thermo-scientific). DNA

concentration was determined using Qubit fluorometer (Life Technologies). RNA integrity was analysed and quantified using 4200-Tapestation (Agilent).

2.4.1.2 Whole exome and panel sequencing

Whole exome sequencing was performed by Paula Proszek at the Clinical Genomics laboratory (RMH, Sutton UK). Libraries were prepared from 50-200 ng of DNA using the Kapa HyperPlus kit and DNA was indexed utilising 8bp-TruSeq-Custom Unique Dual Index Adpaters (IDT). Libraries were pooled in 8-plex, (2500-500ng of each library), by equal mass and normalised to the lowest mass sample. Samples were then hybridised overnight (16-18 hrs) with the xGen Exome Research panel v1 (IDT). An enrichment PCR of seven cycles was performed. Samples were sequenced on an Illumina NovaSeq 6000 system using the S2-200 Reagent kit (Illumina, 20012861) or the SP Reagent kit (Illumina, 20027465).

PDX and CDX models were sequenced using the pHGG-panel following the protocol described in the section 2.2.2 for sample preparation/sequencing. Variants observed in the models were compared against their original tumour sample as well as VAF.

2.4.1.3 RNA sequencing

At least 150 ng of RNA was sequenced at Eurofins Genomics. Strand-specific cDNA libraries were made by purification of poly-A containing mRNA molecules follow by mRNA fragmentation and random primed cDNA synthesis (strand specific). Adapter ligation and adapter specific PCR amplification was performed before sequencing on Illumina sequencers (Hiseq or Novaseq) using 150 bp paired-end reads chemistry according to the manufacturer's instructions.

RNAseq data was aligned with STAR and summarized as gene level fragments per kilobase per million reads sequenced using BEDTools and HTSeq. Following rlog transformation and normalization, differential expression was assigned with DESeq.2. Fusion transcripts were detected using chimerscan ver 0.4.5a filtered to remove common false positives. RNASeq raw count files were uSIsed to construct an expression matrix using Roche's internal pipeline. The expression matrix was normalized using edgeR and Voom in R (cran.rproject.org/), and a heatmap was created from the absolute gene expression data. Pre-ranked gene set enrichment analysis was used to compare gene expression of cultures with different sensitivities

to Trametinib and Dasatinb using the R package fastGSEA (fGSEA) based upon curated canonical pathways (MsigDB, Broad). Pre-ranked gene set enrichment analysis was used from RNA-seq and full proteome analysis to compare gene expression of resistant clones to trametinib compared to the parental line using the R package fastGSEA (fGSEA) based upon curated canonical pathways (MsigDB, Broad).

2.4.1.4 Methylation

A total of 50-500 ng of DNA was bisulphite-modified using the EZ DNA Methylation-Direct kit (Zymo, D5006), loaded onto the Illumina Infinium MethylationEPIC BeadChip and the array intensities were read on the Illumina iScan system at the University College London Genomics Centre, according the manufacturer's instructions.

Methylation data from the Illumina Infinium HumanMethylation850 BeadChip was pre-processed using the minfi package in R (v11b4). DNA copy number was recovered from combined intensities using the coummee package The Heidelberg brain tumour classifier [20] (molecularneuropathology.org) was used to assign a calibrated score to each case, associating it with one of the 91 tumour entities which feature within the current classifier. Clustering of beta values from methylation arrays was performed based upon correlation distance using a ward algorithm. DNA copy number was derived from combined $\log_2$ intensity data based upon an internal median processed using the R packages minfi and conumee to call copy number in 15,431 bins across the genome.

2.5 Liquid biopsy

2.5.1 Plasma and CSF sample preparation

5-10mL of peripheral blood sample were collected into Cell-Free DNA Collection Tubes (Streck) from CRC-4294 and CCR-4873. pHGG blood samples were centrifuged twice for 10min a first centrifugation at 1,600 g and a second centrifugation at 1,600 g or 16,000 g to remove cellular contents and/or debris. For local RMH, CSF was collected using Streck tubes or into sterile tubes and centrifuged at 1,600 g for 10 min to remove cellular contents and/or debris. Samples were stored at -80°C until cfDNA extraction. Local protocols to isolate plasma and CSF were used for the remaining liquid biopsies cases, collected from different site-sources.

cfDNA isolation from plasma and CSF was performed using the QIAamp circulating nucleic acid kit (Qiagen, 55114) following quantification using the Qubit fluorometer (ThermoFisher Scientific, dsDNA HS Assay kit, Q32854) and fragment analysis by 4200-TapeStation (Agilent, Genomic DNA ScreenTape 5067-5366).

2.6 Droplet-digital PCR

Custom TaqMan-based quantitative PCR genotyping assays (Bio-Rad; Applied Biosystems, Thermo Scientific and IDT, Integrated DNA Technologies) were designed to specifically detect genetic abnormalities (mutations and amplifications). The assay limit of detection (LoD) was assessed by performing serial dilution of the mutant DNA in constant concentration of wild-type DNA (1:10, 1:100, 1:1.000 and 1:10.000) and run in duplicate using 5 ng of DNA. The LoD was calculated as the fractional abundance of the neat mutant sample divided by the lowest dilution with detectable mutant copies (at least two mutant droplets). For each assay, three controls were run in duplicates including: one non-template control, one wild-type control (fragmented Promega DNA at 1ng/ul) and one positive control harbouring the alteration of interest.

The Bio-Rad QX200 Droplet Digital PCR (ddPCR) system was used, which allows the detection of rare DNA target copies with high sensitivity. DNA was randomly encapsulated into approximately 15,000 oil nanoliter-sized droplets, using the Automated Droplet Generator (BioRad, QX200 AutoDG), containing ddPCR Supermix for probes (no dUTP) (BioRad, 1863024), genotyping assay (specific per alteration), water and the DNA of interest. The PCR reaction was performed in a thermocycler **{Table 2-1}**, plates were then placed on the droplet reader where the droplets are streamed individually through a detector and signals from mutant positive (FAM), wild-type (VIC/HEX), double-positive (FAM and VIC/HEX) and negative droplets (empty) are counted to provide absolute quantification of DNA in digital form. The mutant allele concentration (C_{MUT}) and wild-type allele concentration (C_{WT}) were calculated with Quantasoft Analysis Pro (BioRad), the mutant allele fraction (AF_{dPCR}) and the concentration of cfDNA in the CSF or plasma (CcfDNA ng/mL) were calculated with the following calculations as previously described in [316, 334]:

$$AF_{dPCR} = C_{MUT}/(C_{MUT}+C_{WT})$$
$$C_{MUT\_ORI} = V_{PCR} \times C_{MUT} \times V_{ELU} / V_{DNA\text{-}PCR} \times V_{SAMPLE}$$
$$C_{WT\_ORI} = V_{PCR} \times C_{WT} \times V_{ELU} / V_{DNA\text{-}PCR} \times V_{SAMPLE}$$

C_{MUT_ori} is mutant allele concentration in original CSF or plasma (copies/mL)
C_{WT_ori} is wild-type allele concentration in original CSF or plasma (copies/mL)
V_{PCR} is volume of final PCR mix (µL)
V_{SAMPLE} is the volume of CSF or plasma used to extract cfDNA (mL)
V_{ELU} is the volume of cfDNA elution generated from DNA extraction (µL)
$V_{DNA\text{-}PCR}$ is the volume of cfDNA used in final PCR mix

$$CcfDNA \approx 0.003 \times (C_{MUT} + C_{WT})$$
the mass of 1 haploid human genome is 0.003 ng

Temperature (°C)	Time (min)	Ramp rate (°C/sec)	Cycles
95	10	2	1
94	0.5	2	40
Variable	1	2	
98	10	2	1
4	∞	1	N/A

Table 2-1 PCR cycling conditions. Annealing temperature was optimised per set of primers/probes. N/A = not applicable.

Analysis of tumour burden from the HERBY cohort by ddPCR at the different time-point and cfDNA was correlated with multimodal radiological indicators of response and progression. These parameters were determined by Dr Tim Jaspan and Dr Daniel Rodriguez and included i) MRI imagining using the Response Assessment in Neuro-Oncology (RANO) criteria, based on T-1 weighted and T2-fluid-attenuated inversion recovery sequences (FLAIR) and ii) diffusion/perfusion imagining for response assessment [321, 335].

2.7 *In vitro* culture

2.7.1 Primary derived patient culture establishment

DIPG patient-derived cultures were established from a small piece of tumour after patients underwent biopsy in the BIOMEDE trial from UK centres. The tissue was either sent in Hibernate A transport media (ThermoFisher Scientific, A12475-01) in a 20 mL tube and shipped at room temperature or minced with a sterile scalpel blade in DMEM/F12 (Life Technologies, 11320-074) supplemented with 0.2% BSA (Sigma-Aldrich, A1595) and 10% DMSO (Sigma-Aldrich, D2650), frozen at -80°C for at least 24h before shipment on dry ice to our laboratory.

Cryopreserved tissue was briefly thawed in the water bath at 37°C followed by addition of 5mL of stem cell media and centrifuged at 1,000 rpm for 10 min at room temperature. When the tissue was received in Hibernate A it was centrifuged at 1,300rpm for 4min and minced with a sterile scalpel blade in a petri dish. The tissue was then transferred to a universal tube and digested using Liberase DL (diluted at 10x in stem cell media) in the incubator for 10 min (gently mixing after 5 min). 10 mL of stem cell media was added follow by centrifugation at 1,300 rpm for 4 min. Supernatant was removed and the tissue/cells were resuspended in stem cell media and continuously pipetted to ensure it was dissociated. If the tissue appeared to contain excessive blood, it was incubated for 1 min at 37°C using red blood cell lysis buffer (ThermoFisher, A1049201), then washed with stem cell media and centrifuged at 1,300 rpm for 4 min. The tissue/cells and supernatants were transferred to culture flasks either to grow attached in laminin-coated flasks (Merck Millipore, CC095, diluted in PBS to coat a surface at 1µl/cm^2, and incubated for 2-4 hrs before use) for 2D-Lam establishment and in suspension for 3D (neutrosphere, NS) formation in ultra-low attachment flasks (Sigma-Aldrich, CLS3815). Primary cultures were monitored for growth under the microscope and fresh stem cell media was added every 2-3 days. Cells were incubated at 37° C, 5% CO_2, 95% humidity.

The cells were grown in stem cell media consisting of Dulbecco's Modified Eagles Medium: Nutrient Mixture F12 (DMEM/F12; Life Technologies, 11330-038), Neurobasal-A Medium (Life Technologies, 10888-022), HEPES Buffer Solution 1M (Life Technologies, 15630-080), MEM Sodium Pyruvate Solution 100 nM (Life Technologies, 11360-070), MEM Non-Essential Amino Acids Solution 10mM (Life Technologies, 11140-050) and Glutamax-I Supplement (Life Technologies, 35050-061). The media was supplemented with B-27 Supplement (Life Technologies, 12587-010), 20 ng/ml recombinant Human-EGF (2B Scientific LTD, 100-26), 20 ng/ml recombinant Human-FGF (2B Scientific LTD, 100-146), 20 ng/ml recombinant Human-PDGF-AA (2B Scientific LTD, 100-16), 20 ng/ml recombinant Human-PDGF-BB (2B Scientific LTD, 100-18), and 2 µg/ml Heparin Solution (Stem Cell Technologies, 07980). For the first passages (around 3-4), primary cultures were maintained with antibiotic-antimycotic at 100x (Invitrogen, 15240-096).

When cultures reached confluency, 90% surface area for 2D-Lam cells and a diameter of 200 µm for 3D-NS cultures, cells were split into new flasks for expansion and maintenance. For 2D-Lam, media was removed from the flask and cells were incubated in accutase dissociation reagent (Sigma, A6964) for 2-3 min at 37°C or until

cells were detached from the flask, media was then added making sure all the cells were in suspension. Cells were then transferred to a universal tube for centrifugation at 1,200 rpm for 5 min and the pellet was resuspended using fresh media. For 3D-NS cultures, cells were centrifuged with the original media at 1,000 rpm for 10 min, the supernatant was discarded and the pellet resuspended in accutase for 5-8 min, followed by pipetting up and down to finish breaking the NS which then were neutralized by adding 5 mL of fresh media and centrifuged at 1,300 rpm for 3 min. Supernatant was removed gentle without disturbing the pellet and 200 µl of media was added follow by pipetting up and down to obtain single cell suspension. When both 2D-Lam and 3D-NS cultures were resuspended, cell count and viability were assessed using the automated cell counter, Countess II FL (Invitrogen, AMQAX1000).

When possible after splitting cultures, cells were banked for long-term storage. Pelleted cells were resuspended in 1mL of StemCell Banker (AMS Biotechnologies, 11890), aliquoted into 2mL cryovials and placed in a Mr Frosty container for slow freezing at -80°C, before transferring to liquid nitrogen storage. To recover cryopreserved cells, they were quickly thawed using the water bath at 37°C and transfer to a conical tube containing 5 mL of stem cell media. Sample were centrifuged at 1300 rpm for 5 min to pellet the cells and supernatant is discarded. The pellet was then resuspended in stem cell media and cells were seeded in a flask.

A culture was considered successfully established if after five passages, cells were proliferating, providing enough cells to perform the experiments required, as well as having the ability to continue expanding the culture to bank stocks of cells for future work. At passage 5-6, two cells pellets were made for genomic characterisation to study how well the primary models recapitulated the original patient-sample disease including methylation arrays profiling, pHGG-panel sequencing, and RNA-seq. Furthermore, cultures were checked for mycoplasma contamination by PCR, as well as for its authenticity which was verified by short tandem repeat (STR) DNA fingerprinting using GenPrint10 (Promega, B9510).

2.7.2 Optimal cell density and doubling times

To determine the optimal number of cells needed for the drug assays, cell density assays were conducted, which consisted on plating five different cell-density conditions per model. 96-well black plates laminin-coated were used for 2D-Lam and cell-repellent surface plates for 3D-NS (Grainer, 655090 for 2D and 655976 for 3D).

To asses optimal cell-density cells were observed under the microscope to determine which condition gave around 90% confluency for 2D-Lam or reached 200-300 µm diameter for 3D-NS, at the end-point (day 11) to ensure logarithmic growth for the duration of the experiment. Cell viability was determined using Cell Titer-Glo (Promega, G7571 for 2D and G7572 for 3D) following the manufacturer's protocol. This was done at the end-point to compare growth linearity according to the number of cells plated. The population doubling time was estimated for the all the DIPG patient-derived lines, in 2D and 3D, by seeding two different cell-density conditions into 96-well plates, using the same plates as above. The number of cells were chosen according to the optimal cell-density assay. Cell viability was measured with Cell Titer-Glo (Promega) every two days or depending on the proliferation rate of the model. Luminescence values were plotted and the exponential part of the curve was used to measure the doubling time according to the following equation:

$$PDT = t \ln 2 / \ln (Xe/Xb)$$

t is the incubation time in any unit
Xb is the cell number at the beginning of the incubation time
Xe is the cell number at the end of the incubation time

2.7.3 Drug assays

Primary cells were plated using black 96-well plates according to the cell density estimated per line, as previously described. Cells were incubated for three days before adding the compounds with serial dilution at ten different concentrations and eight days later, (end-point, day 11) cell viability was measured using Cell Titer-Glo (Promega). Relative luminescence units (RLU) for each well was normalised to the median RLU from the DMSO control wells as 100% viability. At least duplicates per drug condition were performed as well as three independent biological replicates. GI50 values (drug concentration causing 50% inhibition of cell proliferation) were calculated using GraphPad Prism and the curves show the mean ± SD of the replicates per condition measured. All the compounds were purchased from Selleckchem except for CUDC-907 and PTC-209 which were obtained from Apexbio. The compounds were diluted in DMSO to a final concentration of 10mM and made single use drug aliquots of 20-40 µl.

Drug plates were prepared using the acoustic liquid handler Echo 550 (Labcyte) with the help of Mark Stubbs (Higher Scientific Officer from Cancer Therapeutics, ICR) and Diana Carvalho within our laboratory. Each plate included six compounds at eight

different concentrations as well as a cytotoxic chemotherapeutic (camptothecin, a highly potent specific DNA topoisomerase I inhibitor) as positive control, as well as DMSO as a negative control. Two different sets of plates were prepared: Plate 1 was assessed in all the cultures and contained dasatinib, everolimus, erlotinib, crizotinib, panobinostat and olaparib; Plate 2 consisted of specific drugs chosen targeting the molecular alterations detected in the tumour sample by the Paeds-v2 panel. The compounds were mainly FDA-approved drugs and/or in clinical trials (single drug or multi-arm/basket trials) for pHGG patients **{Table 2-2}.** When possible, the drug screen was performed in both conditions, as 3D-NS and as 2D-LAM. Drug assays were performed as previously described in duplicates or triplicates and conducted three biological replicates.

Drug combinations were used to investigate the synergistic effect of: i) trametinib + dasatinib, ii) trametinib + ulixertinib and iii) dasatinib + ulixertinib. Drugs were combined by adding one compound in rows and another in columns with serial dilutions resulting in a 6 x 10 dose matrix using different fold dilutions manually prepared as shown in **{Figure 2-3}**.

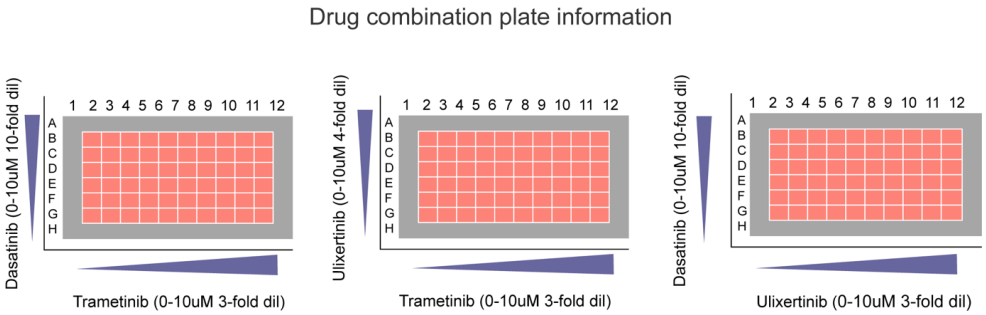


Figure 2-3 Illustration of plate distribution for the three combination assays and the fold-dilution per drug.

Drug combination assays were performed in duplicates with three independent biological replicates. Cells were seeded, drugs were added three days later and at day 11 cell viability was measured using Cell Titer-Glo (Promega). The stand-alone web-application SynergyFinder (https://synergyfinder.fimm.fi) was used for interactive analysis and visualization of multi-drug combination profiling data following the Bliss independence model [336, 337]. This model assumes a stochastic process in which two drugs elicit their effects independently, and the expected combination effect (IBliss) can be calculated based on the probability of independent events as IBliss=IX+IY–IXIY, where IX and IY are the single drug inhibition levels at doses X

and Y. Interaction between two drugs is considered likely to be antagonism if Bliss scores are below -10, likely to be additivity if Bliss scores are -10 to 10 and likely to be synergism if Bliss scores are above 10.

Drug	Dose (μM)	Description	FDA	Clinical Trial
Everolimus	0-10	mTOR inhibitor	YES	NCT02813135/NCT01734512
Dasatinib	0-10	BCR/ABL and Src kinase inhibitor	YES	NCT02233049/NCT02389309/ NCT03352427
Erlotinib	0-10	EGFR inhibitor	YES	NCT02233049
Panobinostat	0-1	DAC inhibitor	YES	NCT02717455
Olaparib	0-30	PARP inhibitor	YES	NCT02813135/NCT03233204/ NCT03155620
Crizotinib	0-10	Tyrosine kinase inhibitor	YES	NCT01644773*
AZD1775	0-10	WEE1 inhibitor	NO	NCT02813135/NCT02095132
ONC201	0-10	Akt/ERK inhibitor	NO	NCT03416530/NCT02525692
Crenolanib	0-10	PDGFR- and Flt3 inihibitor	NO	NCT01393912*/NCT02626364**
Imatinib	0-10	Kinase inhibitor	YES	NCT00021229*/NCT00021229*
Trametinib	0-10	MEK1/2 inhibitor	YES	NCT03593993/NCT03919071/ NCT03363217
CUDC-907	0-10	dual PI3K and HDAC inhibitor	NO	NCT02909777
AZD2014	0-10	dual mTOR1/2	NO	NCT02813135
LY3023414	0-10	dual PI3K/mTOR Inhibitor	NO	NCT03213678
Ulixertinib (BVD-523)	0-10	ERK1/2 inhibitor	NO	NCT03698994
Talazoparib	0-10	PARP inhibitor	NO	NCT02116777*
Sorafenib	0-10	Kinase inhibitor	YES	NCT01338857*
Regorafenib	0-10	Kinase inhibitor	YES	NCT04051606**
Vandetanib	0-10	Kinase inhibitor	YES	NCT00996723*/NCT00472017*/ NCT01582191
Saracatinib (AZD0530)	0-10	Src inhibitor	NO	NCT00704366**
Selumetinib (AZD6244)	0-10	MEK1/2inhibitor	NO	NCT01089101/NCT03095248
Dabrafenib	0-10	MAPK inhibitor	YES	NCT03340506/NCT03919071
Ribociclib (LEE011)	0-10	CDK inhibitor	YES	NCT02813135/NCT02934568/ NCT03387020/
Temsirolimus	0-10	mTOR inhibitor	YES	NCT02420613 (in combination with Vorinostat)
PTC209	0-10	BMI1 inhibitor	NO	NCT03605550 (PTC596, updated molecule, not commercially available)
Tranylcypromine (TCP)	0-10	LSD1 inhibitor	YES	NCT02717884** (Leukemia)

Table 2-2 List of drugs used in the drug screen. The table indicates if the drugs are FDA approved as well the NCT clinical trial. *Completed/terminated clinical trial; ** > 18 years old.

2.7.4 Generation of resistant clones

BIOMEDE-169 resistant cells were established by culturing the parental line in escalating concentrations of trametinib from 0.05 μM (GI50) to 1 μM in an exponential stepwise manner, as well as exposing the cells to a constant concentration of trametinib of 0.5 μM (GI80 value)**.** Both methods were performed using two technical replicates and two independent biological replicates for a total of eight derived cultures plus four replicate control flasks which were treated with the same concentration of DMSO in parallel to the establishment of the resistant cells. Cells were maintained in DMSO only for two passages to keep the baseline clonal population as close as possible to the parental culture. A total of 1-1.4x10^6 cells were seeded in a T75 flask for the resistance assay and 2-3 days later the media with trametinib was added at the appropriate concentration or DMSO to the control flasks. For the escalating-dose approach, ten different concentrations were used with an exponential increment and the same dose was added a total of 6 times. During the generation of resistance, the stem cell media containing the drug or DMSO was replaced three times a week and the cells split when they reached 90% confluency. Every time cells were sub-cultured, pellets were made for DNA/RNA isolation and if possible, cells were cryopreserved for long-term storage. To assess the emergence of resistant clones to trametinib, dose-response assays were performed as previously described.

2.7.5 Protein assays

2.7.5.1 Protein extraction

Cells were scraped from the flask and collected in the media or in cold PBS. Cells were then centrifuged at 1,300rpm for 5min, resuspended in 1 mL of cold PBS and centrifuged at 4,000rpm for 4min. The supernatant was then removed and the pellet was resuspended in cold lysis buffer in 50-100μl depending on the number of cells. Buffer was made from cell lysis buffer (Cell Signalling Technologies, 9803), protease inhibitor cocktail mini-tablet (Roche, Diagnostics, 11836153001) and phosphatase inhibitors (Sigma-Aldrich, P044 and P5726). Cells in the lysis buffer were snap-frozen and stored at -80°C until protein extraction, for which samples were thawed on ice to lyse for 2-4h occasionally mixing gently. Lysates were sonicated for 10 sec at 40% amplitude, spin at 14,000 xg for 10 min in a cold microfuge and then the supernatant was transferred to a new tube following quantification using BCA Protein Assay Kit (Thermo-Fisher, 23225).

2.7.5.2 Capillary-based protein quantification

The following primary antibodies were used: AKT 1:50 (Cell Signalling Technologies, 9272), pAKT-Ser473 1:50 (Cell Signalling Technologies, 4060), ERKp44/42 1:100 (Cell Signalling Technologies, 9102), pERK-Thr202/Tyr204 1:100 (Cell Signalling Technologies, 9101), MEK1/2 1:50 (Cell Signalling Technologies, 9122), MEK1/2-pSer217/221 1:50, α-Actinin 1:200 (Cell Signalling Technologies, 6487) and the AKT/MAPK pathway antibody cocktail 1:25 (abcam, 151279). As secondary antibody Goat Anti-Rabbit HRP conjugate (Protein Simple, 042-206) was utilised. A total of 0.75µg of protein lysate was used. Capillary electrophoresis was conducted using the automated Wes system (ProteinSimple) with the 12-230 kDa Separation module (ProteinSimple, SM-W004) and the anti-rabbit detection module (Proteinsimple, DM-001) following the manufacturer's instructions. Compass software from ProteinSimple was used for the analysis.

2.7.5.3 Proteomics

Full proteome and phospho-proteomics were performed at the Proteomics Core Facility ICR (Chelsea).

Cells were washed in cold PBS three times before being scraped from the flask and collected in cold PBS. Cells were then centrifuged at 1,600rpm for 5min, resuspended in 1mL of cold PBS and centrifuged at 4,000rpm for 4min. The supernatant was then removed and the pellet snap-frozen and stored at -80°C until protein extraction. Cell pellets were lysed in 5% SDS/ 100 mM TEAB buffer with probe sonication and heating at 95ºC for 10min. Protein concentration was measured by Pierce 660 nm Protein Assay and 300 µg of protein were taken for each sample. Proteins were reduced with TCEP (5 mM tris-2-carboxyethyl phosphine) and alkylated by iodoacetamide, and then purified by methanol/chloroform precipitation. Trypsin was added at 1:30 ratio (trypsin:proteins) for 18h digestion at 37ºC. 150 µg of peptides per sample were tandem mass tagged (TMT) labelled as instructed by the manufacturer (Thermo Scientific). The TMT labelled peptide mixture was fractionated on a BEH XBridge C18 column (2.1 mm i.d. x 150 mm) with a 35 min gradient from 5 – 35% CH_3CN/NH_4OH at pH 10. Fractions were collected at every 42s and pooled to 28 fractions.

Phosphopeptide enrichment used the High-Select Fe-NTA Phospho-peptide Enrichment Kit. The enriched phospho-peptides (16 fractions per set) and the

immobilized metal affinity chromatography (IMAC) flow through (28 fractions, used for whole proteome analysis) were analysed on an Orbitrap Fusion Lumos coupled with an Ultimate 3000 RSLCnano System. Samples were loaded on a nanotrap (100 µm id x 2 cm) (PepMap C18, 5 µ) at 10 µL/min with 0.1% formic acid and then separated on an analytical column (75 µm id x 50 cm) (PepMap C18, 2µ) over at 300 nL/min. The gradient was a 120 min of 6.4 – 28% CH_3CN/0.1% formic acid/ 150 min cycle time per fraction for phosphopeptides analysis and 90 min gradient of 5 - 30.4% CH_3CN/0.1% formic acid/ 120 min cycle time per fraction for full proteome analysis. The Orbitrap Fusion was operated in the Top Speed mode at 3 s per cycle. The survey scans (m/z 375-1500) were acquired in the Orbitrap at a resolution of 120K (AGC $4x10^5$ and maximum injection time 50 ms). For the phosphopeptides analysis, the multiply charged ions, above 2e4 counts, were subjected to HCD fragmentation with a collision energy at 38% and isolation width 0.7 Th, and MS/MS spectra were acquired in the Orbitrap (AGC $1x10^5$ and maximum injection time 86 ms) with 50K resolution. Dynamic exclusion width was set at ± 10 ppm for 40 s. For the full proteome analysis, the data acquisition used MS3-SPS5 method, i.e. the MS2 fragmentation was in CID at 35% collision energy for multiple charged ions at 5000 counts. Following each MS2, the 5-notch MS3 was performed on the top 5 most abundant fragments isolated by Synchronous Precursor Selection (SPS), by HCD at 65% CE then detected in Orbitrap at m/z 100-500 with 50K resolution to for peptide quantification data. The AGC was set 1.5e5 with maximum injection time at 86 ms.

All raw files were processed in Proteome Discoverer 2.3 (phsophoproteome) or 2.4 (full proteome) (Thermo Fisher) using the SequestHT search engine. Spectra were searched against fasta files of reviewed Uniprot Homo sapiens entries (November 2019) and an in-house contaminate database. Search parameters for phosphoproteome were: trypsin with 2 maximum miss-cleavage sites, mass tolerances at 20ppm for Precursor, and 0.02Da for fragment ions, dynamic modifications of Deamidated (N, Q), Oxidation (M) and Phospho (S, T, Y), and static modifications of Carbamidomethyl (C) and TMT6plex (peptide N-terminus and K). For full proteome, it was 0.5Da for fragment ions, dynamic modification of Oxidation (M) and acetylation (Protein N-terminus), and static modifications of Carbamidomethyl (C) and TMT6plex (peptide N-terminus and K). Peptides were validated by Percolator with q-value set at 0.01 (strict) and 0.05 (relaxed). Phosphorylation site localization probabilities were computed by the ptmRS node. The TMT10plex reporter ion quantifier included 20 ppm integration tolerance on the most confident centroid peak at the MS3 level. Only unique peptides were used for quantification. The co-Isolation

threshold was set at 100%. Peptides with average reported S/N>3 were used for protein quantification, and the SPS mass matches threshold was set at 50%. Only master proteins were reported.

Normalised and scaled abundances were calculated and the mean between the three biological replicates was performed for the full proteome analysis and phosphoproteomics. Scaled abundances were compared with a t-test adjusted for false discovery rate (Benjamini Hochberg). GSEA was performed as described in 2.4.1.3 section using the fGSEA package from the ranked fold change and the clones were compared against the BIOMEDE-169 parental cells. Unsupervised hierarchical heatmaps for proteome were performed using the median centred values. Gene ontology analysis from shared molecular functions was performed using the package clusterProfiler.

2.8 *In vivo* methods

All the *in vivo* experiments were conducted by Dr Diana Martins Carvalho, Postdoctoral Fellow in our laboratory and my associate supervisor; MRI was done by Dr Jessica Boult, a Senior Scientific Officer from the Division of Radiotherapy and Imaging; and immunohistochemistry carried out by Valeria Molinari, a Higher Scientific Officer in our laboratory.

2.8.1 Intracranial injections

The experiments were carried out in accordance with the local ethical review panel, the UK Home Office Animals (Scientific Procedures) Act 1986, the United Kingdom National Cancer Research Institute guidelines for the welfare of animals in cancer research and the ARRIVE (animal research: reporting *in vivo* experiments) guidelines [338].

DIPG patient-derived xenografts (PDX) were established from the same piece of tissue used for *in vitro* models generation described in 2.7.1 section. When the sample was received as cryopreserved this procedure was performed in parallel to the establishment of *in vitro* models. If the samples were received in Hibernate A and animals were available in the facility, the procedure was also performed in parallel; if not a small piece of tissue was cryopreserved for later injection.

A single cell suspension was prepared from cells from tissue (PDX, n=9) immediately prior to implantation in in NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ (NSG) mice (Charles River, UK). Animals were anesthetized with 4% isoflurane and maintained at 2-3% isoflurane (0.5L/min) delivered in oxygen (1L/min). Core body temperature was maintained using a thermo-regulated heated blanket. A subcutaneous injection of buprenorphine (0.03mg/Kg) and Meloxicam (5mg/Kg) was given for general analgesia. Animals were depilated at the incision site and Emla cream 5% (lidocaine/prilocaine) was applied on the skin. The cranium was exposed via midline incision under aseptic conditions, and a 31-gauge burr hole drilled above the injection site. Mice were then placed on a stereotactic apparatus for orthotopic implantation. The coordinates used for the pons were x=+1.0, z=-0.8, y=-4mm from the lambda. 5µL of cell suspension were stereotactically implanted per animal, using a 25-gauge SGE standard fixed needle syringe (SGE™ 005000) at a rate of rate of 1 µl/min for PDX and 2 µl/min for CDX using a digital pump (HA1100, Pico Plus Elite, Harvard Apparatus, Holliston, MA, USA). At the completion of infusion, the syringe needle was allowed to remain in place for at least 2 minutes, and then manually withdrawn slowly to minimize backflow of the injected cell suspension. Mice were monitored until fully recovered from surgery. 24h post-surgery a subcutaneous injection of buprenorphine (0.03 mg/Kg) was administered. Mice were weighed twice a week and sacrificed by cervical dislocation upon deterioration of condition and tissue taken for further analysis

For the CDX (BIOMEDE-169 2D Parental, A.T1, B.T3, C.T4, D.T5 and D.T6, BIOMEDE-169 3D, BIOMEDE-134 3D, BIOMEDE-181 2D and BIOMEDE-181 3D), 250,000 cells were orthotopically injected into the pons of the NSG mice (3-12 per sample) as described above. Mycoplasma test and STR (as described above) was performed before cells were injected.

At endpoint, animals were sacrificed by neck dislocation and the brains were collected. A sagittal cut along the midline was performed, the right hemisphere was fixed in 10% buffered formalin solution for up to 30 hours and the left hemisphere was cryopreserved in Stem Cell Banker. The pons, cerebellum and thalamus were minced and cryopreserved in 5 vials and the cortex was cryopreserved in 1 vial. A small piece of the pons area was snap frozen for molecular analysis.

For serial xenografting, 1 vial of the pons, cerebellum and thalamus was thawed. A single cell suspension was prepared using Liberase TL and 5ul were injected intracranially as described above.

For serial xenografting, 1 vial of the pons, cerebellum and thalamus was thawed. A single cell suspension was prepared using Liberase TL and 5ul were injected intracranially as described above.

2.8.2 Magnetic resonance imaging

Tumours were identified using ^{1}H MRI performed on a horizontal bore Bruker Biospec 70/20 (Ettlingen, Germany) equipped with physiological monitoring equipment (SA Instruments, Stony Brook, NY, USA) using a 2cm x 2cm mouse brain array coil. Anaesthesia was induced using 3% isoflurane delivered in oxygen (1l/min) and maintained at 1-2%. Core body temperature was maintained using a thermo-regulated water-heated blanket. Following optimization of the magnetic field homogeneity using a localised map shim over the whole brain, a rapid acquisition with relaxation enhancement (RARE) T_2-weighted sequence (repetition time (T_R) = 4500ms, effective echo time (T_{Eeff}) = 36ms, 2 averages, RARE factor = 8, in-plane resolution 98µm x 98µm, 1mm thick contiguous axial, coronal and sagittal slices) was used for localisation and assessment of tumours. All imaging was performed by Dr Jessica Boult.

2.8.3 Immunohistochemistry

The right hemisphere of mouse brains was collected and fixed in 10% buffered formalin solution for 30 hours then embedded in paraffin. 4µm-thick sections were cut and stained with haematoxylin and eosin (H&E) (HD Supplies, HS355-1 and HS260-1). For immunohistochemistry, sodium citrate (pH 6.0) heat-mediated antigen retrieval was performed and staining was carried out using antibodies directed against Human Nuclear Antigen (HNA) (1:100, Millipore, MAB4383). All primary antibodies were diluted into 1% Tris buffer solution with 0.05% Tween-20, and incubated 1 hour at room temperature. Novocastra Novolink Polymer Detection Systems Kit (Leica Biosystem, RE-7150) was used for detection. Slides were mounted using Leica CV Ultra mounting medium (Leica, 070937891). This was performed by a Higher Scientific Officer in the lab, Valeria Molinari.

5 days on, 2 days off, and treatment started at d55 post-injection. MRI was performed pre- and post-treatment.

10mg of trametinib were dissolved in 10mL of DMSO (1mg/mL), then 100 µl aliquots were prepared and stored at -20°C. Every day before treatment 100 µl of 1mg/mL trametinib aliquots were dissolve in 0.9 mL of 10%w/v hydroxyproylbetacyclodextrin (dissolved in PBS). 10µl of drug/vehicle was then administered PO per gram of mouse.Mice were monitored by daily weighing and were sacrificed by cervical dislocation upon deterioration of condition, and brain was taken for further analysis. Effects of drug treatment on survival as the primary endpoint were assessed by GraphPad Prism (Kaplan-Meier survival analysis). Mouse brains collected at the end of the efficacy study were processed for IHC and WES as described above.

2.8.4.2 AZD1775

A small piece of cryopreserved tissue from B193 from serial mouse-xenograft was orthotopically injected into the pons of the NSG mice as described above. Mice (29-33 days old) were randomized into two groups: Group 1– vehicle (0.5%methycellulose) and Group 2 - AZD1775 (60mg/kg, PO, q.d.) Animals were treated for 6.1 weeks, 2 weeks on, 1 week off, 5 days on, 2 days off and treatment started at d103 post-injection. MRI was performed pre-and post-treatment.

500 mg of AZD1775 were dissolved in 83.3 ml 0.5% methycellulose (6 mg/mL) and aliquots were stored at -20°C. Every day before treatment aliquots were thawed and 10µl of drug/vehicle was administered PO per gram of mouse. Mice were monitored by daily weighing and were sacrificed by cervical dislocation upon deterioration of condition and tissue taken for further analysis. Mouse brains collected at the end of the efficacy study were processed for IHC.

2.9 Statistical Analysis

Statistical analysis was carried out using R 3.3.0 (www.r-project.org) and GraphPad Prism 8. Descriptive statistics were performed using GraphPad Prism 8.

FFPE and high molecular weight metrics comparisons were performed by using unpaired Student’s t-test, p-value of less than 0.05 was considered significant. For paired FFPE and FF samples comparisons were performed using paired Student’s t-test, p-value of less than 0.05 was considered significant.

cfDNA concentration comparisons amongst molecular subgroups, and GI50 comparisons from the drug screening, one-way ANOVA, Dunnett's multiple comparisons test. p-value of less than 0.05 was considered significant. ctDNA concentration and VAF comparions amongst sample type (CSF vs plasma and CSF vs serum) was performed by using one-way ANOVA, Tukey multiple comparisons test. p-value of less than 0.05 was considered significant.

For AZD1775 GI50 comparisons, unpaired two-tailed Mann-Whitney test. p-value of less than 0.05 was considered significant.

Effects of drug treatment on survival as the primary endpoint and overall survival in the BIOMEDE orthotopic *in vivo* models were assessed using Mantel Cox log-rank test. p-value of less than 0.05 was considered significant.

Normalised and scaled abundances proteins and phosphopeptides were compared by using a Student's t-test adjusted for false discovery rate (FDR) according to Benjamini and Hochberg.For gene expression data analysis multiple testing corrections were made using FDR according to Benjamini and Hochberg. p-value of less than 0.05 was considered significant.

In all other cases, unless otherwise specified, standard two-tailed t-test was applied.

CHAPTER 3 : Development of targeted sequencing panels to identify prognostic, predictive and diagnostic markers in paediatric solid tumours

3.1 Introduction

The genomic landscape of paediatric cancer is becoming increasingly more well-defined leading to the conclusion that childhood cancers have in general fewer somatic mutations than adults, but that mutations in epigenetic regulators occur at a higher incidence [10, 34, 75, 193, 214, 339-345]. Key studies have identified recurrent mutations in histones 3.3 and 3.1 (*H3F3A* and *HIST1H3B*) as well as in the activin A receptor type I (*ACVR1*) that are unique to pHGG and DIPG [73, 74, 76]. Similarly, *ATRX* and *TERT* alterations in neuroblastoma are associated with poor prognosis, as is *MYCN* amplification [194, 195, 346]. The updated World Health Organization (WHO) classification of tumours of the central nervous system (CNS) based on molecular features is a clear example of the huge impact of applying molecular profiling to guide diagnosis and treatment, with the potential to improve outcomes in childhood cancers [26].

With an extensive list of recurrent alterations with potential actionable and diagnostic value for paediatric solid tumours, sequential testing of single genes using standard methods has become unfeasible due to a lack of available material and high costs. Next Generation Sequencing (NGS) offers a solution to these issues. There are different approaches that can be utilized at several levels of complexity including whole genome, whole exome or targeted sequencing. Currently whole genome sequencing (WGS) and whole exome sequencing (WES) are mostly applied for discovery proposes and they remain challenging to apply in routine clinical practise. This is due to the cost and the large amount of DNA required, often not achievable from small formalin fixed paraffin embedded (FFPE) biopsies, and the lack of sufficient depth needed to investigate clinical samples with low neoplastic cell content. Panel-based NGS assays which simultaneously sequence a targeted set of genes with recurrent alterations of known clinical or biological implications are potentially cheaper, more accurate and more suited to clinical diagnostics than current approaches [347].

Development and validation of high-throughput gene panel sequencing is challenging. Typically, DNA is only available from formalin-fixed, paraffin-embedded (FFPE) samples, which yields relatively poor-quality DNA. DNA extraction and library construction to clinical laboratory standards requires optimisation, and it is necessary to construct a standardised informatics pipeline that identifies and interprets actionable mutations. Appropriate and rapid clinical reporting of identified variants and incorporation of the results into the electronic patient records also need to be considered if molecular stratification of childhood cancer is to be successfully translated to the clinic [348]. There are several examples of validation and implementation of targeted sequencing in adult cancer [349-352]. In the past years, several approaches using high-throughput sequencing have been applied for clinical decision-making in children with solid tumours [254, 267, 270, 353], however a clinically validated panel specifically targeting recurrent alterations in childhood cancers using archival FFPE specimens would significantly assist the development of molecular stratification strategies in paediatric oncology.

In this chapter, I describe the development and validation, within an accredited clinical pathology laboratory (CPA UK), of a paediatric solid tumour sequencing assay for use with either routine FFPE or fresh frozen (FF) samples. As part of the validation, I established overall performance, sensitivity, specificity, repeatability, reproducibility, accuracy and limit of detection, following guidelines previously described for validation of genetic tests. In addition, I describe the development and validation of a panel for the detection of structural variants in paediatric brain tumours, in particular, in paediatric high-grade glioma (pHGG). The results of this chapter have been part of several publications [87, 354, 355].

3.2 Results

3.2.1 Gene selection

Two panels, Paeds-v1, (v1) and Paeds-v2, (v2) were developed for clinical diagnostics applications and validated to Good Laboratory and Clinical Practice (GCLP) standards. These comprised 78 genes (311 kb) in the first version (v1) and 91 (473 kb) for an expanded version of the assay (v2) **{Figure 3-1A}**. The genes were selected in collaboration with national experts in paediatric oncology patient care covering all areas of paediatric solid tumours (glioma, medulloblastoma, bone and soft tissue sarcomas, renal tumours and neuroblastoma among others). Targets were

chosen by consensus, based on the most clinically relevant aberrations. The criteria used was whether the alterations fell within one of the following categories: predictive biomarker (level 1), prognostic biomarker (level 2), diagnostic biomarker (3), potentially targetable biomarkers with inhibitors available or under development (level 4), known germline or high-risk single nucleotide polymorphism (level 5), or of unclear significance, research only (level 6). Factors influencing the choice of targets included the childhood tumour type where alterations have been reported, the availability of molecules targeting these genes and whether there were open clinical trials using such agents for children with solid tumours **{Appendix Table 1}**. A library of customized biotinylated DNA probes complementary to the genes of interest was synthesized by Nimblegen (as described in methods section) for the detection of single nucleotide variants (SNVs), short insertion-deletions (indels), copy number variations (CNVs) and structural variants (SVs). Exons were padded with 5 base pairs (bp) of intronic sequence to increase exon depth and for detection of splice-site variants.

The CNS-fusion detection panel, (fusion-panel), is a novel capture-based assay to detect SV in CNS tumours including 24 genes (pilot version ~ 1.39Mb and final version ~ 2.2Mb) **{Figure 3-1B}.** The genes were selected by reviewing recent studies describing the importance and recurrence of structural variants in childhood brain tumours which in some cases is crucial to determine an accurate diagnosis and prognosis.

In addition, a capture-based panel for the detection of SNVs, indels and CNV in pHGG was designed. The targets were selected according to recurrent altered genes in pHGG identified by our laboratory in two published genomics studies [22, 87]. The pHGG-panel was composed of 330 genes (1.2Mb) **{Appendix Table 2}**.

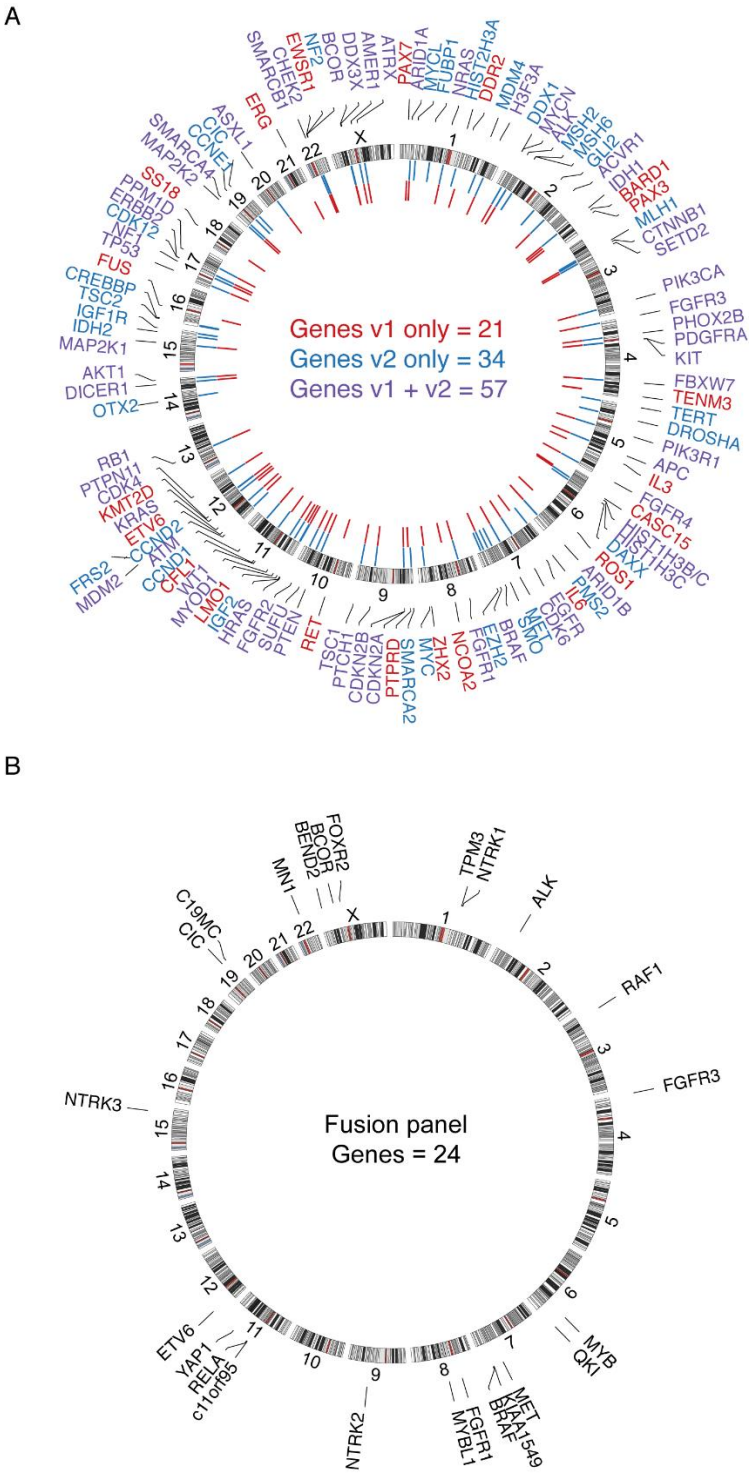


Figure 3-1 Circos plot of the genes included in the targeted panels. (A) The paediatric panels comprise 78 genes in v1 and 91 genes for v2, in red are genes included only in Paeds-v1, in red genes included only in Paeds-v2 and in purple genes included in both versions (B). The fusion-panel is comprised of 24 genes.

3.2.2 Paeds panels validation

To validate the assays in a CPA laboratory, I followed the standardised framework for clinical assay validation set out by Mattocks *et al.* [260]. Overall performance across the target regions was determined, measuring precision, sensitivity and specificity. As a standard, I used a set of four Horizon cell blends previously characterised by NGS and droplet digital PCR (ddPCR) for both Paeds-v1 (v1) and Paeds-v2 (v2). For v1, 15 paediatric cell lines, 83 FFPE and 30 FF clinical samples with known variants were also used. For v2 panel, I used 10 FFPE samples with known variants previously identified on v1.

3.2.2.1 Overall performance

Overall, v1 and v2 panels performed well across the four cell blends and FFPE samples. A total of 24/901 (2.7%) ROI were classified as underperforming for v1 and 14/2330 (0.6%) for v2, with read depth lower than 2 x standard deviation (SD) of the mean based on $\log_2$ [$\log_2$(ROI)>mean($\log_2$(ROI))-2xSD($\log_2$(ROI)). Underperforming ROI were mostly located within highly GC-enriched regions (22/24 for v1 and 11/15 for v2), which are known to be refractory to efficient hybridization and/or amplification **{Figure 3-2}**. The capture from certain regions that underperformed using v1 were significantly improved using v2 such as *TERT* promoter, *CDKN2A* exon 1 *H3F3A* exon 2 and *RB1* exon 1, amongst others **{Figure 3-3}**.

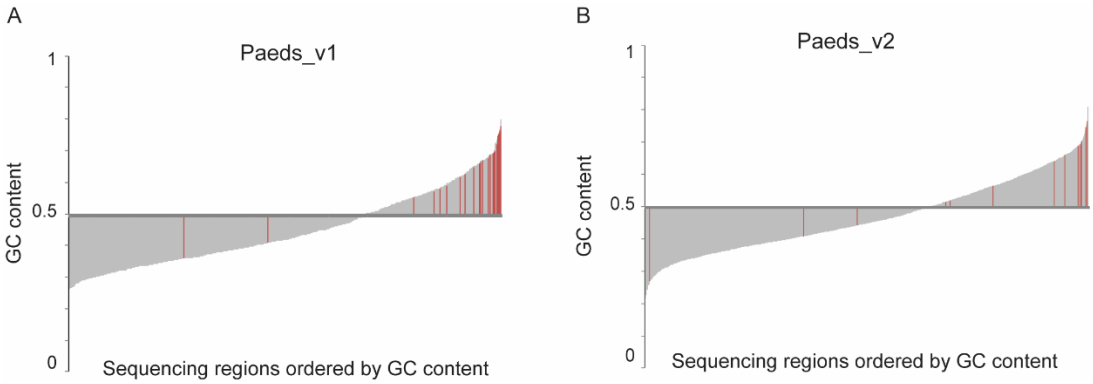


Figure 3-2 Bar-plot showing GC content in the regions capture by the Paeds-v1 and Paeds-v2 panels. The plots are ordered from low to high GC-content of each region capture. Red bars highlight the underperforming regions from (A) v1 (24/901) and (B) from v2 (14/2330). In both cases they were mainly located in high GC-content regions, 22/24 for v1 and 11/14 for v2.

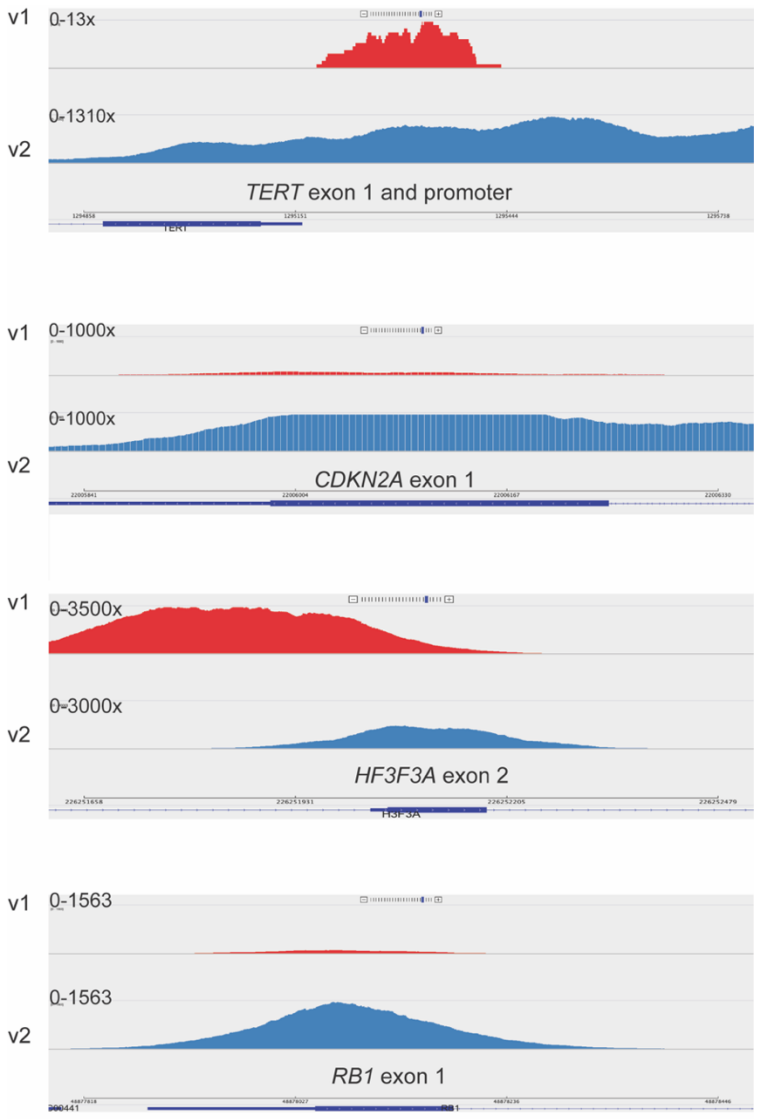


Figure 3-3 Sashimi coverage plots comparing Paeds-v1 and Paeds-v2 underperforming region of interest (ROI). Sample 022 (FFPE) was plot to illustrate the improvement of ROI using Paeds-v2 (blue) compared to Paeds-v1 (red). Underperforming ROI from v1 shown correspond to: *TERT* exon 1 and promoter *CDKN2A* exon 1, *H3F3A* exon 2 and *RB1* exon 1.

Quality and coverage metrics were generated across all samples used as part of the validation. Samples were classified by DNA integrity: fragmented DNA from FFPE cases versus high molecular weight DNA (HMW), including cases from fresh frozen tissue (FF) and cell lines.

For v1, there was no difference for the average total number of reads between FFPE ($8.8x10^6$ ± $3.1x10^6$) and HMW cases ($7.9x10^6$±$3x10^6$) (p=0.0971, unpaired t-test) **{Figure 3-4A}**. The overall mean depth was higher in HMW (899±347) compared to FFPE cases (698±365) (p=0.0022, unpaired t-test) **{Figure 3-4B}**. Duplicates were higher in FFPE (60.2%) compared to HMW cases (36.1%) (p<0.0001, unpaired t-test)

{Figure 3-4C}. The percentage of bases from unique reads on target was 45.9±3 for FFPE and 42.7±2.4 for HMW cases **{Figure 3-4D}**.

For v2, the average total number of reads was lower for FFPE ($9.6x10^6±1.9x10^6$) than for HMW cases ($1.32x10^7±2.1x10^6$) (p=0.0023, unpaired t-test) **{Figure 3-4E}**. The overall mean depth was higher for HMW (752±278) compared to FFPE cases (567±152.6) (p=0.0023, unpaired t test) **{Figure 3-4F}**. As specificity of the probes improved, a single hybridisation was performed in v2 avoiding a PCR step which resulted in reduction of duplicates in the latest v2 panel compared to v1: 20.3% for FFPE (v1=60.2%) and 12.8% for HMW cases (v1=36.1%) **{Figure 3-4G}**. There was no difference in the percentage of bases from unique reads on target (46.6±1.7 for FFPE vs 46.4±2.7 for HMW) (p=0.8912, unpaired t-test) **{Figure 3-4H}**.

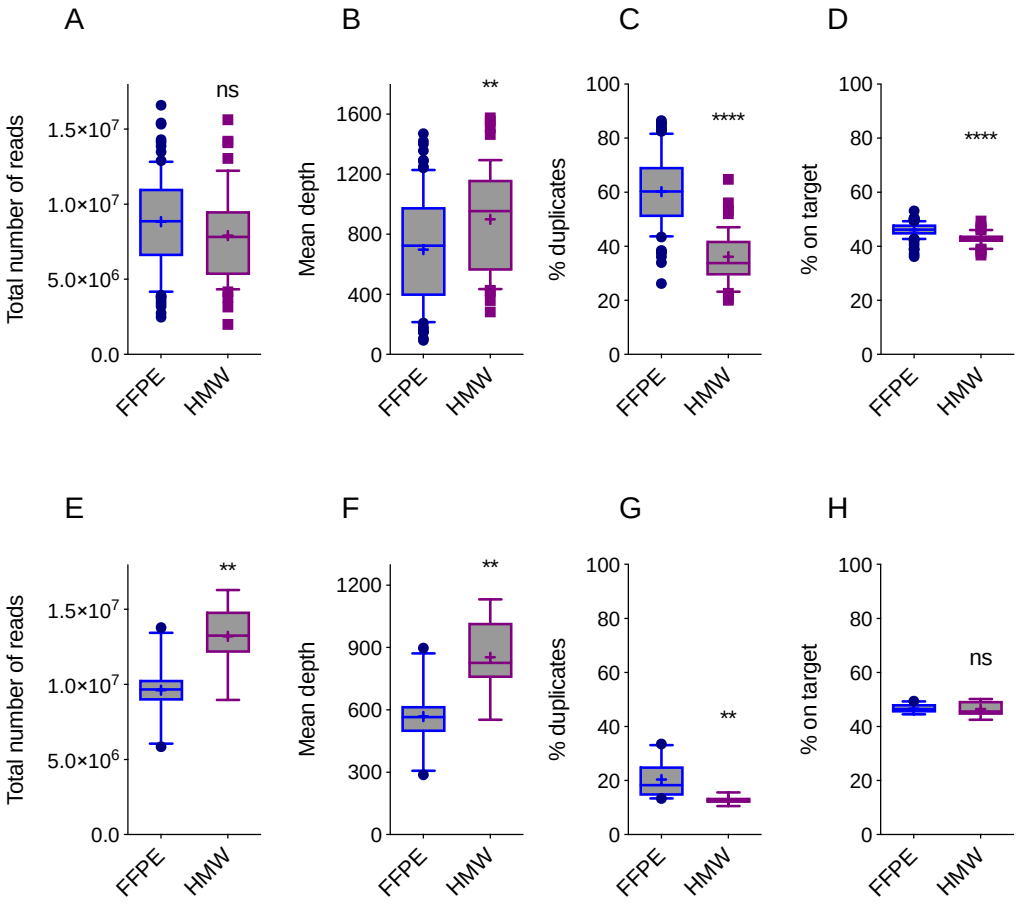


Figure 3-4 Box-plots showing metrics of Paeds-v1 and Paeds-v2 panels. (A-D) Paeds-v1 metrics and (E-H) Paeds-v2 metrics. The parameters shown include quality metrics including number of reads, mean depth, percentage of duplicates and percentage of unique reads on target separated by DNA integrity: fragmented DNA samples from formalin-fixed paraffin embedded samples (FFPE) and high molecular weight DNA from fresh frozen tissue and cell lines (HMW). The thick line within the box is the median, mean is shown as "+", the whiskers are drawn down to the 10th percentile and up to the 90th, points below and above the whiskers are drawn as individual points. All p-values are based upon unpaired t-test ****p<0.00001, ***p<0.0001, **p<0.001, ns p>0.05.

3.2.2.2 Limit of detection

The limit of detection was determined using the cancer SNVs and indels present in the cell blends at known variant allele frequency (VAF) by ddPCR. The pipeline detected all SNVs for both assays (61 for v1 and 76 for v2), including SNVs with an expected VAF of 4-5% (33 for v1 and 40 for v2). Similarly, 15/17 indels were detected for v1 and 25/25 for v2, including indels with an expected VAF of 4-5% (5 for v1 and 9 for v2). Of the two indels not detected using v1, one was 18 bp in length at an expected VAF of 4.2%, whilst the other was 2 bp at 5% VAF. The correlation of VAF for SNVs and indels between ddPCR and the panels was r^2=0.969 [95%CI:0.8910-0.9670] for v1 **{Figure 3-5A}** and r^2=0.9628 [95%CI: 0.8685 to 0.9393] for v2 **{Figure 3-5B}**. Therefore, a minimum threshold of 5% VAF in the analysis pipeline was established, which allows for detection of a heterozygous mutation when >10% neoplastic cells are present in the tumour sample.

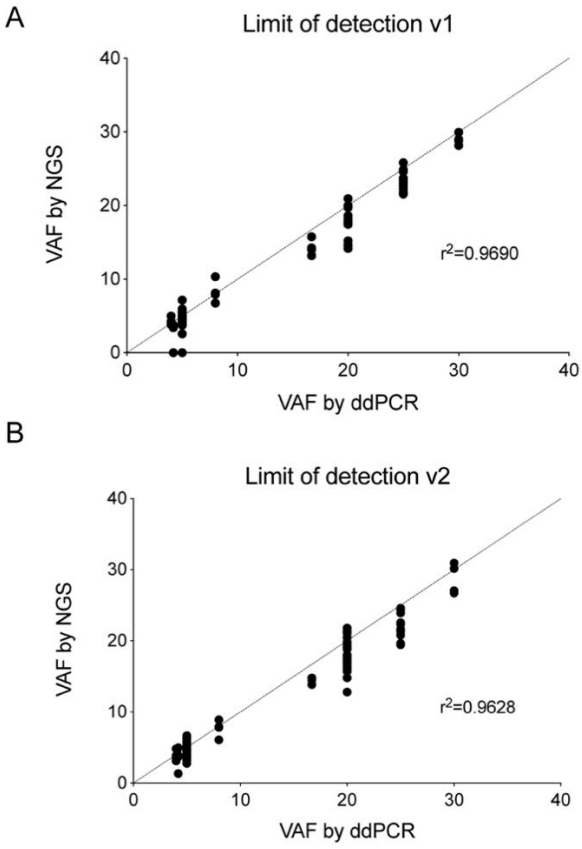


Figure 3-5 Correlation of known variant allele frequencies (VAFs) by droplet digital PCR (x axis) against VAF obtained by NGS (y axis) for all cancer-specific variants using (A) Paeds-v1 and (B) Paeds-v2.

3.2.2.3 Assessment of precision

To measure precision, I took advantage of natural variants present as intrinsic "background" SNVs and indels in the captured regions from the four cell blends. Precision was assessed by comparing the alterations expected with those detected to obtain within-run precision (repeatability), and between-run precision data (intermediate precision). Variants ≤ 5% AF in all four blends and within poor performing regions were excluded leaving a total of 528 SNVs (132 variants in 4 blends) and 108 indels (27 indels in 4 blends) for v1 analysis and 1272 SNVs (318 variants in 4 blends) and 141 indels (47 indels in 4 blends) for v2 analysis. 100% of SNVs were detected for v1 and 99% for v2. By contrast, 83.3% indels were detected for v1 and 96% for v2.

Repeatability. Pairwise correlation of VAF was: v1 r^2=0.994 [95%CI:0.991-0.996] for SNVs **{Figure 3-6A-B}** and $r^2 \geq 0.785$ [95%CI:0.652-0.919] for indels **{Figure 3-6C-D}**; v2 r^2=0.985 [95%CI:0.982-0.988] for SNVs **{Figure 3-6E-F}** and r^2=0.897 [95%CI:0.844-0.949] for indels **{Figure 3-6G-H}**. These results indicate that both panels accurately reproduce data from independently prepared samples on the same run.

Intermediate precision. Pairwise correlation for each of the four samples between two independent runs for v1 was $r^2 \geq 0.995$ [95%CI:0.993-0.997] for SNVs **{Figure 3-7A}** and $r^2 \geq 0.827$ [95%CI: 0.716-0.937] for indels **{Figure 3-7B}** while overall correlation for was $r^2 \geq 0.996$ [95% CI: 0.995-0.997] for SNVs **{Figure 3-7C}** and $r^2 \geq 0.875$ [95%CI:0.829-0.921] for indels **{Figure 3-7D}**. Pairwise correlation for v2 was: $r^2 \geq 0.986$ [95%CI:0.984-0.989] for SNVs **{Figure 3-7E}** and $r^2 \geq 0.916$ [95%CI: 0.874-0.959] for indels **{Figure 3-7F}** while overall correlation was $r^2 \geq 0.990$ [0.989-0.991] **{Figure 3-7G}** for SNVs and $r^2 \geq 0.937$ [95%CI:0.919-0.953] for indels **{Figure 3-7H}**.These results indicate that both panels accurately reproduce data from repeat samples on different runs.

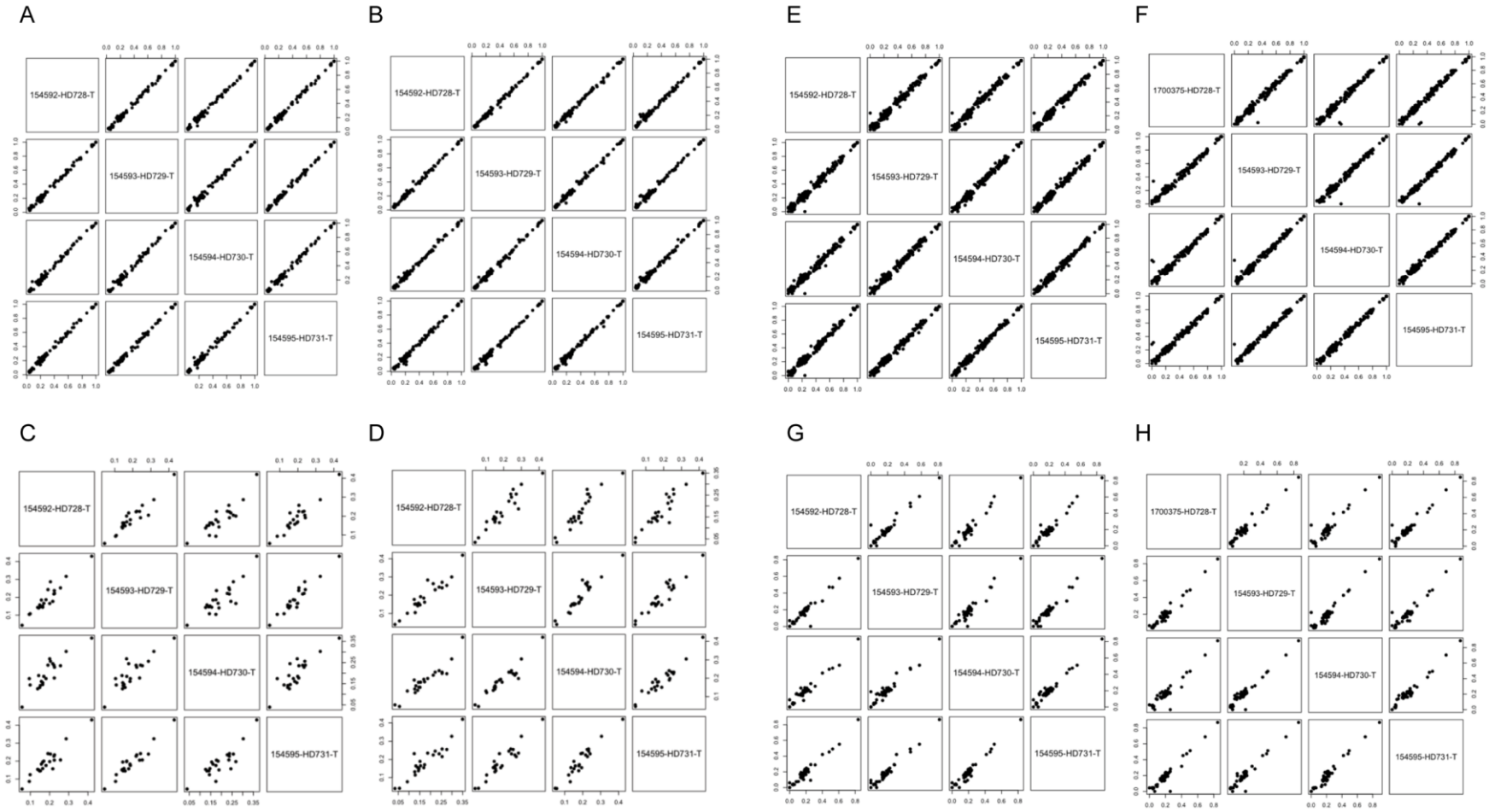


Figure 3-6 Repeatability of Paeds-v1 and Paeds-v2 panels validation. Consistency of single nucleotide variants (SNVs) and insertions-deletions (indels) allele frequency in the four HD blends with identical background variants for (A-D) Paeds-v1 and (E-H) for Paeds-v2: (A) SNVs run_user1_v1, (B). SNVs run_user2_v1 (C), indels run_user1_v1, (D) indels run_user2_v1, (E) SNVs run_user1_v2, (F) SNVs Run_user2_v2, (G) run_user1_v2 and (H) Indels Run_user2_v2.

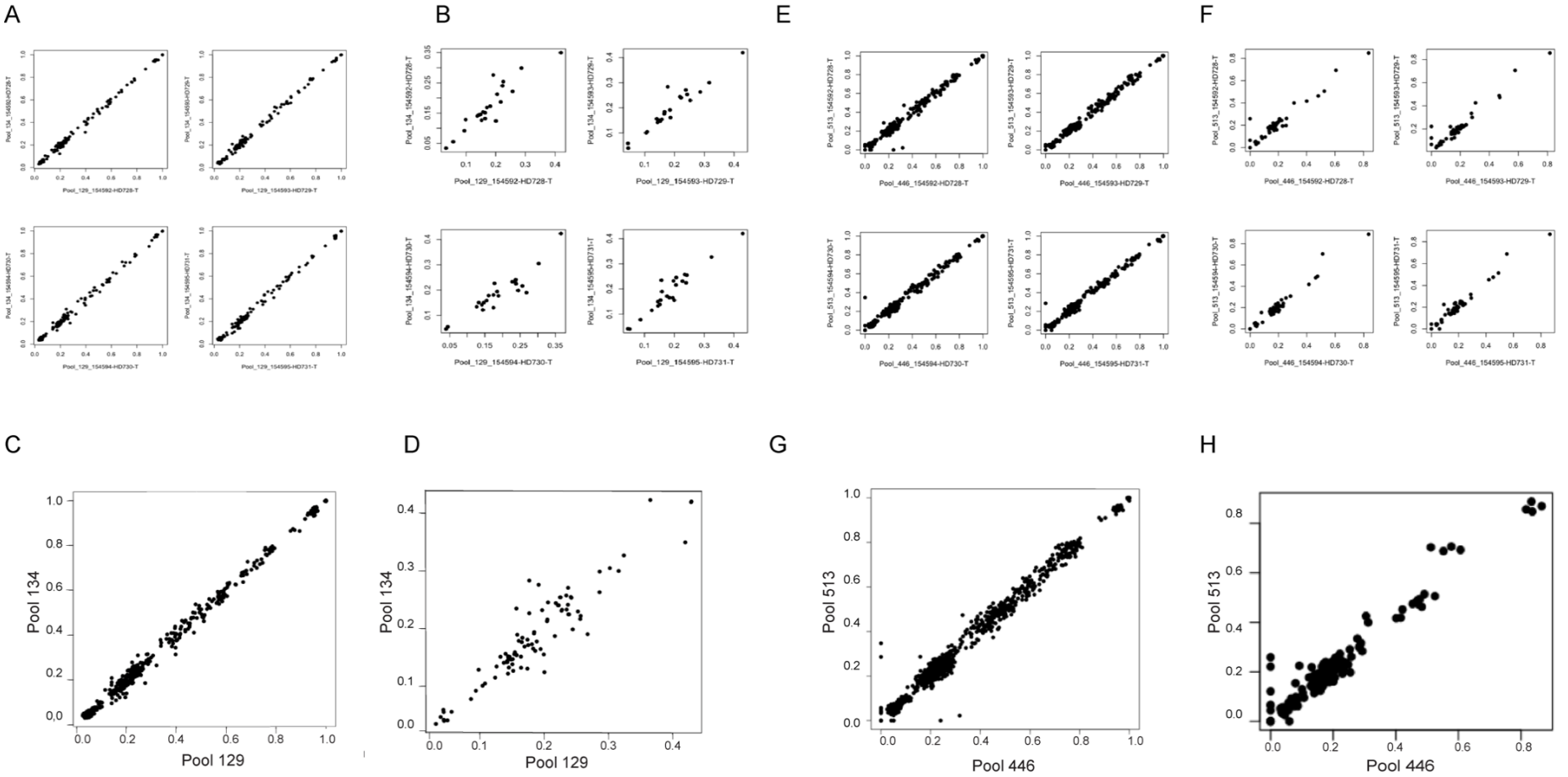


Figure 3-7 Intermediate precision of Paeds-v1 and Paeds-v2 panels validation. Pairwise correlation of variant allele frequency for each of the four HD blends between the two runs (x axis corresponds to run_user1 and y axis to run_user2) (A-D) for Paeds-v1 and (E-H) for Paeds-v2. (A) SNVs for v1, (B) indels for v1, (E) SNVs for v2, (F) indels for v2. Overall correlation of variant allele frequency for the HD blends between the two runs analysing the four samples together (x axis corresponds to run_user1 and axis to run_user2) (C) SNVs for v1, (D) indels for v1, (G) SNVs for v2, and (H) indels for v2.

3.2.2.4 Assessment of sensitivity and specificity

To determine sensitivity, the same background variants (v1=528 SNVs and 108 indels; v2=1272 SNVs and 188 indels) together with the known cancer-specific variants (v1=61 SNVs and 17 indels; v2=76 SNVs and 25 indels) from the four cell blends were used. SNVs and indels were called and their presence was compared to the list of variants expected in the capture regions from the cell blends. The sensitivity for detection of SNVs was ≥98% [95%CI:0.98-1] for v1 and ≥99.8% [95%CI: 0.99=1] for v2. The sensitivity of detection of indels was ≥83% [95%CI:0.761-0.897] for v1 and ≥98.5% [95%CI: 0.97-1] for v2.

For v1, all 589/589 SNVs were detected as True Positives (TP), resulting in the absence of any False-Negatives (FN). For indels, 105/125 TPs were detected, with 20/125 FNs. The undetected indels were manually checked on Integrative Genomics Viewer (IGV). I observed that 12 of 20 were located +4 bp upstream of the exon (our bed file covers ±5 bp), four had poor coverage, two fell in highly repetitive regions and one was a long indel (18 bp). For v2, the FN rate was 17/1331 for SNVs, and 15/213 for indels. The undetected indels and SNVs were manually checked on Integrative Genomics Viewer (IGV). 25/35 of the cancer-specific SNPs and indels present at 5% by ddPCR by random sampling effect were detected <5% by sequencing. Using a cut-off of 5% VAF these variants would not be reported.

To determine specificity, the cancer-specific data from the four cell blends were used harbouring a total of 61 TP and 87 true negative (TN) SNVs for v1 and 76 TP and 100 TN for v2. There were insufficient TN for v1 and v2 to determine specificity for indels (n=3). SNVs were called and their presence was compared to the list of variants expected in the capture regions from the cell blends. The specificity of cancer-specific SNVs was ≥98% [95%CI:0.946-1] for v1 and ≥99% [95%CI: 0.953-1] for v2. Positive-Predictive Value (PPV) was ≥98% [95%CI:0.926-1] for v1 and ≥99% [95%CI:0.94-1] for v2. Negative-Predictive Value (NPV) was ≥98% [95%CI:0.946-1] for v1 and ≥99% [95%CI:0.953-1] for v2.

3.2.2.5 Performance and variant detection comparison in paired FF-FFPE clinical samples

To assess the performance of v1 panel on FFPE clinical samples, 15 paired DNA samples isolated from both FF and FFPE paediatric samples were sequenced. This

was done by comparing samples QC, variant detection as well as correlation of VAF. Due to poor quality of one FFPE sample, a paired case was excluded for metric analysis. There was no difference between the number of reads for FFPE ($9.1x10^6 \pm 3.1 x10^6$) compared to FF cases ($7.8x10^6 \pm 1.2 x10^6$) (p=0.2047, paired t-test) **{Figure 3-8A}**. There was no difference for overall mean depth between FFPE (829±297) and FF cases (985±145) (p=0.1344, paired t-test) **{Figure 3-8B}**. As expected, percentage of duplicates were substantially higher in FFPE (53%±8.5%) compared to FF cases (30%±7.1%) (p<0.0001, paired t-test) **{Figure 3-8C}**. The overall percentage from unique reads on target was lower for FFPE (47.5%±2.3%) compared to FF cases (43.8%±2.1%) (p=0.0009, paired t-test) **{Figure 3-8D}**.There were no difference for the average of targeted positions covered at 250x depth between FFPE (84.6%±11.7%) and FF cases (91.1%±1.9%) (p=0.0758, paired t test) **{Figure 3-8E}**. Pre-capture DNA library-prep fragment size was lower for FFPE (280bp±19bp) compared to FF (325bp±25bp) (p=0.0003, paired t-test) **{Figure 3-8F}**.

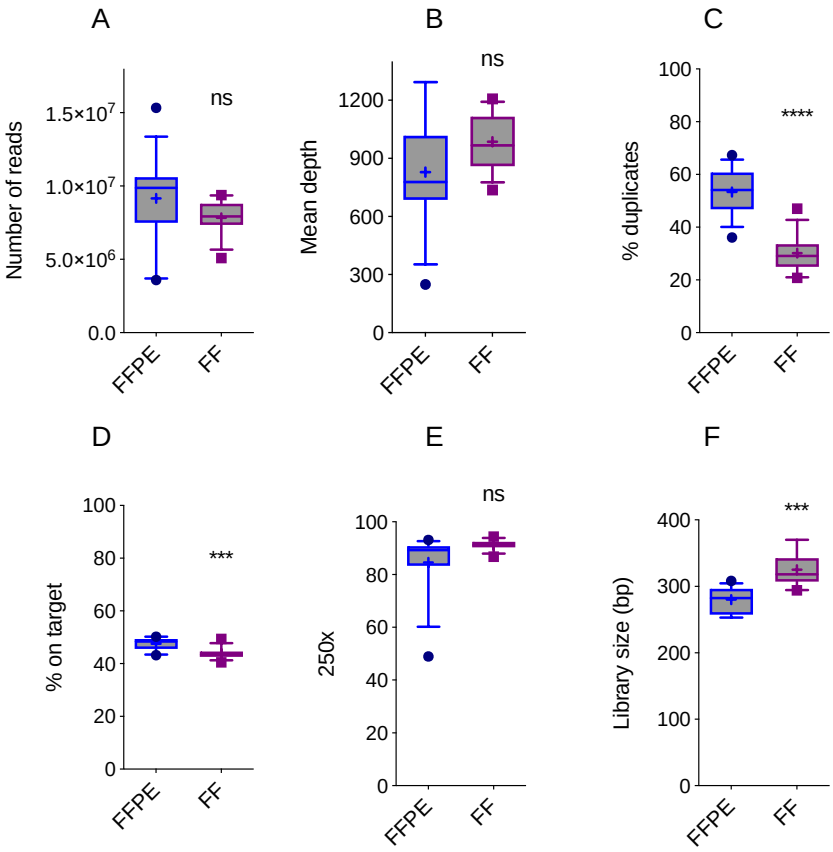


Figure 3-8 Box-plots showing quality metrics of paired FFPE and FF samples. Quality control metrics are separated by DNA integrity formalin-fixed paraffin embedded (FFPE) and fresh frozen (FF) matched samples (n=15). Metrics included (A) number of reads, (B) mean depth, (C) percentage of duplicates, (D) percentage of unique reads on target, (E) percentage of targeted positions covered at 250X, and (F) insert size of pre-capture library in base-pairs. The thick line within the box is the median, mean is shown as "+", the whiskers are drawn down to the 10th percentile and up to the 90th, points below and above the whiskers are drawn as individual points. All p-values are based upon paired t-test. ****p<0.00001, ***p<0.0001, ns p>0.05.

VAFs found in the paired FF-FFPE samples were compared, obtaining an overall correlation of r^2=0.983 (95%CI: 0.984-0.985; p<0.0001) **{Figure 3-9}.** A total of 42.3% (5562/13146) variants were detected in FF but not in FFPE, of which 78.1% (4346/5562) had VAF below 5%, with 17.6% (982/5562) having VAF between 5-10%. Less than 5% variants missed in FFPE samples were present in FF at VAF above 10%. Conversely, a total of 8.2% (1084/13146) variants were detected in FFPE but not in FF, of which 50.8% (551/1084) had VAF below 5%, with 33.2% (360/1084) having VAF between 5-10%, and the remaining 16.0% (173/1084) were present in FFPE only at VAF above 10%.

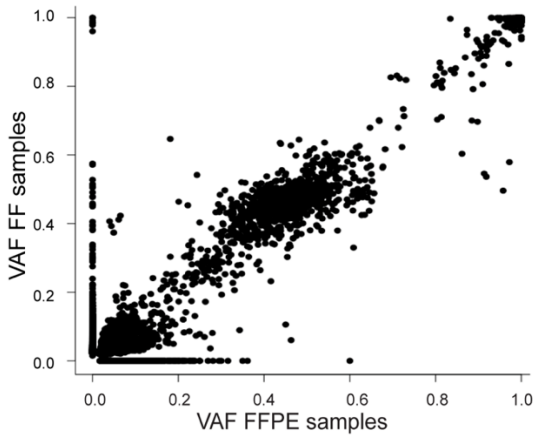


Figure 3-9 Overall correlation of variant allele frequencies (VAFs) found between the 15 formalin-fixed paraffin embedded (x axis) and fresh frozen (y axis) paired samples.

3.2.2.6 Detection of known variants in paediatric samples

To assess the ability of the v1 to detect known variants in clinical samples, variant analysis of 41 paediatric samples with 90 known genetic abnormalities was performed. These samples presented 30 variants in 13 cell lines, and 60 alterations in 14 FFPE and 14 FF samples with genetic alterations identified by routine testing. Of these 50 were SNVs, including point mutations in *TP53, ALK, CTNNB1, DDX3X, SMARCA4* among others, one duplication (*BRAF*-T599dup), 7 indels including *DDX3X* and *TP53*, 13 were amplifications including *MYCN* and *CDK4*, and 19 chromosome losses such as chromosome 9q loss including LOH of *PTCH1* and *TSC1*. 100% of the variants interrogated by the panel were successfully detected **{Table 3-1A}** and **{Table 3-1B}**. Furthermore, to assesses the ability of v2 to detect known variants in clinical samples, ten paediatric FFPE cases were sequenced, harbouring 62 alterations (60 SNVs and 2 indels), previously run on v1. All the mutations were detected obtaining a correlation of VAF of r^2=0.9301 [95% CI:0.91-1]. **{Figure 3-10}**.

A

Cell line ID	Gene	Alteration	Detected	Allele frequency expected	Allele frequency observed
Be(2)C	*TP53*	p.C135F	YES	no data available	100%
Be(2)C	*MYCN*	AMPLIFICATION	YES	no applicable	no applicable
CCA	*KRAS*	p.Q61L	YES	no data available	29%
IMR32	*ATM*	p.V2716A	YES	59%	59%
IMR32	*MYCN*	AMPLIFICATION	YES	no applicable	no applicable
KELLY	*ALK*	p.F1174L	YES	39%	32%
KELLY	*MAP2K1*	p.A390T	YES	48%	47%
KELLY	*TP53*	p.P177T	YES	93%	99%
KELLY	*MYCN*	AMPLIFICATION	YES	no applicable	no applicable
LAN1	*ALK*	p.F1174L	YES	no data available	47%
LAN1	*TP53*	p.C182*	YES	no data available	99%
LAN1	*MYCN*	AMPLIFICATION	YES	no applicable	no applicable
LAN5	*ALK*	p.R1275Q	YES	no data available	50%
LAN5	*MYCN*	AMPLIFICATION	YES	no applicable	no applicable
NBLS	*NF1*	splice_acceptor_variant c.6705-1G>T	YES	no data available	42%
RD	*ATM**	p.D273N	YES	17%	2%
RD	*NF1*	p.E977*	YES	56%	59%
RD	*NRAS*	p.Q61H	YES	68%	61%
RD	*TP53*	p.R248W	YES	100%	100%
RH30	*CDK4*	AMPLIFICATION	YES	no applicable	no applicable
RH41	*APC*	p.M526L	YES	60%	59%
RH41	*TP53*	p.P152fs	YES	100%	100%
RMS559	*FGFR4*	p.V582L	YES	no data available	76%
SKNAS	*NRAS*	p.Q61L	YES	45%	46%
SKNAS	*RB1*	p.L477P	YES	47%	31%
SKNAS	*TP53*	DEL exons 10,11	YES	no applicable	no applicable
SKNSH	*NRAS*	p.Q61L	YES	15%	23%
SKNSH	*SMARCA4*	p.R973T	YES	32%	45%
SKNSH	*CHEK2*	p.T410fs	YES	59%	44%
SKNSH	*ALK*	p.F1174L	YES	no data available	36%

**ATM* mutation in this cell line is subclonal and variation in AF is expected with on-going passages

B

Genes with alterations detected by other methodologies	Alteration	Tumour Type	Total cases expected	% of cases detected
DDX3X	*SNV and indel*	*Medulloblastoma*	6	100
PTCH1	*SNV and indel*	*Medulloblastoma*	5	100
TP53	*SNV and indel*	*Medulloblastoma*	3	100
MYCN	*SNV*	*Medulloblastoma*	2	100
MYCN	*Amplification*	*Neuroblastoma (n=3) Medulloblastoma (n=4)*	7	100
CTNNB1	*SNV*	*Medulloblastoma*	5	100
H3F3A	*SNV*	*Glioma*	3	100
SMARCA4	*SNV*	*Medulloblastoma*	3	100
BRAF	*SNV*	*Glioma*	2	100
ALK	*SNV*	*Neuroblastoma*	1	100
HIST1H3B	*SNV*	*Glioma*	1	100
AKT1	*SNV*	*Medulloblastoma*	1	100
ACVR1	*SNV*	*Medulloblastoma*	1	100
PIK3CA	*SNV*	*Medulloblastoma*	1	100
MLL2	*SNV*	*Medulloblastoma*	1	100
chr 9q - (PTCH1, TSC1)	*loss*	*Medulloblastoma*	5	100
chr 10- (PTEN, SUFU, FGFR2)	*loss*	*Medulloblastoma*	4	100
chr 6- (HIST1H3B, HIST1H3C, ROS1, ARID1B)	*loss*	*Medulloblastoma*	2	100
chr 9- (CDKN2A/B, PTHC1, TSC1)	*loss*	*Medulloblastoma*	2	100
chr12- (MLL2, CDK4)	*loss*	*Medulloblastoma*	1	100
ATM LOH	*loss*	*Medulloblastoma*	1	100
chr 3p- (CTNNB1, STED2)	*loss*	*Medulloblastoma*	1	100
chr17- (TP53, NF1, HER2, PPM1D)	*loss*	*Medulloblastoma*	1	100
chr17p- (NF1, TP53)	*loss*	*Medulloblastoma*	1	100
Total			60	

Table 3-1 Molecular alterations in samples used for Paeds-v1 panel validation. A) Known variants in paediatric cancer cell lines were compared against capture sequencing from the Cancer Cell Line Encyclopaedia and other published data. (B) Known variants in paediatric formalin-fixed paraffin embedded (FFPE, n=14) and fresh frozen (FF, n=14) samples were compared against other platforms such as RNA seq, 450k array, Sanger Sequencing and *FISH.*

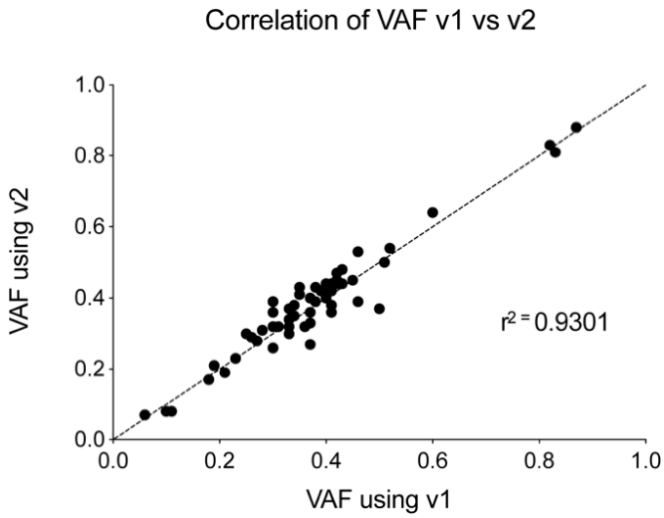


Figure 3-10 Correlation of variant allele frequencies (VAFs) obtained by Paeds-v1 (x axis) and Paeds-v2 (y axis) for cancer variants in FFPE patient samples (60 SNPs and 2 indels).

3.2.2.7 Detection of rearrangements

Five sarcoma FFPE cases were included in the analysis where translocations had previously been detected by RT-qPCR involving *EWSR1*. Rearrangements in *EWSR1* were detected in four out of the five FFPE cases (80%) leading to fusion genes of *EWSR1* with partners *ATF1* (detected in two samples), *FLI1* and *CREB1* **{Figure 3-11}**. The fusion not detected was *EWSR1-NR4A3*.

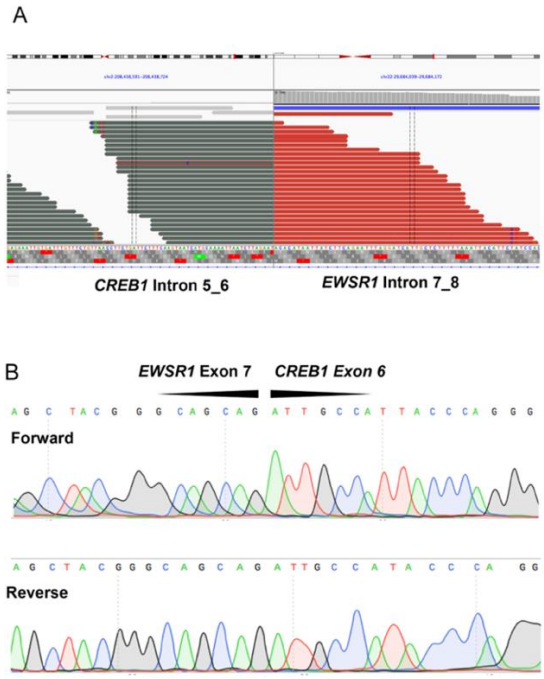


Figure 3-11 *EWSR1-CREB1* translocation in a sarcoma sample known to be positive for this translocation. (A) Integrative Genomics Viewer plot (IGV) identified by the panel showing DNA supporting reads of the fusion gene between *CREB1* intron 5 to 6 and *EWSR1* intron 7 to 8 and (B) electropherogram by Sanger sequencing showing cDNA forward sequence in the top and reverse sequence in the bottom for the fusion between *EWSR1* exon 7 and *CREB1* exon 6.

3.2.3 Fusion-panel validation

3.2.3.1 Overall performance and SV identification in trial fusion-panel

To validate this assay for research purposes, I first tested a pilot panel with a small number of reactions to determine if the use of this technology was adequate to detect SV in FFPE and FF from pHGG tumours, including large genes such as *NTRK1-3*. A total of eight (FFPE=4, HMW=4) cases, were assessed in two MiSeq runs. Overall the fusion-panel had a good performance with over 94.6% (585/618) of ROI achieving an average coverage >100 in the 4 HMW and 4 FFPE.

There was no difference for the average total number of reads between FFPE 1x5x10^7 ±9.8x10^6) and HMW cases (1.17x10^7 ±1.5x10^6) (p=0.4990, unpaired t-test) **{Figure 3-12A}**. The overall mean depth was lower for FFPE (296±138) compared to HMW cases (597±83) (p=0.0097, unpaired t-test) **{Figure 3-12B}**. Duplicates were higher in FFPE (62.2%±12.09) compared to HMW cases (11.5%±0.4%) (p=0.0002 unpaired t-test) **{Figure 3-12C}**. Amongst the eight samples included, four (two FFPE and two HMW) harboured known translocations including *ETV6:NTRK3* (n=2), *MN1:BEND2* (n=1) and *ALK* with unknown partner (n=1). Using MANTA, we were able to detect three fusions with high confidence and one (*ALK*) with medium confidence (3 spanning read pairs and 13 split-reads) **{Table 3 2A}.**

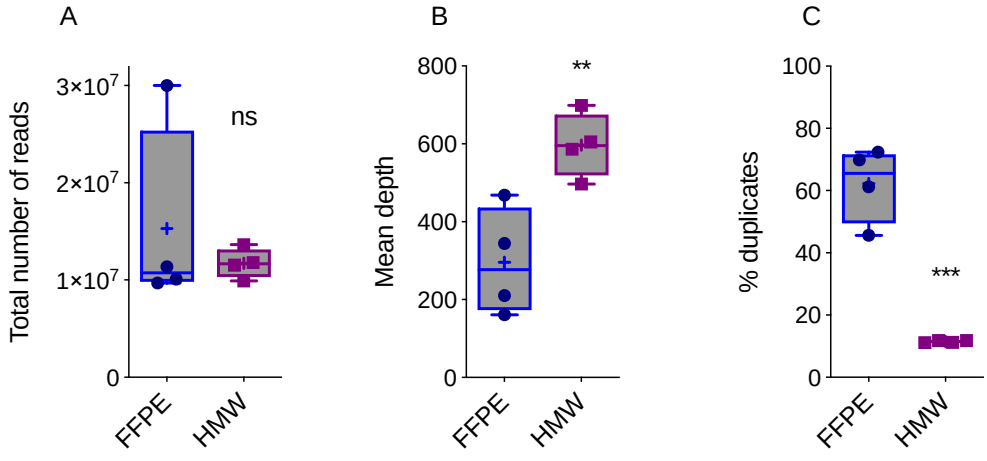


Figure 3-12 Box-plots showing quality metrics for fusion-panel validation. Quality control metrics are separated by DNA integrity formalin-fixed paraffin embedded (FFPE) and high molecular weight DNA (FF and cell lines). Metrics included (A) number of reads, (B) mean depth and (C) percentage of duplicates. The thick line within the box is the median, mean is shown as "+", the whiskers are drawn down to the 10th percentile and up to the 90th, points below and above the whiskers are drawn as individual points. All p-values are based upon unpaired t-test ***p<0.0001, **p<0.001, ns p>0.05.

3.2.3.2 Fusion-panel optimisation

There were some intronic regions associated with repetitive sequence elements in which probes were not able to be designed, therefore coverage-gaps were observed when looking at IGV as well as some regions with bad performance. To overcome this issue and avoid false negative detection, optimisation of the library-prep and sequencing was performed as well as improvement in the panel design as described in Methods section 2.3.2. Longer library fragment sizes (250-500 bp) were generated by reducing the fragmentation time (from 30 min to 15 min) as well as modifying the ratio of Ampure beads and DNA in the dual-SPRI size selection step. The latest step was modified by decreasing the ratio of AMPure-XP beads and DNA in the upper-size cut, from 0.7x to 0.5x, as well as by increasing the ratio in the lower-size cut, from 0.9x to 0.7x **{Figure 3-13}**.

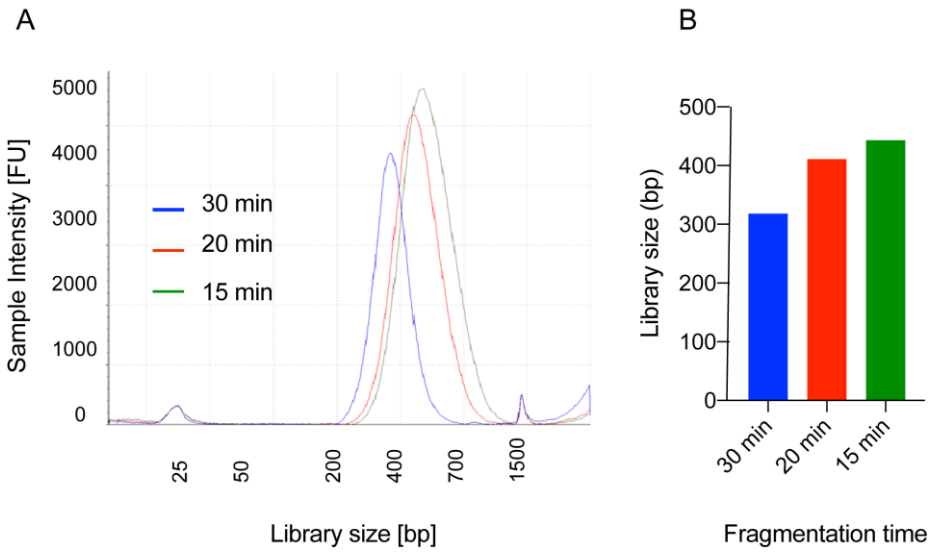


Figure 3-13 Library fragment size optimisation for fusion-panel validation. (A) Library-preparation fragment size distribution assessed using Tapestation per enzymatic time condition. (B) Bar-plot of average fragment library size per enzymatic time condition. Optimisation was performed in (R077 cell line) same sample with the aim to achieve longer fragment size using different enzymatic time: in blue 30 min, red 20 min and green 15 min.

To assess if longer library fragment size and longer sequence reads improved the performance of the pilot fusion-panel, I used a total of 20 samples consisting of 19 glioma cases and one non-small cell lung carcinoma (NSCLC). The cases included ten FFPE and ten FF, from which seven samples harboured known fusions (*MN1:BEND2* n=4, *ETV6:NTRK3* n=1, *ALK* n=2 with unknown partners) **{Table 3 2B}.** Longer library-pep average size fragments were obtained for both FFPE (317±57bp) and FF (466±17bp). Two FFPE samples had a mean depth <100, and one >1800, and were excluded for metrics analysis. The average mean depth was improved for

both FFPE (458±565) and FF (857±149). 6/7 expected fusions were detected by MANTA as well as one *FGFR1* intragenic tandem duplication (ITD) in a patient with HGG. The SV not detected was in the NSCLC patient harbouring an *ALK* fusion. The panel covers *ALK* from exon 20 to 28, with the most common breakpoint occurring within intron 19 to 20; the mean depth for *ALK* intron 19-20 was 1149 and no capture gaps were observed **{Figure 3-14}**. It is possible that the tumour content was too low for the fusion detection or the breakpoint was outside the region covered by the panel. We subsequently decided to cover the whole gene in the updated version.

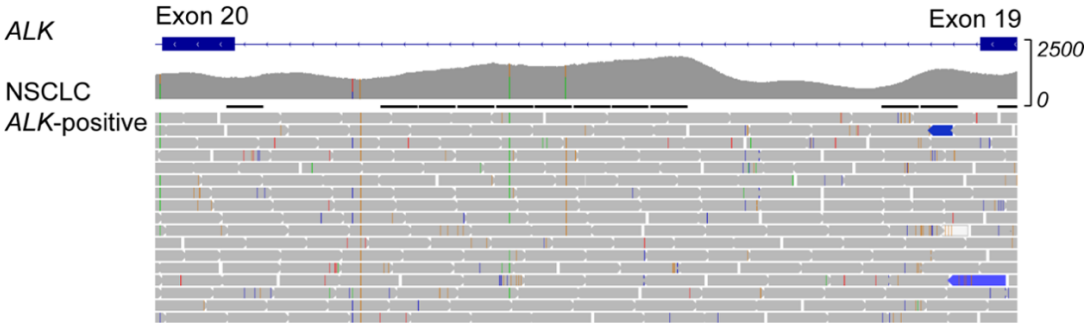


Figure 3-14 Integrative genomes viewer (IGV) plot of NSCLC with expected *ALK* fusion showing the coverage for *ALK* intron 19-20.

To improve the design, ten underperforming ROI assigned for copy number detection were removed and replaced with probes in different genomic regions. We increased the capture probe ratio in regions which were observed to be poorly covered 753/2670 (28,3%) and we included two additional genes (*BEND2* and *QK1*). To test the optimised panel, I performed a run with four FFPE cases, harbouring known SVs, in a MiSeq sequencer by using the long fragment protocol and one overnight hybridisation. Only 2% (14/696) of the ROI had a mean depth of <100x in the four FFPE cases. The Inter-Quartile Range (IQR) was calculated and compared with the pilot version. The optimised fusion-panel presented a lower IQR indicating a better sequencing coverage uniformity (IQR=194 versus IQR=364) **{Figure 3-15}** and **{Figure 3-16}**. Duplicates were reduced compared to the two-night hybridisation protocol from 58.8% to 29.9%. The four structural variants were detected, including three with high confidence calls: *FGFR1* tandem duplication, *MN1:BEND2*, *BRAF:KIAA1549* and one with medium confidence (as previously) *ALK:SPECC1L-ADORA2A* **{Table 3 2C}**.

A

Pilot-trial first test runs

Sample ID	Sample	Tumour type	Gene	Chr	Alt	Fusion
QCTB-R061	FF	Astroblastoma	*MN1*	chr22	]chrX:18207283]C	*MN1:BEND2*
QCTB-R077	CELLS	Infant Glioma	*NTRK3*	chr15	AGAT]chr12:12027420]	*ETV6:NTRK3*
RMH7975	FFPE	Infant Glioma	*NTRK3*	chr15	T]chr12:12033406]	*ETV6:NTRK3*
RMH7977	FFPE	Infant Glioma	*ALK*	chr2	T]chr22:24732484]	*ALK:SPECC1L-ADORA2A*

B

Pilot-trial second test runs

Sample ID	Sample	Tumour Type	Gene	Chr	Alt	Structural variant
QCTB-R061	FF	Astroblastoma-Multiregion	*MN1*	chr22	]chrX:18207283]C	*MN1:BEND2*
QCTB-R077	CELLS	Infant Glioma	*NTRK3*	chr15	AGAT]chr12:12027420]	*ETV6:NTRK3*
RMH7716	FFPE	Astroblastoma-Multiregion	*MN1*	chr22	]X:18207283]C	*MN1:BEND2*
RMH7719	FFPE	Astroblastoma-Multiregion	*MN1*	chr22	]X:18207283]C	*MN1:BEND2*
RMH7724	FFPE	Astroblastoma-Multiregion	*MN1*	chr22	]X:18207283]C	*MN1:BEND2*
RMH7981	FFPE	Infant Glioma	*ALK*	chr2	T]chr22:24732484]	*ALK:SPECC1L-ADORA2A*
HERBY-2411	FFPE	High-grade Glioma	*FGFR1*	chr8	chr8:38271163-38277125	*ITD- FGFR1*
1609648	FFPE	NSCLC	*ALK*	chr2	*not detected*	*not detected*

C

Optimised-panel test run

Sample ID	Sample	Tumour type	Gene	Chr	Alt	Fusion
HERBY-2411	FFPE	High-grade Glioma	*FGFR1*	chr8	chr8:38271163:38277125	*ITD- FGFR1*
RMH7719	FFPE	Astroblastoma-Multiregion	*MN1*	chr22	]X:18207283]C	*MN1:BEND2*
RMH7977	FFPE	Infant Glioma	*ALK*	chr2	T]chr22:24732484]	*ALK:SPECC1L-ADORA2A*
RMH9154	FFPE	Pilocytic Astrocytoma	*BRAF*	chr2	chr7:138538295-140493672	*BRAF:KIAA1549*

Table 3 2 List of samples harbouring fusions used to validate the fusion-panel. Type of sample is listed as well as gene nomination by MANTA and Breakdancer, breakpoint chromosome position. (A) Pilot trial first runs, (B) pilot trial optimisation steps and (C) updated panel test run.

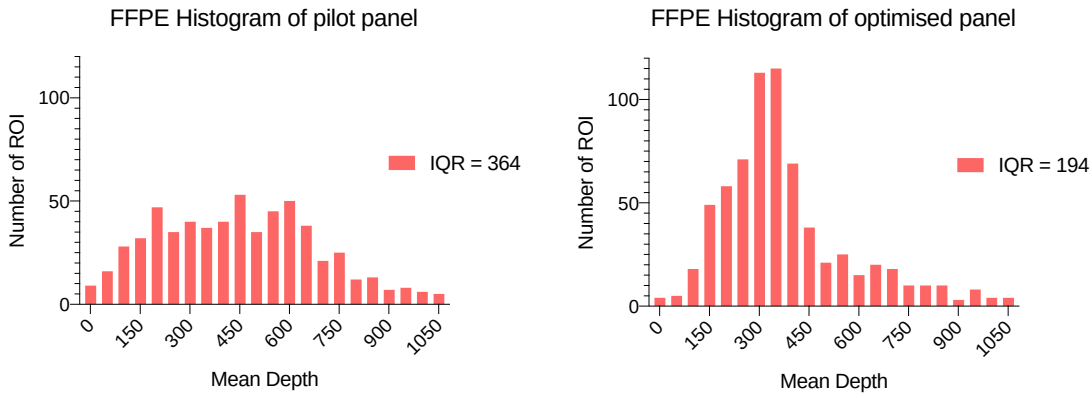


Figure 3-15 Coverage histograms of the pilot fusion-panel (left) and optimised fusion-panel (right) fusion capture panel. The histograms illustrate the overall coverage distribution by displaying the number of ROI covered by mapped sequencing reads at different depths. Read depths (x-axis) and ROI covered at each depth on the y-axis. (IQR = Inter-Quartile-Range).

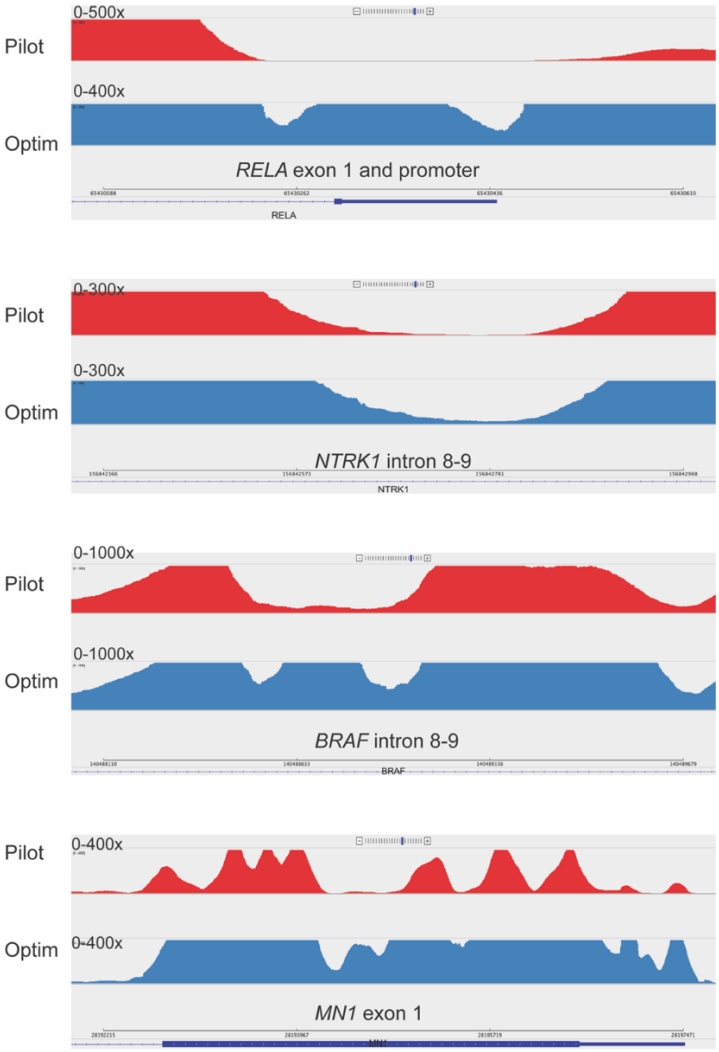


Figure 3-16 Sashimi coverage plots of pilot fusion-panel and optimised fusion-panel. The plots compare underperforming ROI for H2411 (FFPE) using pilot fusion-panel (red) against optimised fusion-panel version (blue) of *RELA* exon 1 and promoter, *NTRK1* intron 8-9, *BRAF* intron 8-9 and *MN1* exon 1.

3.2.4 Paediatric high-grade glioma panel validation

To validate this assay for research purpose, nine pHGG cases were used across two Miseq runs (run 1, HMW =5, run 2, FFPE=5). Overall the pHGG-panel had an excellent performance with 99.3% (5086/5125) of ROI achieving an average coverage of >100 in the 5 HMW and 4 FFPE cases.

There was no difference for the average total number of reads between FFPE 1.8x5x10^7 ±7.4x10^6) and HMW cases (1.2x10^7 ±1.3x10^6) (p=0.1069, unpaired t-test **{Figure 3-17A}**. There was no differences in the overall mean depth for FFPE (400±237) compared to HMW cases (332±27) (p=0.5370, unpaired t-test) **{Figure 3-17B}**. Duplicates were higher in FFPE (24.7%±7.3) compared to HMW cases (4.5%±0.2%) (p=0004 unpaired t-test) **{Figure 3-17C}**. The mean depth and number of reads was slightly higher for FFPE compared to HMW and this was because there was one more case included in the HMW run. The pHGG-panel presented a good coverage uniformity as shown in the histogram. As expected, FFPE samples presented a higher variation in coverage across the ROI compared to HMW (FFPE-IQR=277 vs HMW-IQR=191) **{Figure 3-18}**.

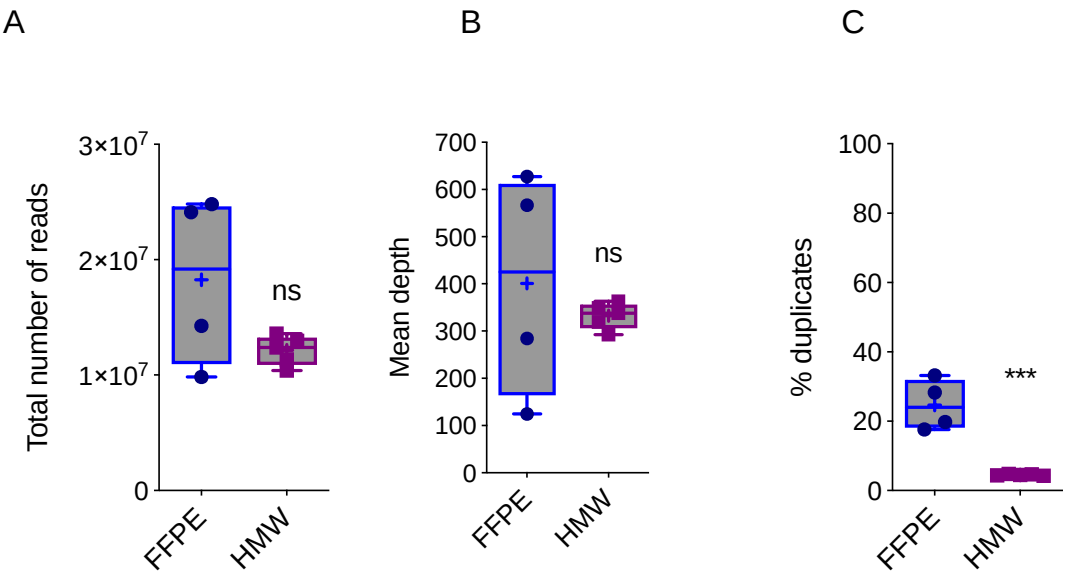


Figure 3-17 Box-plots showing quality metrics for pHGG-panel validation. Quality control metrics are separated by DNA integrity, formalin-fixed paraffin embedded (FFPE) and high molecular weight DNA (FF and cell lines). Metrics include (A) number of reads, (B) mean depth and (C) percentage of duplicates. The thick line within the box is the median, mean is shown as "+", the whiskers are drawn down to the 10th percentile and up to the 90th, points below and above the whiskers are drawn as individual points. All p-values are based upon unpaired t-test. ****p<0.00001, ns p>0.05.

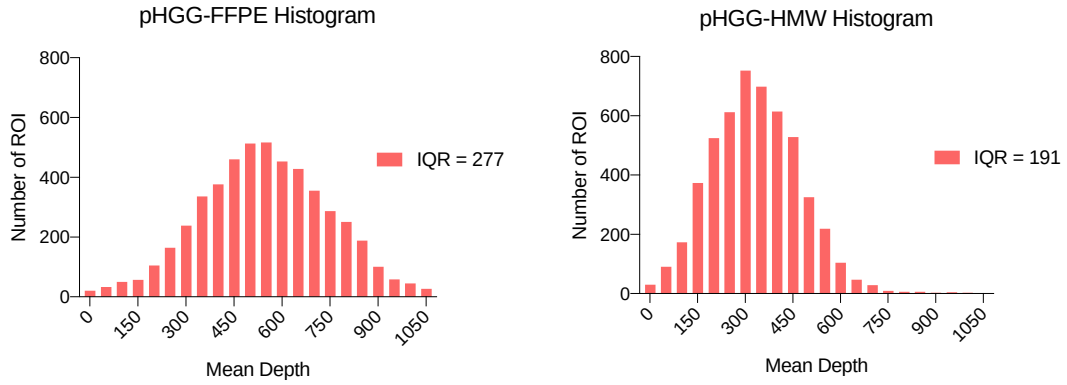


Figure 3-18 Coverage histograms of FFPE (left) and HMW (right) validation samples used for pHGG-panel. The histograms illustrate the overall coverage distribution by displaying the number of ROI covered by mapped sequencing reads at different depths. Read depths (x-axis) and ROI covered at each depth on the y-axis. (IQR = Inter-Quartile-Range).

Eight out of the nine cases had WES data and harboured 317 SNVs, 7 indels and 1 deletion overlapping molecular alterations between the pHGG and WES. The presence of the alterations was manually inspected on IGV and VAF was compared. CXJ0016 was a hypermutator case with a total of 269 overlapping variants with WES. All the SNPs and cancer variants were detected, including H3.3-K27M (n=2), *BRAF*-V600E (n=2), *TP53* (7), *CDKNA*-deletion (n=1) amongst others, obtaining an overall VAF correlation of r^2=0.7655 [95% CI:0.8789-0.9925] **{Figure 3-19}**.

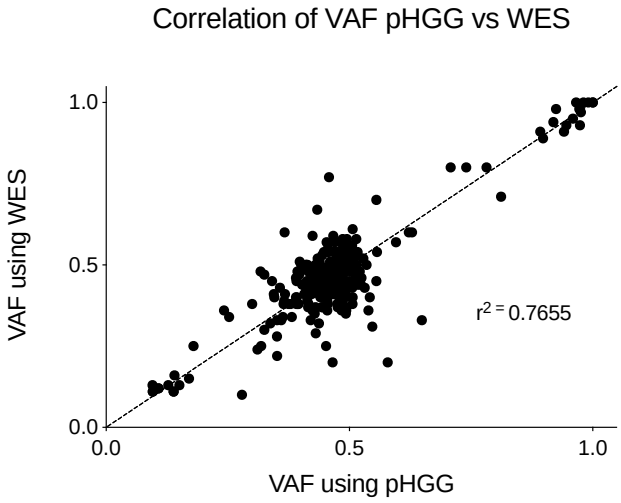


Figure 3-19 Correlation of variant allele frequencies (VAFs) between pHGG-panel (x axis) and whole exome sequencing (y axis) of cancer variants in FFPE and HMW patient samples (317 SNPs and 7 indels).

3.3 Discussion

The implementation of NGS assays in the clinic requires a robust validation in clinically-accredited laboratories. In this chapter, I provide evidence of the development and implementation of paediatric NGS capture-based assays for diagnostic test to accurately detect clinically relevant genomic alterations in FFPE as well as FF. The design and validation of Paeds-v1 assay has been part of a publication on which I am first author [354].

AF for known SNVs and indels were very similar in within-run and between-run replicates, demonstrating that the paediatric panel assays are repeatable and reproducible. SNVs were detected at a wide range of VAFs simulating the heterogeneity expected in cancer samples, including 33 SNVs with an expected VAF of 4-5%. The detection of variants at low VAF is crucial, especially in samples with a low neoplastic cell content. False-negative calls for v1 were mainly indels at low AF and occurring at splice sites (+4bp upstream of the exon). There is currently no consensus as to what the most appropriate minimum region of interest to extend the intronic sequencing for clinical reporting may be, and in many cases the biological meaning of these mutations is unknown. Nevertheless, the latest guidelines recommend to assume a disrupted gene function in certain types of variants (e.g., nonsense, frameshift, canonical ±1 or 2 splice sites, initiation codon, single exon or multiexon deletion) which would be covered with our current pipeline [356].

I also compared the performance of paired FFPE-FF specimens obtaining comparable quality metrics between both tissue types, as well as a high overall correlation when comparing AF of the variants within samples (r^2=0.985 95%CI:0.984-0.985; p<0.0001). The discrepancies of the variants may be explained by variation in neoplastic cell content between FF-FFPE and intra-tumour heterogeneity leading to sub-clonal alterations [102, 103, 357]. The good correlation is particularly important when a large proportion of clinical samples routinely available are derived from FFPE, where nucleic acid quality is generally compromised and chemically challenged, leading to DNA degradation and potential deamination or oxidation artefacts.

I verified the accuracy of the NGS capture-based approach in cell lines and clinical specimens (FFPE and FF) containing known genetic abnormalities previously characterised by other methodologies. We obtained a high concordance when we compared the expected alterations to those detected using our assay. Furthermore,

our data also shows that this method can detect structural variants, including amplifications, deletions and chromosomal rearrangements. Only one out of 5 SVs involving *EWSR1* was not identified by the assay which could be due to the lack of coverage at the intronic genomic location of the breakpoint. As expected, this is one of the limitations of the methodology, as capturing intronic regions poses challenges associated to the presence of repetitive sequence elements that lead to the chromosomal breaks. This can be partially overcome by including the breakpoint regions of the most common partner genes involved in the translocations.

In addition, I described optimisation for suitable detection of SV including improvements in the capture-based fusion-panel design as well as library preparation. Accurate detection of SV in paediatric brain tumours has become indispensable for clinical decision-making. For instance, *BRAF-KIAA1549* duplication correlates with diagnosis (low grade glioma) [358], *c11orf95-RELA* fusions are presented in a subset of supratentorial ependymoma, correlating with dismal prognosis [24] and *BCOR* duplications defining CNS high-grade neuroepithelial tumours [21]. Furthermore, the presence of structural variants involving genes such as *NTRK1-3*, *ALK* and *FGFR1/3* can help to guide treatment by the use of targeted agents such as larotrectinib (NCT02637687), entrectinib (NCT02650401) and erdafitinib (NCT03155620) amongst others. The fusion-panel would be able to detect these genetic abnormalities and guide clinicians to make an accurate diagnosis and potential use of targeted therapies when appropriate.

In summary, I have developed robust clinical assays that can detect SNVs, small indels, copy number variation and SV with high reproducibility and repeatability in routine clinical FFPE samples from a variety of centres. These data show that these capture DNA-based assays can be an accurate and practical platform for molecular stratification and identification of actionable targets required to accelerate personalised medicine clinical trials in childhood solid tumours. The paediatric v1 and v2 panels were incorporated into a pilot molecular profiling study for paediatric patients at the Royal Marsden Hospital (London, UK) and was further extended across the UK. The results of this study together with the applications of the fusion-panel and pHGG-panel are shown in the next chapters.

CHAPTER 4 : Implementation of panel targeted sequencing for paediatric solid tumours into clinical practise to guide targeted treatment

4.1 Introduction

In adult malignancies, precision medicine initiatives enabling standardised, high-throughput molecular profiling and predictive biomarker-based stratification have been implemented to maximise clinical efficacy of targeted therapeutics [359-362]. There is an urgent need to translate such opportunities to the treatment of childhood cancer, which remains the primary cause of death in children after infancy [2].

While there are now a number of world-wide strategies to support precision medicine in paediatric cancer, that was not the case at the start of this PhD. Specifically, in the UK the implementation of personalised medicine in childhood cancers had been limited by a lack of clinically validated multi-target sequencing approaches specific for paediatric solid tumours. This is particularly important to support clinical trials that match molecular alterations to targeted therapies, including umbrella and basket designs trials such as paediatric eSMART (NCT02813135) and INFORM in Europe, and NCI-MATCH (NCT03155620) in US [267, 270]. These trials are supported by comprehensive molecular profiling programmes which include whole-exome sequencing (WES) and RNA-seq and, in some cases copy number analysis, whole-genome sequencing (WGS), expression microarray or methylation arrays [254, 265, 267, 269]. However, logistical and financial practicalities limit large-scale implementation of this approach in most health-care settings. Targeted Next-Generation Sequencing (NGS) panels are typically more cost-effective, can be tailored to the study population, and standardised according to regulatory requirements. Therefore, this may present a more suitable alternative for implementation into health-care systems by achieving deeper sequencing increasing sensitivity at lower cost than WES and WGS.

One of the differences between adult and paediatric cancer is that paediatric cancers harbour fewer mutations, however the proportion of fusion genes is higher than their older counterparts, providing an excellent opportunity for diagnostic and predictive biomarkers to targeted agents [10, 11, 363]. Structural variants (SVs) are defined as large deletions and insertions, duplications, inversions and translocations of at least

50 bp. Array Comparative Genomic Hybridization (aGGH) and NGS-based assays are typically the methods of choice for SV identification which usually requires the use of high-molecular weight DNA and/or RNA [333]. Moreover, the detection of SV can be complex and needs specific pipeline algorithms to identify candidates from abnormally mapped reads including paired-end mapping, split read, read depth and the novo assembly [333, 364].

The identification of SVs is of particular importance in CNS tumours, in which recent studies described them to be diagnostic biomarkers of certain tumour subgroups correlating with a distinct prognosis and of potential response to targeted agents [24, 34, 35, 220, 365-367]. This is exemplified by the new brain tumour entities that have emerged from molecular classification of primitive neuroectodermal tumours of the central nervous system (CNS-PNETs) [368]. Furthermore, although rare there some studies describing the presence of SV genes in pHGG and DIPG including *NTRK1/2/3* (*e.g. TPM3:NTRK1*, *AGBL4:NTRK2,* and *ETV6:NTRK3*), and other receptor tyrosine kinase (RTK) genes (*e.g. ALK*, *MET, PDGFRA* and *ROS1)* [22, 75, 369-374]. Nevertheless, the identification of SVs might be underpowered due to the detection challenges mentioned above. In particular, these studies have shown that pHGG of younger age (under 3-5 age), wild-type for the typical driver genes (H3.3/1*, BRAF, IDH1/2*), seem to be driven by single fusion gene events. Interestingly, infant HGG appear to have better outcome compared to older children even in patients with incomplete resection or without radiation [22, 375, 376].

In order to address the lack of molecular profiling in the UK in children with cancer, I describe the implementation in the National Health Service (NHS) of the paediatric NGS panels developed in Chapter 3. To better integrate the panel sequencing results into clinical practise for children with cancer, the study offered panel sequencing in prospective or archived tumour tissue, with clinical reporting of results via a formal Molecular Tumour Board (MTB). Additionally, the use of the fusion-panel in two different cohorts of pHGG led to the discovery of novel internal tandem duplication of *NTRK2* gene in a non-brainstem pHGG clinical trial (HERBY), which was a phase-II trial evaluating bevacizumab in addition to radiotherapy and temozolomide after initial surgery [69, 87, 321, 335]. Moreover, we used the fusion-panel in the largest series of infant gliomas assembled to date, in order to assess the presence of structural variants and identify driving alterations [89].

4.2 Results

4.2.1 Patient samples and Overall Performance

A total of 255 cases were submitted from 223 patients enrolled in ten different institutions in the UK. DNA isolated from tissue and matched germline DNA was subjected to panel sequencing using the Paeds-v1/v2 capture panels described in Chapter 3. An overview of the study is given in **{Figure 4-1}**. Although patients were eligible to enrol at any time, 90% of evaluable patients had at least one episode of progression/relapse before study enrolment. FFPE tissue from the most recent surgery was requested for all but three patients, in which fresh frozen tissue was used.

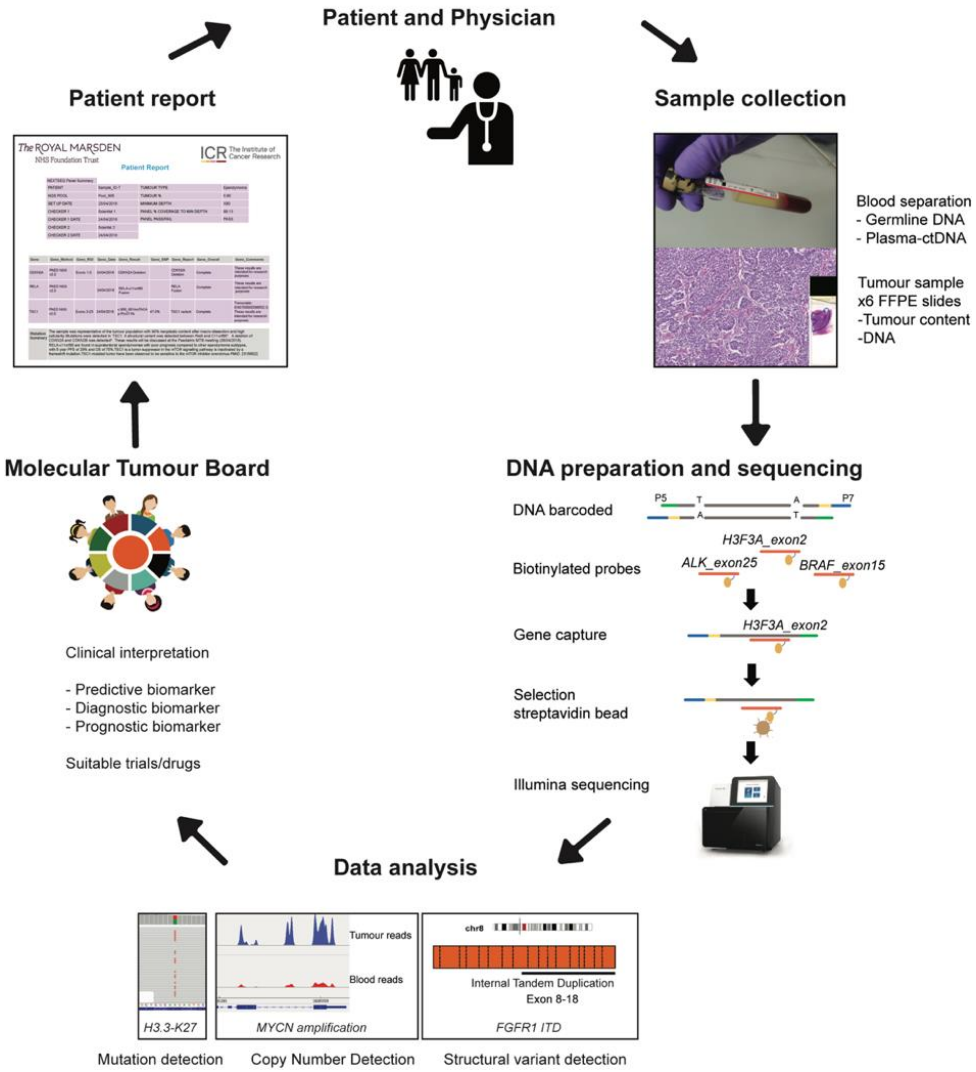


Figure 4-1 NGS study overview. Following informed consent, tumour and blood samples were collected. DNA was extracted and sequence libraries were prepared using the capture-based paediatric solid tumour panel. Following sequencing, samples underwent an in-house data analysis pipeline that detects mutations, structural variants and copy number changes. Genomic alterations were manually reviewed by two independent scientists and then discussed in a Molecular Tumour Board before a clinical report was issued.

Adequate coverage for clinical reporting of results was obtained in 82% of submitted samples **{Figure 4-2A}**. Reasons for sample rejection or failure were as follows: tumour content less than 10%, DNA less than 20 ng and/or excessive DNA fragmentation. The median depth of coverage for all reported cases was 495 (interquartile range: 264-868). The most common cancers sequenced were glioma (38), neuroblastoma (27) and rhabdomyosarcoma (26) **{Figure 4-2B}**.

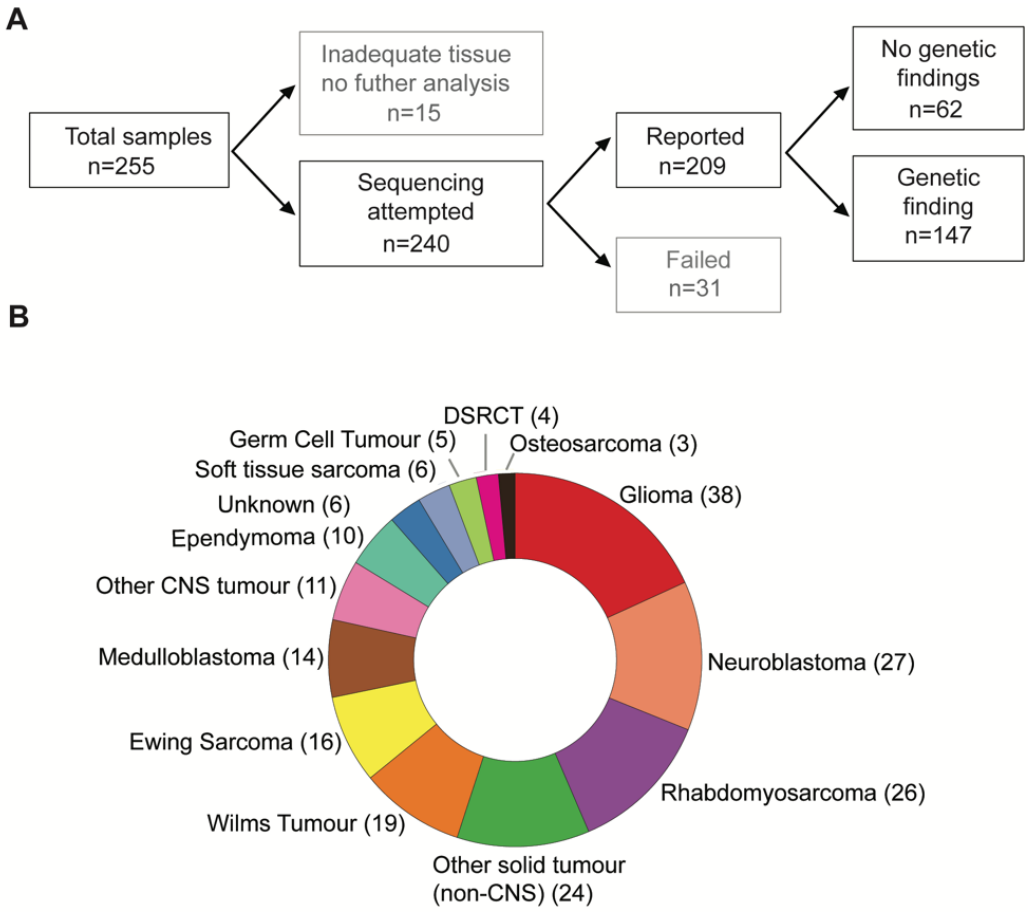


Figure 4-2 Tumour samples submitted for sequencing. (A) Summary of sample flow and total number of samples successfully sequenced. (B) Distribution of tumour types among reported cases.

4.2.2 Genetic Findings

A monthly Molecular Tumour Board (MTB) was established for discussion of findings from the CCR-4294 pilot study; interpreted results from panel sequencing were then reported to the treating clinician. The MTB core members included paediatric/adolescent oncologists, experts in early clinical trials, molecular pathologists, bioinformaticians and paediatric tumour biologists, from RM, Great Ormond Street Hospital and The Institute of Cancer Research, London. OncoKB was used as a basis to define tiers of actionability [377]. In addition, COSMIC defined mutations/SNV, genetic amplifications, gains or losses, for which a paediatric clinical trial was currently recruiting, were also considered, as well as alterations where compelling pre-clinical paediatric data existed for that target [378]. Heterozygous gene loss and missense mutations outside of defined hotspot regions were defined as not actionable.

At least one genetic alteration was detected in 70% (145/209) samples at an allele frequency ≥ 5%. The somatic genetic alterations detected, grouped according to underlying diagnosis, are summarised in **{Figure 4-3}** and **{Figure 4-4}**. In keeping with other studies [363], the most frequently mutated gene was *TP53* in 36/209 (17%) cases; in addition high frequencies of alterations in genes known to be recurrently altered in paediatric malignancies such as *ATRX*, *CDKN2A*, *CTNNB1* in 12/209 (5.7%), *MYCN* in 11/209 (5.2%) and *H3F3A, PIK3CA* in 10/209 (4.3%) were detected.

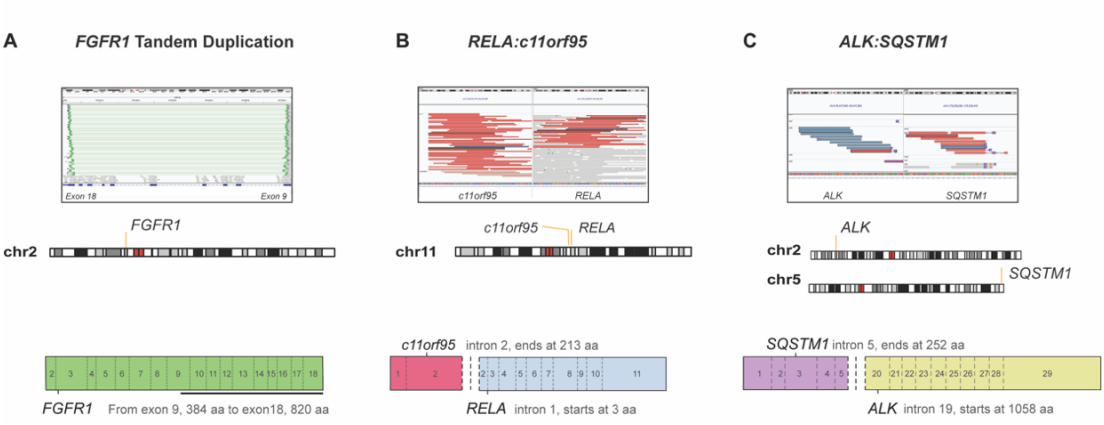


Figure 4-3. Illustration of structural variants detected by the Paeds-panels. Snapshot of IGV showing spanning reads covering the structural variant region are shown on the top and a cartoon illustration of the structural variant is displayed in the bottom. (A) *FGFR1* tandem duplication from exon 9 to exon 18 detected in a patient with glioma. (B) Fusion between exon 2 of *c11orf95* and exon 2 of *RELA* detected in a patient with ependymoma. (C) Fusion between exon 5 of *SQSTM1* and exon 20 of *ALK* detected in a patient with an inflammatory myofibroblastic tumour.

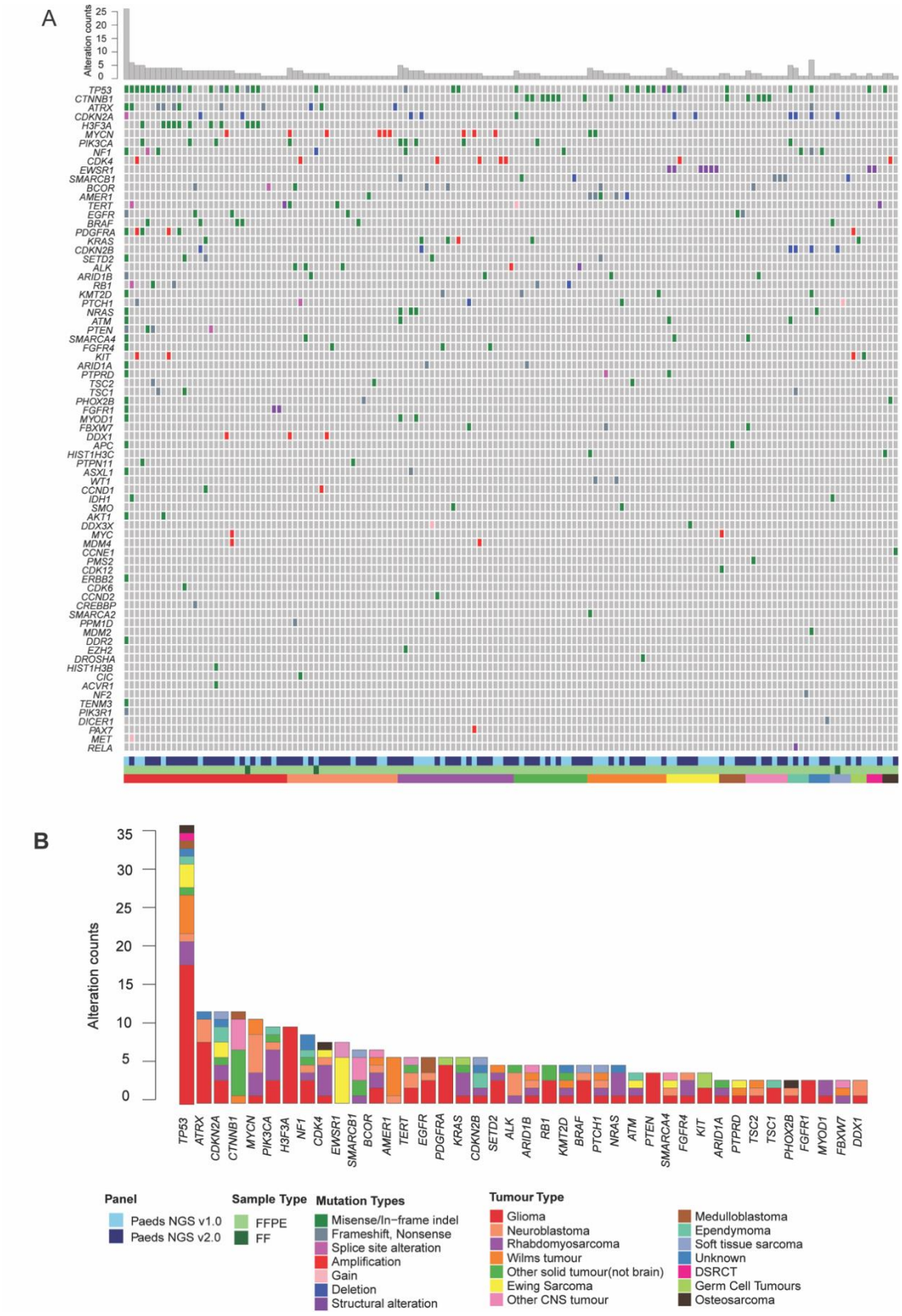


Figure 4-4 Overview of Paeds-panels clinical sequencing results. (A) Oncoprint represents somatic mutations and DNA copy number findings. Samples are grouped in columns with genes displayed along rows. Samples are arranged according to the tumour type and genes sorted by frequency. Panel version, sample type, molecular annotations and diagnosis are provided as bars according to the included key. (B) Bar-plot of most recurrent altered genes, sorted by frequency and colour coded according to the tumour type.

4.2.3 Clinical Actionability

Potentially targetable alterations, defined by OncoKB tiers of actionability in addition to predictive biomarkers for currently recruiting paediatric clinical trials, were detected in 51% of sequenced samples **{Figure 4-5A}**. Of the 107 tumour samples classified as potentially actionable, 42 (39%) had greater than one such alteration detected. For each tumour sample, only the alteration for which there was the highest tier of evidence for actionability was included. Glioma was the tumour type with more defined actionable alterations found, followed by osteosarcoma and rhabomyosarcoma **{Figure 4-5B}**. No tier 1 alterations (FDA recognised biomarker predictive of response to an FDA approved drug) were detected, indicative of the lack of regulatory approvals for paediatric indications. Only one patient had a tier 2A alteration: a patient with an inflammatory myofibroblastic tumour, harbouring an *ALK:SQSTM1* translocation. The patient had a complete surgical resection and did not require systemic therapy.

As a feasibility study, follow-up data was not routinely collected for all patients. Of the 57 patients with a tier 2B or 3 alteration and available follow-up data, only four (7%) received targeted therapies: Three patients with *BRAF*-V600E mutations were treated with dabrafenib/trametinib combination therapy: patient 1 had a pleomorphic xanthoastrocytoma and was commenced on dabrafenib/trametinib following third disease progression. The patient remains on treatment with stable disease after 9 months. Patient 2 had glioblastoma multiforme, and was commenced on dabrafenib/trametinib after disease progression. The patient had stable disease for 13 months before further progression. Patient 3 had multiply-relapsed metastatic ameloblastic fibro-odontosarcoma [379]; by day 28 of treatment, there had been a partial response but asymptomatic cardiac toxicity required discontinuation of both drugs. Upon normalisation of the shortening and ejection fractions, the patient was recommenced single-agent dabrafenib and had sustained partial response for 15 further months **{Figure 4-6}.** A patient with multiply-relapsed metastatic germinoma and *PDGFRA*/*KIT* amplification was given dasatinib, who progressed on treatment.

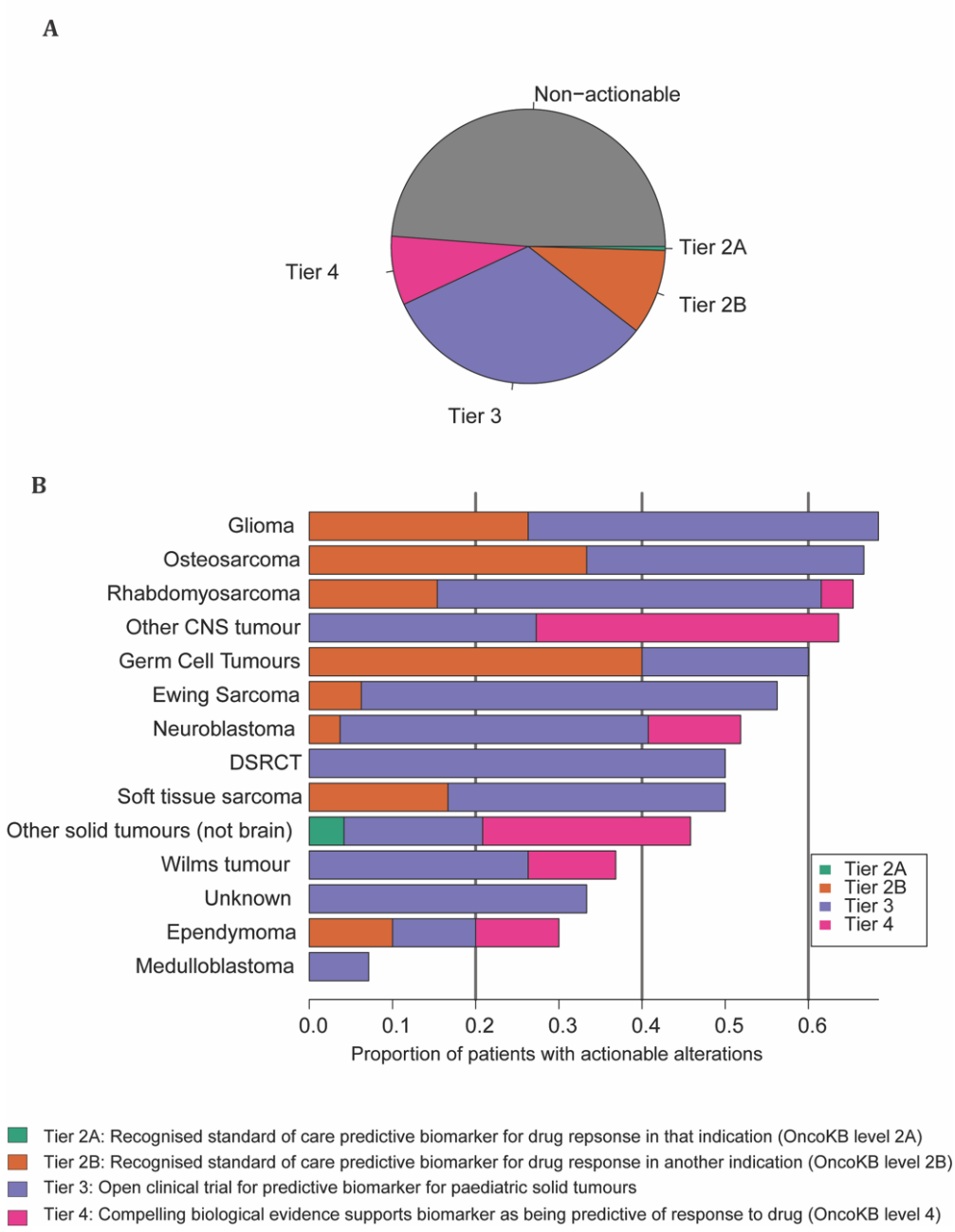


Figure 4-5 Clinical actionability of Paeds-panels clinical sequencing results. (A) Somatic alterations were defined according to OncoKB levels of evidence. Actionability tiers are described in the key. Distribution of actionability tiers for entire sequenced cohort. (B) Distribution of actionability tiers across common tumours, colour coded according to tumour type.

One patient with high grade glioma (patient ID 045-T) had a total of 49 somatic mutations (in ~0.18 Mb) consistent with a hypermutator phenotype [380], associated with mismatch repair deficiency and predictive of potential sensitivity to immune checkpoint blockade [138]. However, the patient was not fit for clinical trial enrolment by the time sequencing results were available.

Other patients had findings that informed prognosis: a mutation in *CTNNB1* was found in a patient originally diagnosed with a supra-tentorial primitive neuroectodermal tumour (PNET), biologically more in keeping with a WNT-activated medulloblastoma. Other examples included a *MYOD1* mutation in a patient with embryonal rhabdomyosarcoma tumour, associated with distinct clinical features and poor prognosis [216], and a *RELA:c11orf95* fusion in a supratentorial ependymoma, associated with high-risk disease [24].

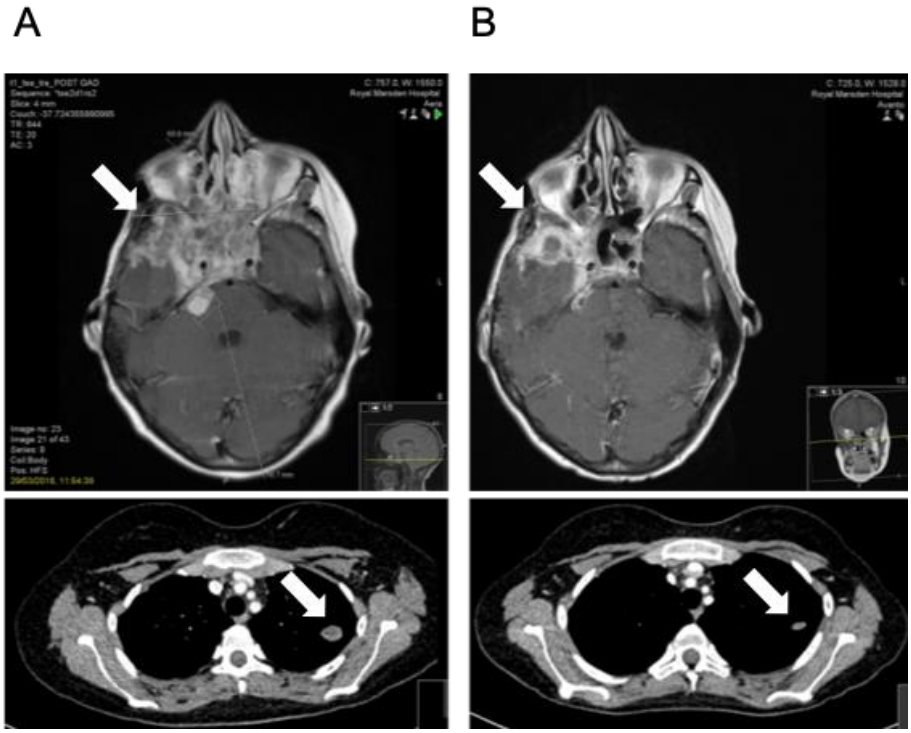


Figure 4-6 Magnetic resonance imaging (MRI) from patient with metastatic ameloblastic fibro-odontosarcoma. MRI showing of the tumour (A) before and (B) after 8 weeks of BRAF inhibitor treatment. A reduction of primary tumour is seen in the skull base (top) and the pulmonary metastases (bottom). Tumour areas are highlighted by white arrows. Image kindly provided by Dr Sally George.

4.2.4 Analysis of paired samples

For eight patients, paired samples were sequenced at different stages of treatment **{Figure 4-7}**. In six of these, there were differences between the variants detected at different time-points. Mutations in *PTEN, NF1* and *TP53* were observed in a patient with high grade glioma (patient 2) after dabrafenib/trametinib treatment but not in the pre-treatment sample. The patient subsequently received everolimus but progressed after 3 months on treatment. The acquisition of *NF1* mutations as a resistance mechanism after BRAF inhibition is consistent with findings in *BRAF*-V600E mutant melanoma [381, 382]. Another child with glioma harboured shared alterations in

H3F3A and *TP53* both at diagnosis and progression, whereas *PTEN* was only present at diagnosis and *PIK3CA* at progression. In a patient with Wilms' tumour, a potentially targetable *TSC2* mutation was found in the 3rd relapse sample, which was not present in the previous sample.

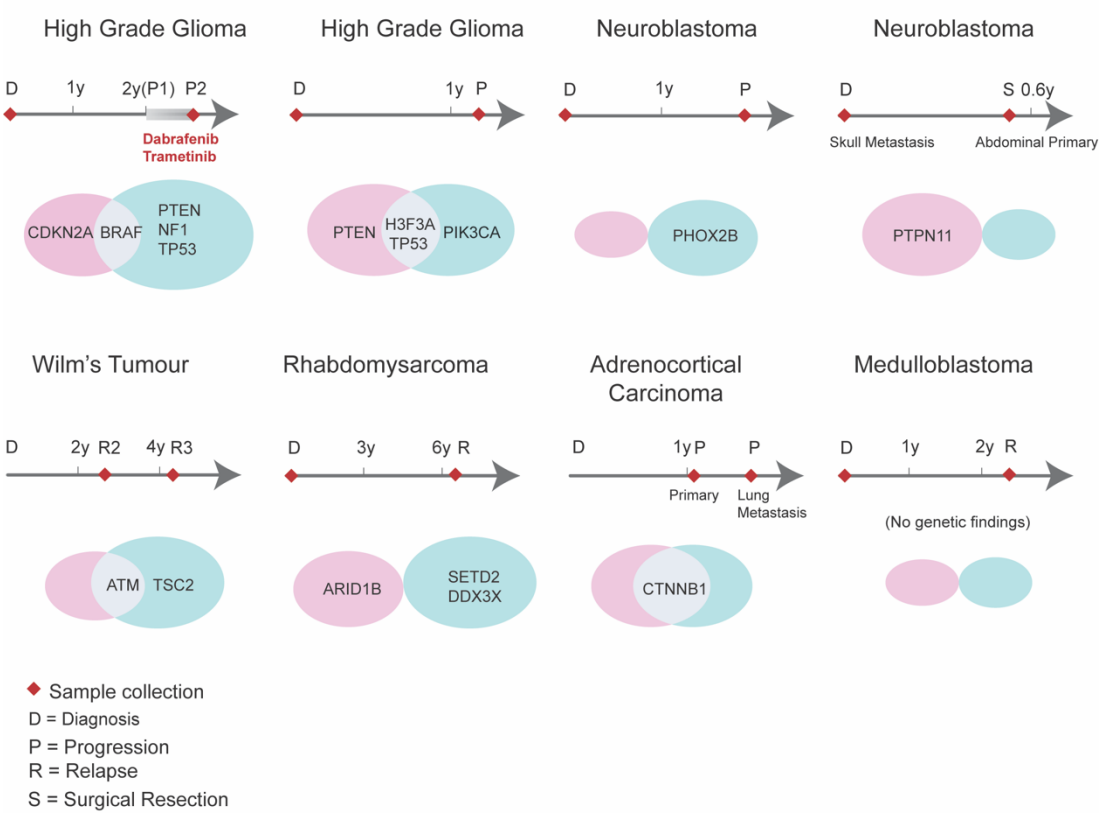


Figure 4-7 Comparison of results from paired samples, sequenced at different time points of Paeds-panels clinical sequencing samples. Venn diagrams compare the genetic findings in eight patients. Shared alterations are illustrated by the intersection of the two ovals. Alterations detected at only the 1st time point are represented in the pink oval, and alterations identified at the 2nd time point only are represented in the blue oval. The size of the oval represents the number of variants identified in each patient.

4.3 Fusion-panel application

4.3.1 Sample cohort and overall performance

We next applied the fusion-panel assay to detect structural variants in the context of a clinical trial in non-brainstem pHGG (HERBY). This cohort included a total of 68 tumours (33 FFPE and 35 FF) and 24 blood samples.

Quality control metrics were performed to assess panel performance. The average number of reads was slightly lower for FFPE ($1.7x10^7 \pm 9.4x10^6$) compared for FF

cases ($2.2x10^7$±$9.1x10^6$) (p=0.0187, unpaired t-test) **{Figure 4-8A}**. The overall mean depth was lower for FFPE (642±406) compared for FF cases (995±282) (p=0.0001, unpaired t-test) **{Figure 4-8B}**. Duplicates were higher in FFPE (44%±20%) compared to FF cases (14%±2%) (p<0.0001, unpaired t-test) **{Figure 4-8C}.** There was no difference for the percentage of bases from unique reads on target between FFPE (46%±13.3%) and FF (48.3%±13.4%) (p=0.5123, unpaired t-test) **{Figure 4-8E}**. Only two samples (FFPE) had an average mean coverage below 100 (88 and 90). The average of targeted positions covered at 100x depth was lower for FFPE (85%±19%) compared to FF cases (95.1%±1.9%) (p=0.0039, unpaired t-test) **{Figure 4-8F}**. As expected, library-prep fragment length was lower for FFPE (346bp±47bp) than for FF cases (413bp±35bp) (p=0.001, unpaired t-test **{Figure 4-8G}**.

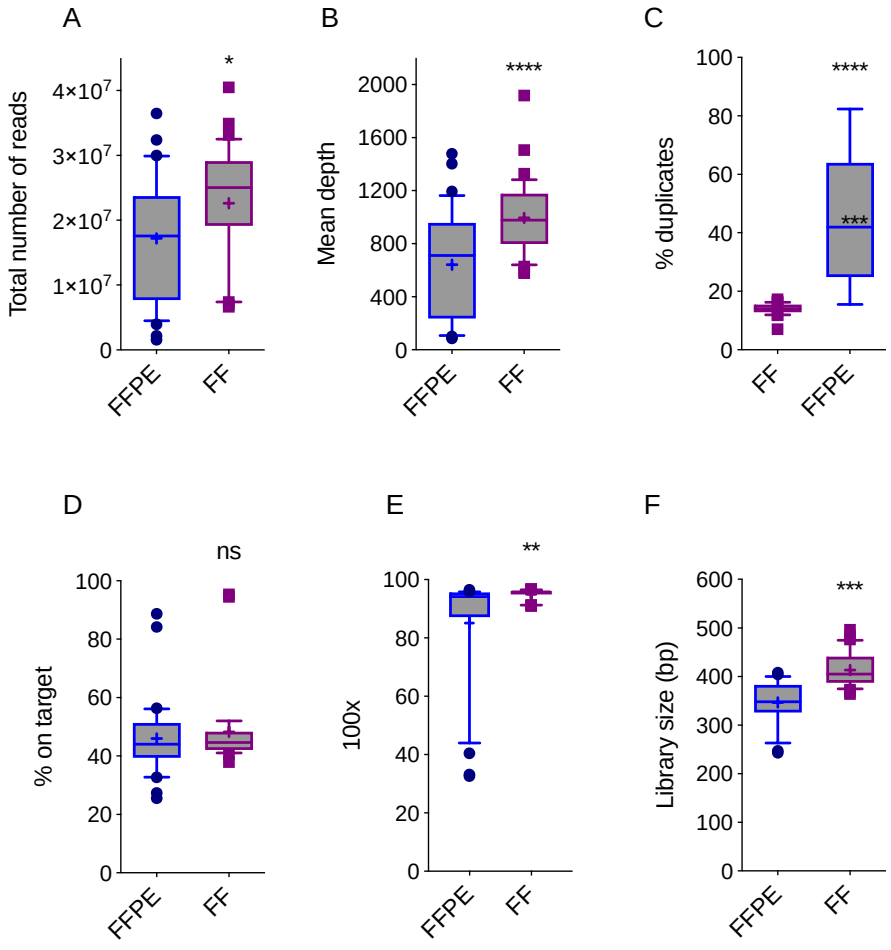


Figure 4-8 Box-plots showing quality metrics of Paeds-panel clinical sequencing results samples. Quality metrics are separated by DNA integrity, formalin-fixed paraffin embedded (FFPE) and fresh frozen (FF) HERBY samples. Metrics include (A) number of reads, (B) mean depth, (C) percentage of duplicates, (D) percentage of unique reads on target, (E) percentage of targeted positions covered at 100X and (F) library fragment size in base-pairs. The thick line within the box is the median, mean is shown as "+", the whiskers are drawn down to the 10th percentile and up to the 90th, points below and above the whiskers are drawn as individual points. All p-values are based upon unpaired t-test ****p<0.00001, ***p<0.0001, **p<0.001, *p<0.05, ns p>0.05.

4.3.2 Applying the fusion-panel for fusion identification in HERBY

We identified SVs in 12 samples **{Table 4-1}** and **{Figure 4-9}**, eight of which presented novel structural variants involving *NTRK2*. In particular, 10.3% (7/12) harboured novel ITD-*NTRK2* and HERBY_106 presented *SLC4A4:NTRK2* fusion between exon 3 of *SLC4A4* (chr4q13.3) to exon 11 of *NTRK2* (chr9q21.33). We confirmed the ITD-*NTRK2* was a somatic event as no reads were present in any of the germline DNA samples. The size of the ITD varied from sample to sample (~28-356 kb) involving exons 11 to 3'-UTR **{Table 4-2}** and **{Figure 4-10}**. Six out of seven ITD-*NTRK2* cases were H3.3 K27M and presented mutations in *TP53* as well as *ATRX* truncating mutations in 5/7 cases. On the basis of these findings, David Jones (DKFZ/KiTZ, Heidelberg) provided data from a cohort of 59 pHGG samples in which RNAseq and/or WGS was available. Two cases (3.4%) were found to harbour ITD-*NTRK2*, which represented 2/14 (14.3%) midline H3.3 K27M samples. These findings support our discovery of novel ITD-*NTRK2* with a potential enrichment in K27M positive midline pHGG.

We identified two cases with novel structural variants involving *MN1*. HERBY_066 presented a fusion between exon 1 of *MN1* (22q12.1) to exon 2 of *ARHGEF6* (Xq26.3) and HERBY_052 harboured a fusion between exon 1 of *MN1* (22q12.1) to exon 3 of *CARD6* (5p13.1) **{Figure 4-11A}**. HERBY_052 was a compact and necrotic tumour with perivascular radiating arrangements **{Figure 4-11B}**, classified by the methylation-based classification of central nervous system tumours [20], as high-grade neuroepithelial tumour with *MN1* alteration (CNS HGNET-MN1, score = 0.713) using the Illumina 450K methylation array **{Figure 4-11C}**. An ITD-*FGFR1* of exons 8 to 18 was identified in HERBY_049 **{Figure 4-11D}**, this finding is in line with the methylation classification result of pilocytic astrocytoma and supported by histology as anaplastic features were observed. In another sample. Furthermore, we identified a duplication between exon 11 of *PTPRZ1* (chr7q31.32) to exon 4 of *MET* (chr7q31.2) **{Figure 4-11E}** which also harboured *MET* amplification and classified as paediatric glioblastoma with receptor tyrosine-kinase activation (GBM_pedRTK).

Sample ID	H3.3 status	Location	Fusion
HERBY_013	WT	Hemispheric	*ITD-NTRK2*
HERBY_015	K27M	Midline	*ITD-NTRK2*
HERBY_037	K27M	Midline	*ITD-NTRK2*
HERBY_082	K27M	Midline	*ITD-NTRK2*
HERBY_084	K27M	Midline	*ITD-NTRK2*
HERBY_097	K27M	Midline	*ITD-NTRK2*
HERBY_118	K27M	Midline	*ITD-NTRK2*
HERBY_106	WT	Hemispheric	*SLC4A4:NTRK2*
HERBY_049	WT	Hemispheric	*ITD-FGFR1*
HERBY_052	WT	Hemispheric	*MN1:CARD6*
HERBY_066	WT	Hemispheric	*ARHGEF6:MN1*
HERBY_094	WT	Hemispheric	*PTPRZ1:MET*

Table 4-1 List of patient samples from HERBY positive by the fusion-panel.

Sample ID	Duplicated region	Coding bp size	Total genome bp size
HERBY_013	exon 15 to 17	368	146056
HERBY_015	exon 12 to 14	237	27826
HERBY_037	exon 13 to 16	438	153746
HERBY_082	exon 12 to 16	474	135829
HERBY_084	exon 13 to 21	6954	365628
HERBY_097	exon 12 to 16	474	195552
HERBY_118	exon 11 to 16	780	182962
GBM_012	exon 12 to 21	1013	228798
GBM_027	exon 12 to 16	474	169734

Table 4-2 Summary of *NTRK2* tandem duplication positive cases, including HERBY and two from Heidelberg cohort. Table includes duplicated exonic region and size as bp for the coding and total genomic region involving the alteration.

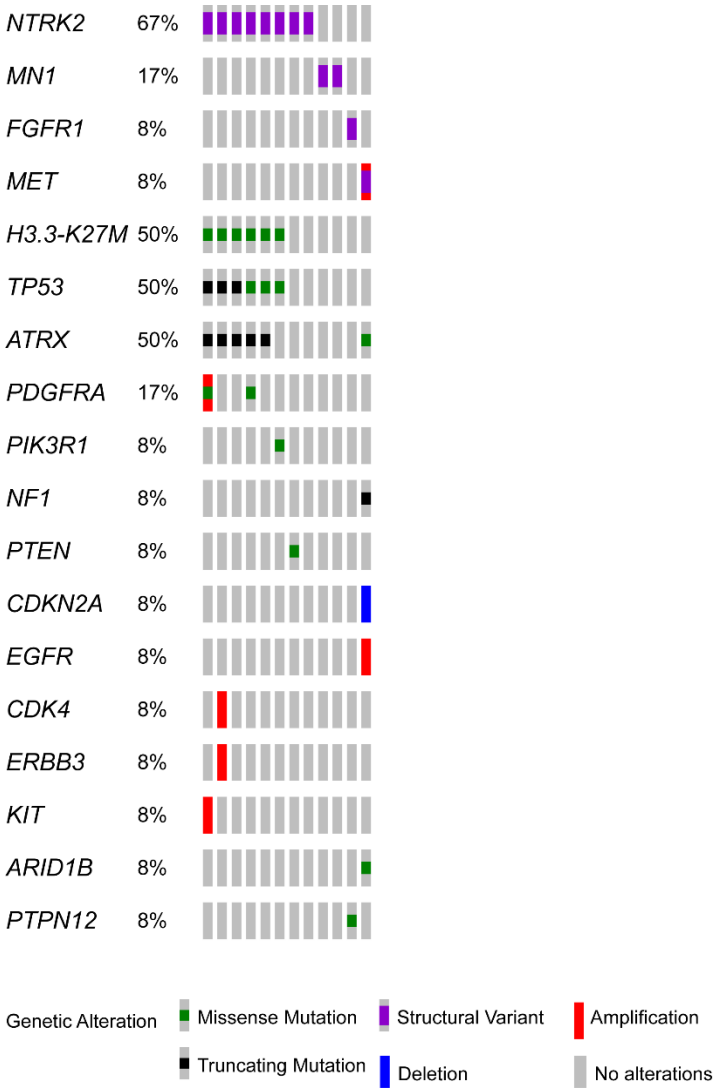


Figure 4-9 Oncoprint representation of an integrated annotation of WES and the fusion-panel. Single nucleotide variants, truncating mutations, copy number alterations and structural variants are shown in the 12 HERBY cases with structural variants.

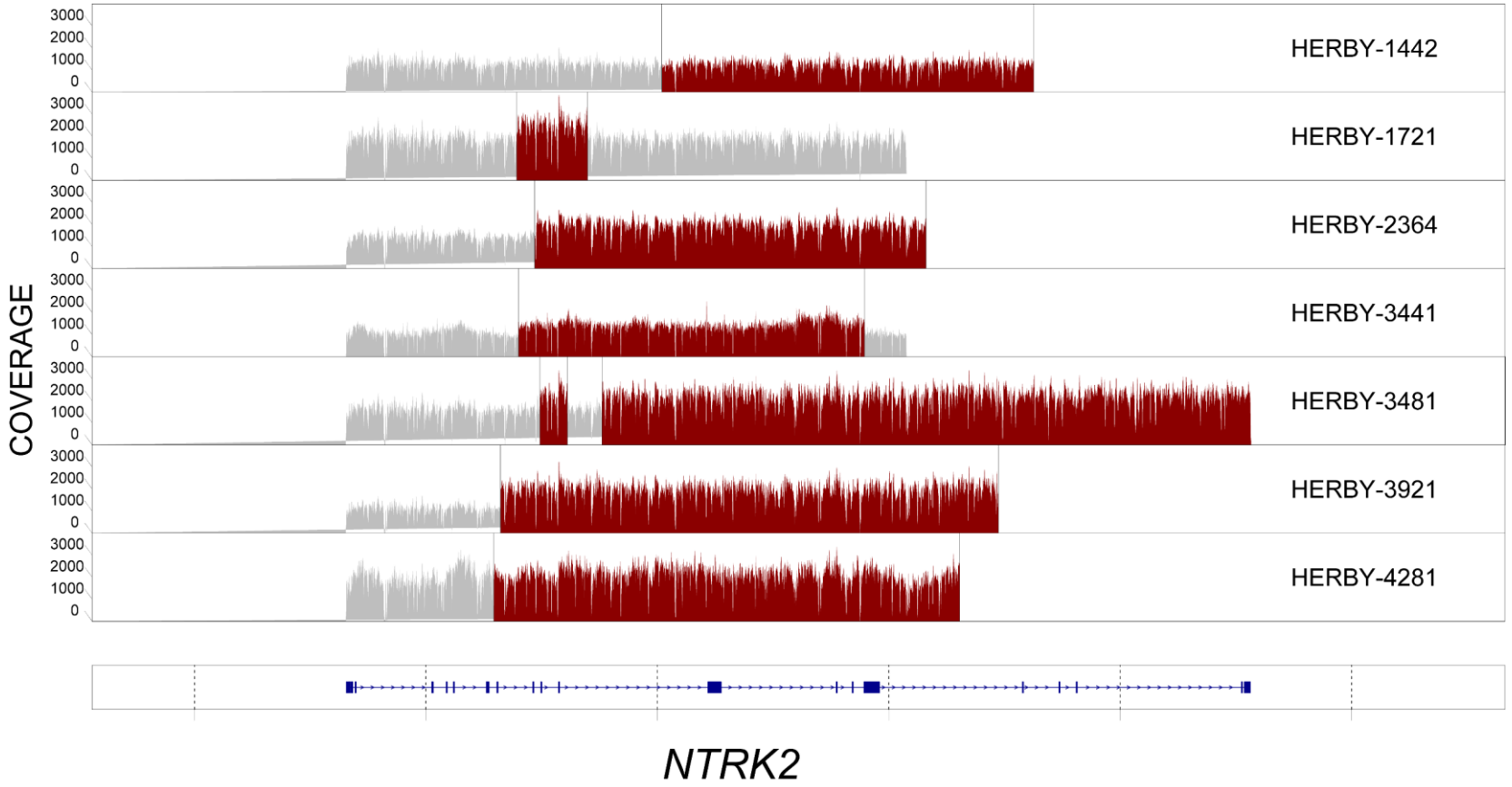


Figure 4-10 Coverage plot for *NTRK2* positive cases. The plot includes the seven HERBY cases harbouring internal tandem duplication. Every patient is divided by line. x-axis show *NTRK2* genomic location where exons are the blue squares and the lines that joined them represent the intronic region, y-axis represent the coverage, number of reads sequenced at a given genomic location. Duplicated region is highlighted in dark red.

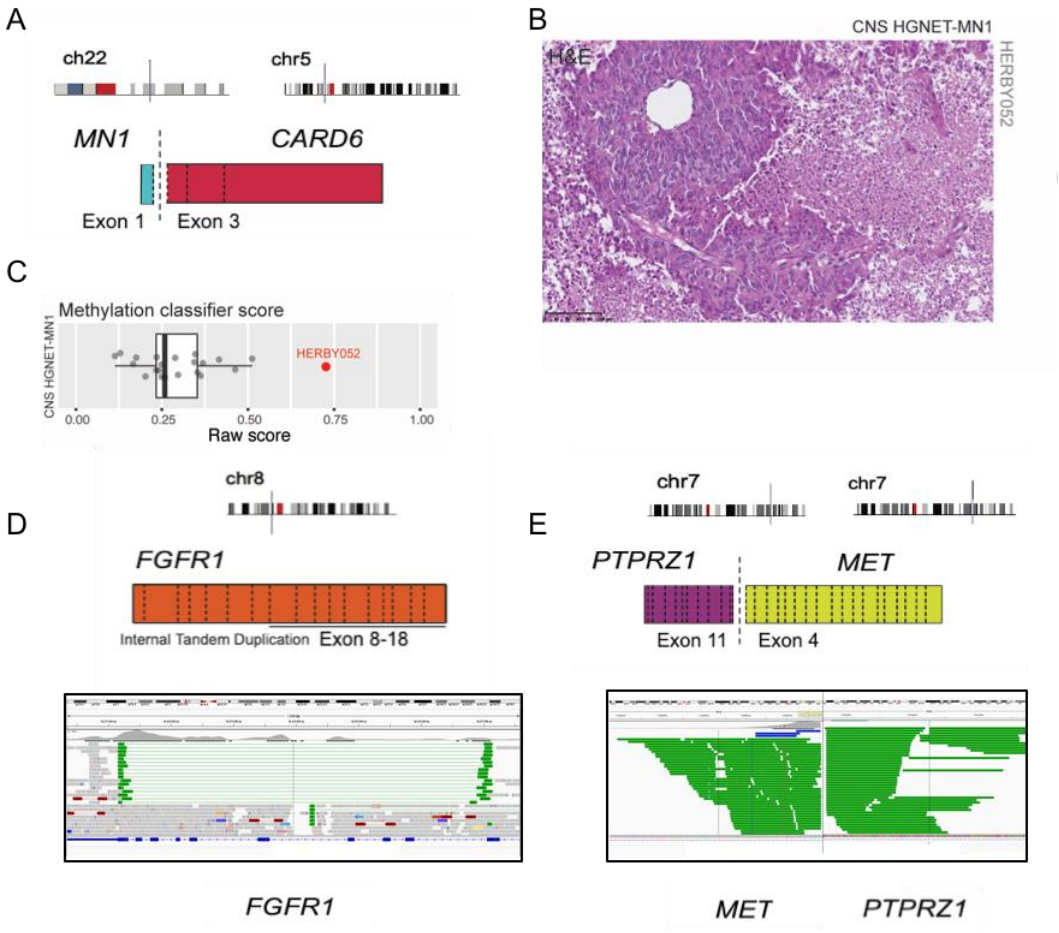


Figure 4-11 HERBY fusion-positive cases examples. (A) Cartoon of a novel *MN1:CARD6* gene fusion by capture panel sequencing between exon 1 of *MN1* and 3 of *CARD6*. (B) Haematoxylin and eosin staining of the case most closely resembling CNS HGNET-MN1. (C) Box-plot of reference methylation classifier score for HGNET-MN1 case. (D) Cartoon of the ITD-*FGFR1* on the top and below an integrative genomic viewer (IGV) snapshot showing the ITD supporting reads in green. (E) Cartoon of the *PTPRZ1:MET* fusion on the top and below an integrative genomic viewer (IGV) snapshot showing the fusion supporting reads in green.

4.3.3 Applying the fusion-panel for fusion identification in infant HGG

As part of the PhD project of Matthew Clarke in our laboratory, focused on the genomic landscape of infant HGG, we have used the fusion-panel together with other methodology such as methylation, RNA-seq and WGS to identify SVs in this particular cohort of patients [89]. Enough material to perform the fusion-panel was available in 114/241 of the infant glioma cases included in this study. Out of the 114 cases (n=9, cerebral hemispheres n=37, midline n=21 and unknown location n=47), 45% (51/114) harboured SVs identified by the fusion-panel. Of these, tumours with genetic alterations consistent with distinct CNS entities, were excluded including *KIAA1549:BRAF* (n=22) strongly associated with pilocityc astrocytoma [358], ITD-

FGFR1 (n=4) as a highly recurrent alteration in dysembryoplastic neuroepithelial tumour (DNET) [37], *MYB/MYL1* (n=2) commonly found in two distinct tumour entities angiocentric gliomas and isomorphic diffuse gliomas [38-40], and *MN1* (n=1) which defines the novel entity of HGNET-MN1 [21, 355]. Furthermore, cases were excluded based on clear Heidelberg classifier matches to other non-glioma CNS or existing defined glioma subgroup tumours from methylation array profiling data, as well as, expected genetic alteration defining glioma subgroups by panel sequencing or WES (*IDH1*-R132H, H3 K27M). After sample exclusions and integration of the fusion-panel with WGS and RNA-seq, 25/41 cases (61%) harboured fusions in either *ALK* (n=10), *NTRK1/2/3 (n*=1, 5 and 8, respectively), *ROS1* (n=2) or *MET* (n=1), usually in the absence of other alterations.

By using methylation array most of the fusion-positive cases were classified as infant hemispheric glioma (IHG, n=21) or low scoring desmoplastic infantile ganglioglioma astrocytoma (DIGG/DIAs, n=4) which clustered apart from other glioma subgroups in a tSNE projection. When possible, fusions were confirmed by other methods such as whole genome sequencing, RNA and/or Sanger sequencing and were frequently accompanied by detectable focal DNA copy number breakpoints within the fusion partners **{Figure 4-12A-D}**. There was a trend towards the presence of any fusion conferring a longer overall survival compared to those without (p=0.0687, log-rank test) **{Figure 4-12E}**.

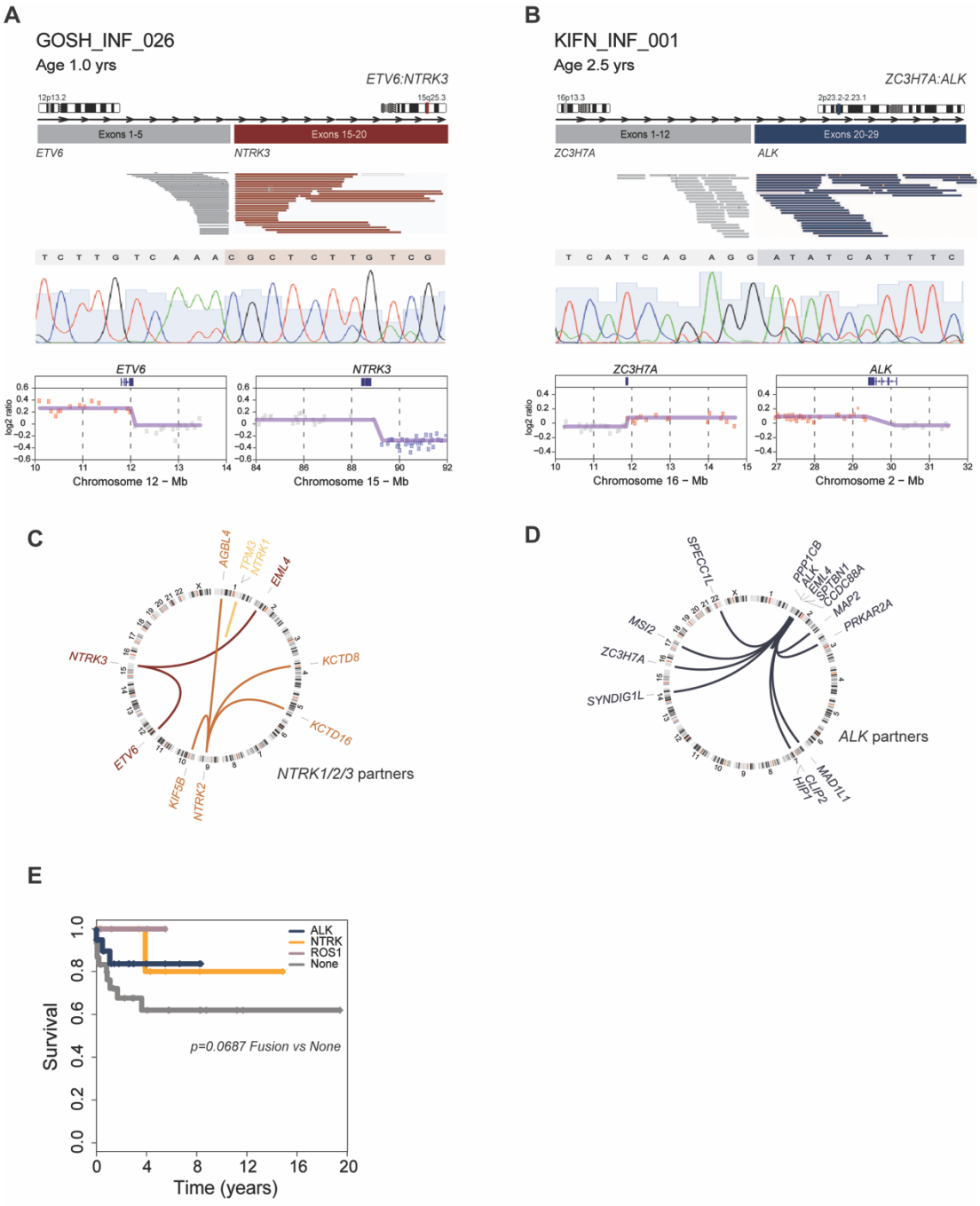


Figure 4-12 Infant HGG fusion-posiitve CNV overview and overall survival comparison. (A) Cartoon representation of the fusion *ETV6:NTRK3* and (B) *ZC3H7:ALK,* with reads on either side of the breakpoint coloured by gene partner taken form the integrated genome viwer (IGV) on the top. Sanger sequencing trace spanning the breakpoint is shown in the middle and on the bottom are copy number plots (log2 ratio y-axis) for chormosomal regions spanning the breakpoints (x-axis). Points representing copy number gain are coloured in red, blue for loss and grey for no change. The smoothed values are overlaid by the purple line. (C) Circos plot of gene fusions targeting *NTRK1* (light orange), *NTRK2* (organge) and *NTRK3* (dark orange). Lines linking fusion genes parters are represented by ideograms arranged around the circle. (D) Circos plot of gene fusions targeting *ALK* (dark blue). Lines linking fusion genes parters are represented by ideograms arranged around the circle. (E) Kaplan-Meir plot of overall survival of cases separated by fusion event (n=63) p-value was calculated by the log-rank test (p=0.0687 for any fusion *versus* none).

4.4 Discussion

Comprehensive molecular profiling strategies have been shown to be feasible in children with cancer and have revealed encouraging results [254, 267, 269]. However, wide-scale implementation is impractical in most health-care settings and even if resources were unlimited, it is also restricted by the availability of biopsy material. In this chapter, I demonstrate the clinical application of capture-based panel sequencing assays developed and described in Chapter 3. To do this, I have used three unique sample cohorts: an RMH pilot-study CCR4294, a multicentre clinical trial HERBY and a retrospective set of infant hemispheric gliomas. From these studies there have been three publications where I am co-author [87, 89, 383].

Despite the high detection rate of potentially actionable alterations, few patients received treatment with targeted agents. The reasons for this were multifactorial and include the following: lack of available clinical trials, difficulties accessing novel drugs on a compassionate-use basis, and/or clinical deterioration of the patient. In addition, although many patients had relapsed/refractory disease, a considerable proportion of patients were still on either first-line therapy or proven standard relapse therapies at the time of sequencing. A number of patients were also enrolled in available phase I/II trials that did not require biomarker screening.

With the sequencing of paired tumour samples at different times, I show the importance of tumour heterogeneity and evolution, adding to the mounting literature in support of the clinical importance of biopsy at relapse for children with cancer [63, 102, 384]. Notably, many tumour mutations emerging at the time of relapse (*PTEN, NF1, PIK3CA* and *TSC2)* are recognised predictive biomarkers of a targeted therapeutic response [385-389]. Cells derived from a *BRAF*-V600E HGG patient with *PTEN, NF1* and *TP53* mutations at relapse after dabrafenib/trametinib treatment, were shared as an international collaboration, with Theodore Nicolaides, a principal investigator whose one of his research projects is studying resistance to BRAF inhibitor in *BRAF*-V600E mutant glioma [390].

Although capture-based panel sequencing is an excellent tool, it has limitations and more comprehensive approaches need to be taken in consideration for the design of new clinical trials. Capture NGS panels are able to detect translocations in DNA with the ability to determine the single-nucleotide breakpoint, so long as those breakpoints occur in or close to a targeted region. We used MANTA to detect spanning pair reads

and split reads, thereby identifying fusion gene partners. However, detection of fusion genes is inevitably restricted. A more extensive method for detection of SVs is required such as the fusion-panel adjusting the region of interest capture to the tumour(s) type evaluated. If RNA is available, a promising strategy is the use of anchored multiplex PCR-based enrichment to detect fusions, eliminating the need to sequence long and complex intronic regions. In addition, methylation profiling is particularly relevant for a precise diagnostic classification of CNS tumours [391]. The paediatric panel is already routinely carried out for all paediatric solid patients in the NHS and supported by the National Institute for Health and Care Excellence (NICE, nice.org.uk/guidance/mib133). Finally, the paediatric panel will fit into a more modern molecular diagnostics service, Stratified Medicine Paediatrics (SMPaeds), where a pipeline to perform RNA-seq, low copy number WGS and methylation is currently being validated to support clinical trials including paediatric solid tumours at relapse in the UK, such as eSMART [269].

By using the fusion-panel in the samples from the HERBY clinical trial, I was able to identify novel ITD-*NTRK2* predominantly found in thalamic H3.3 K27M tumours. To the best of my knowledge this is the first time ITD-*NTRK2* have been described in paediatric or adult tumours. Similar findings were observed by David Jones, in an independent cohort of samples using different sequencing methodology. Interestingly, in a pan-cancer study published in 2018 in nature by Gröbner *et al.,* they reported SVs in paediatric cancers, and found that 25% of HGG_K27M, 5.6% of HGG_other and 2.9% OF PA, presented *NTRK2* disrupted by SV breakpoints (Supplementary Table19) from [12]. Further work needs to be done to elucidate the functional implication of these duplications. Unfortunately, for none of the ITD-*NTRK2* positive samples live tissue was available, however throughout an international collaboration with James Olson (from Seattle), we will evaluate efficacy of TRK inhibitors and pathway modulation in a patient-derived *in vitro* model which is ITD-*NTRK2* positive. Other findings from this study included a sample carrying an ITD-*FGFR1*, and two patients with *MN1* fusion. These patients are likely to be different tumour entities, corresponding to pilocytic astrocytoma and CNS HGNET-MN1, respectively. Moreover, a patient with a *PTPRZ1:MET* fusion was identified, an alteration previously seen in adult and paediatric HGG and associated with poor prognosis and tumour shrinkage upon crizotinib treatment [367, 392].

In the infant cohort, a total of 45% patients were found to have SVs detected by the fusion-panel. Of those, 29 were consistent with established brain tumours entities and

the remaining infant HGG patients appear to have a prevalence of fusions involving the receptor tyrosine kinases *NTRK1-3* and *ALK.* Further work, carried by Matt Clarke *et al*, showed that additional fusions were found in *NTRK1-3* and *ALK* but also in *MET* and *ROS1* using RNA-seq and WGS [89]. *ALK* fusion positive cases presented sensitivity *in vitro* and *in vivo* to ALK inhibition, resulting in tumour volume reduction and extension of survival compared to standard treatment of TMZ (data not shown in this chapter but the available in [89]). Excitingly, a patient diagnosed of HGG at 1 month old, was confirmed to harbour a *MAD1L1:ALK* fusion and at disease progression started ceritinib, resulting in stable residual disease for nearly two years to date. In addition, through the fusion-panel an *NTRK2* fusion was identified in a 1.5 years old girl diagnosed with an infiltrative high-grade glioma in the spine. The patient underwent resection and several cycles of different chemotherapeutic agents (following the protocol Baby SFOP). After this, on the basis of the *NTRK2* fusion found, the patient is being treated with the TRK inhibitor larotrectinib. Currently the tumour remains stable (data obtained throughout personal communication from the treating cliniclan Dr Lynley Marshall).

The findings from the fusion-panel represent a unique opportunity that can directly be translated in the clinic for the use of targeted agents, and potentially avoid the aggressiveness and long-term side-effects of chemotherapy and radiotherapy. Such inhibitors are currently in clinical trials and accessible to children (NCT02637687, NCT04094610, NCT0265040). Entrectinib and larotrectinib are kinase inhibitors FDA-approved molecules prescribed for the treatment of any solid cancer harbouring NTRK1/2/3 fusion protein, which have shown clinical responses in *NTRK* fusion-positive tumours [366, 393].

In summary, I have shown that using as little as 50 ng of DNA, this technology is accurate, reproducible and a practical platform for molecular stratification and identification of actionable targets, required to accelerate precision medicine clinical trials in childhood tumours.

CHAPTER 5 : Exploring the use of liquid biopsy in paediatric high-grade glioma and DIPG

5.1 Introduction

The incorporation of tissue molecular profiling in patients with pHGG and DIPG into clinical practise has been demonstrated to be essential to guide treatment decision of these patients [15, 394]. However, this implies invasive neurosurgical procedures which frequently are associated with a risk of morbidity or mortality [272, 273, 395]. This is particularly important for tumours located within the brainstem, such as DIPG, where biopsy is technically very challenging and is associated with considerable complications [272, 273].

The study of liquid biopsy has emerged as an alternative and/or complementary approach to tumour biopsy. Liquid biopsy comprises the study of tumour derived material from any biological fluids including blood, cerebrospinal fluid (CSF), urine, and saliva. In this context, cell-free DNA (cfDNA), extracted from different sources of biofluids, is used to assess tumour-specific alterations in a less invasive manner. The fraction of cfDNA derived from tumour cells is known as circulating-tumour DNA (ctDNA). One of the benefits of liquid biopsy is its utility to correlate the presence of driver mutations with tumour burden and therapy response at multiple time-points, preventing the risks, costs and skilled expertise of surgical intervention. In this context, many paediatric gliomas are characterised by hotspot driver mutations (H3.1/H3.3 K27M, H3.3 G34R, *BRAF*-V600E, *IDH1*-R132X) [22] or by single fusion events (*NTRK:ETV6, BRAF:KIAA1549*) [220, 358]. This makes them perfect candidates for the use of ctDNA to monitor treatment response enabling early detection of tumour progression over the course of the disease. The presence of ctDNA from plasma has been demonstrated in paediatric solid tumours including neuroblastoma, osteosarcoma, Ewing sarcoma, Wilms tumour and rhabdomyosarcoma [304, 305, 307, 312, 320, 396]. At the start of this PhD, the use of ctDNA derived from plasma in tumours involving the CNS appeared to be more challenging compared to other paediatric non-CNS tumours [288]. Despite this, several studies have shown the detection of CSF from ctDNA from a mixture of paediatric and adult brain tumours [289-292, 334]. In particular, Wang and colleagues found molecular alterations in 74% of patients from ctDNA derived from CSF, obtaining an average of 417 ng of ctDNA in an average of 4.8 mL of CSF [290]. By using

amplicon NGS methodology (SafeSeqS), they revealed that ctDNA was found in all CNS tumours whose tumours were directly adjacent to a CSF reservoir [290].

In the last three years, similar results have been found in ctDNA derived from CSF in patients with pHGG and DIPG by using nested PCR, ddPCR and panel sequencing [315-317, 397]. Besides detecting molecular alterations in ctDNA derived from CSF, these studies have also demonstrated that CSF-ctDNA levels increase during disease progression and at autopsy, as well as after radiation, suggesting the role of radiotherapy to disrupt the BBB [316, 317]. The largest study evaluating ctDNA-CSF in adult glioma was carried by Miller and colleagues at MSKCC (Memorial Sloan Kettering Cancer Center) [398]. Using the MSK-IMPACT capture-based NGS assay, they identified ctDNA in 49.4% of patients, with ctDNA levels correlating with disease burden and poor outcome [398]. Importantly, they were able to track tumour evolution through longitudinal CSF samples, showing loss of *EGFR* amplification and emergence of *PDGFRA* amplification over the course of the disease [398].

Less is known about the utility of plasma ctDNA in brain tumours and in particular in pHGG and DIPG. Pan and colleagues showed the detection of ctDNA derived from 3 mL of plasma in 3/7 paediatric patients with brainstem tumours; of those three, two had undetectable mutations in the plasma ctDNA compared to the ctDNA derived from the CSF [397]. Conversely, a study from Panditharatna and colleagues showed detectable levels of ctDNA derived from 1 mL of plasma in 16/20 patients with diffuse midline glioma at diagnosis/upfront therapy [317].

To this end, I sought to explore whether molecular alterations could be identified in liquid biopsy samples from pHGG and DIPG patients. ddPCR assays were validated and applied to quantify ctDNA levels derived from plasma, serum and CSF. A secondary aim of this chapter was to correlate DNA concentrations and the presence of genetic alterations with tumour burden and multimodal radiological indicators of response and tumour progression. For this purpose, longitudinal plasma samples from HERBY were used.

5.2 Results

5.2.1 ddPCR assay validation for the detection of ctDNA from liquid biopsy

Liquid biopsy from 32 pHGG and DIPG patients were collected from three different cohorts of patients with known molecular alterations, and utilised for this feasibility study. These patients harboured genetic alterations in *H3F3A* (K27M and G34R), *BRAF* (V600E), *ACVR1* (G328V), *IDH1* (R132H and R132S), *TP53* (C238Y and R282W) and *PIK3CA* (E542K and H1047R) as well as one patient with *MYCN* amplification. The first goal was to develop a robust detection method for the genetic alterations. To do this, customised and commercially available assays for ddPCR were validated for the detection of patient specific molecular alterations. A summary of ddPCR workflow is shown in **{Figure 5-1}**. ddPCR primer and probe sequences are shown in **{Appendix Table 3}**.

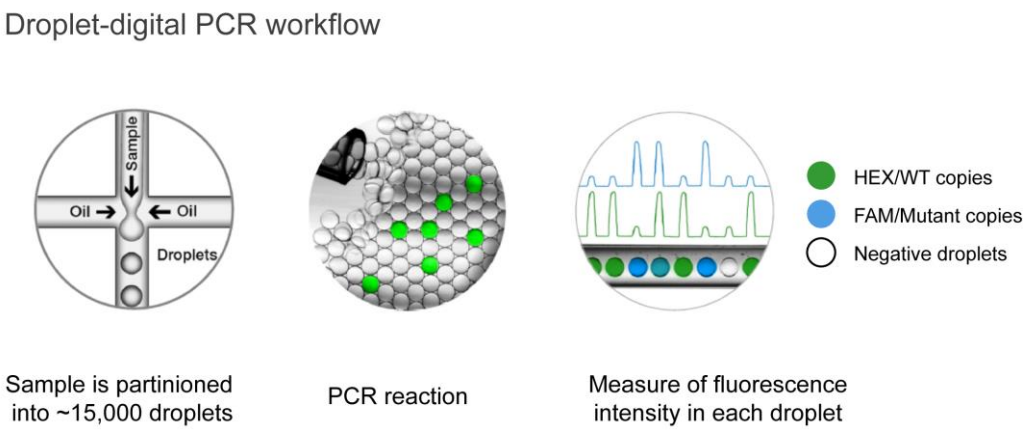


Figure 5-1 ddPCR sample workflow. Droplet generation is performed in a reaction mix containing DNA, primers and probes as well as ddPCR master mix, then the reaction is amplified by PCR and loaded in the QX200 droplet reader where the droplets are read.

Each genotyping assay was tested by using a positive sample harbouring the specific alteration of interest, and VAF was compared between ddPCR and NGS, with an observed correlation of r^2=0.9543 **{Figure 5-2}.**

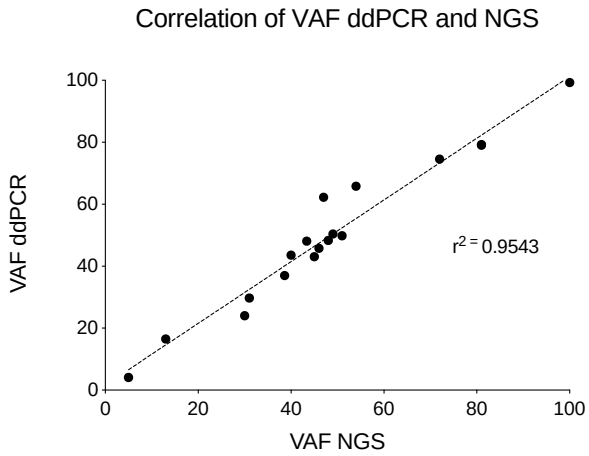


Figure 5-2 Correlation of variant allele frequencies (VAFs) by NGS (x-axis) and ddPCR (y-axis) for ddPCR assay validation. 13 assays were tested in samples positive for the mutations analysed (n=18). ddPCR assays performed well resulting in a correlation of r^2=0.9543.

The *MYCN* amplification assay contained two probes, one within the *MYCN* gene (Hs00201049 SNP genotyping assay ID) and one in a control region at chromosome 5p15.33. The amplification assay was tested by comparing the ratio of copies/µl of *MYCN* to the control gene. Three ctDNA-plasma positive samples by NGS from *MYCN*-neuroblastoma patients were used for the assay validation. Tissue samples for each patient were used as a positive control and were run in duplicate. *MYCN* amplification was detected in the DNA derived from the tissue and the ctDNA isolated from plasma. The ctDNA samples taken at diagnosis from the three patients showed a fold-amplification of 32.7, 115.2 and 110.8 **{Figure 5-3}.**

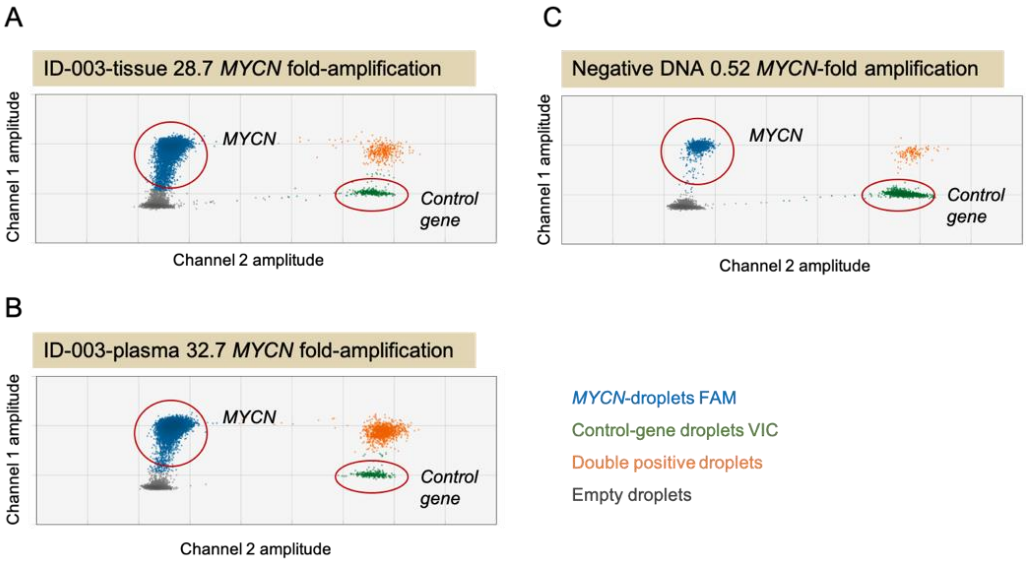


Figure 5-3 *MYCN* ddPCR assay validation. Example of a neuroblastoma patient with known *MYCN* amplification (previously detected by NGS using the Paeds-v1) was used to validate the ddPCR assay. Representative droplet digital plots from the neuroblastoma patient with high level amplification of *MYCN* in (A) FFPE tissue at diagnosis, (B) cfDNA at diagnosis and (C) Non-amplified *MYCN* DNA sample.

To assess limit of detection (LoD) of point mutation detection assays, mutant DNA samples were serially diluted in 10-fold dilution with wild-type genomic DNA (1/10, 1/100, 1/1,000 and 1/10,000). Genomic DNA from tissue was fragmented and a total DNA input of 5 ng was utilised to simulate the anticipated low amount of ctDNA. LoD was calculated as VAF of neat sample divided by the lowest dilution with detectable signal for mutant, with at least two droplets containing mutant DNA. Two different *H3F3A*-K27M assays were assessed, one commercially available from BioRad and one reported by Stallard and colleagues [316]. Both assays performed well obtaining a good droplet separation between FAM and VIC/HEX labels, with a similar LoD (BioRad = 0.793% and custom = 0.791%) **{Figure 5-4}**. In addition, no mutant droplets were observed in any of the wild-type template control DNA included per assay in each run. By using 5ng of DNA, LoD ranged from 0.041% to 0.993%, with an average of 0.330% **{Table 5-1}** and **{Figure 5-5}**.

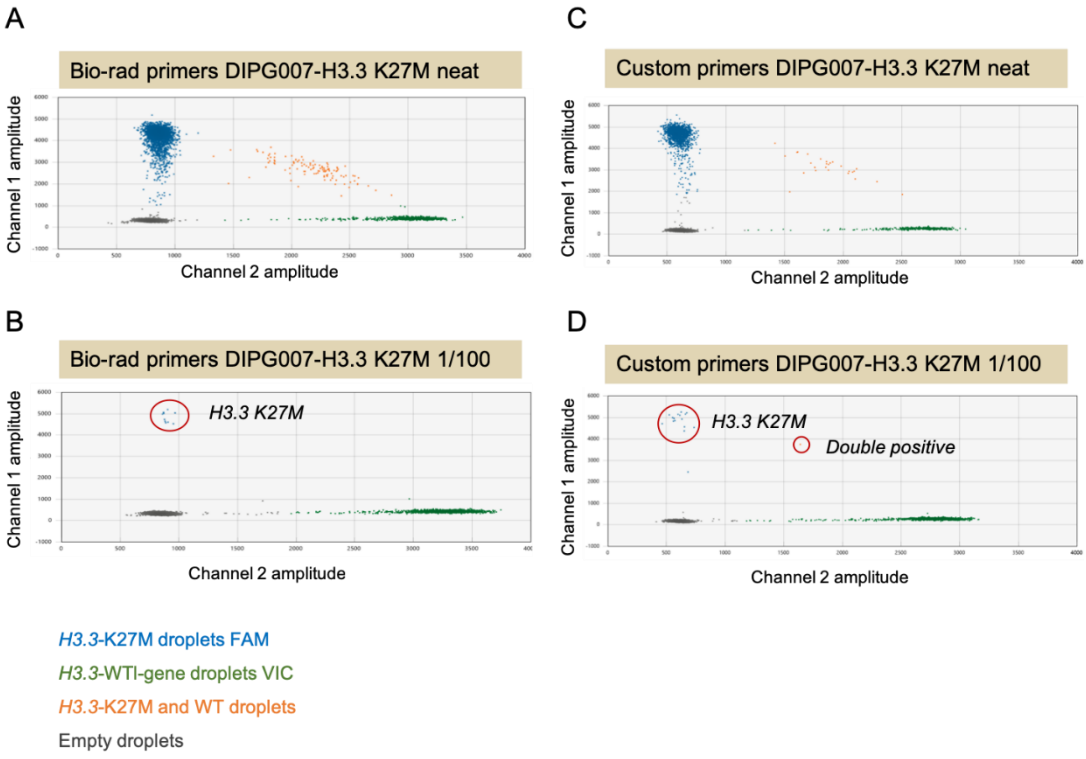


Figure 5-4 *H3F3A* ddPCR assay limit of detection summary. Droplet digital PCR 2D amplitude plots of (A) *H3.3*-K27M tested in HSJD-DIPG007 as positive control by using Bio-rad assay on undiluted (neat) *H3.3*-K27M DNA (1760/2226 VAF of 79.3%). (B), Bio-rad assay on 1/100 dilution of *H3.3*-K27M DNA with wild-type DNA (10/1564 droplets, VAF of 6.4%). (C) Custom assay on neat *H3.3*-K27M DNA (1586/2014 VAF of 79.1%) (D) Custom assay n 1/100 dilution of *H3.3*-K27M DNA with wild-type DNA (17/1613 droplets, VAF of 7.7%). H3.3-K27M droplets are shown in blue, H3.3 wild-type droplets are shown in green, double positive droplets containing H3.3 K27M and wild-type DNA are shown in orange and empty droplets with no DNA are shown in grey.

ddPCR assay	NGS VAF	ddPCR VAF	Dilution	LoD (%)	Mutant droplets	Mutant copies/ul	WT copies/ul
H3F3A-K27M-BioRad	81	79.3	1/100	**0.793**	10	0.33	53.22
H3F3A-K27M-custom*	81	79.1	1/100	**0.791**	17	0.66	63.63
*H3.3*_G34R-BioRad	48	50.3	1/1000	**0.050**	2	0.07	50.30
*BRAF*_V600E-custom	49	50.4	1/1000	**0.050**	2	0.07	65.30
IDH1-R132S-custom	13	16.5	1/100	**0.165**	2	0.07	60.60
IDH1-R132H-custom	40	43.6	1/1000	**0.044**	2	0.07	63.20
IDH1-R132G-custom	45	43.1	1/1000	**0.043**	4	0.13	72.00
TP53-C238Y-custom	99.3	100	1/100	**0.993**	19	0.71	180.61
TP53-R282W-custom	38.6	37.0	1/100	**0.370**	2	0.07	74.30
ACVR1-R328V-custom	54	64.3	1/100	**0.643**	14	0.52	181.14
PIK3CA-E542K-custom	5	4.1	1/100	**0.041**	2	0.07	145.10
PIK3CA-H1047R-custom	30	24.0	1/100	**0.240**	5	0.03	92.60

Table 5-1 ddPCR assays limit of detection (LoD) description. Samples were run in duplicate and merged data is shown. Variant allele frequency (VAF) identified by NGS and ddPCR is shown. LoD is calculated by dividing ddPCR VAF by the lowest dilution with at least two positive droplets present. * Assay described in [316]

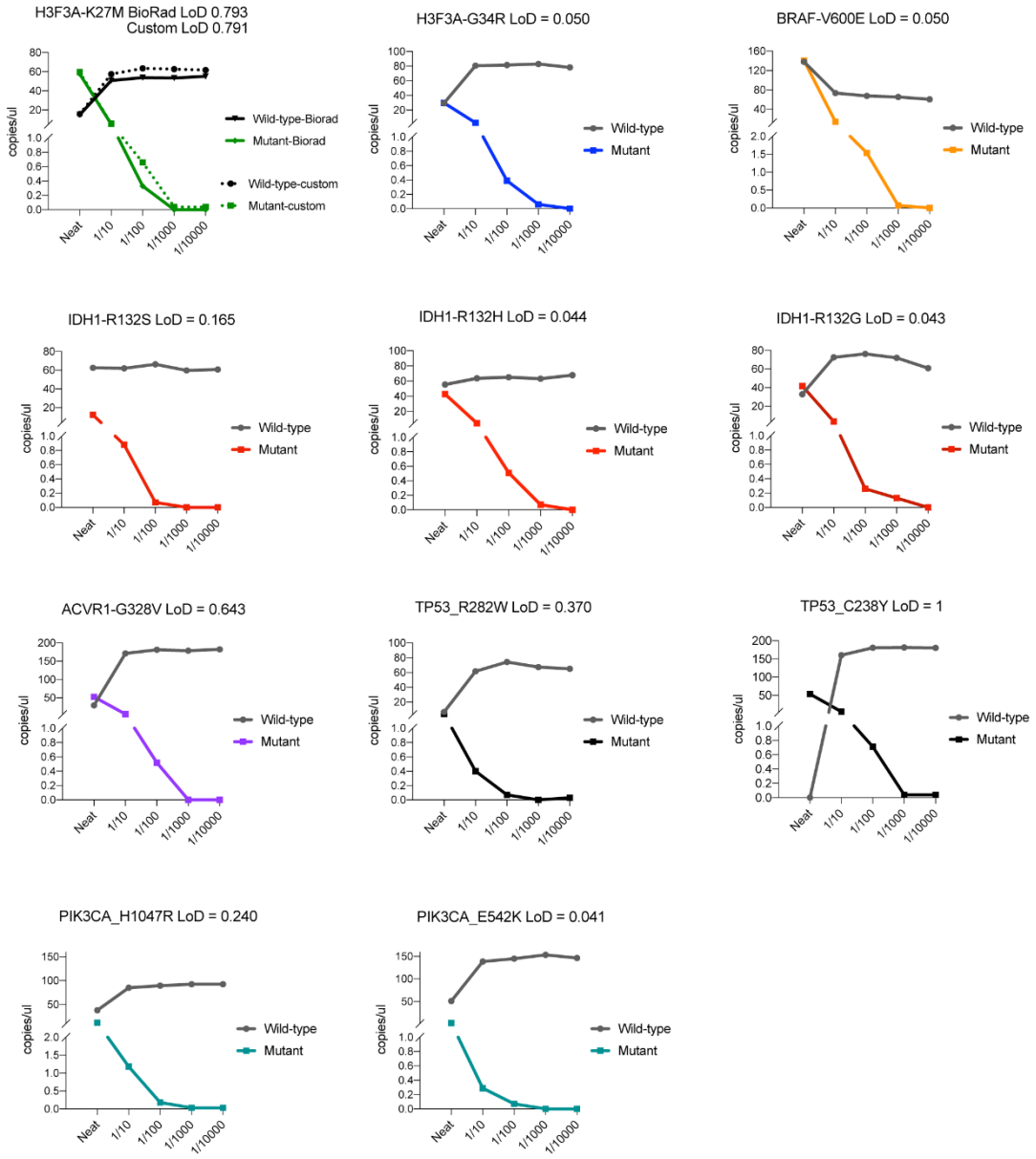


Figure 5-5 ddPCR assay limit of detection plots. Linear dilutions of mutant DNA in constant background of wild-type DNA are shown against number of copies/ul for wild-type and mutant alleles. A total of 5ng of total DNA was load in each PCR well. Samples were run in duplicates and merged data is shown.

5.2.2 Genetic alterations can be detected in ctDNA from CSF and plasma in pHGG and DIPG

To test the feasibility of ctDNA detection in pHGG and DIPG, the validated ddPCR methodology was applied in a cohort of 43 liquid biopsy samples from 32 patients, which included 27 plasma, 6 serum, 9 CSF and one cyst fluid sample **{Figure 5-6A}**. The average volume of liquid biopsy was 3.14 mL (SD=1. 2) for plasma, 2 mL (SD=0.4) for serum, and 1.744 mL (SD=1.5) for CSF **{Figure 5-6B}**. From patient CXJ024, a large volume (350mL) of cystic fluid was collected at time of resection and 35 mL were used for cfDNA extraction.

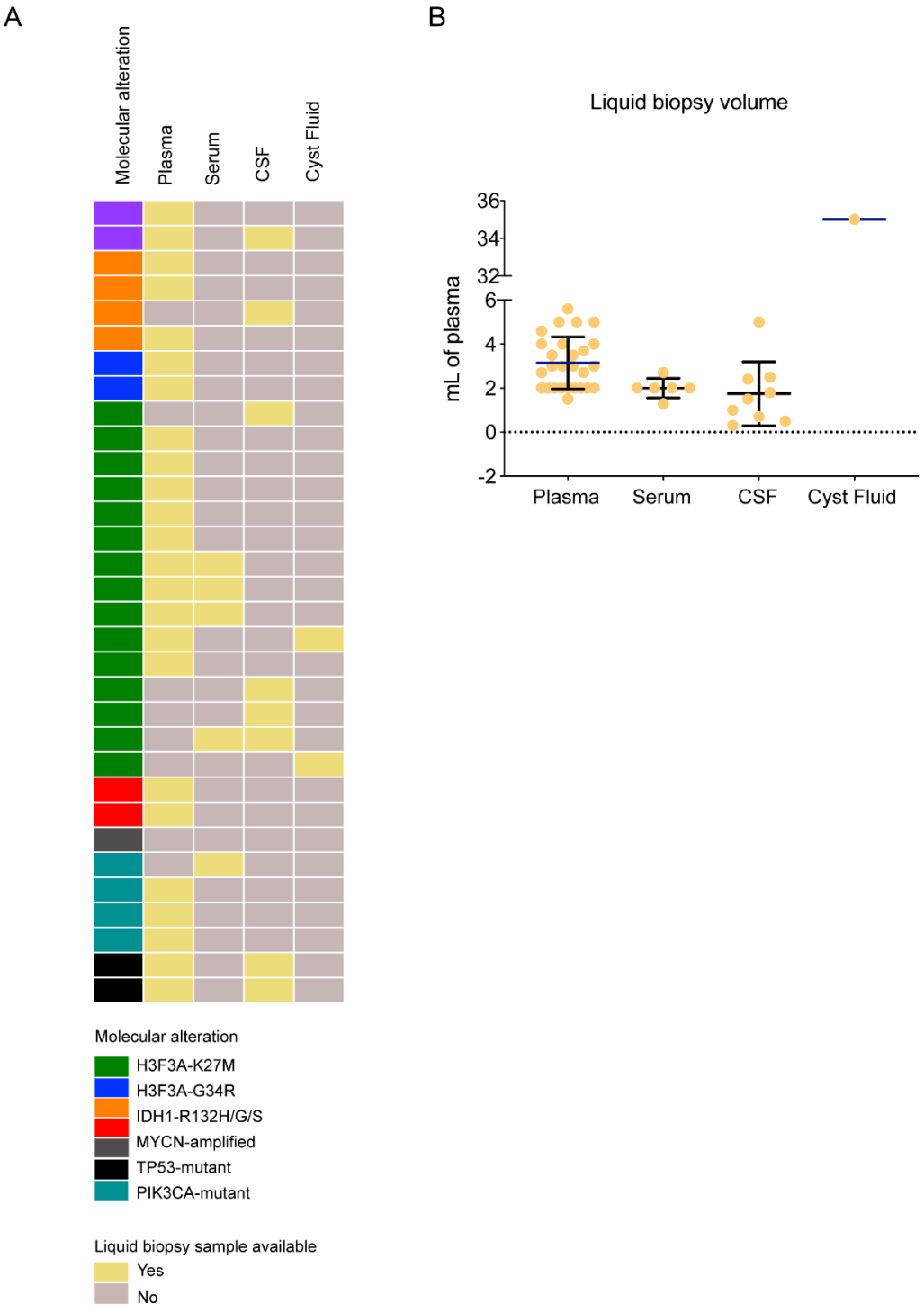


Figure 5-6 Cohort of pHGG and DIPG samples used for liquid biopsy feasibility study. (A) Summary table of samples available where each row represents a patient and each column a sample, cells are coloured by molecular alteration assessed and sample availability. (B) Dot plot of liquid biopsy in mL used for cfDNA extraction, each sample is represented by a dot, middle line represents the mean and upper and bottom line the SD.

Molecular alterations were found in a total of 16 ctDNA samples, including those derived from plasma (7/27, ~26%), CSF (6/9, ~67%), serum (2/6, 33.4%) and the only cystic fluid specimen available. Variants included *H3F3A*-G34R (n=2), *H3F3A*-K27M (n=7), *IDH1*-R132H (n=1), *PIK3CA*-H1047R (n=1), *PIK3CA*-E542K (n=1), *ACVR1*-G328V (n=1), *TP5*3-C238Y (n=1) and *TP53*-R282W (n=2) **{Table 5-2}.**

Patient ID	ddPCR assay ID	Sample type	Volume for ctDNA	VAF	Mutant droplets	Wild-type droplets	cfDNA Mutant ng/mL	cfDNA Wild-type ng/mL
013-T	H3F3A-K27M	Plasma	2.7	2.451	10	398	0.028	1.131
045-T	*TP53*-R282W	CSF	5	49.343	3945	4050	271.913	279.585
045-T	*TP53*-R282W	Plasma	5	0.118	3	2540	0.021	18.313
054-T	*H3F3A*-G34R	Plasma	5	0.468	2	425	0.007	1.561
106-T	*IDH1*-R132H	Plasma	3.5	0.335	11	3272	0.029	8.844
120-T	*H3F3A*-G34R	Plasma	3	1.143	2	173	0.005	0.422
131-T	*H3F3A*-K27M	Cyst	38.5	42.68	10944	14698	367.594	541.671
131-T	*H3F3A*-K27M	Plasma	2	0.85	3	350	0.009	1.048
15-3381	*H3F3A*-K27M	CSF	1	3.646	7	185	0.044	1.177
16120B	*PIK3CA*-E542K	Serum	2	0.072	7	9674	0.291	472.795
BIOMEDE-134	*H3F3A*-K27M	Serum	1.3	0.39	2	511	0.015	3.934
BIOMEDE-276	*PIK3CA*-H1047R	Plasma	2	0.116	2	1727	0.006	5.098
CXJ026	*TP53*-C238Y	CSF	2.5	2.073	4	189	0.013	0.633
CXJ028	*ACVR1* G328V	CSF	1.8	35.737	446	802	1.472	2.662
I-16-3200	*H3F3A*-K27M	CSF	0.7	0.051	3	5916	0.071	159.274
I-16-855	*H3F3A*-K27M	CSF	1.5	1.149	9	774	0.054	4.703

Table 5-2 Summary table of liquid biopsy with detectable ctDNA levels. In the table are shown the mutation assessed by ddPCR, type of liquid biopsy and volume used for extraction, variant allele frequency (VAF), mutant and wild-type number of droplets as well as mutant and wild-type ng/ml.

Although not formally significant due to small numbers and high degree of variability, average of positive droplets was higher in ctDNA derived from CSF 735.7 (SD=1582), than from plasma 4.7 (SD=3.9) and serum 4.5 (SD=3.53) (p=0.5879 and p=0.8167 respectively, one-way ANOVA, Tukey's multiple comparisons test) **{Figure 5-7A}**. Equally, the average VAF was higher in ctDNA derived from CSF 15.33% (SD=21.54%) than from plasma 0.78% (SD=0.31%) and serum 0.22% (0.16%)

(p=0.2867 and p=0.5633 respectively, one-way ANOVA, Tukey's multiple comparisons test) **{Figure 5-7B}**. The highest number of positive droplets (10,944, VAF=42.68%) was found in the cystic fluid. Paired CSF/cyst fluid and plasma/serum were available for five patients - of these, two alterations were detected in both liquid biopsy sources and for the remaining three cases variants were only identified in the CSF. For patient 045-T, who presented with a hemispheric HGG with hypermutator phenotype, described in Chapter 4, *TP53*-R282W was identified in ctDNA derived from CSF (VAF=49.34%) and the plasma (VAF=0.12%). In addition, patient-131-T, with a right thalamic glioma, *H3F3FA*-K27M was identified in the cystic fluid (VAF=42.68%) and the plasma (VAF=0.85%).

Although the formal threshold for a positive sample was set as at least two positive droplets, it is worth noting that a single positive droplet was found in seven cases, including five cfDNA derived from plasma (*H3F3FA*-K27M n=4, and *ACVR1*-G325V n=1) and two CSF (*H3F3FA*-K27M n=2).

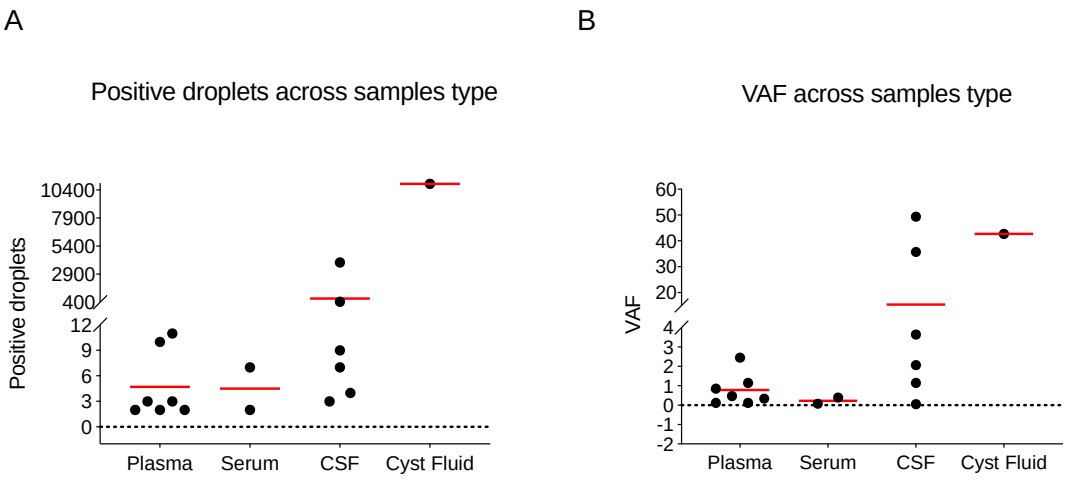


Figure 5-7 Dot plot of detectable ctDNA from liquid biopsy samples in pHGG and DIPG samples. Samples are separated by ctDNA source, plasma, serum, cerebrospinal fluid (CSF) and cystic fluid. (A) Positive droplets (>2) across sample type are shown and (B) variant allele frequency across sample type. Each sample is represented by a dot and the red line represents the mean.

The concentration of cfDNA was calculated from the number of copies/µl (mutant + wild-type) detected by ddPCR, following published methods [289]. Mean of cfDNA concentration per mL was 5.2 ng/mL (SD=4.4) from plasma samples, 110.8 ng/mL (SD=179.9) from serum and 80.33 (SD=184.2) from CSF **{Figure 5-8}**. 1012 ng/ml were obtained from the cyst fluid sample. By assessing the DNA integrity with a Tapestation electrophotometric analyser, it was found that 4/6 cfDNA extracted from serum presented a smear of fragmented DNA including genomic DNA (gDNA)

{Figure 5-9A}. In addition, ctDNA was found in 8/13 samples with a detectable cfDNA **{Figure 5-9B}**. Of note, ctDNA-CSF sample ID I-16-3200 *H3.3*-K27M positive, showed gDNA contamination and had the lowest VAF of CSF-ctDNA samples (0.05%) **{Figure 5-9C}**.

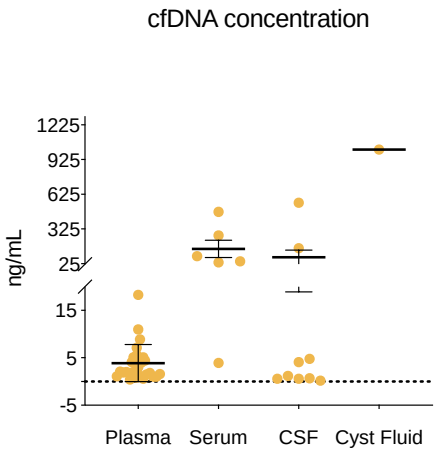


Figure 5-8 Dot plot of cfDNA concentrations (ng/mL) of liquid biopsy samples. pHGG and DIPG and samples are separated by cfDNA source, plasma, serum, cerebrospinal fluid (CSF) and cystic fluid. Each sample is represented by a dot, middle line represents the mean and upper and bottom line the SD.

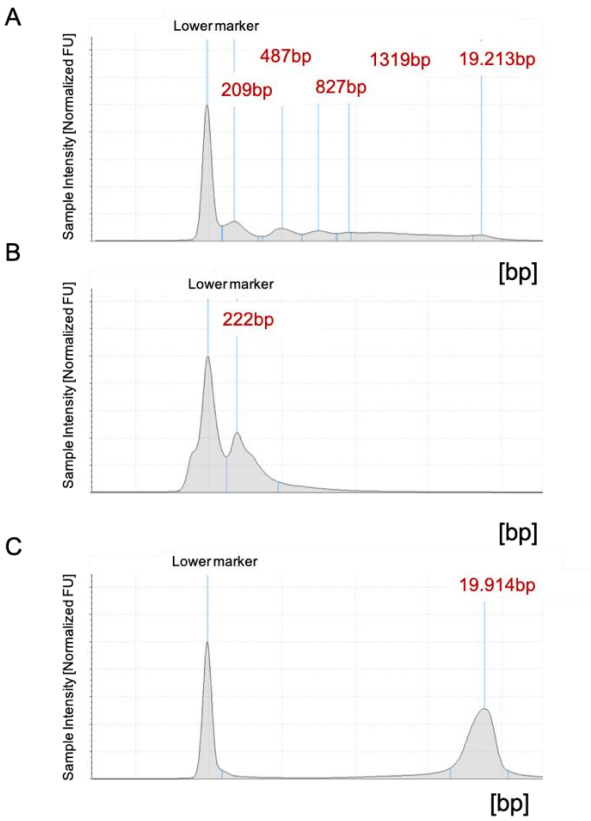


Figure 5-9 Electropherogram of cfDNA size distribution obtained by using the Tapestation. Genomic DNA ScreenTape assay was utilised. The y-axis shows the signal intensity (FU) and the x-axis shows the DNA fragment size is represented in base pairs (bp). (A) Example of a highly fragmented DNA-serum sample (ID-BIOMEDE-118). (B) Example of a cfDNA-CSF with a high peak of an average size of 222bp (ID-045-T). (C) Example of a DNA sample derived from CSF (ID-C15-654) with genomic DNA contamination of 19.914bp average size.

Moreover, there was a CSF available for an infant glioma (OPBG_INF_035) with a known fusion in *ETV6:NTRK3*. cfDNA was extracted from 4.5mL of CSF and 30 ng of cfDNA were used to run the fusion-panel (described in Chapter 3). Library preparation and hybridisation, using the fusion-panel, were conducted by Matthew Clarke in my laboratory. ctDNA was detected from the infant glioma case exhibiting the *ETV6:NTRK3* fusion **{Figure 5-10}.**

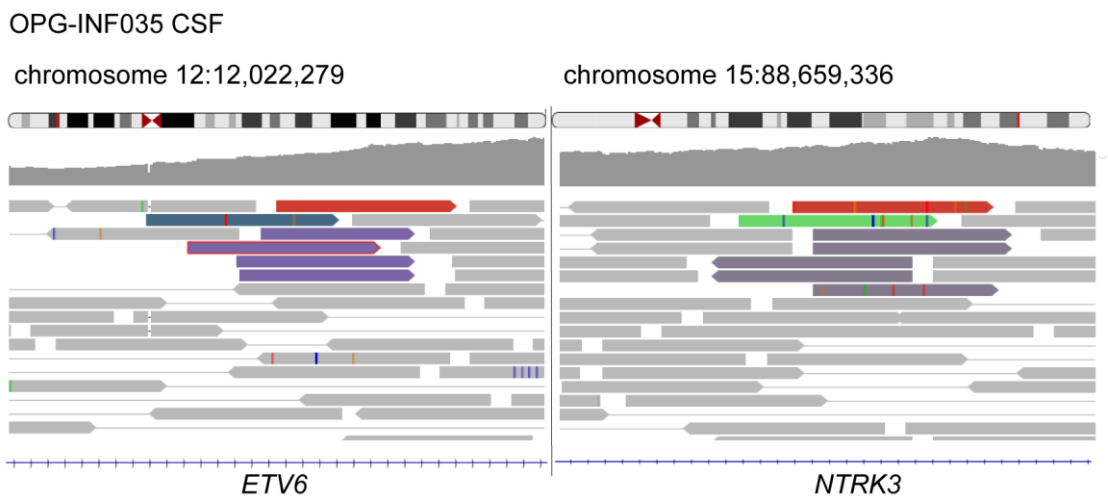


Figure 5-10 Integrative Genomics Viewer plot (IGV) of *ETV6:NTRK3* translocation detected from ctDNA-CSF. The translocation was originally found in the tissue sample from an infant HGG patient by the fusion-panel. The translocation in the ctDNA-CSF was also identified by the fusion-panel showing DNA supporting reads of the fusion gene between *ETV6* coloured in purple and *NTRK3* in grey.

5.2.3 Exploring the use of liquid biopsy in the HERBY clinical trial cohort

To assess genetic alterations in cfDNA derived from plasma within a clinical context, longitudinal samples from the HERBY trial were used. Blood samples were taken at 5 different time-points during the course of the patient treatment **{Figure 5-11}.** Plasma was isolated at the local centres and sent to our laboratory. cfDNA was extracted from 127 plasma aliquots from different time-points, corresponding to 41 HERBY patients, who harboured genetic alterations in *H3F3A, IDH1*, *BRAF* and *MYCN* **{Figure 5-12A}**.

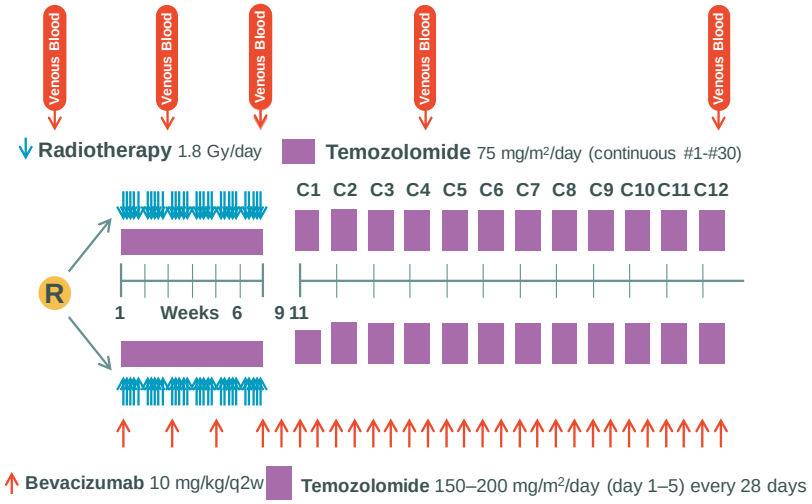


Figure 5-11 HERBY sample collection overview. The illustration shows the two treatment arms in the HERBY phase II study of the addition of bevacizumab to radiotherapy and temozolomide in non-brainstem pHGG. The cartoon includes the different time-point where blood was withdrawn during the course of treatment.

.

The mean of plasma from which cfDNA was extracted was 0.49 mL (SD=0.35, excluding one sample from which 4mL of plasma were used for extraction) **{Figure 5-12B}**. DNA integrity was measured by using Tapestation showing four different types of DNA size distribution: 32 cases presented a detectable cfDNA peak (~160bp) **{Figure 5-13A}**, five cases contained high amount of gDNA (>45.000bp) **{Figure 5-13B}**, 13 cases showed detectable cfDNA and genomic DNA peak **{Figure 5-13C},** and the remaining 75 cases no DNA size peak was detectable **{Figure 5-13D}**. Of these, the four cases with high amount of gDNA contamination were excluded from the DNA concentration comparison.

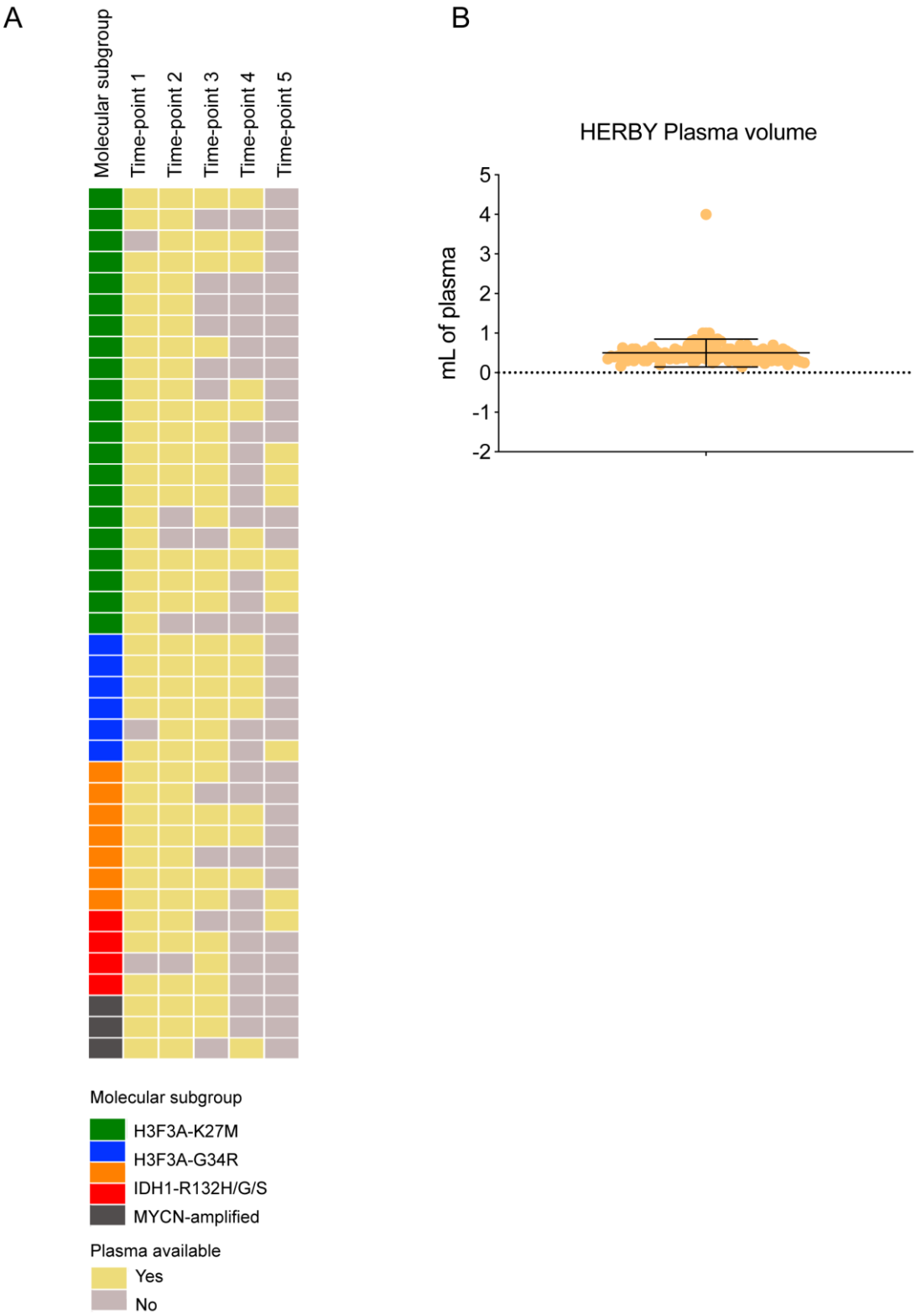


Figure 5-12 Plasma samples available from the HERBY trial. (A) Heatmap of plasma samples available sorted by mutation *H3.3, IDH1, BRAF* and *MYCN.* (B) Dot plot of the volume of plasma used for cfDNA extraction in mL. Each sample is represented by a dot, middle line represents the mean and upper and bottom line the SD. Time-point:1=Baseline, 2= week 3, 3=week 7, 4=month 6 and, 5=End of treatment.

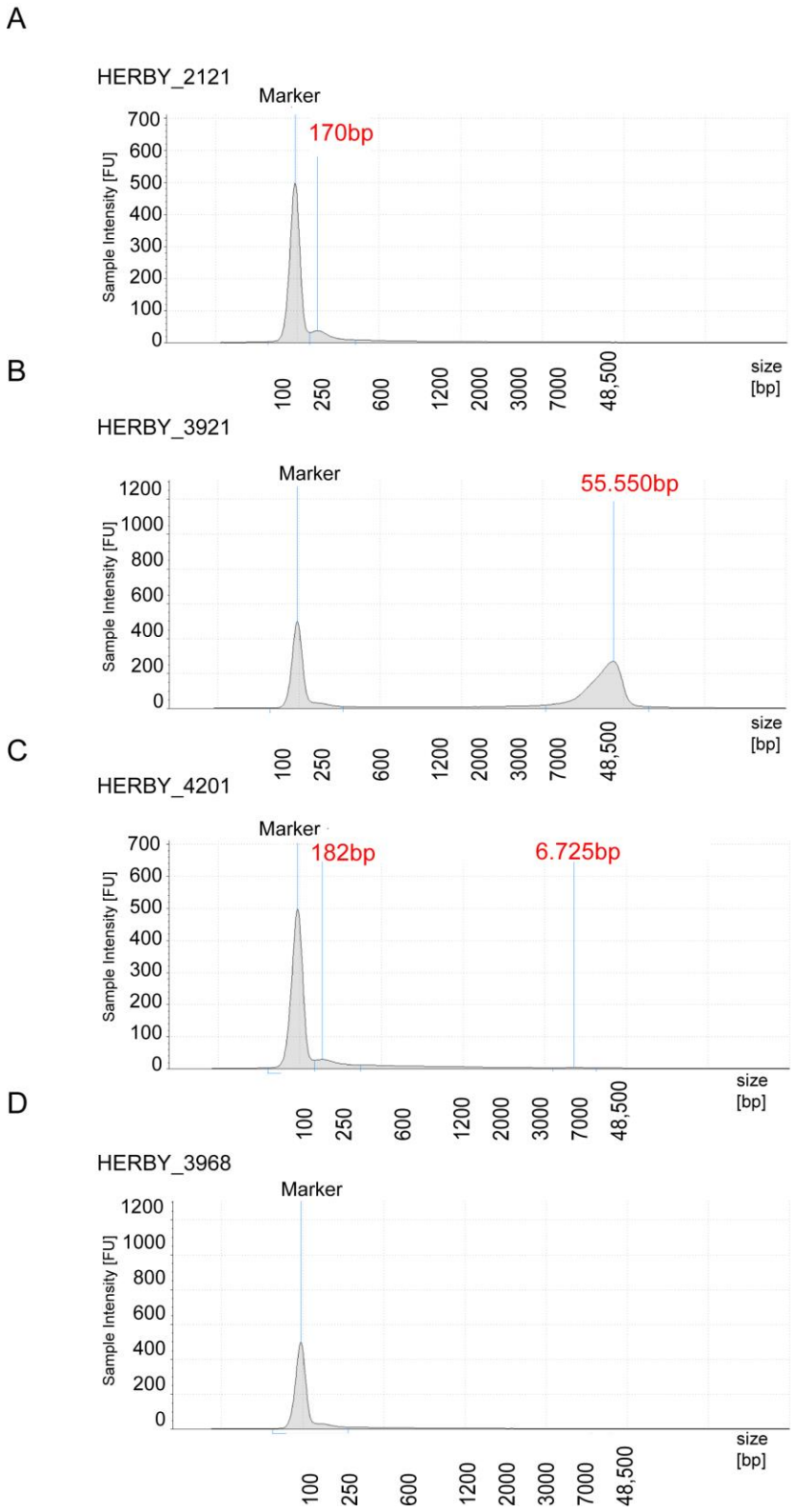


Figure 5-13 Electropherogram of size distribution from HERBY cfDNA plasma. cfDNA and cellular genomic DNA (gDNA) was detected by the Tapestation with the Genomic DNA ScreenTape assay. The y-axis shows the signal intensity (FU) and the x-axis shows the DNA fragment size is represented in base pairs (bp). (A) DNA with detectable levels of cfDNA (~170bp). (B) DNA with only detectable levels of gDNA (~55.500bp). (C) DNA with detectable levels of cfDNA (~182bp) and gDNA (~6.725bp). (D) DNA with no detectable peak. RFU, relative fluorescence unit.

The mean of total DNA extracted from plasma was 2.52 ng (SD=2.83, excluding the four cases with high levels of gDNA) **{Figure 5-14A}**. The mean of total DNA extracted per mL of plasma was 5.25 ng (SD=5.21, excluding the four cases with high levels of gDNA) **{Figure 5-14B}**. The DNA samples were run neat and the mean of DNA ddPCR input was 1.76 ng (SD=2.04) **{Figure 5-14C}**.

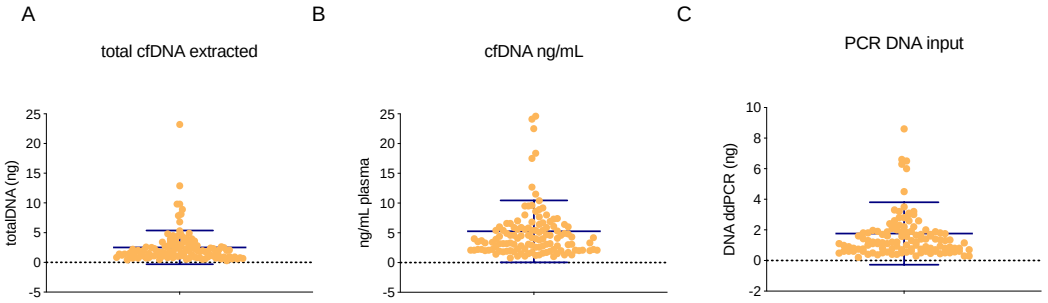


Figure 5-14 Dot plot showing cfDNA levels from the HERBY plasma samples. (A) Total cfDNA extracted, (B) ng/mL of cfDNA and (C) ddPCR DNA input in ng. Each sample is represented by a dot, middle line represents the mean and upper and bottom line the SD.

Disappointingly, none of the cfDNA HERBY samples tested for *H3F3A*-K27M, *H3F3A*-G34R, *IDH1*-R132H/S/G, *BRAF*-V600E or *MYCN* amplification, were positive (>two mutant droplets for point mutations and >4-fold for *MYCN* amplification). However, there were four cases where one positive droplet was found (*H3F3A* -K27M n=1, *H3F3A* -G34R n=1 and *BRAF*-V600E n=2) **{Table 5-3}.**

HERBY_ID	Mutation	Time-point	Positive droplets	Negative droplets	cfDNA ng/mL	cfDNA ddPCR input ng
H3481_RMH9012	H3.3_K27M	1	1	48	0.96	0.46
H1765_RMH8078	H3.3_G34R	1	1	112	2.15	0.77
H2921_RMH8095	BRAF_V600E	1	1	609	22.47	6.29
H3361_RMH8087	BRAF_V600E	1	1	2510	49.4	23.7

Table 5-3 Summary table of the HERBY sample where one positive droplet was detected by ddPCR. Amongst the samples were one H3.3_K27M, two H3.3_G34R and 2 BRAF_V600E. Droplets are divided by positive droplets (droplets with mutant), negative droplets (droplets with wild-type DNA only). In addition, concentration is shown as ng/mL and the cfDNA ddPCR input in ng.

cfDNA concentration was compared between molecular subgroups. Although there was no significant difference between subgroups at baseline (p=0.1026, one-way ANOVA), there was a trend of higher concentration of cfDNA in *BRAF*-V600E positive patients compared to *H3F3A*-K27M and G34R (p=0.0547 and p=0.0661, respectively, one-way ANOVA, Dunnett's multiple comparisons test) **{Figure 5-15}**. cfDNA per mL

was associated with patient response and tumour burden. To do this, cfDNA concentrations were compared across longitudinal specimens over the course of the patient disease **{Figure 5-16}.** Anecdotal variations across time-points were observed in four patients. HERBY 2121 and HERBY-3762, both *H3F3A*-K27M were found to have high levels of cfDNA at later time-point (3 and 4 respectively), compared to cfDNA-plasma specimens taken at earlier time-points (1 and 2 for both samples). In both *H3F3A*-K27M cases formal progression was observed in the MRI scans correlating with higher cfDNA levels. EFS was 5.5 months for HERBY-2121 and 4 months for HERBY-3762 patients **{Figure 5-17A-B}**. The opposite was observed for two *BRAF*-V600E patients, HERBY-2921 and HERBY-3361. The levels of cfDNA were decreasing at later treatment time-points (3 and 4 for both samples) compared to earlier time-points (1 and 2 respectively). In both *BRAF*-V600E patients, stable disease was observed correlating with lower levels of cfDNA in plasma and progression was observed at later time-point in the MRI scans than the two *H3F3A*-K27M patients. EFS was of 8 months for HERBY-2921 and 10 months for HERBY-3361 patients **{Figure 5-17C-D}**.

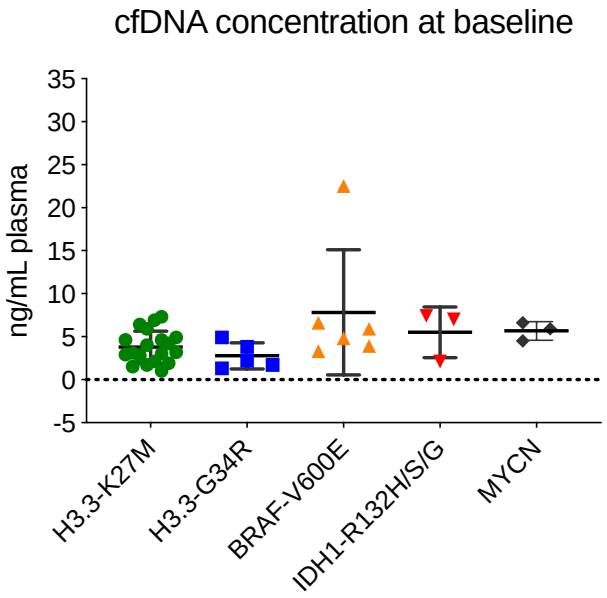


Figure 5-15 Dot plot of cfDNA concentration in plasma (ng/mL). Samples taken at baseline are shown divided and colured by molecular subgroup. Each sample is represented by a dot, middle line represents the mean and upper and bottom line the SD.

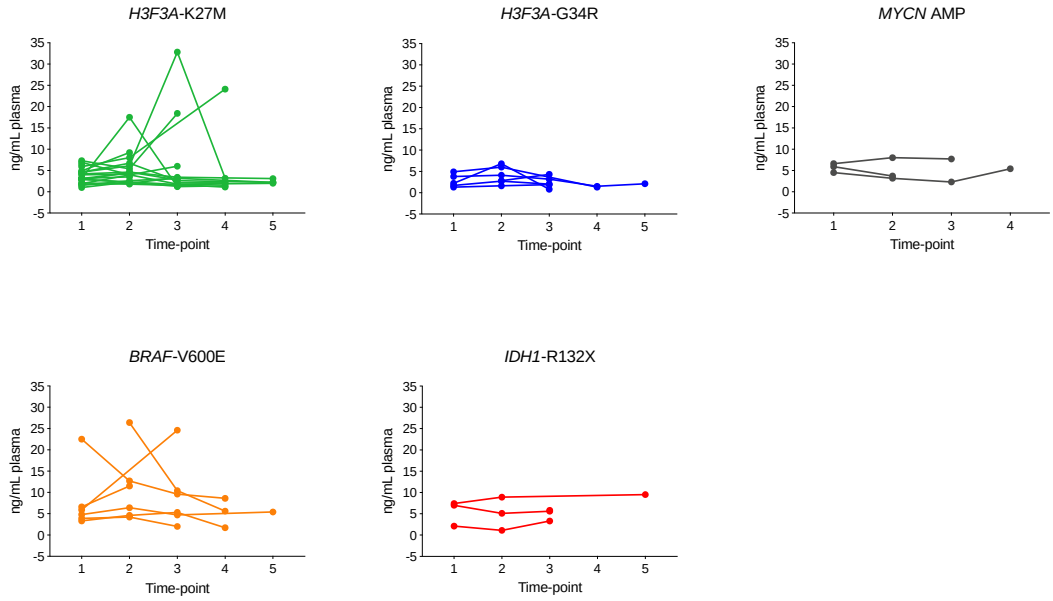
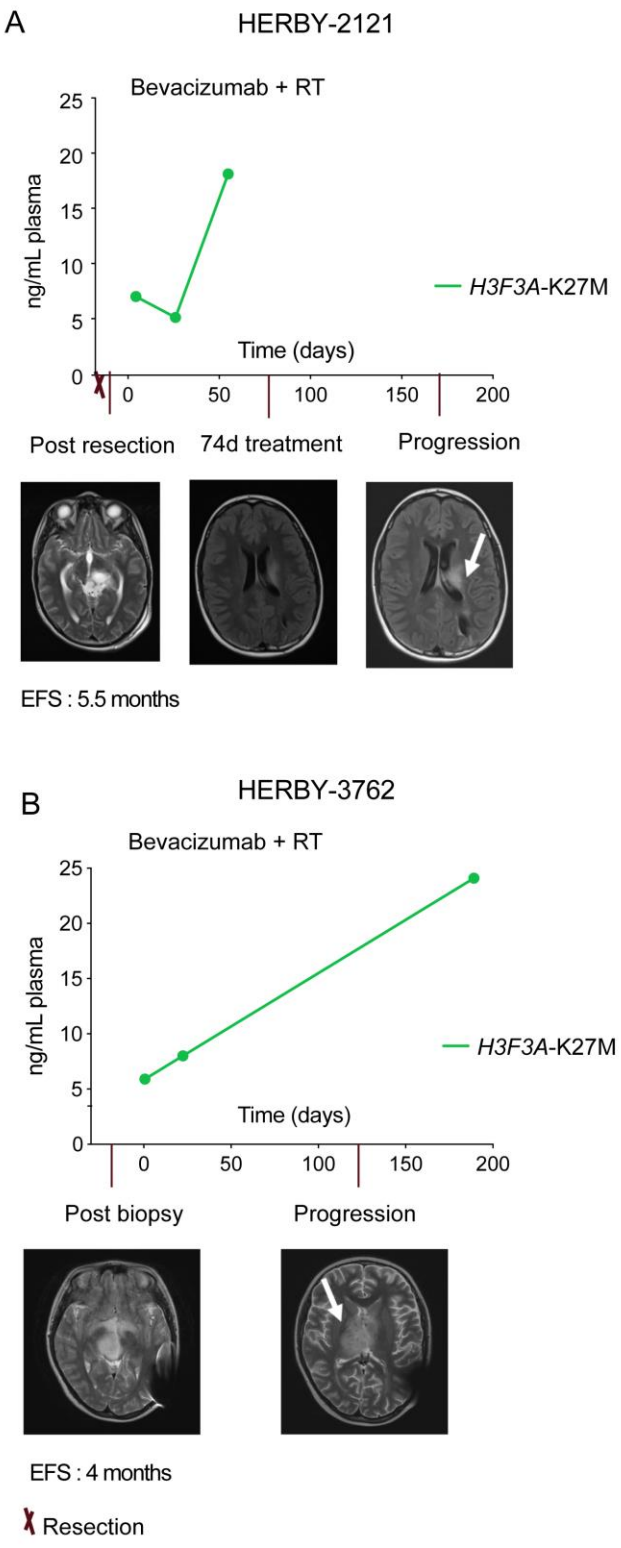


Figure 5-16 Graphs showing cfDNA concentration (ng/mL) per sample at different treatment time-points. Each dot represents a sample and specimens from the same patient are joined by a line. Data is plotted and coloured by molecular subgroup. Time-point:1=Baseline, 2= week 3, 3=week 7, 4=month 6 and, 5=End of treatment.


A
HERBY-2121
Bevacizumab + RT
ng/mL plasma
25
20
15
10
5
0
H3F3A-K27M
Time (days)
0
50
100
150
200
Post resection
74d treatment
Progression
EFS : 5.5 months
B
HERBY-3762
Bevacizumab + RT
ng/mL plasma
25
20
15
10
5
0
H3F3A-K27M
Time (days)
0
50
100
150
200
Post biopsy
Progression
EFS : 4 months
Resection

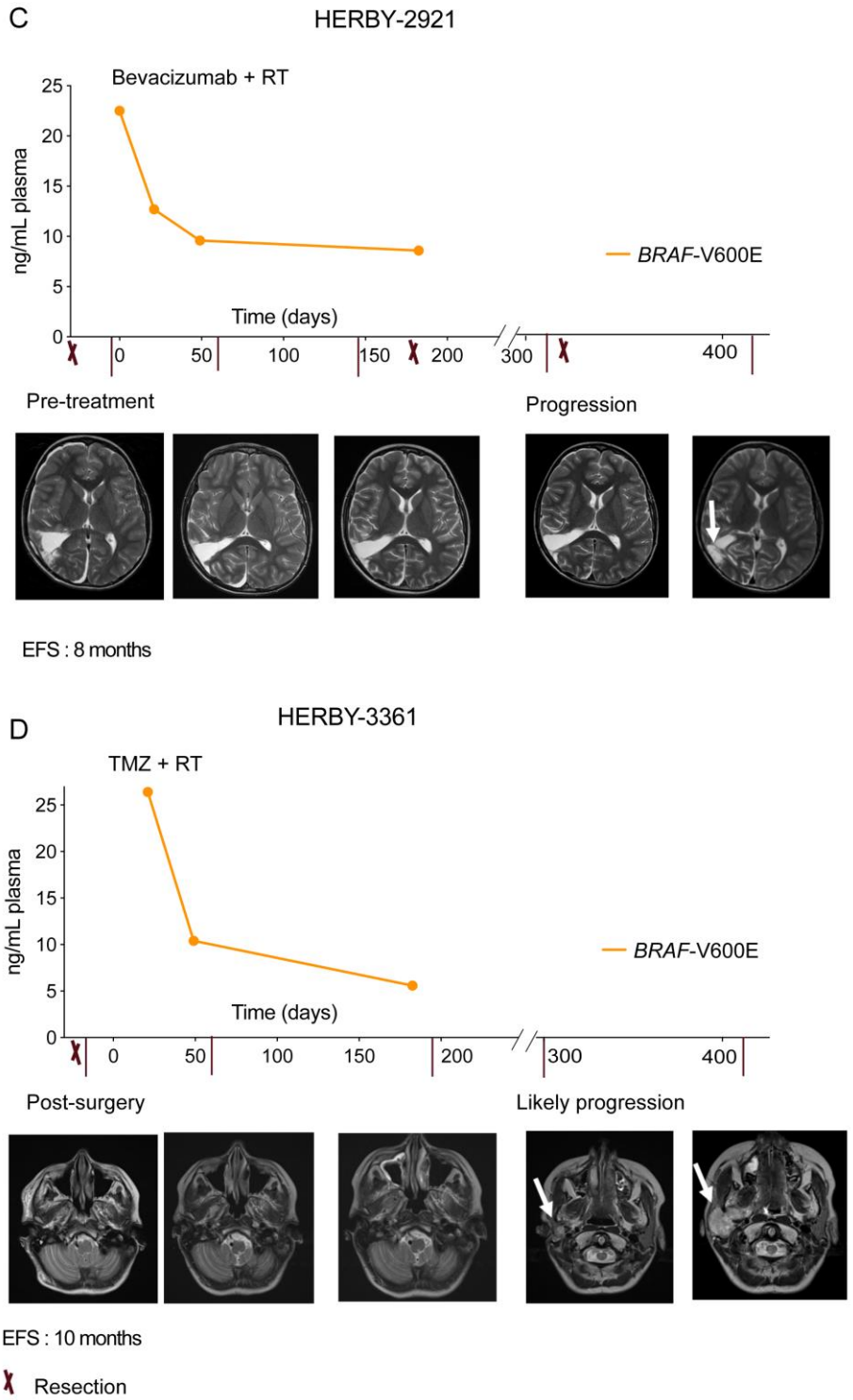


Figure 5-17 HERBY cfDNA concentration levels at longitudinal samples correlating with clinical response. (A-B) Increased levels of cfDNA were observed in two *H3F3A*-K27M at time 3 (HERBY-2121) and time 2 and 4 (HERBY-2762) compared to baseline correlating with early progression and EFS of 5.5 and 4 months respectively. (C-D) Decreased levels of cfDNA were observed in two *BRAF*-V600E at time 2, 3, 4 (HERBY-2921) and time 3 and 4 (HERBY-3361) compared earlier time-points correlating with stable disease and EFS of 8 and 10 months respectively. Plasma cfDNA levels in ng/ml are shown in y-axis and time of treatment in days in x-axis. Axial T2-weighted scans are shown below at different time-points of the patient disease. Disease progression is stated for each patient and the white arrows highlight increased of the tumoural area at progression. Event free survival is stated below the MRI scans for each patient. MRI scans were kindly provided by Dr Tim Jaspan and Dr Daniel Rodriguez.

5.3 Discussion

In this chapter, I describe the validation of different ddPCR assays for the detection of point mutations in *H3F3A, IDH1, PIK3CA, BRAF, ACVR1* and *TP53,* as well as amplification of *MYCN*. By applying this methodology to cfDNA, tumour mutations were detectable in CSF, cystic fluid, plasma and serum DNA derived from pHGG and DIPG patients. In concordance with other studies [289, 292], it was found that ctDNA was present at higher percentage and VAF in the CSF compared to plasma and/or serum specimens.

Although this study was done in a limited number of samples, the results support the use of CSF over plasma as source of tumour DNA for molecular profiling, which has been previously reported in other studies [291, 315, 397-399]. Detectable ctDNA derived from CSF was found in ~67% of samples compared to 26% and 33% derived from plasma and serum respectively. A similar percentage of ctDNA-CSF has been published in pHGG and DIPG, with a detection range of ~66-84% (75% (3/4) [317], 66% (4/6) [315], 83.8% [397]). The reported levels of ctDNA-plasma in CNS tumours present less concordance amongst studies with a wide detection range of 16%-80% (16% 3/19 [398], 33% 4/12 [291], 37.5% 3/8 [397], 45% 302/665 [400], 16/20, 80% [317]).

Increasing number of clinical trials require molecular characterisation for specific biomarker detection as inclusion criteria. For example, *H3F3A*-K27M and *BRAF*-V600E need to be confirmed in patients to be eligible for ONC201 (NCT03416530) or dabrafenib in combination with trametinib (NCT02684058), respectively. This is of particular importance for patients such as those with DIPG, where tissue biopsy can be a very invasive and often an unsafe procedure [272, 273]. In addition, the use of ctDNA can provide a unique opportunity to assess therapeutic response to targeted agent. This can be done by collecting liquid biopsies at multiple time-points and tracking the VAF through the course of the patient disease.

Longitudinal plasma samples from HERBY, the largest randomised clinical trial in non-brainstem pHGG, represented a unique cohort of trial samples. Due to an extraordinary multidisciplinary effort, comprehensive tumour characterisation involving pathology, molecular, clinical and radiological features were performed as part of the trial [69, 87, 321, 335, 401]. It was unfortunate that no ctDNA could be detected in any of the samples assessed as part of my study. At the time the clinical

trial began in 2011, the field of cfDNA was just emerging and the sample collection was not intended for this type of analysis, hence the volume taken was that small. While we were aware of this limitation, it was yet considered to be a valuable resource and important question to be addressed. Nevertheless, concentration of cfDNA was evaluated and ng/mL were compared across time-points showing and increment of cfDNA concentrations in two patients *H3F3A*-K27M compared to baseline correlating with early progression. Contrary to two *BRAF*-V600E patients in which cfDNA concentrations were found to decreased at later-time point treatment compared to earlier time-point correlating with stable disease and longer EFS compared to *H3F3A* -K27M tumours. Although this data is limited, cfDNA levels in four patients were aligned and correlated with clinical response as seen in the MRI scans.

Circulating tumour DNA represents a small fraction of total cfDNA, and the low yields seen in pHGG and DIPG patients represent a major challenge for the detection of this potentially useful biomarker. It is thought that the low permeability of the brain-blood barrier might prevent ctDNA from spreading into the bloodstream. This is supported by the fact that higher ctDNA levels derived from plasma are observed in diffuse midline glioma after radiation (72-100h), suggesting that radiotherapy might disrupt the BBB allowing the ctDNA to get in the bloodstream [316, 317]. Another reason for lower levels of ctDNA isolated from plasma and or serum is the presence of background genomic DNA from non-malignant cells. In particular, it was observed that no ctDNA was detected in samples presenting highly fragmented cellular DNA, presumably derived from cells undergoing necrosis.

To increase the detection rate of ctDNA, different considerations should be evaluated. Firstly, as it was observed from the HERBY cohort, small volumes of plasma have major implications in the detection of ctDNA. At least 4 mL of plasma are required for clinically approved liquid biopsy test such as Guardant360 ctDNA [402]. Although it is understandable that large amounts of whole-blood are not always feasible to drawn from young patients, when possible, 9-10 mL of whole-blood to obtain 4-5 mL of plasma would be ideal. Furthermore, as cellular DNA contamination can affect the sensitivity of ctDNA detection, some studies have applied *in-silico* and *in vitro* size selection to achieve higher sensitive evaluation of ctDNA [403]. However, this needs to be further verified as size selection after cfDNA extraction might contribute to potential loss of cfDNA material. Another strategy that Panditharatna and colleagues used in their study, which detected ctDNA in 80% of diffuse midline gliomas at diagnosis/upfront therapy, was a preamplification step of 9 cycles [317]. This could

explain their high detection rate and should be further validated to assess the potential false positive rate introduced by pre-amplification.

While ddPCR is a broad approach used in the clinic to monitor cfDNA, new methods need to be assessed to obtain lower limit of detection and therefore increase sensitivity of ctDNA detection. To do this, new strategies combining the use of unique molecular identifiers (UMIs), to facilitate the identification of single DNA molecules from PCR duplicates, with deep sequencing are promising strategies to detect ctDNA [404, 405]. In addition, this strategy sequences a list of genes that can be customised allowing the detection of multiple genes, which can be valuable to track emergence of resistance alterations. In this context, Cell3 Target (Nonacus, oncology) offers a confident and sensitive calling of mutations down to 0.1% of VAF from as little as 10 ng ctDNA input by incorporating UMIs into targeted NGS customised gene panel. This methodology is currently been validated by a PhD student in the ICR, Reda Stankunaite, who has designed a paediatric panel for ctDNA detection. In particular, longitudinal samples from BIOMEDE-257 cfDNA derived from plasma, were run using the paediatric ctDNA panel. Mutations and VAF were able to be tracked over the course of the patient disease from plasma ctDNA.

I have shown the potential use of ctDNA to detect molecular alterations in CSF, cyst fluid and less often in plasma and serum. The implementation of ctDNA from plasma and CSF in routine clinical practice, still requires larger studies to validate these and previous studies [289, 315-317, 397]. Further work needs to be done evaluating the correlation of ctDNA levels with tumour stage, molecular entity and tumour location. In addition, the use of ctDNA as therapeutic biomarker to track tumour evolution in response to treatment represents an attractive strategy to identify potential resistance mechanisms to targeted therapies.

CHAPTER 6 : Molecular characterization and target identification in BIOMEDE, a co-clinical trial in DIPG

6.1 Introduction

The survival of children with DIPG remains dismal, and new treatments are desperately needed. Until recently the incorporation of molecular biomarkers into clinical practise for DIPG was not contemplated, partly due to the high risk associated with performing a biopsy, in turn limiting our understanding of the biology of the disease. As stereotactic biopsy in DIPG was shown to have a low morbidity [406, 407], tissue acquisition allowed for the use of multi-omics profiling techniques leading to the identifying of molecular alterations that could potentially be targeted by therapeutic agents [22, 76, 128, 408].

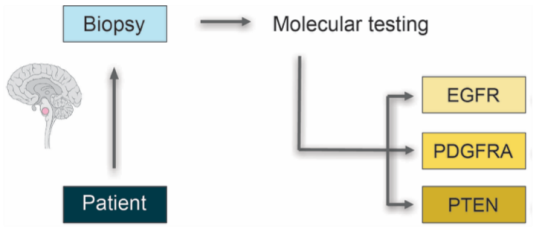
The emergence of patient-specific *in vitro* and *in vivo* models represents an excellent tool to guide clinical decision-making in the context of personalised medicine. In adult oncology, such models, also called “avatars”, are being utilised in co-clinical trials to predict patient response and understand chemosensitivity and chemoresistance mechanisms (NCT03170180, NCT03283527, NCT03890614, NCT03979170) [409-415]. In paediatric oncology, recent efforts to establish patient-derived tumour models have been made [416-418]. Such initiatives are providing an extraordinary resource to characterise cancer driver genes and evaluate treatment strategies by using high-throughput screening [419, 420]. But yet, the translational use of these models in the clinic for paediatric tumours is not as advanced as for adult cancers.

DIPG are a group of heterogeneous tumours highlighting the need for collecting frozen and live tissue at diagnosis. This will allow the profiling of treatment-naive tumours, and will enable the establishment of “avatars” that recapitulate intrinsic tumour sensitivity capturing the heterogeneity of the original tumour patient. This represents a unique opportunity to guide clinical decision-making at diagnosis, and additionally allows for the possibility of modelling resistance in parallel, in order to anticipate treatment strategies at relapse based upon the biology of the patient.

The utility of targeted agents in early-phase clinical trials for patients with DIPG is currently being explored. This includes the HDAC inhibitor, panobinostat (NCT02717455), the dual HDAC and PI3K inhibitor, fimepinostat (CUDC-907), PTC596, an inhibitor of the polycomb ring finger oncogene BMI1 (NCT03605550),

and the PI3K/AKT inhibitor GDC-0084 (NCT03696355). The use of GD2.CAR-T cells (NCT04099797) or the oncolytic adeno-virus DNX-2401 (NCT03178032) are also promising treatment strategies in DIPG patients [132, 140]. Furthermore, there are clinical trials in DIPG which as inclusion criteria require molecular profiling, such as the combination of dasatinib with everolimus in gliomas harbouring *PDGFRA* alterations (NCT03352427), or the dopamine receptor 2/3 antagonist, ONC201, in H3 K27M positive gliomas (NCT03416530) [130, 421]. Additionally, the clinical benefit of using molecular profiling to determine the specific targeted treatment is under evaluation in DIPG, including several clinical trials such as DIPG-BATS (NCT01182350) and PNOC008 (NCT03739372) [273, 318].

BIOMEDE is an adaptive, multicentre, phase-II clinical trial in newly diagnosed DIPG comparing the efficacy of three targeted agents, everolimus, dasatinib and erlotinib, in combination with radiotherapy (NCT02233049) **{Figure 6-1}**. The target and drug of choice for BIOMEDE were based upon previous studies which had identified *EGFR* gain/amplification or more frequently overexpression, activation of the mTOR pathway and *PDGFRA* gain or amplification in DIPG patients [76, 422-427]. In line with these findings, therapeutic agents against those targets have been evaluated in children with DIPG and adult glioblastoma including the drugs selected for evaluation in the BIOMEDE trial [422, 428, 429]. In particular, in a European study of the ITCC (Innovative Therapies for Children with Cancer) consortium, a trend towards an improvement in progression free survival and overall survival was observed in EGFR positive DIPG tumours treated with erlotinib compared to EGFR negative tumours [422].



Biomarker	EGFR Positive	EGFR Negative	EGFR Unknown
PTEN Loss	R3	R2	R3
PTEN Positive	R1	dasatinib no randomisation	R1
PTEN Unknown	R3	R2	R3

R1	Erlotinib vs Dasatinib
R2	Everolimus vs Dasatinib
R3	Erlotinib vs Everolimus vs Dasatinib

Figure 6-1 BIOMEDE clinical trial overview. BIOMEDE eligibility criteria for the different sub-trials in newly diagnose DIPG confirmed by central pathology review. Based on the overexpression of EGFR and PTEN loss detected by immunohistochemistry patients were randomised and stratified to study de efficacy of three investigational compounds: erlotinib an EGFR-inhibitor, everolimus an mTOR inhibitor and dasatinib a multi-kinase inhibitor.

Our laboratory was the assigned as the biology reference in the United Kingdom, receiving the tissue from the patient biopsy to undertake comprehensive molecular characterization. With this unique opportunity, we sought to explore genetic dependencies in DIPG patients in the context of a co-clinical trial, with the aim to identify rational therapeutic options using individualised preclinical evidence as to their efficacy. Throughout the course of my PhD, I have established novel patient-derived *in vitro* cultures from biopsy specimens of 11 patients, in both 2D (laminin matrix) and 3D (neurosphere) conditions, as well as orthotopic xenografts *in vivo*, with a high concordance in their molecular profile compared to the original tumour specimen. Cells were screened against a series of common and bespoke FDA approved drugs based upon previous evidence in DIPG and/or the specific molecular alterations detected in the original patient sample. In this chapter, I show the feasibility in generating patient-specific testable hypotheses that have been clinically translated in a subset of patients to guide therapeutic choice of DIPG patients at relapse in a prospective clinical trial in DIPG.

6.2 Results

6.2.1 Clinical trial samples received

Up to 14th December 2018 a total of 33 DIPG patients were consented to the in BIOMEDE-UK trial. Of these, 27 were randomised into one of the three trial treatments (dasatinib n=12, everolimus n=10 and erlotinib n=5). The remaining six patients were not allocated to any treatment; of those, four patients were not eligible and two patients withdrew consent. Tumour biopsy was collected from 13 different hospital sites including 23/27 fresh frozen tissue (FF) and 18/27 live tissue (Hibernate A n=7, cryopreserved n=10 and as cells n=1) **{Figure 6-2}**.

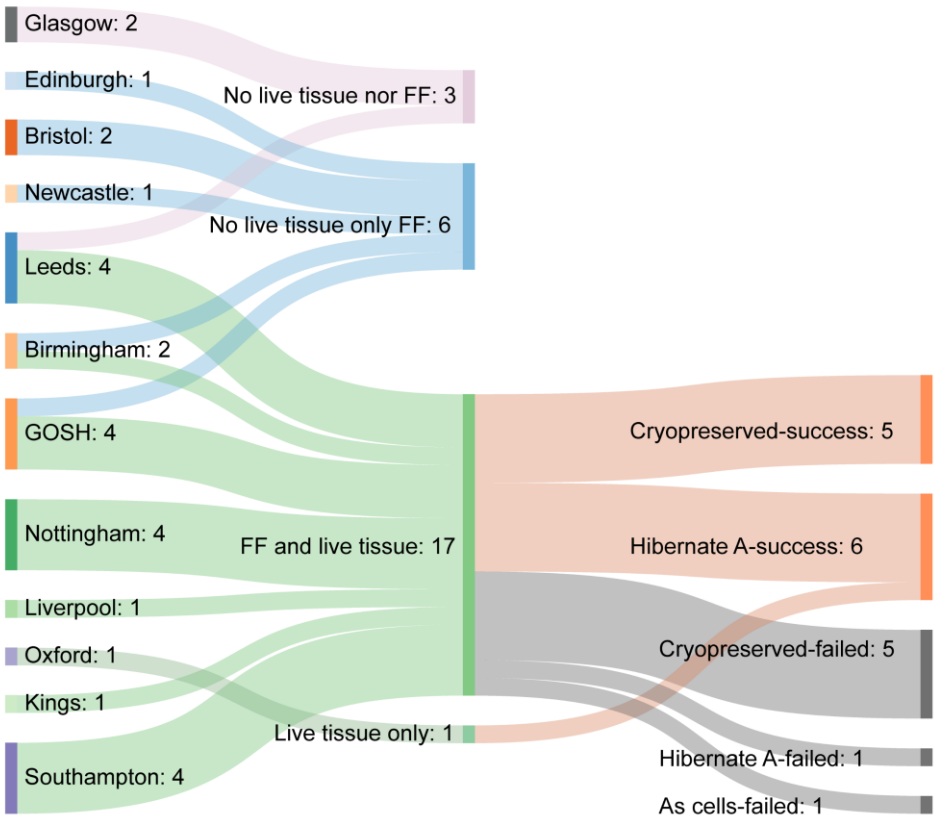


Figure 6-2 Sankey plot showing BIOMEDE-UK samples. Specimens were received as FF and/or live tissue (Hibernate A n=7, cryopreserved n=10 or as cells n=1) from 13 different centre-sites.

6.2.2 Clinical reporting in BIOMEDE-UK

Taking advantage of the technology developed in a CPA accredited laboratory, the Paeds-v2 panel was carried out on the 24 biopsy samples (from BIOMEDE-181 a cell pellet from p0 was used as fresh frozen sample was not available). The variants were reported to the treating clinician with a quick turnaround for clinical-decision making of BIOMEDE patients at relapse.

At least one alteration was detected in all the samples at an allele frequency ≥ 5%. K27M mutations in the histone H3 genes (20/24) (83%, H3.3 n=17 and H3.1 n=3) were the most common alterations, followed by mutations in *TP53* 19/24, (79%). *PDGFRA* alterations (6/24, 25%) included four amplifications, one of which harboured concomitant mutation (*PDGFRA*-N468S), another with a large in-frame insertion in exon 11 of 28 amino acids, and a final patient with a missense mutation (*PDGFRA*-N659K). In addition, alterations in the phosphatidylinositol 3-kinase (PI3K)/mammalian target of rapamycin (mTOR) (*PIK3CA* n=4, *PTN11* n=3, *PIK3R1* n=3) and MAP-kinase (*NF1* n=1 and *BRAF* n=1) pathways were common, as well as

alterations in chromatin modifiers (*ATRX* n=3, *ARID1A*/*B* n=2, *KMT2B/C* n=2 and *ASXL1* n=1) **{Figure 6-3A}**. For BIOMEDE-125 two biopsy samples were profiled, overlapping variants in *H3.3*-K27M, *PPM1D*-C478* and *ATRX*-D2136Y were found in both tissues, however only in one of the tissue samples a mutation in *PTEN* (Q171P) was identified with an allele frequency of 67%.

Potentially targetable alterations, classified in accordance to the tier-criteria established in Chapter 5, were identified in the 24 patients; of these, 79% (19/24) had more than one tier variant. As previously, no Tier 1 alterations were identified *(FDA recognised biomarker predictive of response to an FDA approved drug)*. Six BIOMEDE patients had Tier 2B alterations, all six involving *PDGFRA;* four of these harboured *PDGFRA* and *KIT* co-amplifications *(recognised standard of care predictive biomarker for drug in response another indication).* 18 patients presented Tier 3 alterations, the most common being *H3F3A*, *TP53* and *PIK3CA* (open clinical trial for predictive biomarker in paediatric solid tumours) **{Figure 6-3B}**. Recommendations for clinical decision making were reported to the treating clinician **{Table 6-1}**.

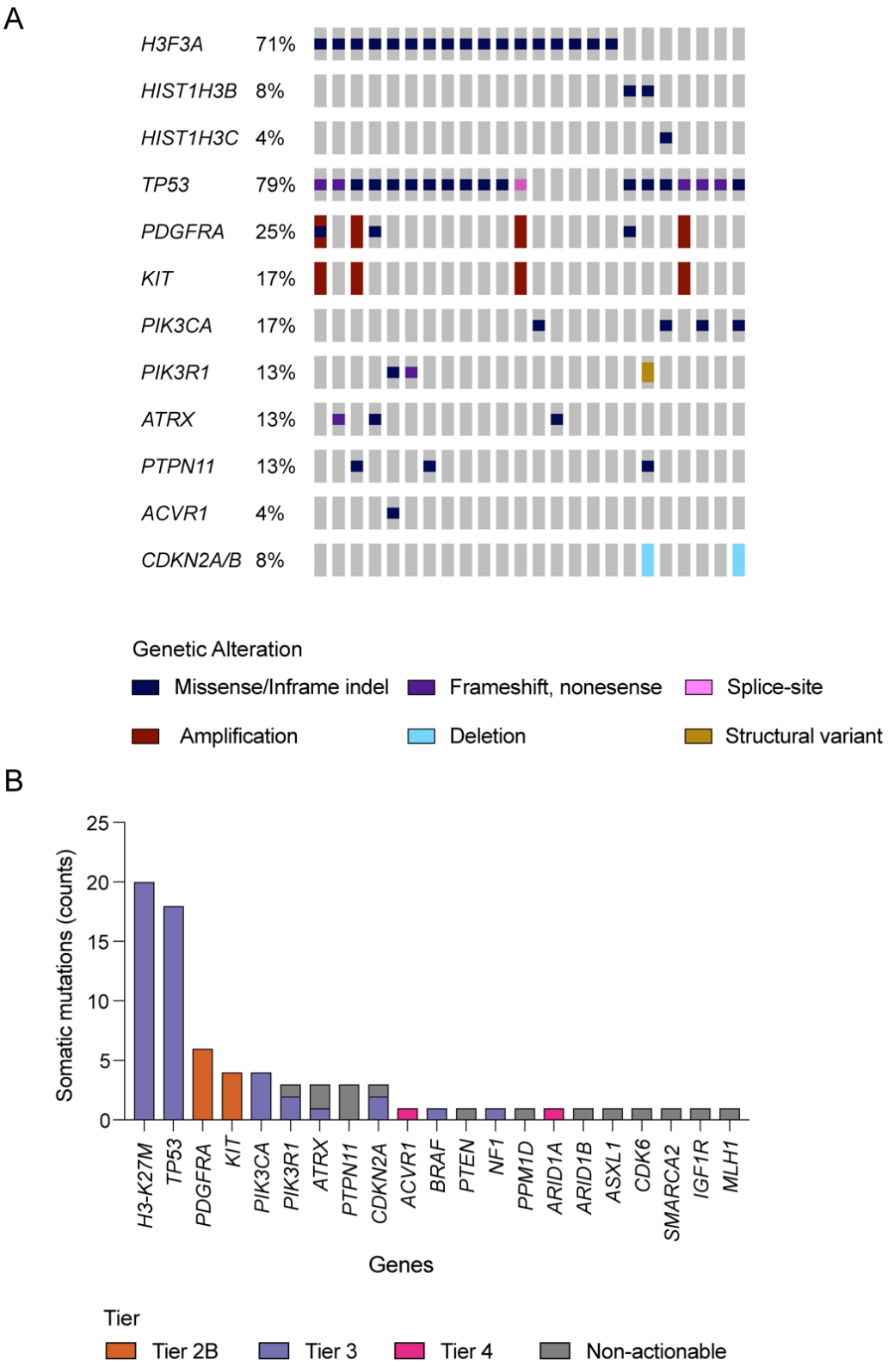


Figure 6-3 Molecular alterations summary of BIOMEDE patients. (A) Oncoprint representation for 24 DIPG patients profile using the Paeds-v2 panel. Samples are arranged by columns clustered by gene and coloured by genetic alteration. (B) Bar-plot of the somatic alterations coloured type of alteration. Tier 2B recognised standard of care predictive biomarker for response in another indication, Tier 3: open clinical trial for predictive biomarker for paediatric solid tumours and Tier 4: compelling biological evidence supports biomarker as being predictive of response to drug.

Genetic alterations	Targeted Agent	Clinical Trial	Patients	Tier
PDGFRA	Dasatinib Crenolanib	NCT03352427 NCT02626364	6	2B
H3.3/1 K27M	Panobinostat ONC201	NCT02717455 NCT03416530	20	3
TP53	AZD1775	NCT02813135	18	3
PIK3CA	AZD2014 LY3023414	NCT02813135 NCT03213678	4	3
CDKN2A	Ribociclib	NCT02813135	2	3
PIK3R1	AZD2014 LY3023414	NCT02813135 NCT03213678	2	3
ATRX	Olaparib or AZD1775	NCT02813135	1	3
BRAF	Trametinib Binimetinib	NCT03363217 NCT02285439	1	3
NF1	Trametinib Binimetinib	NCT03363217 NCT02285439	1	3
MYCN	BMS-986158 AZD1775	NCT03936465 NCT02813135	1	3
ARID1A	VX-970 AZD6738	NCT03718091 NCT04065269	1	3
ACVR1	Vandetanib + Everolimus	N/A	1	4

Table 6-1 Targetable genetic alterations in the BIOMEDE-UK cohort. Tier 2B, 3 and 4 identified in the BIOMEDE-UK cohort with a summary of the clinical recommendations and clinical trials associated. Tier 2B recognised standard of care predictive biomarker for response in another indication, Tier 3: open clinical trial for predictive biomarker for paediatric solid tumours and Tier 4: compelling biological evidence supports biomarker as being predictive of response to drug.

Furthermore, samples were subjected to Illumina methylation EPIC BeadArray (n=23). The Heidelberg brain tumour classifier on the methylation data was used to assign a molecular subgroup to each of the 23 cases. As expected, the majority of the tumours classified as diffuse midline glioma H3 K27M mutant (DMG-K27, n=19), with 16 harbouring K27M mutations in the histone genes (*H3F3A* n=15 and *HIST1H3C* n=1) and three histone wild-type. Three cases classified as glioblastoma MYCN-subgroup (GBM_MYCN), of these, BIOMEDE-118 presented a co-amplification of *MYCN/ID2* and surprisingly the other two, BIOMEDE-182 and BIOMEDE-184, were K27M positive (*HIST1H3B* and *HIST1H3C* respectively). BIOMEDE-128 classified as glioblastoma, subclass RTKIII (GBM_RTKIII) with a low methylation score (0.11).

6.2.3 Establishment of primary patient-derived DIPG models

Tissue from newly diagnosed DIPG samples was used to generate patient-derived primary cell cultures and xenografts (PDX) by implantation of the dissociated tissue into the pons of the mice. Established models were subjected to comprehensive sequencing as well as a personalised drug screen **{Figure 6-4}**.

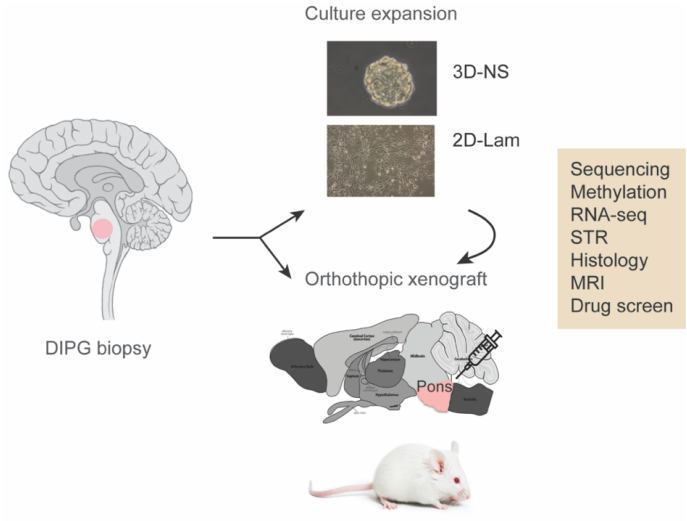


Figure 6-4 BIOMEDE-UK translational-research schematic illustration. The figure shows the generation and characterisation of patient-derived cell cultures (3D as neurospheres and 2D as laminin) and patient derived xenograft (PDX) from DIPG patients. Comprehensive models characterisation was performed including panel and exome sequencing, methylation arrays, RNA-sequencing, STR profiling, histology, magnetic resonance imaging of *in vivo* models and *in vitro* drug screen.

6.2.3.1 *In vitro* models

In vitro models were generated from 61% (11/18) live tissue samples including 6/7 (86%) received in Hibernate A media and 5/10 (50%) as cryopreserved tissue. One tissue sample received in Hibernate A, one received as cells and five as cryopreserved tissue failed to grow. From 64% (7/11) samples, cultures were initiated in two conditions: adherently on a laminin matrix (2D), and in suspension growing as neurospheres (3D) **{Figure 6-5}** and **{Figure 6-6}.**The remaining four samples only generated successful initial cultures under 2D conditions. Subsequent attempts to establish these in 3D after expansion on laminin allowed for some loose NS structures to form (BIOMEDE-117 and BIOMEDE-121) **{Figure 6-6A-B}**, however only BIOMEDE-186 was successfully maintained in 3D **{Figure 6-6H}**. The seven successfully established 3D cultures consisted of tight uniform spherical neurospheres {**Figure 6-6D-G}** and **{Figure 6-6I-K}**. All established cultures were mycoplasma negative and authenticity was confirmed by STR profiling and SNP genotype from methylation arrays **{Appendix Table 4}** and **{Appendix Figure 1}**.

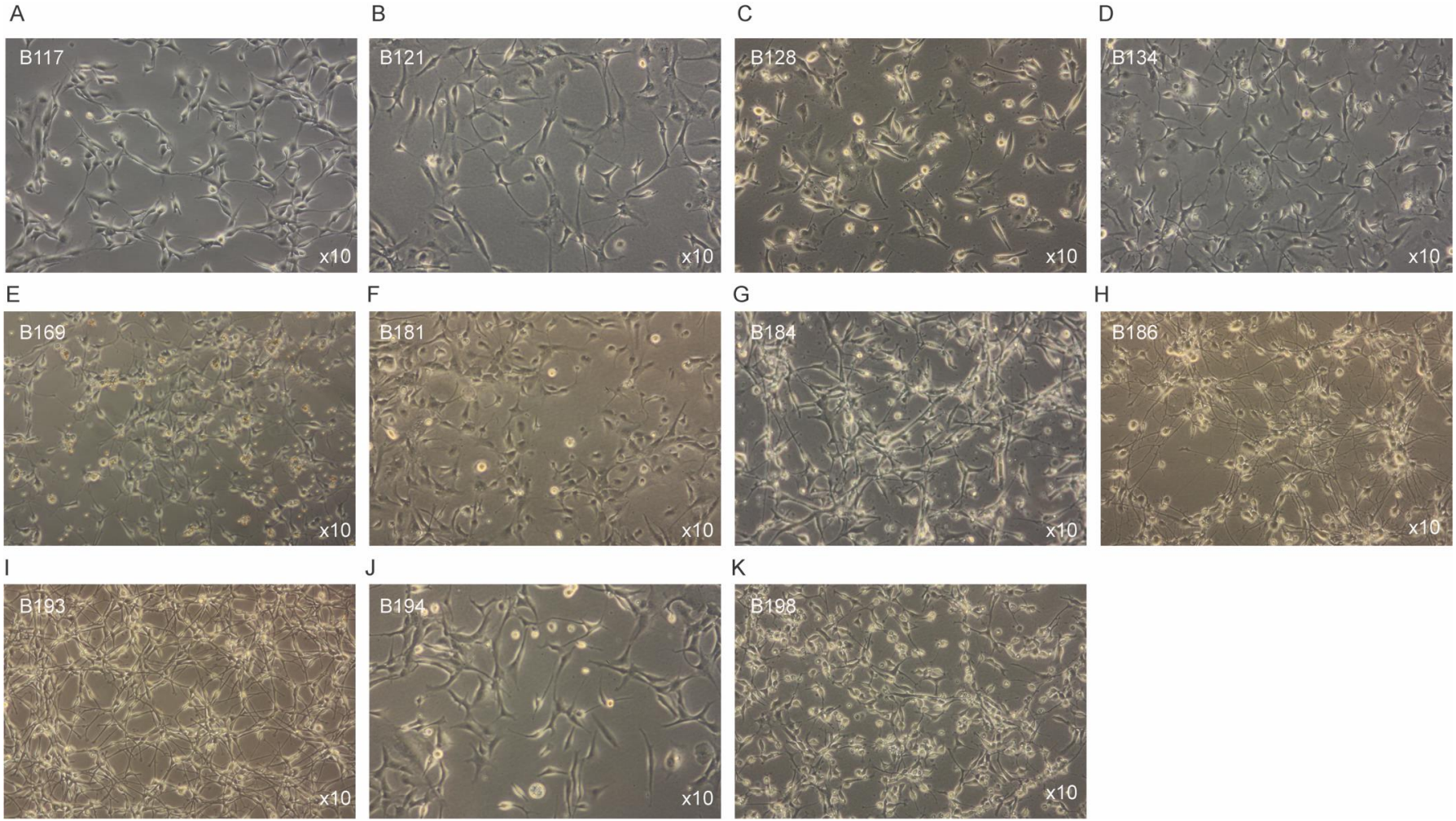


Figure 6-5 Light microscopy images of eleven primary DIPG cultures derived from BIOMEDE-UK patients grown adherently into laminin coated flasks.

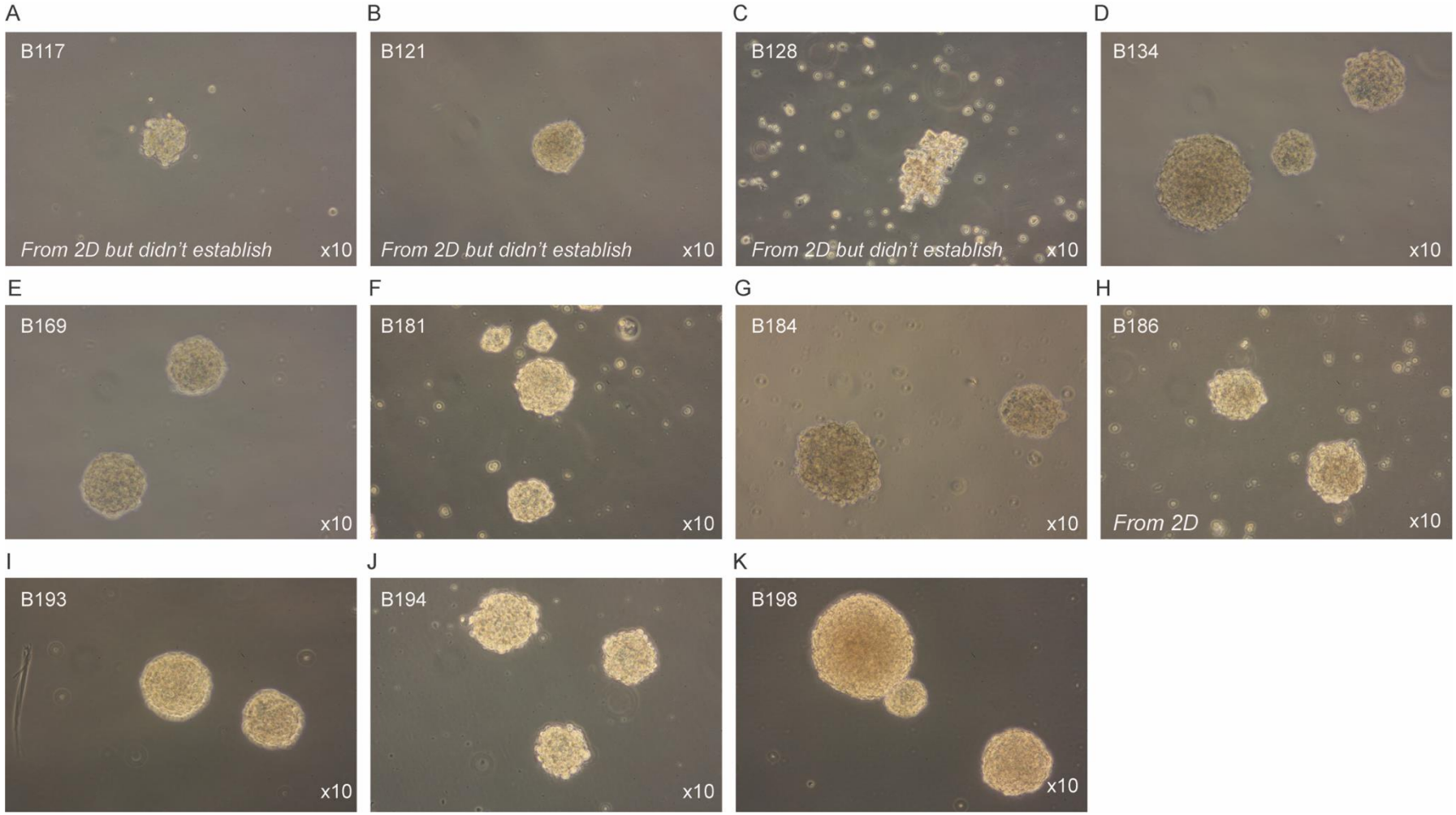


Figure 6-6 Light microscopy images of eleven primary DIPG cultures derived from BIOMEDE-UK patients grown in suspension as neurosphere (3D). Seven cultures were originally originated as 3D from live tissue and four from monolayer culture (2D) from which three failed to be established as 3D.

Doubling times ranged from 3 to 10 days, and the optimal number of cells seeded to perform drug assays varied with a range of 3000-10,000 cells/well **{Table 6-2}** and **{Figure 6-7}** and **{Figure 6-8}**. Factors influencing the number of cells seeded per well to obtain, at end-point assay, 90% confluency for 2D or 200-300 μM diameter for NS were doubling times/proliferation rate and space the cells occupied in the well in 2D. As an example of the former, BIOMEDE-193 3D (doubling time of 6-8 days) required 10,000 cells/well, compared to BIOMEDE-181 3D (doubling time 3-4 days) which required 4000 cells/well. For the latter, BIOMEDE-169 2D and BIOMEDE-181 2D had the same doubling time (3 days) however required a different number of cells (5000-6000 and 3000-4000 cells/well, respectively) as the BIOMEDE-169 cells of in monolayer were substantially smaller than for BIOMEDE-181 **{Figure 6-5E-F}**.

Sample ID	Condition	Cell density (cells/well)	Proliferation rate (days)
BIOMEDE-117	2D	5000	3-4
BIOMEDE-121	2D	8000	4-6
BIOMEDE-128	2D	4000-5000	4
BIOMEDE-134	2D	5000	3-4
	3D	5000	4-5
BIOMEDE-169	2D	5000-6000	3
	3D	3000	4
BIOMEDE-181	2D	3000-4000	3
	3D	4000	3-4
BIOMEDE-184	2D	3000	3
	3D	3000-5000	5
BIOMEDE-186	3D	4000-5000	6-10
BIOMEDE-193	2D	8000-10000	6
	3D	10000	6-8
BIOMEDE-194	2D	8000-10000	3-4
	3D	10000	5-6
BIOMEDE-198	2D	10000	4
	3D	8000-10000	6-7

Table 6-2 BIOMEDE-UK derived cultures summary. Table indicates if the models were successfully established as 2D and/or 3D, number of cells required to be seeded per well as well as proliferation rate in days.

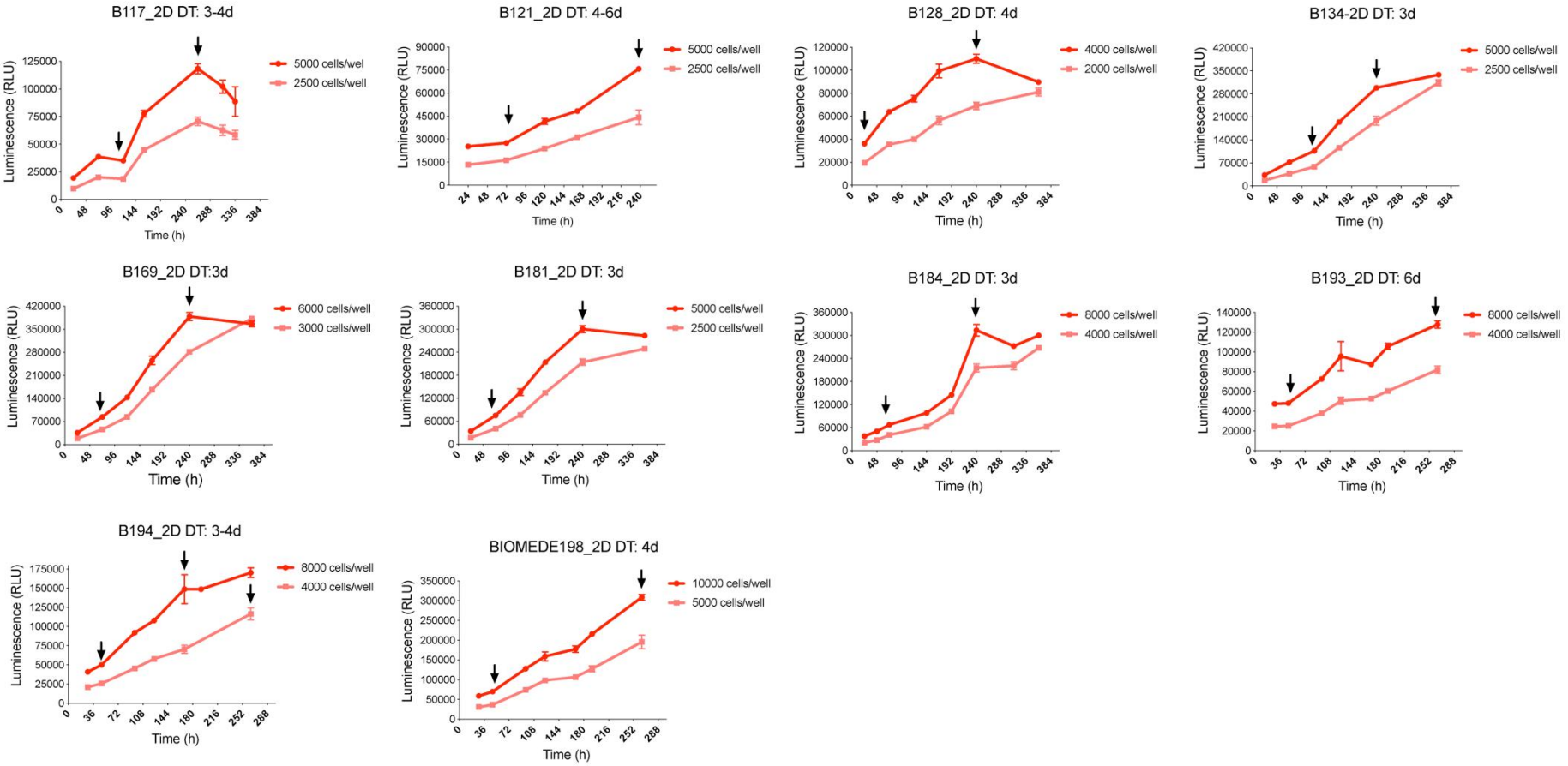


Figure 6-7 Primary DIPG cultures growth in 2D from BIOMEDE-UK patients. Doubling times (DT) were calculated at two different densities (replicates of 6) on 96-well plates, cell viability was measured using CellTiter Glo. Arrows represent the time-points, during the phase of growth was selected to estimate the DT using the formula: ***PDT = t ln2/ln (Xe/Xb),*** where t is the incubation time in hours, xb luminescence at early incubation time and xe luminescence at late time before the stationary phase.

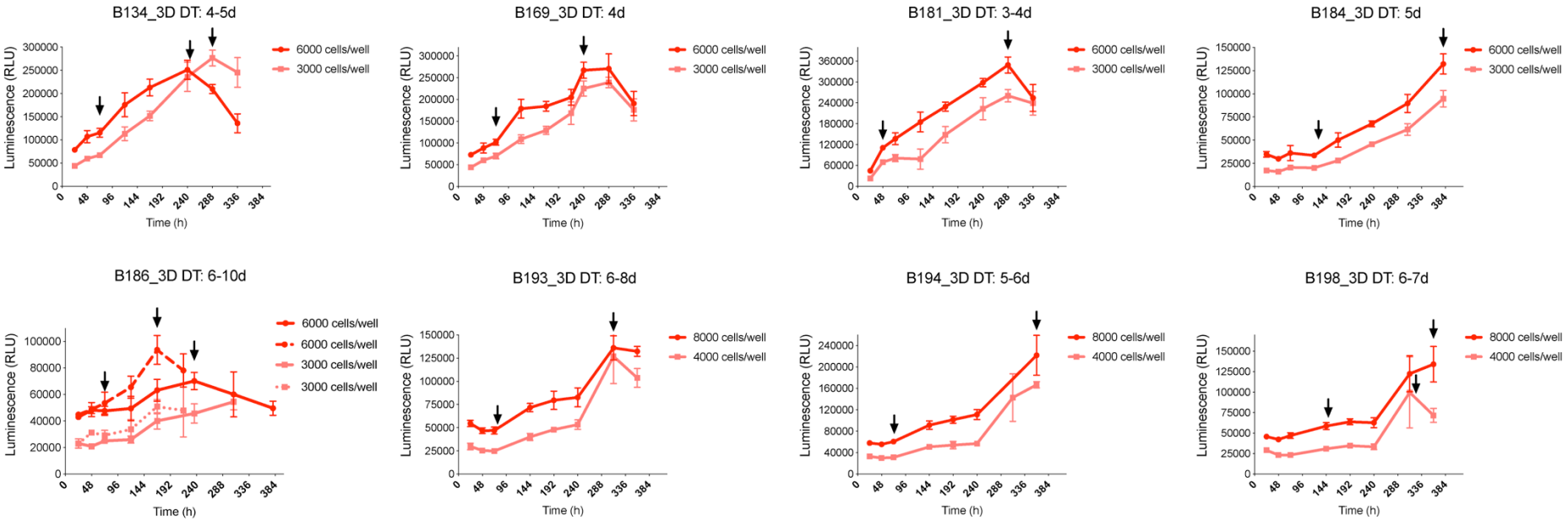


Figure 6-8 Primary DIPG cultures growth in 3D from BIOMEDE-UK patients. Doubling times (DT) were calculated at two different densities (replicates of 6) on 96-well plates, cell viability was measured using CellTiter Glo. Arrows represent the time-points, during the phase of growth was selected to estimate the DT using the formula: ***PDT = t ln2/ln (Xe/Xb),*** where t is the incubation time in hours, xb luminesce at early incubation time and xe luminescence at late time before the stationary phase.

6.2.3.2 *In vivo models*

Direct orthotopic injection of live tissue samples into the pontine region of mouse brains was carried out by Dr Diana Carvalho, in our lab. Successful PDX establishment was achieved for 64% (9/14) samples implanted in NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ (NSG) mice within 12 months, the limit of our Home Office project licence at the time. Tumour engraftment rate was 100% (5/5) from tissue received in Hibernate A, and 44% (4/9) for tissue received cryopreserved. For BIOMEDE-134 and BIOMEDE-169 there was insufficient tissue to establish PDX, therefore cells were first expanded *in vitro* in order to generate cell-derived xenografts (CDX). The mice were monitored for the presence of tumour by magnetic resonance imaging (MRI) by Dr Jessica Boult, a Senior Scientific Officer in the Division of Radiotherapy and Imaging. Overall mice survival varied depending on the models with a range of 83 to 225 days **{Figure 6-9}**.

Histologically the PDX and CDX were highly heterogeneous **{Figure 6-10}**. Most of the models presented high cellularity with the exception of BIOMEDE-134 CDX which did not present neurological symptoms nor presence of tumour by MRI; as per our license guidelines, the experiment had to be terminated at day 364. BIOMEDE-134 was a hypocellular tumour with extensive infiltration to the cerebral hemispheres and showed presence of satellitosis **{Figure 6-10D}**. This model had a long latency, hence if the experiment had continued a higher cellularity might have been observed. Some models such as BIOMEDE-128, BIOMEDE-193, BIOMEDE-198 and BIOMEDE-184 showed an extensive infiltration throughout the brain with high number of tumour cells present in the cerebral fibre tracts and the hindbrain, but also tracks of cells invading to more distant areas such as the midbrain, the hypothalamus, olfactory bulb and periventricular cortex were observed. Conversely, BIOMEDE-118 and BIOMEDE-181 presented more confined tumours into the pons and cerebellum with clearly defined margins.

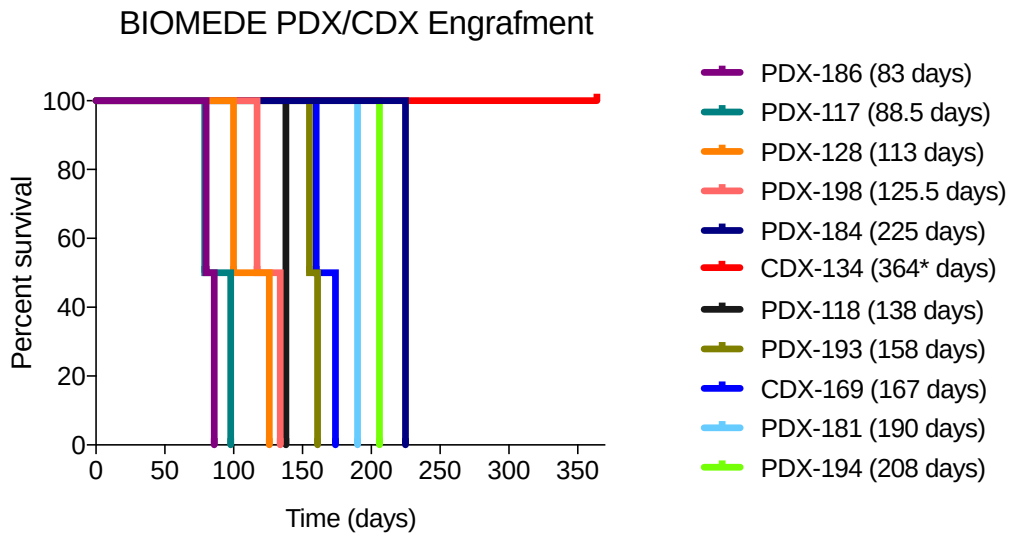


Figure 6-9 Overall survival for the patient derived xenografts (PDX) and cell derived xenograft (CDX). Engrafted models are coloured per patient from whom they were derived. *CDX-134 were not culled at endpoint because e experiment had to be terminated at day 364 as this was the end-point time in the licencse protocol.

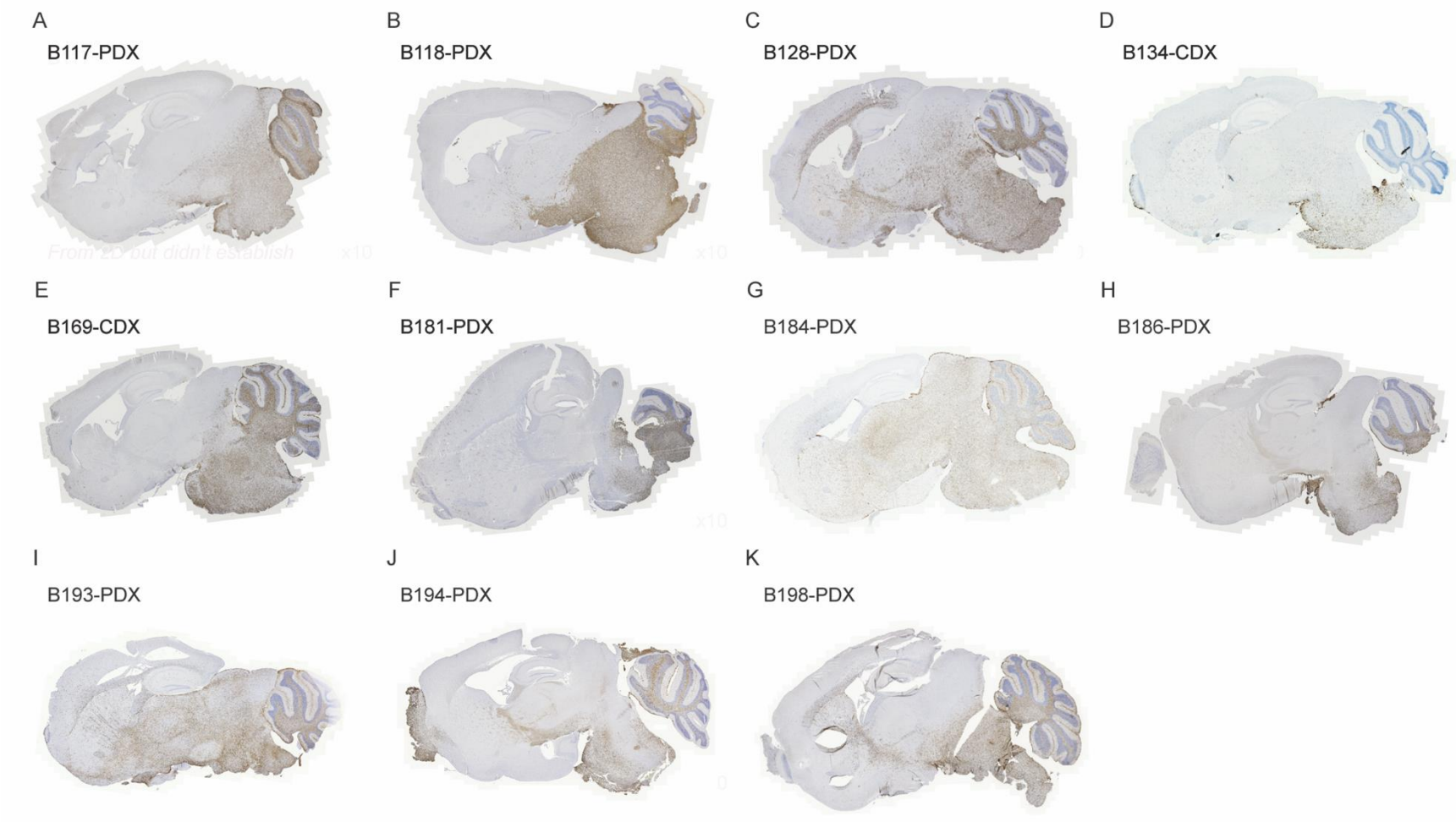


Figure 6-10 Immunohistochemical staining for anti-human nuclei antibody (HNA) of each of the orthotopic patient or cells derived xenograft bearing tumours from BIOMEDE-UK patients.

6.2.4 Credentialing of models

In order to assess how similar our models were to the original tumour sample from which they were derived, DNA and RNA was isolated from cell pellets and PDX/CDX, and targeted sequencing using the pHGG-panel, Illumina EPIC methylation array analysis, and RNA-seq was performed. The Heidelberg methylation classifier v11b4 was used to assign scores and a tSNE projection with a pan-glioma reference set of 1652 brain tumour samples were used.

Established cell cultures as well as PDX/CDX models clustered closely with the methylation classification of the original tumour sample **{Figure 6-11}.** Nine models classified as DMG-K27, seven of which harboured H3.3-K27M mutations and two were histone wild-type. BIOMEDE-193 had a low DMG-K27 classification score (tumour = 0.52, PDX = 0.39, 2D cells = 0.13 and 3D cells = 0.09), the tumour and the models clustered as GBM_MYCN in the tSNE projection. In the same way BIOMEDE-128 classified poorly with all samples most closely corresponding to a paediatric GBM_RTK/MYCN like group. BIOMEDE-118-PDX harbouring MYCN/ID2 amplification classified and clustered as the original tumour sample corresponding to GBM_MYCN. BIOMEDE-184 models (2D-cells and PDX) classified and mapped in the same cluster as their original tumour, corresponding to the GBM_MYCN subgroup, however no *MYCN* amplification was present in these samples.

The mutational profiles were also largely concordant, with models retaining the key characteristic genetic alterations present in the tumours **{Figure 6-12}**. There were nonetheless some discrepancies. BIOMEDE-134 tumour harboured a *PDGFRA* amplification, however none of the independent established cultures as 2D-Lam or 3D-NS presented the gene amplification **{Figure 6-12E}**. *PDGFRA* status in the BIOMEDE-134 FFPE sample was also assessed by the central trial laboratory in Paris using FISH, identifying 43/100 cells to be *PDGFRA* amplified. Similarly, BIOMEDE-186 tumour harboured an amplification as well as a mutation in *PDGFRA* which was observed in the 2D culture and the PDX but not in the 3D-NS culture. In addition, I identified a *MYCN* amplification in 3D-NS culture in this sample which was not present in the original tumour sample nor the PDX nor the 2D-Lam, at least at the limit of detection of the assay **{Figure 6-12I}**. Interestingly, the 3D-NS culture was established from 2D-Lam, which harboured the *PDGFRA* amplification and did not carry the *MYCN* amplification. In addition, later-passage sequencing was performed to assess tumour evolution and potential genomic changes that could take place

under culturing conditions. For 9/11 models, no discrepancies were observed in the mutations reported from the tissue biopsy by panel sequencing between early (~passage 5) and later passage (~passage 11-25). BIOMEDE-184 2D did not harbour the *PIK3R1*-N564D nor the *SF3B2*-R330G mutations at later passage (passage 12) **{Figure 6-12G}** and *NF1*-E78fs was not identified in BIOMEDE-184 2D (passage 5, passage 14) nor 3D (passage 14) **{Figure 1-12H}**. These three alterations appear to be sub-clonal in the patient samples with a VAF of 18%, 17% and 20% respectively. By contrast, a missense mutation in *NF1* emerged in BIOMEDE-184 3D (I1824S) found at passage 5 at VAF of 65 % and at passage 14 at a VAF of 75%.

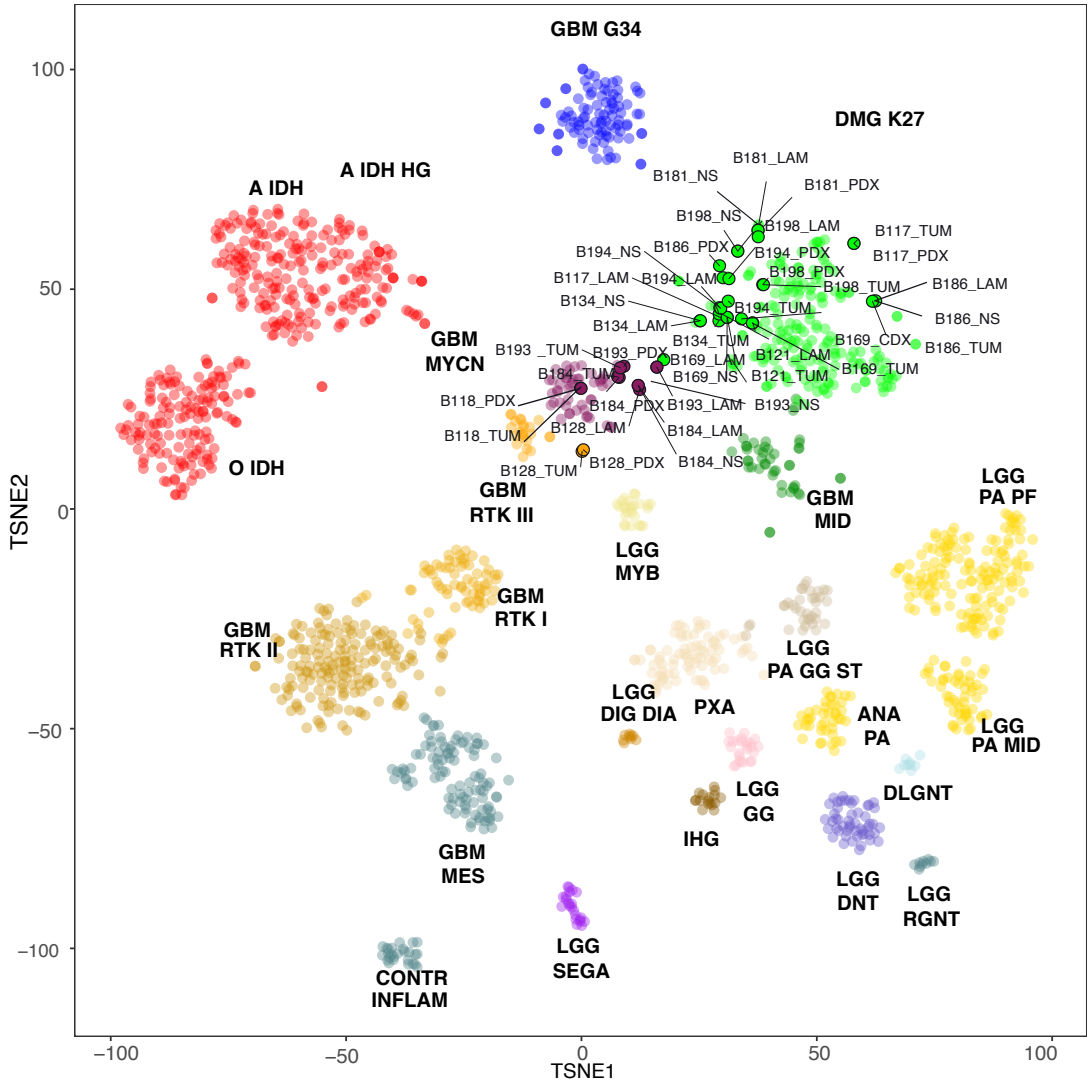


Figure 6-11 t-SNE projection of BIOMEDE-UK models t-static based stochastic neighbour embedding (t-SNE) projection of a combined methylation dataset comprising of the tumour sample, *in vivo* models (PDX/CDX) and *in vitro* models (2D-cells as LAM and 3D shown as NS) (circled) plus a reference set of glioma sub-types (n=1652). The first two projections are plotted on the x and y axes, with samples represented by dots coloured by sub-type according to the key provided.

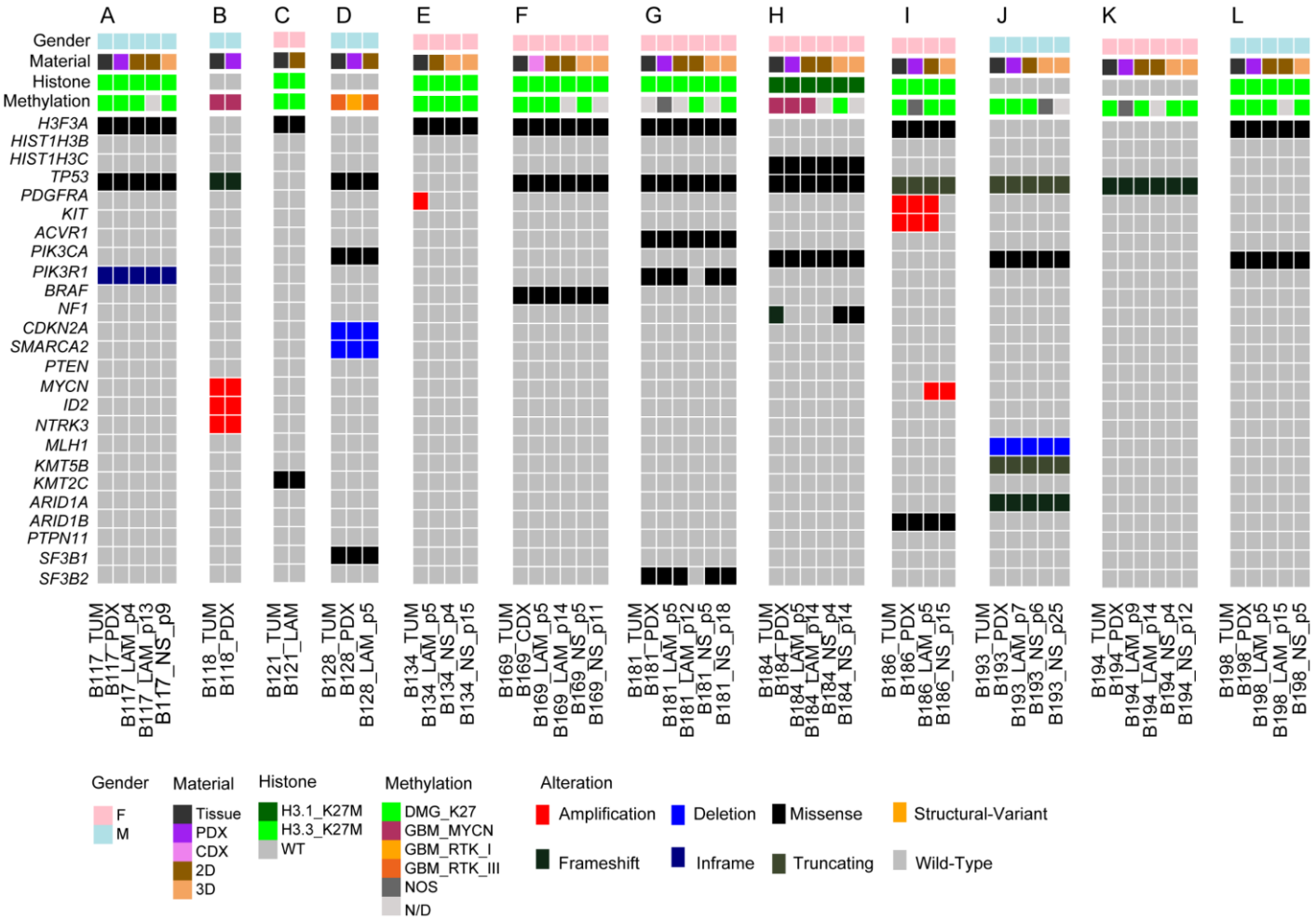


Figure 6-12 Oncoprint of BIOMEDE-UK models. Representation of an integrated annotation of targeted and whole exome sequencing, methylation arrays. Single nucleotide variants, indels, copy number alterations and structural variants are shown. Original tumour and matched models established are arranged by columns and coloured by genetic alteration. Clinicopathological, type of sample, molecular annotations are provided as bars according to the included key.

6.2.5 In vitro drug screening

From the available models established *in vitro*, 17 cultures from 11 patients were subjected to screening against a range of targeted inhibitors, including six models as 3D and 2D (n=12), four as 2D-only and one as 3D-only. The screening was comprised of six common drugs across the cultures (erlotinib, dasatinib, everolimus, panobinostat, olaparib and crizotinib) as well as six compounds chosen targeting molecular alterations identified in the original tumour sample. The assay was performed as described in the summary workflow **{Figure 6-13}**. All drugs were tested up to 10 µM except for panobinostat, which had a maximum dose of 1 µM, and olaparib which was tested up to 30 µM. A heatmap summarising the response of the cultures to each of the common compounds tested was generated **{Figure 6-14}** as well as individual drug response curves per patient for all the compounds **{Appendix Figure 2}**. These data were reported to the UK trial PI for feedback to the treating physician at relapse.

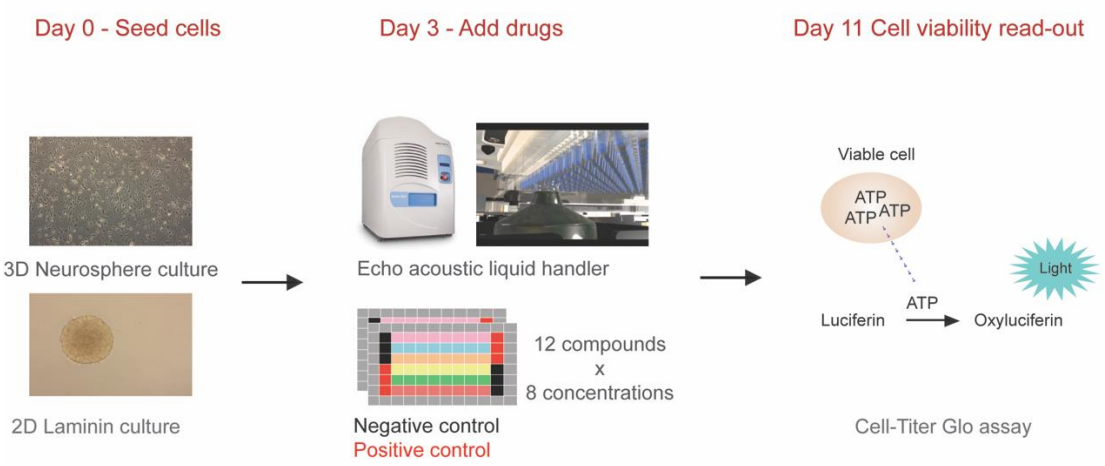


Figure 6-13 BIOMEDE-UK drug screen workflow. 2D and/or 3D cells were seeded in 96-well plates according to the cell densities previously estimated (***In vitro* models 1.2.3.1)**, three days later cells were screened against twelve compounds at eight different concentrations. Drug-plates were prepared using the Echo acoustic liquid handler. At day eleven cell viability was measured with Cell-Titer-Glo.

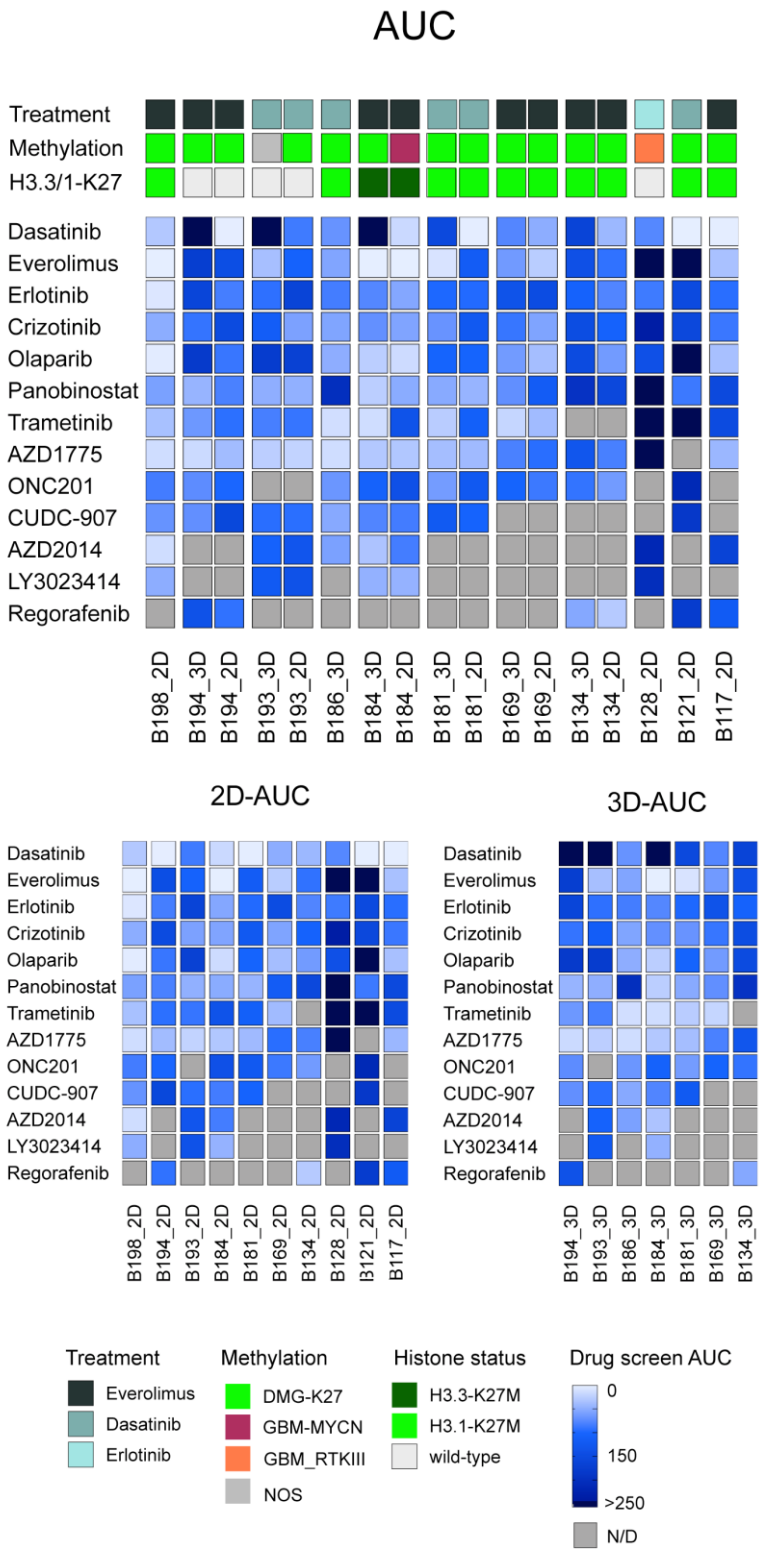


Figure 6-14 Heatmap representation of drug response from the drug screen in the DIPG cultures derived from BIOMEDE-UK patients. The drug screen was performed in 17 cultures from 11 patients including ten as 2D and seven as 3D versus the 13 most common drugs used in the screening. Activity scores are based on AUC normalised values, each drug-response was normalised to the mean of the corresponding drug for all the cultures. Treatment allocation, methylation and histone status are provided coloured according to the key.

Of the drugs used clinically in the BIOMEDE trial, only BIOMEDE-198 showed any degree of sensitivity to erlotinib (GI50 of 0.9110 µM compared to an average GI50 of 9.602 µM for the rest of the cultures), however this patient was randomised to everolimus **{Figure 6-15A}**. BIOMEDE-128 was the only patient from which a model was established and was randomised to erlotinib, for this patient, the GI50 was above the highest concentration tested (10 µM). The two samples which were most sensitive to everolimus, BIOMEDE-184 (2D and 3D) and 198 2D (GI50 of 0.9776 µM, 0.0305 µM and <0.0083 µM respectively) **{Figure 6-15B}** did receive everolimus in the trial. There was a differential degree of sensitivity for dasatinib in the models, in which BIOMEDE-117 BIOMEDE-2D, BIOMEDE-121 BIOMEDE-2D, BIOMEDE-181 2D and BIOMEDE-194 2D (GI50 of 0.1127µM, 0.0255µM, 0.09169µM and 0.5569µM, respectively) showed a better response than the rest of the cultures (mean GI50-sensitive cultures 0.07140 µM versus GI50-insensitive cultures 6.947 µM, p-value=0.0002 unpaired t-test) **{Figure 6-15C}**.

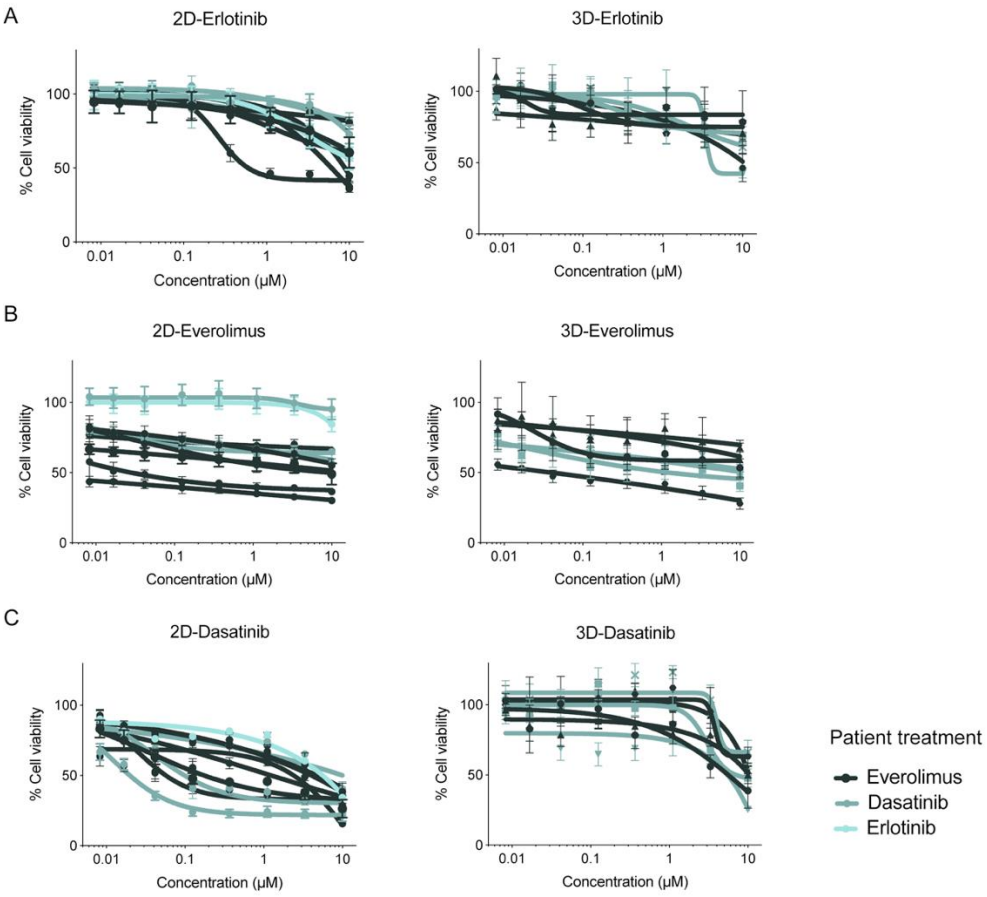


Figure 6-15 Dose-response curves to the BIOMEDE trial drugs. The curves show surviving fraction on y-axes and drug range 0-10 µM on x-axes. The curves across the models established, are separated by drug (A) erlotinib (B), everolimus and (C) dasatinib (and by condition (2D left and 3D right).

Out of the four dasatinib-sensitive cultures, BIOMEDE-121 and BIOMEDE-181 received dasatinib in the trial. Dasatinib-sensitive cultures were characterise by a distinct RNA gene expression pattern with increased gene expression of COL51, COL14A1, ANXA1, FGF7, CTGF, DEXI, KCNJ6, THSB1, and down-regulation of OLIG1/2, SOX1, C2CD4A/B, THSB4 amongst others **{Figure 6-16}**. Comparing my GI50 values with the Genomics of Drug Sensitivity in Cancer (GSDC2) dataset generated at the Wellcome Sanger Institute [430], I observed that the BIOMEDE sensitive cultures appeared at the more sensitive end of the spectrum of response across 760 cell lines from 50 cancer types **{Figure 6-17}.**

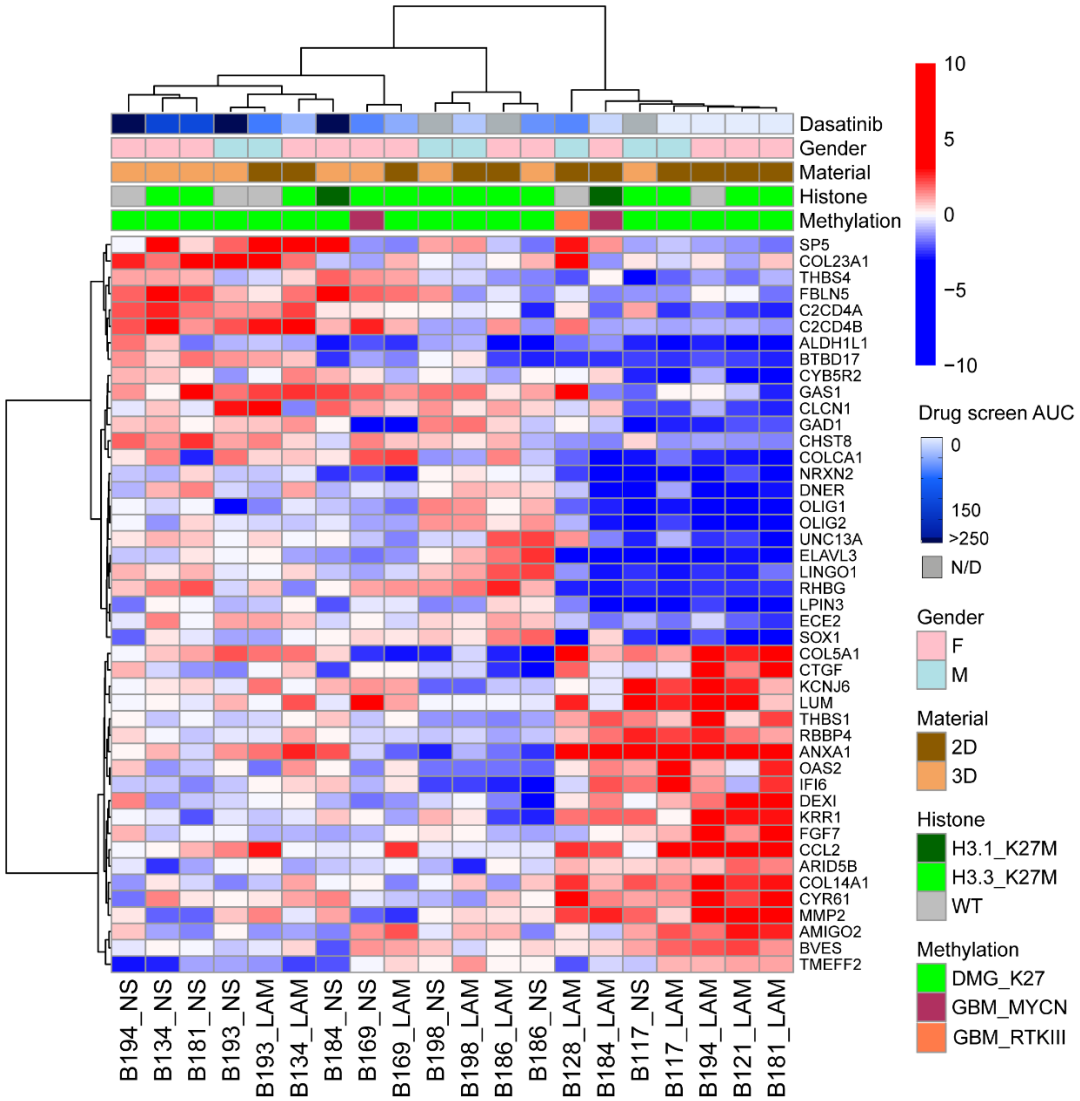


Figure 6-16 Differentially gene expression in dasatinib sensitive versus dasatinib resistant BIOMEDE-UK cultures. Dasatinib sensitivity, gender, material (2D and 3D), histone status and methylation subclass are provided as bars according to the including key.

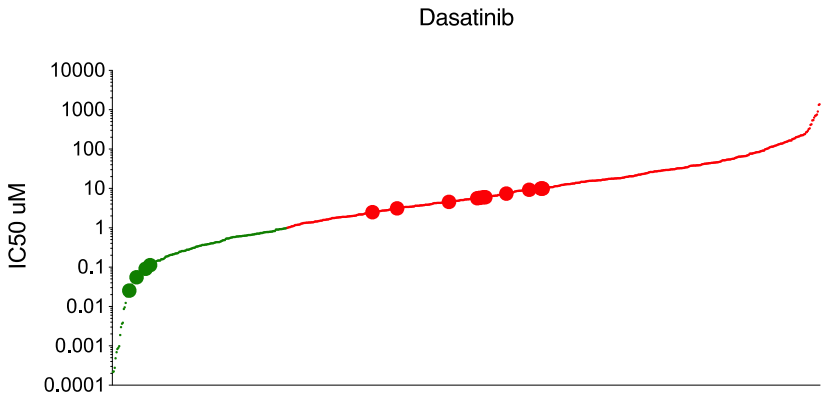


Figure 6-17 Integration of cell line GI50 values ranked by sensitivity of the Wellcome Sanger Institute (n=760) and the BIOMEDE cohort (n=17). Dots are coloured by sensitivity, green being sensitive and red resistant. Bigger size dots represent the BIOMEDE samples.

6.2.5.1 *TP53* truncating mutations confer sensitivity to AZD1775 in DIPG *in vitro* models

TP53 mutations were found in 8/11 (73%) models - five missense mutations, two nonsense and one frameshift **{Figure 6-18A}**. *TP53* gene expression was significantly reduced in tumours and corresponding models harbouring stop-coding and frameshift mutations compared to wild-type and missense mutation samples **{Figure 6-18C}**. These cultures (stop-coding: BIOMEDE-186 3D and BIOMEDE-193 2D and 3D; frameshift: BIOMEDE-194 2D and 3D) showed a high degree of *in vitro* sensitivity to the tyrosine kinase WEE1 inhibitor AZD1775 compared to the wild-type (GI50 of 0.1194-0.2128 µM vs GI50 of 0.2597-9.5780 µM, excluding BIOMEDE-198, p=0.0015, unpaired two-tailed Mann-Whitney test) {**Figure 6-18B}.** Of note, one *TP53* wild-type culture (BIOMEDE-198 3D) was also sensitive (GI50 of 0.1345µM). BIOMEDE-193 PDX (*TP53*-R146*) was used to evaluate the efficacy of AZD1775 *in vivo* by serial xenografting a p0 PDX into the pons of NSG mice (n=4 vehicle controls, n=4 treated). A tolerability study was performed in healthy NSG animals to determine the optimal dose for treatment. Animals were treated continuously for 14 days (60 mg/Kg, PO, q.d.) and no signs of toxicity were observed. We started treating animals 103 days post-implantation with 60 mg/kg of AZD1775 once daily for 6.1 weeks (2 weeks on, 1 week off, 5 days on, 2 days off. Although not significant, a small extension of survival in the treatment arm was observed (155 days versus 160 days, p=0.2129 long-rank test) **{Figure 6-18D}**.

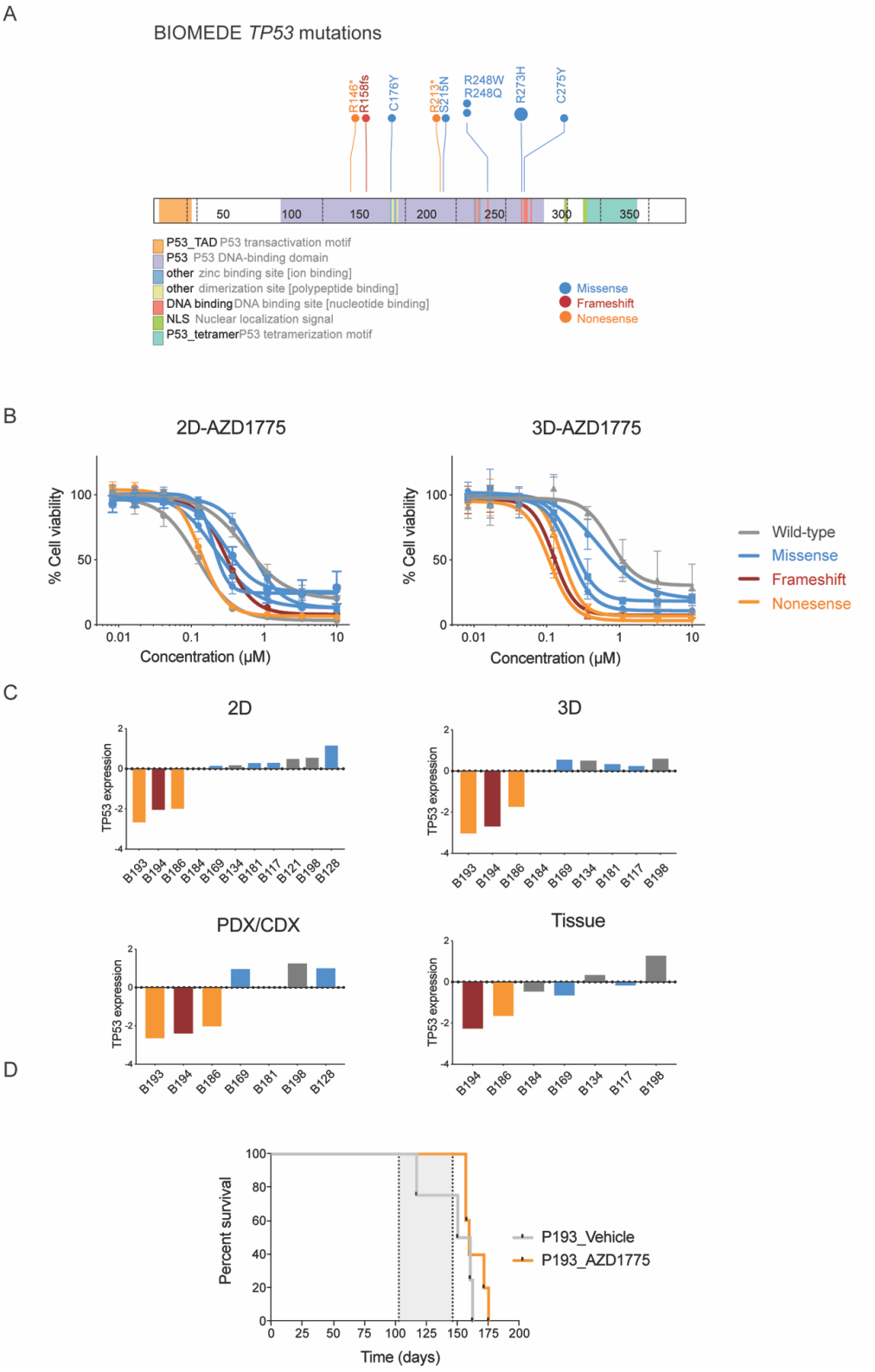


Figure 6-18 *TP53* truncating mutations conferring sensitivity to WEE1-inhibitor AZD1775 in DIPG derived cells from BIOMEDE-UK patients. (A) Lollipop showing mutations present in the BIOMEDE models. (B) Dose-response curves to AZD1775 separated by 2D and 3D and coloured per *TP53* mutation status. (C) TP53 ranked expression bar plots for 2D, 3D, PDX/CDX and tissue samples coloured by *TP53* mutation status (D) Survival for BIOMEDE-193 PDX in NSG (n=4 vehicle and n=4 AZD1775) (p-value = 0.2129, log-rank Mantel Cox test).

6.3 Discussion

There is an urgent need to find effective treatments and to predict response for children with DIPG in order to improve clinical outcomes for this disease. In this chapter, I report the use of real time molecular profiling of newly diagnosed DIPG patients from the BIOMEDE clinical trial. By using a CPA-accredited targeted panel, potential targetable alterations for clinical decision making of DIPG patients at relapse were reported. The results revealed alterations in genes for which targeted agents are available in early phase clinical trials or with pre-clinical evidence [41, 128, 130, 270, 431]. Moreover, from limited tumour material faithful *in vitro* and *in vivo* models were established. The models retained the key driver alterations, resembling the (epi)-genetics of the original patient tumour sample. This allowed us to identify rational therapeutic options by screening the patient-derived primary cells in 2D and 3D against a series of common and bespoke FDA-approved drugs.

A recent study published by A. Tauziède-Espariat and others, in a small number of *MYCN/ID2* co-amplified HGG-MYCN, suggested these tumours represent a different entity with a dismal overall survival, even shorter than H3 K27M [432]. Furthermore, they preserved the expression of H3K27me3 and presented distinct clinico-radiological and phenotypic features compared to classical DIPG tumours. In this context, BIOMEDE-118 tumour cells harboured co-amplification of *MYCN* and *ID2* and was classified as HGG-MYCN by methylation, indicating that this patient might not represent a 'true' DIPG and therefore the inclusion in the trial might be questionable. Moreover, BIOMEDE-128 harboured *CDKN2A/B* and *SMARCA2* deletions, atypical alterations in DIPG, and classified as GBM_pedRTK by methylation. BIOMEDE-193 and BIOMEDE-194 were histone wild-type, however they classified as H3 K27M by methylation and the pathology reports confirmed H3K27me3 loss by IHC. Interestingly, Castel and colleagues have recently reported a new subgroup of diffuse midline glioma lacking H3 K27M mutation but presenting loss of H3K27me3 and EZHIP overexpression associated with dismal prognosis [433]. In this context, overexpression of EZHIP was seen in BIOMEDE-193 but not in BIOMEDE-194. Conversely, BIOMEDE-182 (*HIST1H3B* K27M) and BIOMEDE-184 (*HIST1H3C* K27M) classified as HGG-MYCN by methylation. Unpublished data from our laboratory have identified this to be a common feature observed in some H3.1 K27M DIPG, suggesting that this might represent a distinct epigenetic profile to H3.3 K27M. In this line, Castel and colleagues have reported differences at the

transcriptome and epigenome level between H3.1 and H3.3 K27M diffuse midline glioma, suggesting these tumours might arise from different precursor cell or from distinct epigenetic cell state [95].

Double minutes or extrachromosomal DNA (ecDNA) are commonly seen in glioblastoma, playing an important role in tumour development and evolution [434-438]. Previous studies have shown loss of extrachromosomally-amplified genes in glioblastoma and neuroblastoma cultures, in particular in *MYC* and *MET* genes [439, 440]. In addition, three patient-derived *in vitro* DIPG models were established from the PNOC003 trial, with one failing to maintain a *PDGFRA/KIT* copy number gain observed in the original tumour patient [318]. In the BIOMEDE *in vitro* models, discrepancies between the cells and the original tumour sample were observed specifically with amplification of the oncogenes *PDGFRA* and *MYCN*. One hypothesis could be that the amplified genes localize to acentric extrachromosomal elements and/or were sub-clonal diverting from the original tumour population by clonal selection dynamics during cell culture. In addition, two variants in *NF1* and *PIK3R1,* which appear to be sub-clonal alterations by the allele frequency observed, were negatively or positively selected during cell expansion. These results highlight the importance of studying clonal composition during tumour propagation in heterogenous tumours, such as DIPG, which can be critical to uncover patient-specific drug response and resistance development.

TP53-inactivated tumours lack the G1/S checkpoint and instead rely on G2/M control for DNA repair and survival. AZD1775 is a selective inhibitor of the WEE1 kinase, which is an important regulator of the G2/M checkpoint, and has been shown to be upregulated in pHGG and DIPG [441, 442]. AZD1775 crosses the brain blood barrier and preclinical efficacy has been demonstrated in DIPG [442, 443]. Numerous studies have shown WEE1 inhibition to be dependent of p53 status [444-447], however, Sabine Muller and colleagues did not find such a correlation in HGG [442]. A high degree of *in vitro* sensitivity to AZD1775 (GI50 0.11-0.17 µM) in models harbouring frameshift or truncating mutations in *TP53* was observed, in addition to a modest *in vivo* efficacy in PDX-BIOMEDE-193 (TP53-R146*). It is to be explored whether different treatment window and/or higher dose regimen would show a significant prolongation of survival in this type of model. Although these results should be further tested by using larger *in vivo* studies as well as in combination with radiotherapy, it provides an excellent opportunity to test whether TP53 loss is a potential biomarker

of response to the WEE1 inhibitor in DIPG. AZD1775 is currently being tested in early phase clinical trials in DIPG and other paediatric tumours (NCT01922076 and NCT02813135), in this context it would be very valuable to correlate the clinical response with *TP53* status.

BIOMEDE-PDX models showed a wide range of survival from 70 to over 365 days, which creates limitations in establishing the most appropriate treatment window. Performing *in vivo* efficacy studies in order to treat the patient from which the model was derived from, can be challenging due to the latency of some these models and the lethality of DIPG. Nonetheless these tumours recapitulate what is observed in DIPG patients as they are very invasive, highly diffuse and very heterogeneous from patient to patient. BIOMEDE PDX models represent a valuable source of well-characterised tumours and represent a unique opportunity to perform single-patient derived multi-arm trials to explore drug sensitivity link to genetic dependencies in serial xenografts. In this context, nine PDX have been shared with ITTCC-P4 platform (Innovative Therapies for Children with Cancer Preclinical Proof-of-concept Platform) to expand the repertoire of DIPG tumours and provide the scientific community with the advantage of working with models which have not been exposed to cell culture conditions.

In a prospective clinical trial in DIPG, we are combining comprehensive molecular profiling linked to drug screening of patient-derived *in vitro* and i*n vivo* models. These data show the feasibility in generating patient-specific, testable hypotheses that may be clinically translated in a subset of DIPG patients.

CHAPTER 7 *: MEK1/2* mutations confer resistance to trametinib in a DIPG *BRAF*-G469V model link to proneural-mesenchymal transition

7.1 Introduction

In the era of precision medicine, targeted therapies represent an exciting opportunity in the clinical management of patients with cancer. In the last decade, clinical responses to targeted agents have been demonstrated in many patients, yet resistance can rapidly emerge, playing an important role in treatment failure [448]. Therapy-induced resistance can occur from the acquisition of *de novo* mutations [449, 450] or from expansion of rare pre-existing resistant cells, also called "persister clones" [451-453]. Many mechanisms of resistance have been identified and can be described in four main categories. Of these, modifications affecting the gene targeted are one of the most common mechanisms. These alterations typically act by impeding the specific binding of the small molecules to its target or by increasing the levels of the oncogene (e.g. *ALK*-F1174L mutation in neuroblastoma patients treated with crizotinib) [198, 454-457]. Resistance might also occur by restoration of signalling pathways through upstream or downstream compensatory activation (e.g *KRAS* amplification in *ALK:EML4*-positive non-small cell lung carcinoma (NSCLC) treated with crizotinib) [458, 459]. Similarly, activation of parallel oncogenic pathways can bypass specific pathway blockage (e.g *MET* amplification in *EGFR*-mutant NSCLC) [460-462]. A different and less understood mechanism is the role of cell plasticity enabling reversible lineage transformation in cancer cells in response to treatment (e.g. the shift from epithelial to mesenchymal phenotype seen in multiple carcinomas defined by their transcriptional state) [463-465]. Additionally, intra-tumour heterogeneity has an important role in resistance, making it difficult to discriminate passenger mutations or neutral clones from functional mutations which through selective pressure can drive population expansion and lead to resistance [466, 467].

To bypass resistance and compensatory signalling networks, numerous strategies have been employed, with the ultimate goal of prolonging durable drug response. In this context, new generations of kinase inhibitors, with increased potency, are constantly been developed to tackle "gatekeeper" mutations, which modulate the accessibility of the kinase ATP-binding pocket. One example is the FDA-approved

kinase inhibitor ponatinib, which was developed and successfully used in leukaemia patients to overcome resistance in patients harbouring the *BCR-ABL* T315I gatekeeper mutation [468]. In cases where parallel/downstream/upstream pathway activation are the responsible for the resistance, a combination of inhibitors simultaneously targeting the primary mutation and the secondary compensatory mechanism can be used. One of the challenges of combination therapy is the associated toxicity with the use of multiple drugs [469].

Although targeted agents against the mitogen-activated protein kinase (MAPK) pathway have revolutionised the field of precision oncology, they are frequently associated with the emergence of resistance. The MAPK pathway plays an important role in signal transduction regulating cell proliferation, differentiation and cell death [470]. Dysregulation of the MAPK signalling pathway is implicated in a wide range of cancers as a result of genetic and epigenetic alterations. In adults, *BRAF* mutated tumours include 60% of melanomas, 60% of thyroid cancers, 15% of colorectal and 5-8% of non-small cell lung cancer, with the most prevalent mutation being *BRAF*-V600E [471]. In addition, non-V600E mutations have been identified to be oncogenic and can be classified in three types based on their effect of BRAF activity [472, 473]. The use of MAPK inhibitors, including BRAF inhibitors (vemurafenib and dabrafenib) and MEKi (trametinib), in *BRAF*-V600E positive melanoma resulted in a moderate success of targeted therapies by showing tumour shrinkage and improving patient survival [361, 474, 475]. However, durable responses were limited, due to resistance to single-agents often mediated by re-activation of MAPK through amplification or splice variants in *BRAF*, mutations in the upstream oncogene *NRAS* or the downstream kinase *MAP2K1,* as well as PI3K-PTEN-AKT upregulation, amongst others [450, 476-479]. This led to the combination of BRAFi and MEKi in clinical trials which showed a significant improvement in overall survival compared to single agent therapy and subsequently FDA-approved treatment for advanced *BRAF*-V600 positive melanoma [480, 481]. However, acquired resistance was observed in patients under BRAF and MEK inhibitors, with the identification of *MAP2K1* and *MAP2K2* mutations as one of the main mechanisms [482-485].

MAPK pathway alterations are also commonly found in a considerable number of childhood tumours and in particular in paediatric low and high-grade gliomas. These include pilocytic astrocytoma (*KIAA1549:BRAF* tandem duplication, *RAF* fusions, *NF1*, *FGFR1*, *BRAF*-V600E) [21, 33, 35], mixed glioneuronal tumours (*FGFR1,*

BRAF-V600, *KIAA1549:BRAF*) [36, 486], pleomorphic xanthoastrocytomas (*BRAF*-V600E) [45, 487], infant pHGG (*NTRK1/2/3, ROS1, ALK, MET* fusions) [22, 370], non-brainstem pHGG (*FGFR1, NF1*, *BRAF*-V600E, ITD-*NTRK2, MET*) [22, 87, 367, 488] and DIPG (*PIK3R1, NF1*) [22, 75].

Despite the recent availability in the clinic of BRAF and MEK inhibitors for pHGG, efficacy studies testing the utility of these targeted agents in primary patient-derived models linked to genetic vulnerabilities need to be further explored. For instance, first generation of BRAF inhibitors are not effective in *BRAF* non-V600 nor in *KIAA1549-BRAF* tumours. Similarly, patients with RAF fusions remain unresponsive to both first and second generation of BRAF inhibitors. As previously described, resistance to MAPK inhibitors will ultimately emerge, highlighting the need to explore this phenomenon in order to offer patients an alternative treatment when disease progression occur.

Interestingly, results from the drug screen described in Chapter 6 showed a differential range of sensitivities in response to trametinib in the BIOMEDE patient-derived cell cultures. I used these models to explore for the first time predictive and resistance biomarkers of response to trametinib in the context of DIPG.

7.2 Results

7.2.1 MAPK pathway alterations confer sensitivity to trametinib in DIPG *in vitro* models

The personalised drug screen conducted in Chapter 6 led to the identification of a wide range of responses to trametinib in the BIOMEDE patient-derived cell cultures. Dysregulation of the MAPK pathway was observed in 3/5 of the trametinib-sensitive models **{Figure 7-1}**. BIOMEDE-169 (2D and 3D) harboured a class II *BRAF* oncogenic mutation (G469V, GI50 of 51 nM and 23 nM, respectively) **{Figure 7-2A**} BIOMEDE-181 3D presented a *PIK3R1* hotspot mutation (N564D, GI50 of 50 nM) **{Figure 7-2B}**, and BIOMEDE-184 3D a missense mutation in *NF1* (I1824S, GI50 of 16 nM). In addition, BIOMEDE-198 2D and BIOMEDE-186 3D showed high sensitivity to trametinib (GI50 of 50 nM and 16 nM respectively).

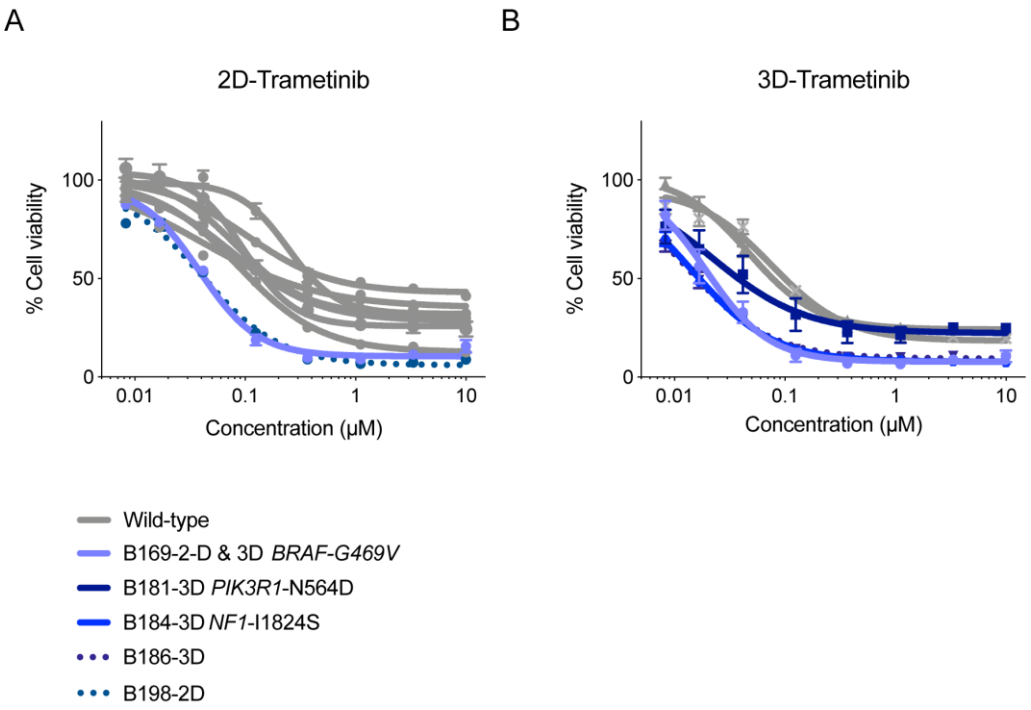


Figure 7-1 MAPK pathway alterations and sensitivity to trametinib *in vitro* in the DIPG cultures derived from BIOMEDE-UK patients. The effect on cell viability (surviving fraction on y-axes) of trametinib with a range of 0-10 μM (x-axes) was assessed (A) in the 2D and (B) in the 3D cultures. The errors bars represent the standard error of the mean. Sensitive cultures with MAPK pathway dysregulation are highlighted in blue.

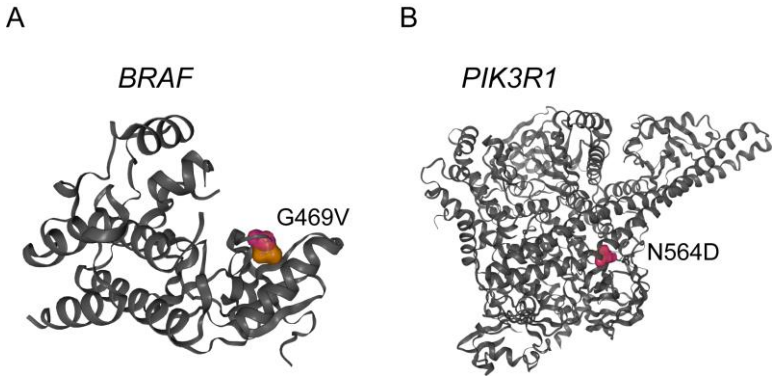


Figure 7-2 Protein structural visualization download from COSMIC-3D web interface (https://cancer.sanger.ac.uk/cosmic3d) [489] showing (A) the protein structure of *BRAF*-G469V mutation (B) and *PIK3R1*-N564D.

An allelic imbalance of *PIK3R1*-N564D and *NF1*-I1824S was observed in the 3D but not in the 2D cultures. Both 2D and 3D conditions were established and grown independently from the original tissue sample taken at diagnosis. The mutant cultures for *PIK3R1* and

NF1 showed a greater response to trametinib than their paired wild-type cultures (3.7-fold and 14.2-fold respectively) **{Figure 7-3}**.

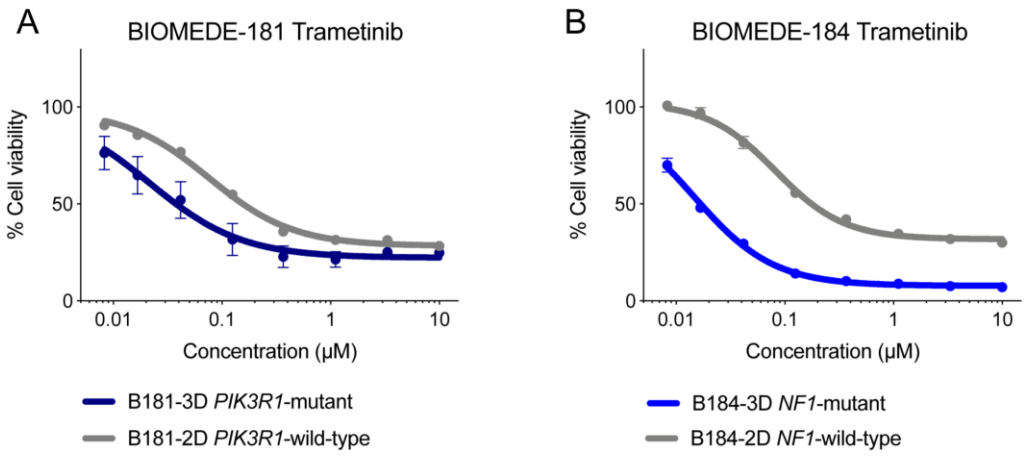


Figure 7-3 BIOMEDE-184 2D and 3D response to trametinib *in vitro*. The effect on cell viability (surviving fraction on y-axes) of trametinib at a range of 0-10 μM (x-axes) was assessed in (A) BIOMEDE-181 (B) and BIOMEDE-184. The errors bars represent the standard error of the mean. Sensitive cultures with MAPK pathway alterations are highlighted in blue.

Panel sequencing was performed at early and later passage in BIOMEDE-184 2D and 3D. *NF1*-I1824S was present in BIOMEDE-184 3D at p5 at a VAF of 65% (93/142 reads) and at later passage, p14, at a VAF of 75% (43/57 reads). However, the mutation was not present in BIOMEDE-184 2D at p4 (0/201) nor at p14 (0/131).

Panel sequencing showed that *PIK3R1*-N564D was present in both 2D and 3D cultures at p5, exhibiting a VAF of 16% (36/227 reads) and 33% (100/301 reads) respectively. Notably, panel sequencing conducted at later passage (p12 for 2D and p18 for 3D), showed the 2D culture did not harbour the *PIK3R1* mutation (0/342 reads) whereas the 3D culture was heterozygous for N564D with a VAF of 49% (195/396 reads). To confirm these findings, I performed ddPCR at longitudinal passages assessing the presence of the *PIK3R1* mutation over time (p5, p13, p15 for both 2D and 3D as well as p18 for 2D and p17 for 3D). The negative and positive selection of *PIK3R1*-N564D in 2D and 3D independent cultures, respectively, was confirmed **{Figure 7-4}**.

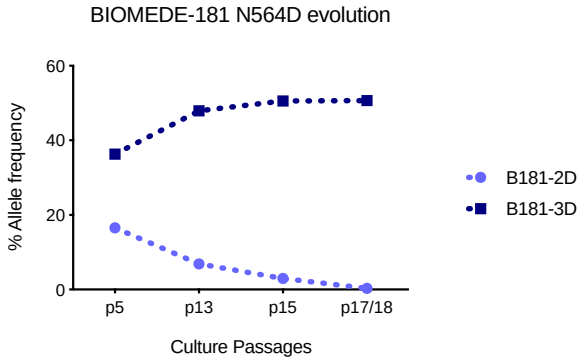


Figure 7-4 Variant allele frequency (VAF) of *PIK3R1*-N564D in BIOMEDE-181 2D and 3D over time. The graph shows VAF of *PIK3R1*-N564D (c.1690 A>G) mutation by droplet digital PCR from 2D (light blue) and 3D (dark blue) cultures at different cultures passages.

To evaluate if the presence of the *PIK3R1* mutation affected the tumourigenicity after orthotopic engraftment *in vivo,* BIOMEDE-181 2D p18 (*PIK3R1*-N564D, VAF of 0.32% 20/5476 droplets) **{Figure 7-5A}**, and BIOMEDE-181 3D p17 (*PIK3R1*-N564D, VAF of 50.6% 4919/9727 droplets) **{Figure 7-5B},** were injected in 10 NSG mice per condition. This was done by Dr Diana Carvalho in the Jones laboratory**.** No tumours were observed in the NSG injected with BIOMEDE-181 2D cells **{Figure 7-5C}**. Conversely 10/10 tumours were generated in the mice injected with BIOMEDE-181 3D p17 cells **{Figure 7-5D}**, with a median overall survival of 310 days.

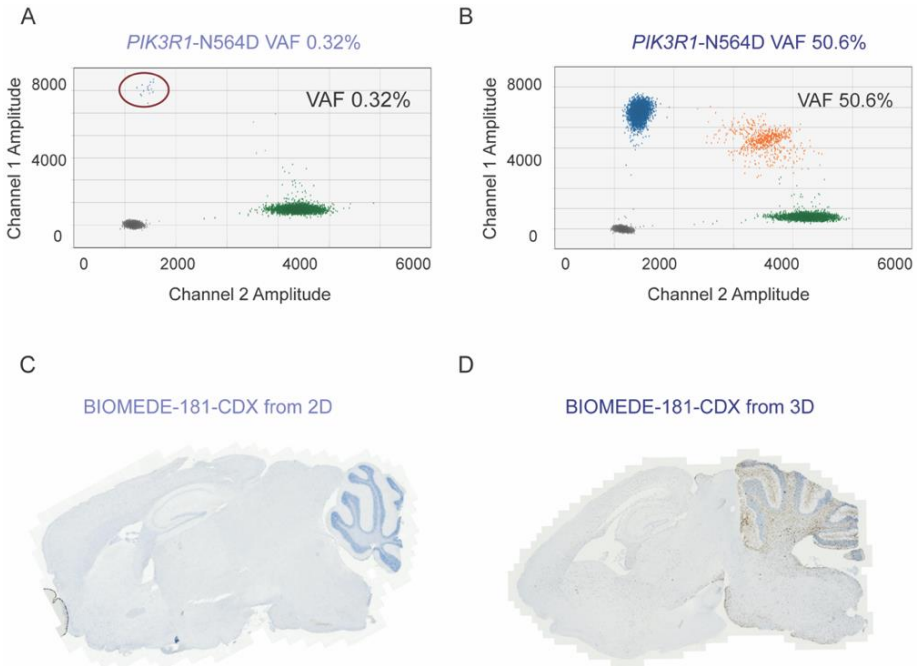


Figure 7-5 *PIK3R1*-N564D tumourigenesis assessment *in vivo.*(A) Droplet digital PCR 2D amplitude plot for *PIK3R1* wild-type (x-axes) and *PIK3R1*-N564D mutation (y-axes) for BIOMEDE-181 2D cells, with 20 mutant droplets out 5476 and a VAF of 0.32% and (B) BIOMEDE-181 3D cells, with 4919 mutant droplets out of 9727 and a VAF of 50.6%. (C) Anti-human nuclei antibody (HNA, brown staining) for BIOMEDE-181 CDX derived from 2D cells showing the lack of tumour and (D) the presence of tumour generated from BIOMEDE-CDX derived from 3D cells.

BIOMEDE-169 patient was assigned to the mTOR inhibitor everolimus, a summary of the patient clinical history is shown in **{Figure 7-6A}**. After ~7 weeks of everolimus in combination with radiotherapy, the patient showed a reduction in tumour size (MRI not available). However, at week 12 the patient presented with pseudo-progression, and at week 16 progression was confirmed on the basis of symptoms and clear MRI changes compared to diagnosis **{Figure 7-6B-C}**. The patient was given several courses of steroids at pseudo-progression and at progression. According to the presence of the *BRAF*-G469V mutation and the sensitivity found to trametinib *in vitro*, the patient was then treated with trametinib on the basis of a compassionate use program (0.025 mg/kg/day once daily). During 11 weeks on trametinib, the patient was able to be taken off steroid treatment, however at week 12, MRI showed progression with the appearance of new metastatic lesions within the brainstem and in the lateral ventricles **{Figure 7-6D}**.

In parallel, BIOMEDE-169 CDX was used to evaluate the efficacy of trametinib *in vivo* into NSG mice (n=5 vehicle controls, n=5 treated). Non-tumour bearing animals were treated continuously for 14 days (1mg/Kg, PO q.d.) and no signs of toxicity were observed. Tumour-bearing animals started treatment at day 55 post-implantation with 1mg/kg of trametinib orally once daily for 53 days (5 days on, 2 days off). There was no difference in overall survival observed (p-value=0.7123 log-rank test) **{Figure 7-7}**.

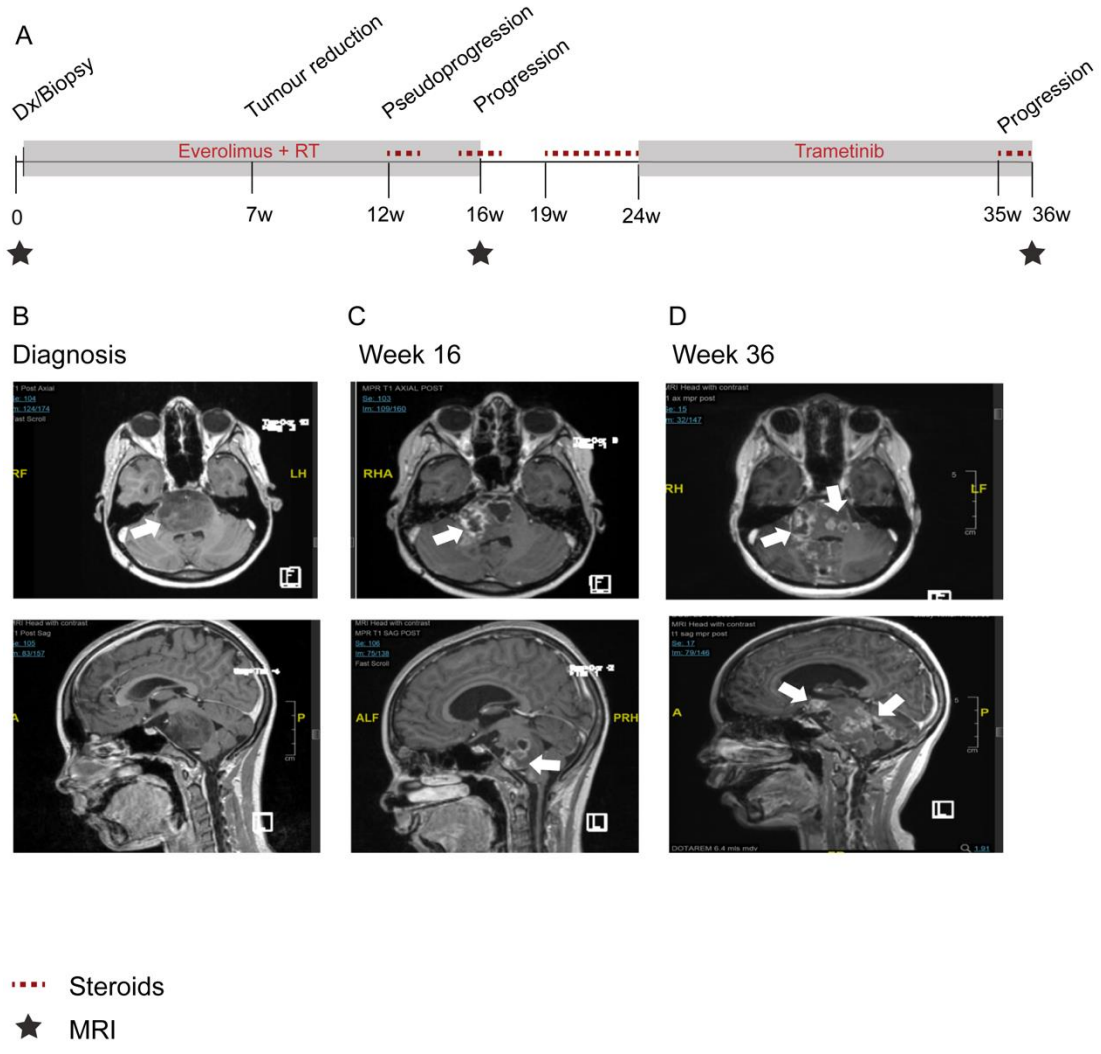


Figure 7-6 BIOMEDE-169 *BRAF*-G469V clinical intervention (A) Timeline of clinical history, with treatment shaded in grey. The red dotted line represents the intervals when the patient was given steroid pulses and the star symbol the time-point where magnetic resonance imaging (MRI) are shown. Axial (top) and sagittal (bottom) T1-weighted post-gadolinium MRI at (A) diagnosis, (B) week 16 showing enlargement of the tumour enhancement area after 15 weeks of everolimus and radiotherapy and (C) week 36 after 12 weeks of trametinib, showing the development of new metastatic disease within the brainstem and the lateral ventricles as well as diffuse leptomeningeal enhancement of the cord. The white arrows show high signal intensity highlighting the tumoural area.

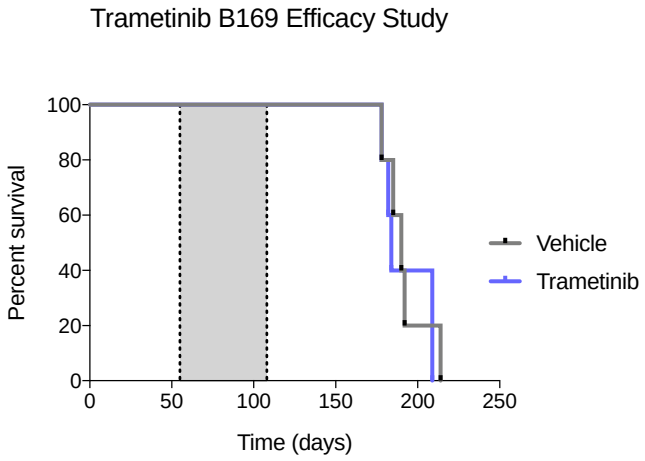


Figure 7-7 Survival for BIOMEDE-169 CDX in NSG. (n=5 vehicle and n=5 trametinib) (p-value=0.7123 log-rank Mantel Cox test).

7.2.2 *MEK1/2* mutations drive resistance to trametinib

To explore the possible resistance mechanisms that may occur in DIPG, trametinib-resistant clones were generated in the *BRAF*-G469V model by exposing the cells to trametinib *in vitro*. Two different strategies were employed for resistance generation, either via increasing concentrations of trametinib in an exponential manner, starting from the GI50 value (0.05 µM) and ranging up to 1.0 µM (A.T1, A.T2, B.T3, B.T4 clones), or via exposure to a constant GI80 dose (0.5 µM) (C.T5, C.T6, D.T7, D.T8 clones) **{Figure 7-8}**.

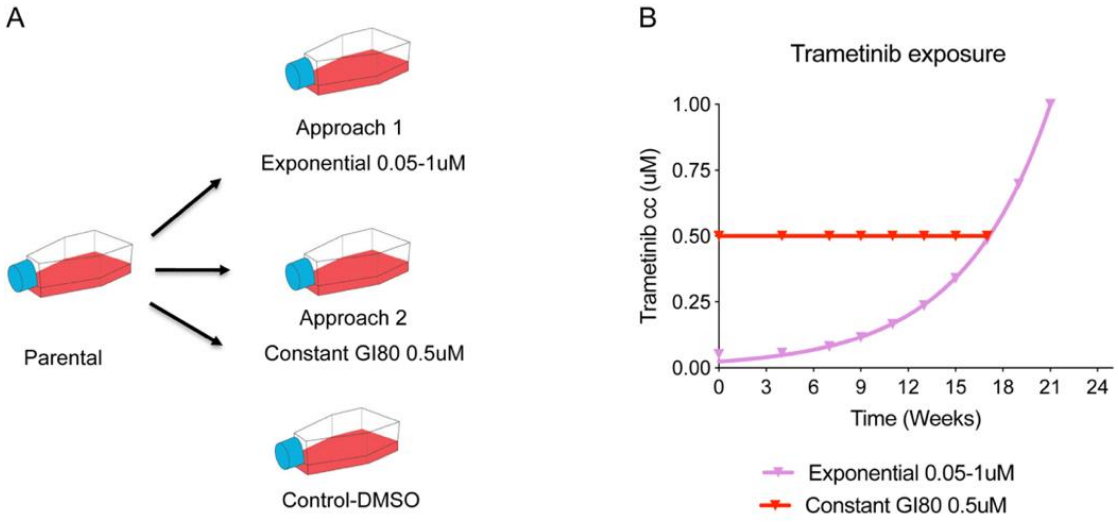


Figure 7-8 Summary of resistance generation methodology. (A) Experimental outline and (B) graph representing the concentration of trametinib over time to which the cells were exposed to in the different approaches utilised 1) exponential concentrations of trametinib from GI50 value 0.05 µM to 1 µM and 2) constant dose corresponding to GI80 value 0.5 µM.

Trametinib sensitivity was assessed at several time-points. By following approach 1, a shift in GI50 was observed in A.T1, A.T2 and B.T3 after ~5 months of trametinib exposure compared to the BIOMEDE-169 parental line **{Figure 7-9A-C}**. The GI50 shift continued increasing after ~7-9 months of treatment, resulting in a fold-shift in GI50 values of 64, 167, and 97 in A.T1, A.T2 and B.T3, respectively, compared to the parental line. Whilst B.T4 did not show a shift in GI50 value after five months of trametinib, sensitivity was re-assessed after ~7 and 9 months of drug exposure detecting a GI50 fold-shift of 161 **{Figure 7-9D}**. After 4-5 months, by using approach 2, a fold-shift in GI50 values of 7, 44, 5.6 and 31 was observed in clones C.T5, C.T6, D.T7 and D.T8, respectively, compared to the parental line **{Figure 7-9E-H}**. By continuing trametinib exposure up to 7 months in C.T6, a GI50 fold-shift of 56 was achieved compared to the parental line.

In order to explore what might be underlying the lack of sensitivity in the resistant cells, whole exome sequencing was conducted in A.T1, A.T2, B.T3, B.T4 and C.T6. Mutations seen in BIOMEDE-169 parental such as *H3F3A*-K27M, *BRAF*-G469V and *TP53*-C176Y amongst others were also present in the derived resistant cells at similar VAF **{Figure 7-10}**. There were three variants present in the original tumour sample that were not seen in the parental cells, these include *SYNDIG1L*-G236S, *CDN2C*-L58LX and *PI4KA*-A1727V. The presence of private mutations in each clone was seen with 3, 8, 12 and 11 variants identified A.T1, B.T3, B.T4 and C.T6. In particular, mutations in *MAP2K1* (*MEK1)* and *MAP2K2* (*MEK2)* were identified in A.T1 (*MEK2*-I115N), B.T3 and B.T4 (*MEK1*-I141S) and C.T6 (*MEK1*-K57N).

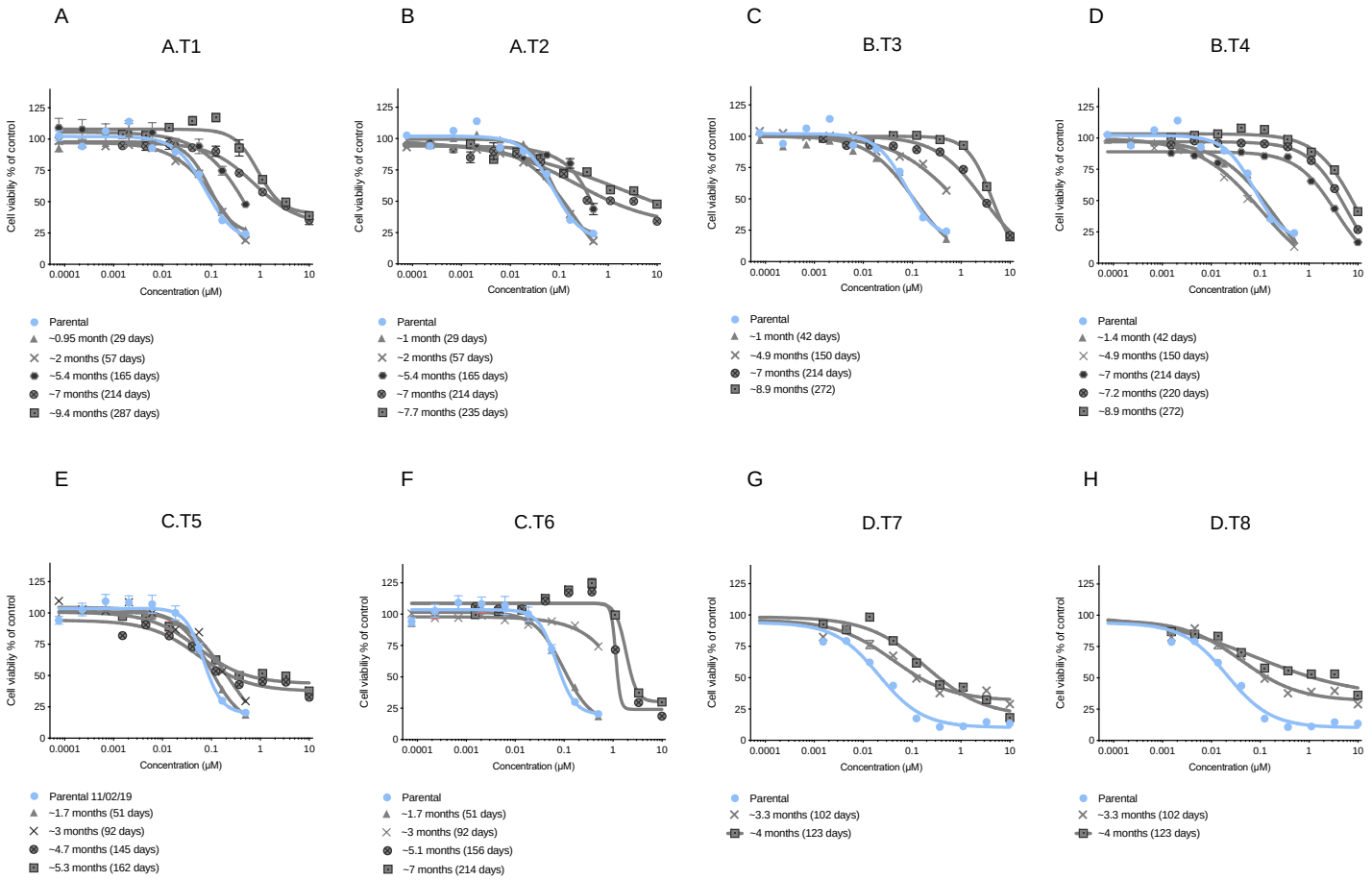


Figure 7-9 Trametinib response in BIOMEDE-169 clones. The effect on cell viability (surviving fraction on y-axes) of trametinib (x-axes) was assessed at several time points for each of the eight clones by using two different approaches. (A-D) Cell viability of clones under exponential concentration of trametinib with a range of 0-1 µM and (E-H) cell viability of clones exposed to constant GI80 (0.5 µM). The errors bars represent the standard error of the mean.

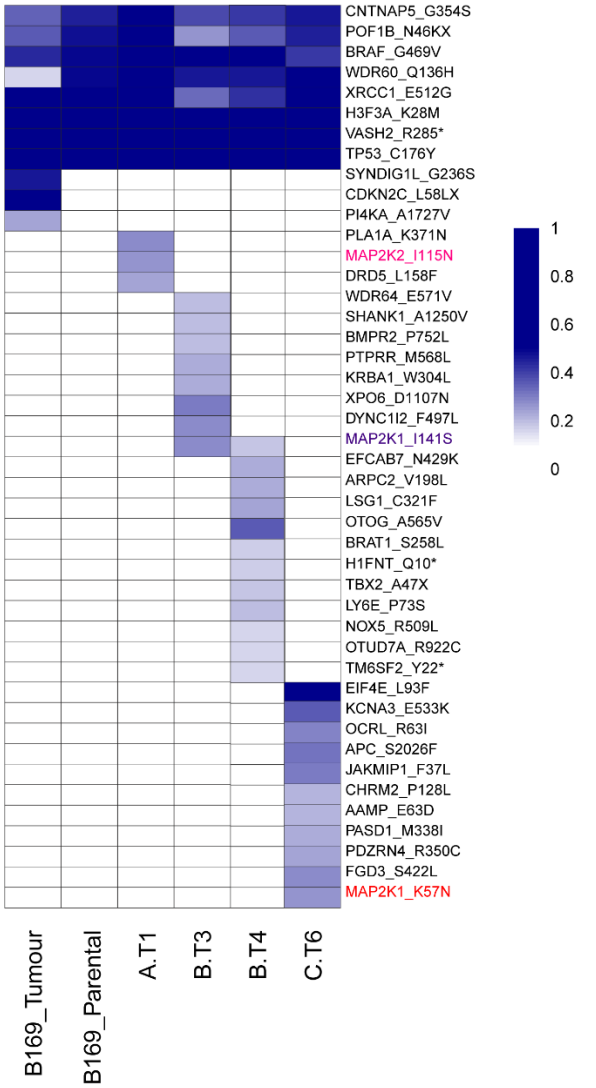


Figure 7-10 Heatmap representation of mutations in BIOMEDE-169 tissue, BIOMEDE-169 parental cells and the resistant clones A.T1, B.T3, and C.T6. Each column represents one sample and each row is one of mutation. The blue intensity is proportional to the variant allele frequency. The presence of *MEK1*/2 mutations is coloured per mutation type in the clones A.T1 *MEK2*-I115N (pink), B.T3 and B.T4 *MEK1*-I141S (purple) and C.T6 *MEK1*-K57N (red).

To explore whether the resistance was reversible, trametinib was withdrawn for two months in A.T1, B.T3 and C.T6 clones. In the absence of the MEK inhibitor, sensitivity to trametinib was re-assessed and allele-specific ddPCR assays for *MEK1/2* mutations were performed to evaluate potential changes in *MEK1/2* VAF. After the "drug holiday" period, resistance was maintained **{Figure 1-11A-B}**, and the *MEK1/2* VAF did not experience a dramatic variation (A.T1 *MEK2*-I115N VAF from 12% to 14.6%, B.T3 *MEK1*-I141S VAF from 28.2% to 30.5% and C.T6 *MEK1*-K57N from 26% to 24.9%) **{Figure 1-11C}**.

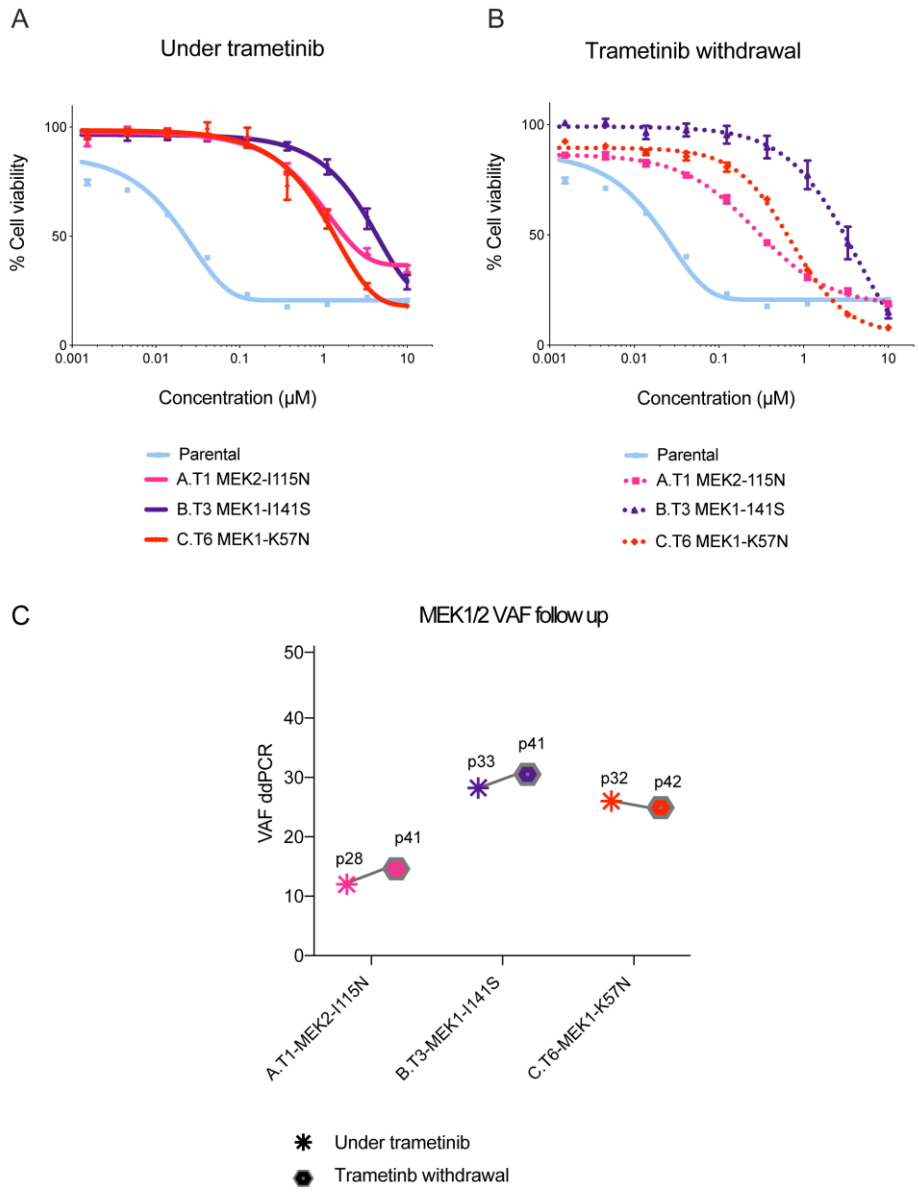


Figure 7-11 Trametinib response and VAF in BIOMEDE-169 resistant clones under trametinib and post-trametinib withdrawal. Cell viability (y-axes) was evaluated against trametinib concentration (x-axes) for BIOMEDE-169 parental cells and the *MEK1*/2 mutant clones (A.T1, B.T3, B.T4 and C.T6). The errors bars represent the standard error of the mean. The effect on cell viability was assessed (A) under trametinib exposure and (B) after ~7 months of trametinib treatment and after 2 months of drug holiday. (C) *MEK*1/2 variant allele frequency was evaluated by ddPCR in A.T1, B.T3 and C.T6 clones pre-trametinib and post-trametinib drug removal.

To evaluate if resistance was selected from pre-existing clones or was acquired in response to trametinib, allele-specific ddPCR assays for *MEK1/2* resistant mutations were conducted in the parental cells as well as in the original tissue sample, and in both vehicle- and trametinib-treated BIOMEDE-169 CDX. No mutant droplets were found in the parental nor in the original patient sample for any of the three mutations,

with an average of droplets containing DNA of 32,165 and 41,083 respectively **{Figure 1-12A-C}**. No mutant droplets were found in any the of xenografts, with an average of droplets containing tumour DNA across the ten samples of 6352 **{Figure 1-12D-E}**. Whole exome sequencing was conducted to determine potential resistant alterations responsible of trametinib failure *in vivo*. While the mutations shared between the BIOMEDE169 parental cells and original tumour sample were retained in all CDX at similar VAF including *H3F3A*-K27M, *BRAF*-G469V and *TP53*-C176Y, no canonical mutations in the MAPK pathway were detected in the treated group.

Allele-specific *MEK1/2* ddPCR assays were then used to track the emergence of *MEK1/2* mutations from longitudinal passages over the course of the continuous exposure experiments. Under both experimental conditions, *MEK1/2* VAF increased in an exponential manner over-time, however earlier emergence was observed in the clone exposed to constant GI80 trametinib concentration (C.T6 *MEK1*-K57N) **{Figure 7-13A}**. GI50 values overtime started shifting, for A.T1, B.T3, B.T4 and C.T6, when the resistant mutations emerged and the shift was more pronounced as VAF increased **{Figure 7-13B}**.

I next sought to evaluate MAPK pathway modulation by using a quantitative capillary protein platform (WES protein simple). The MAPK pathway showed a sustained up-regulation in each of the resistant clones, including pMEK1/2 (p217/221) in A.T1 *MEK2*-I115N ($p<0.0001$, one-way ANOVA) **{Figure 7-14A}**, pERK1/2 (p202/204) in B.T3 *MEK1*-I141S and C.T6 *MEK1*-K57N ($p<0.0044$ and $p<0.0006$, respectively one-way ANOVA) **{Figure 7-14B}**, and pAKT (pS473) in B.T3-*MEK1*-I141S and C.T6-*MEK1*-K57N ($p<0.0166$ and $p<0.0074$, respectively one-way ANOVA) **{Figure 7-14C}**.

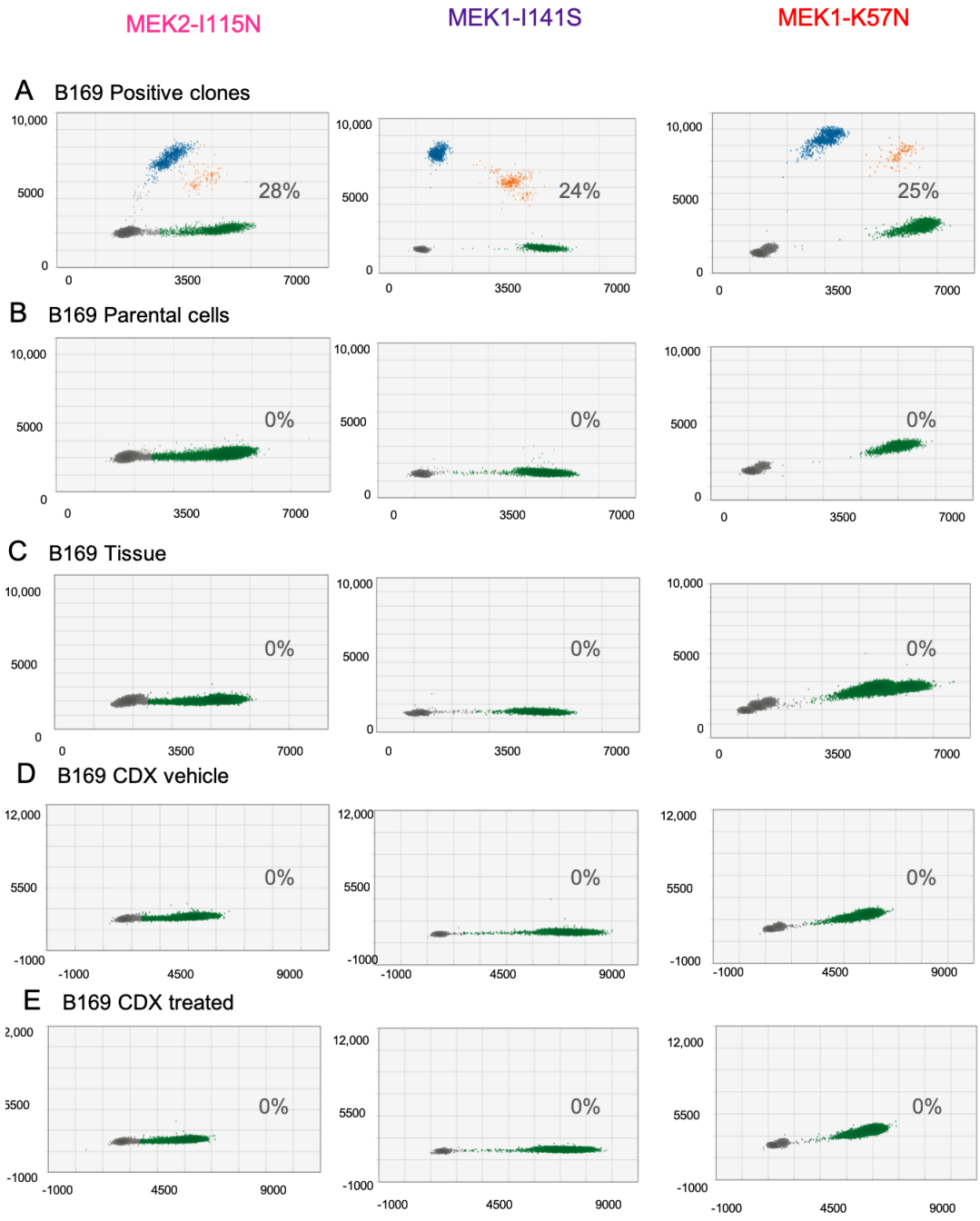


Figure 7-12 *MEK1/2* allele specific mutation were assessed by droplet digital PCR in BIOMEDE-169. 2D amplitude plot for *MEK1/2* wild-type (x-axes) and mutant (y-axes) are shown for the three different mutations *MEK2*-I115N (left), *MEK1*-I141S (middle) and *MEK1*-K57N (right). Each mutation was assessed in the respective mutant clone (A) A.T1 *MEK2*-I115N, B.T3 *MEK1*-I141S and C.T6 *MEK1*-K57N) (B), BIOMEDE-169 parental cells (C) BIOMEDE-169 original tissue sample (D), BIOMEDE-169 CDX vehicle group and (E) BIOMEDE-169 CDX treated group.

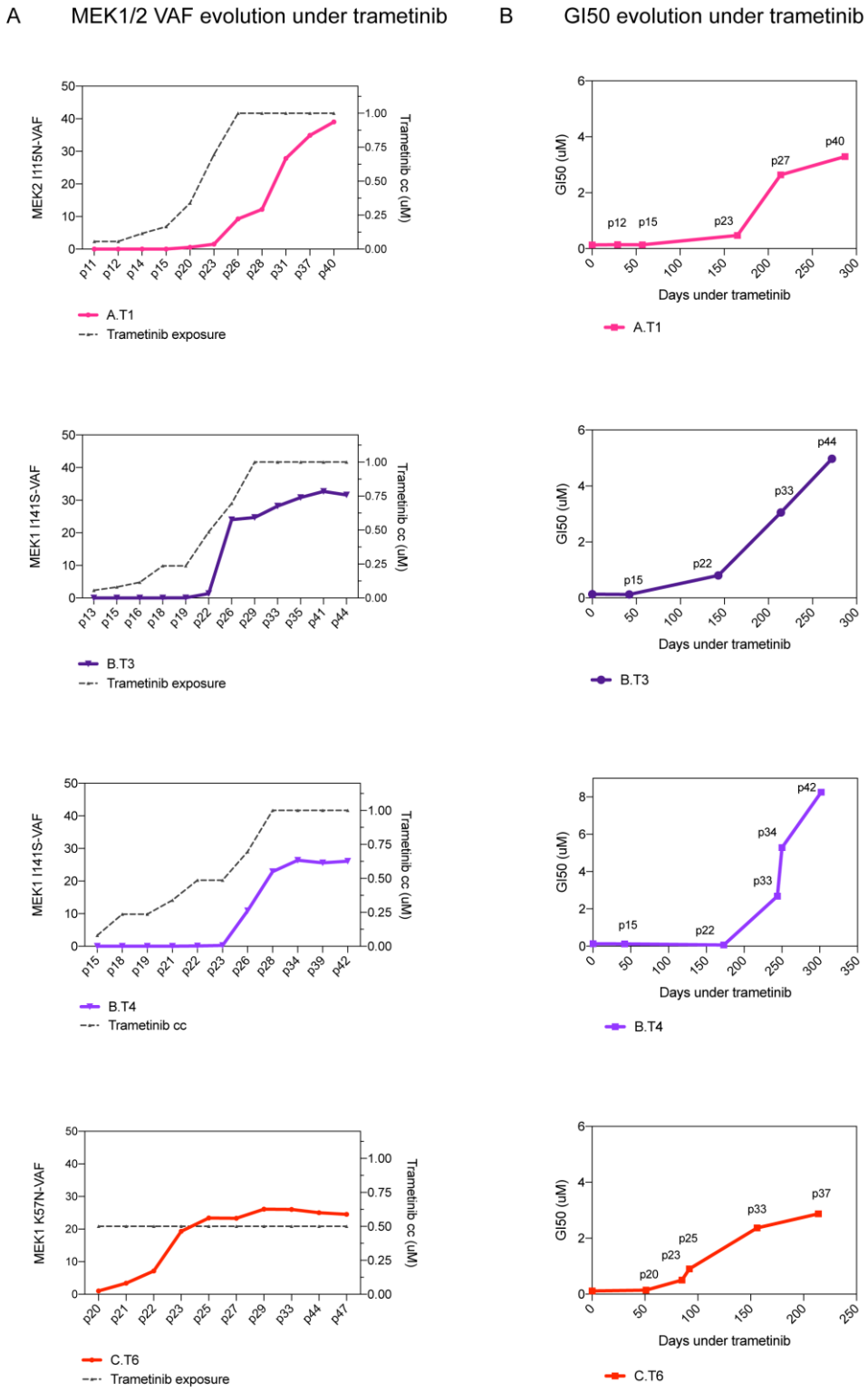


Figure 7-13 Evolution of *MEK1/2* resistant clones over time. (A) *MEK1/2* variant allele frequency (VAF) was assessed by droplet digital PCR (left y-axes) showing the trametinib concentration exposure (right y-axes) at each of the multiple longitudinal passages evaluated (x-axes). (B) Sensitivity to trametinib was assessed to evaluate the emergence of resistance and GI50 (μM) (y-axes) were calculated different time-points of trametinib exposure, shown in days (x-axes). Each clone (A.T1, B.T3, B.T4 and C.T6) is plotted separately and coloured by *MEK1/2* mutation type.

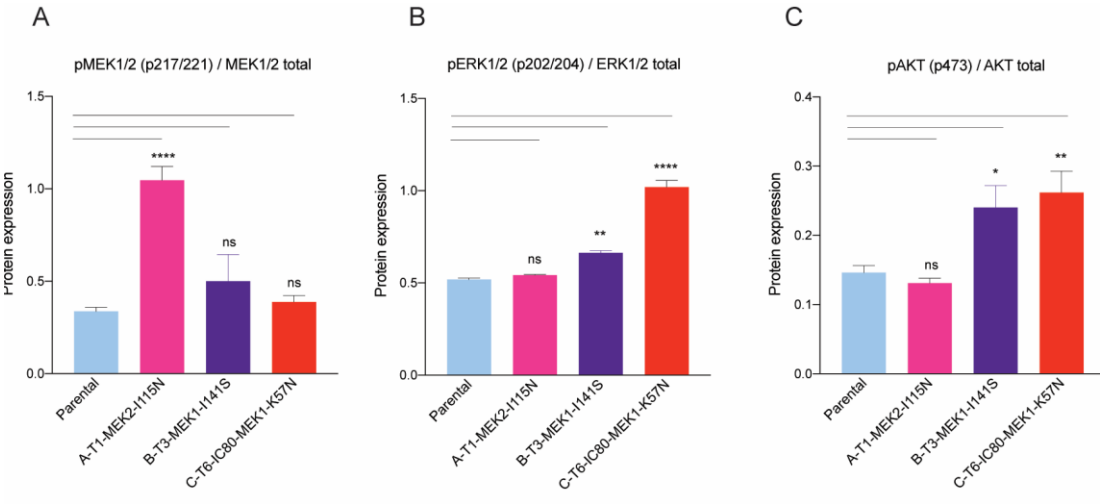


Figure 7-14 Bar-plot showing protein expression of MAPK pathway in BIOMEDE-169 parental cells and *MEK1/2* resistant clones by western-protein simple. (A) pMEK1/2 (p217/221) normalised to total MEK/1/2, (B) pERK1/2 (p202/204) normalised to total ERK1/2 and (C) pAKT (p473) normalised total AKT. All p-values are based upon one-way ANOVA test ****p<0.00001, **p<0.001, *p<0.01 and ns p>0.05.

7.2.3 Trametinib-resistant clones showed mesenchymal transition phenotype

To evaluate the expression landscape of the *MEK1/2* resistant clones (A.T1, B.T3, C.T6), RNA-seq and phospho-/total proteomics were conducted and compared to the parental cells. Massively parallel proteome and phosphoproteome quantification was performed by using tandem mass tag (TMT) labelling, and phosphopeptide enrichment was done using immobilised metal affinity chromatography (IMAC). Liquid chromatography with tandem mass spectrometry (LC-MS/MS) analysis identified a total of 8,277 proteins and 9,444 unique phosphopeptides that mapped 3609 distinct master proteins. The three biological replicates samples for full proteome showed an a very good overall correlation (mean r^2 Parental=0.80, T1=0.85, T3=0.80 and T6=0.88).

RNA-seq showed 277, 127 and 241 differentially up-regulated genes, and 232, 175 and 214 down-regulated in A.T1, B.T3 and C.T6 respectively **{Figure 7-15A}**. Of those, a total of 41 genes were shared to be up-regulated and 56 genes to be down-regulated amongst the three clones **{Figure 7-15B}** and **{Appendix Figure 3A}**. From the total proteome, 447, 378 and 435 proteins were differentially up-regulated and 626, 673, 729 down-regulated in A.T1, B.T3 and C.T6 respectively. **{Figure 7-15C}.** Of these, a total of 137 up-regulated and 225 down-regulated proteins were

common amongst the three clones **{Figure 7-15D}** and **{Appendix Figure 3B}**. Phosphoproteomics identified a total of 212, 37 and 167 phosphopeptides from 124, 22 and 106 proteins with increased phosphorylation and 561, 171 and 395 phosphopeptides from 373, 117 and 221 proteins with decreased phosphorylation in A.T1, B.T3 and C.T6 respectively **{Figure 7-15E}.** Of these, a total of 8 proteins with increased phosphorylation and 57 with decreased phosphorylation were shared amongst the three clones **{Figure 7-15F}** and **{Appendix Figure 3C}**.

Intersection of shared differentially expressed genes / proteins and differentially phosphorylated proteins between any of the three clones from RNA-seq, full proteome and phosphoproteome is shown in **{Figure 7-16}** Genes / proteins associated with cytoskeleton re-organisation, cell migration, cell polarisation, and cell matrix remodelling were found consistent co-activated in the clones. These included the cell-surface glycoprotein CD44, the p21 activated-kinase PAK1, the tumour endothelial marker TNS3, and the transmembrane (type I) heparan sulfate proteoglycan SDC2, amongst others (GALS1, ANXA1, PLE, RIN1, LMNA, FLNA, MYH9) **{Figure 7-16A}**. In addition, the mitogen-activated protein kinase 5 (MAP3K5) was found activated in A.T1 and B.T3. Depletion of neural and oligodendrocyte markers was seen including the neural progenitor stem marker NESTIN the transcription factor SOX10, in addition to the glial cell adhesion protein HEPACAM, actin-filament associated protein AFAPIL2, the kinesin family member KIF26B and the myelin transcription factor 1 MYT1 **{Figure 7-16B}**.

Gene set enrichment analysis (GSEA) from RNA-seq and proteomics showed a high concordance of enrichment signatures amongst the clones **{Figure 7-17}**. Many of these were linked with mesenchymal transition (VERHAAK GLIOBLASTOMA MESENCHYMAL, ONDER CDH1 TARGETS 2 UP, JECHLINGER EPITHELIAL TO MESENCHYMAL TRANSITION UP, REACTOME EXTRACELLULAR MATRIX ORGANIZATION), migration (WU CELL MIGRATION) and invasion (SCHUETZ BREAST CANCER DUCTAL INVASIVE UP, ANASTASSIOU MULTICANCER INVASIVENESS). Archetypal differentially expressed genes / proteins associated with the phenotype included CD44, POSTN, MYOF, MXRA5, CDH11, PDGFRB, S100A4, EPHA2, collagen family, MMP, FBN1 **{Figure 7-18A}**. By contrast, proneural signatures (VERHAAK GLIOBLASTOMA PRONEURAL) were down-regulated in the resistant clones associated with decreased gene / protein expression of OLIG2, SOX10 and MYT1 **{Figure 7-18B}**. Notably, the HUANG DASATINIB RESISTANCE

UP signature was identified in the resistant clones revealing an increased expression of biomarkers of response to dasatinib including CAV1, ANXA1, PDGFC, EPHA2 **{Figure 7-18C}**. Gene ontology analysis showed shared common dysregulated molecular functions across the three *MEK1/2* resistant clones and the most significant are shown in **{Figure 7-19}**. Some of these were extracellular matrix structural constituent, collagen binding, integrin binding, protease binding, calcium ion binding and signalling receptor activity amongst others.

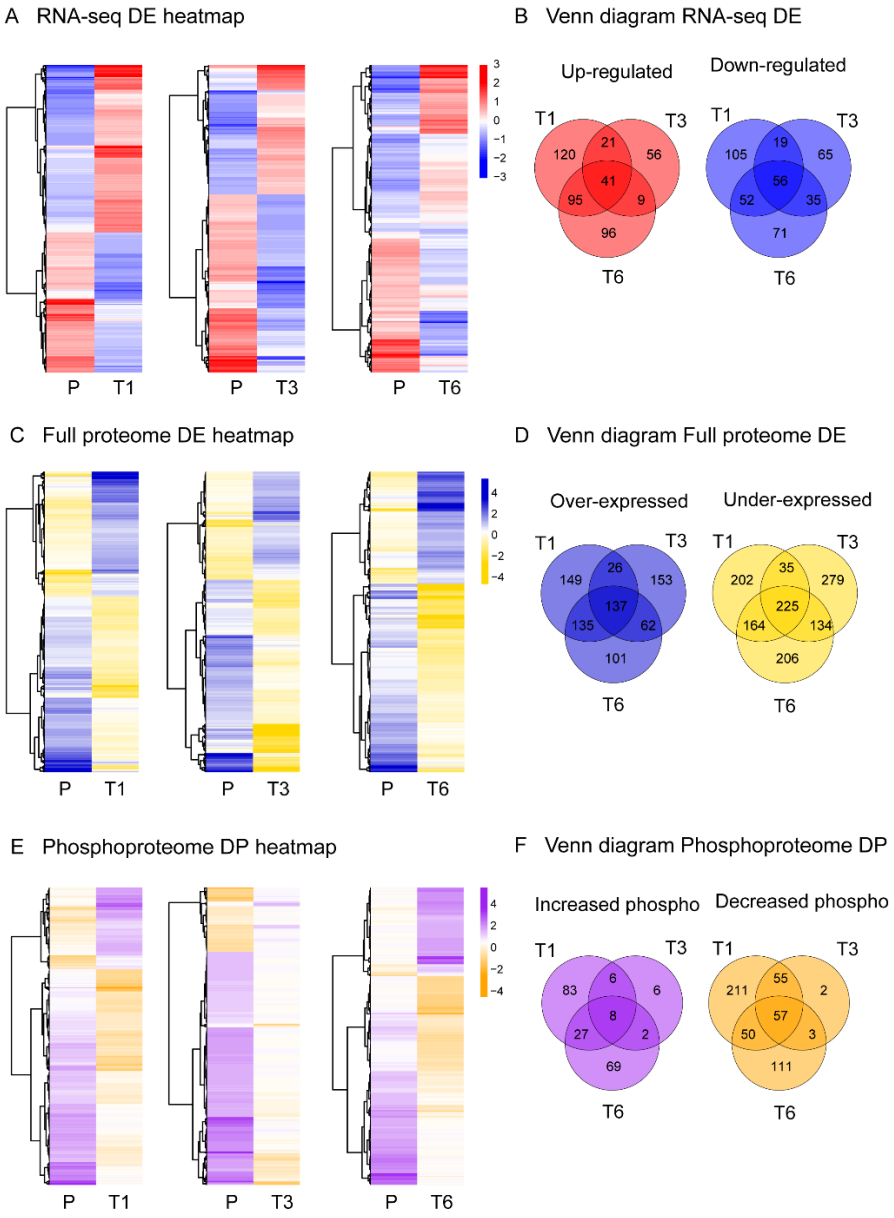


Figure 7-15 Unsupervised hierarchical cluster heatmaps of the MEK1/2 resistant compared to the BIOMEDE-169 parental cells. Venn diagrams of differentially expressed genes / proteins and phosphorylated proteins are shown on the right. Panels show (A-B) RNA-sequencing, (C-D) Full proteome, and (E-F) Phosphoproteome. Colour keys show the log-fold change in each case.

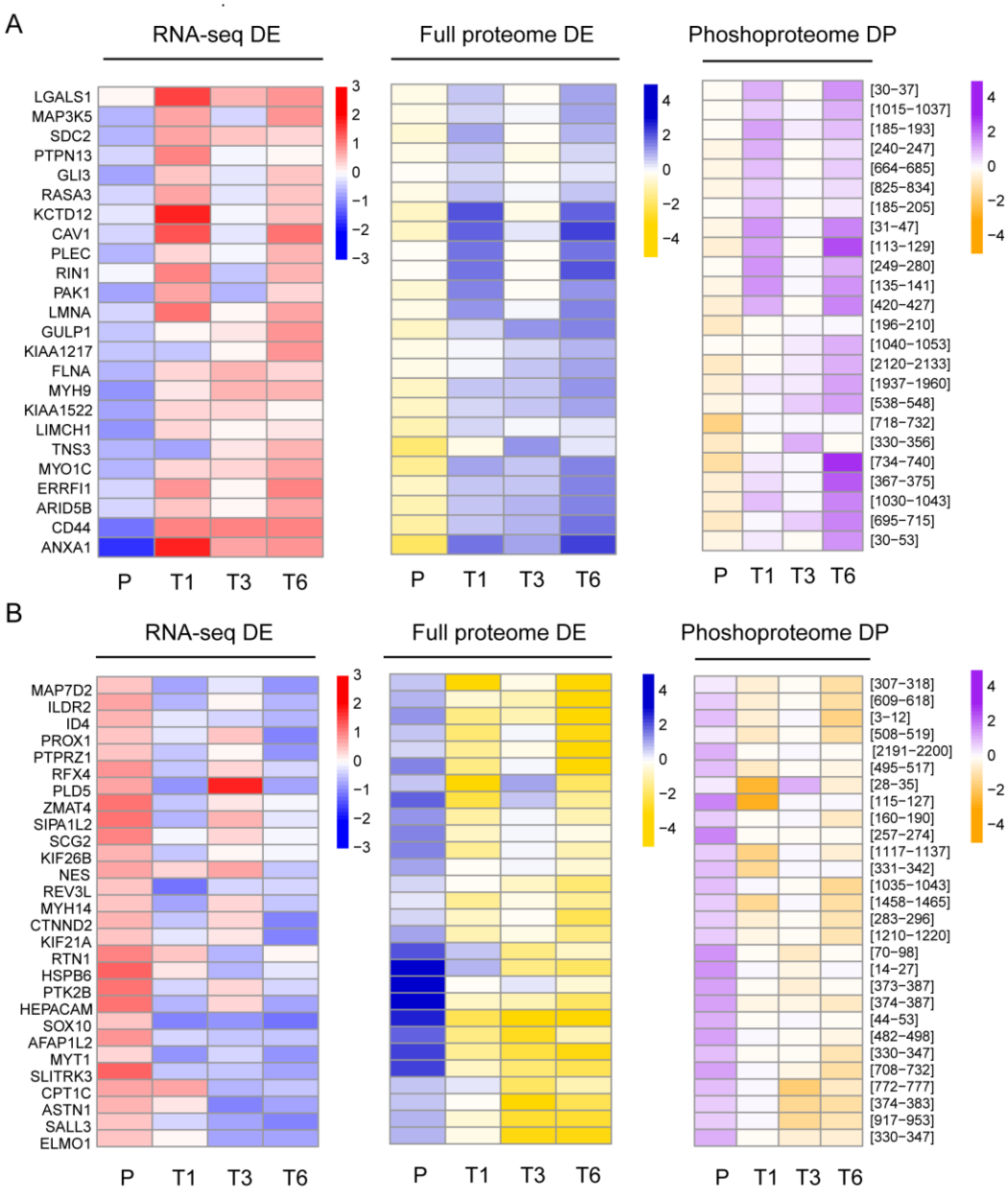


Figure 7-16 Integration of RNA-seq (left), full proteome (middle) and phosphoproteome (right) showing the most common changes in expression, translation and phosphorylation for each target, both for (A) activation and (B) depletion. Targets were selected by log-fold-change greater or less than 1 in all the three platforms in any of the three MEK1/2 clones compared to the parental. Colour of the key shows the log-fold change in each case.

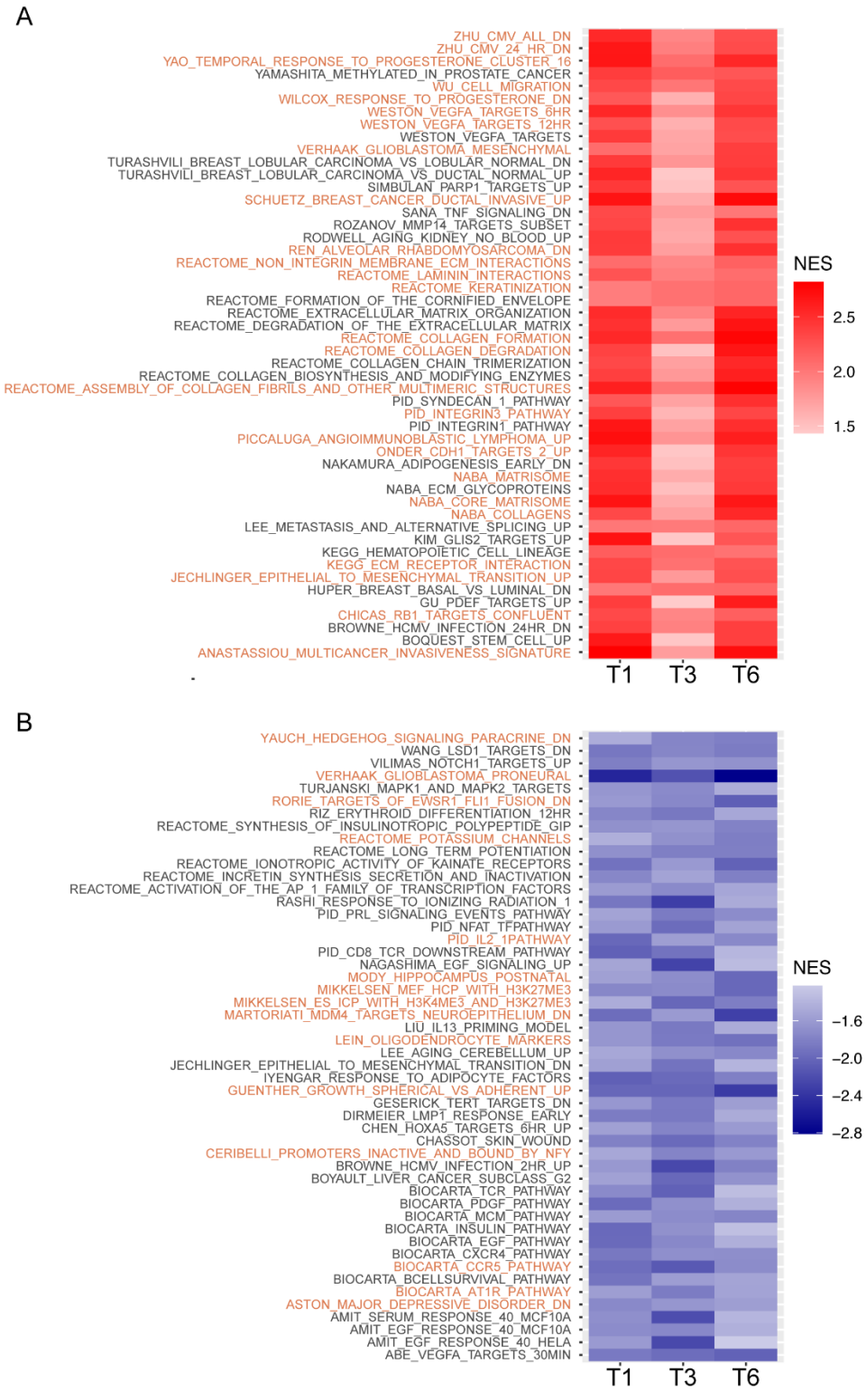
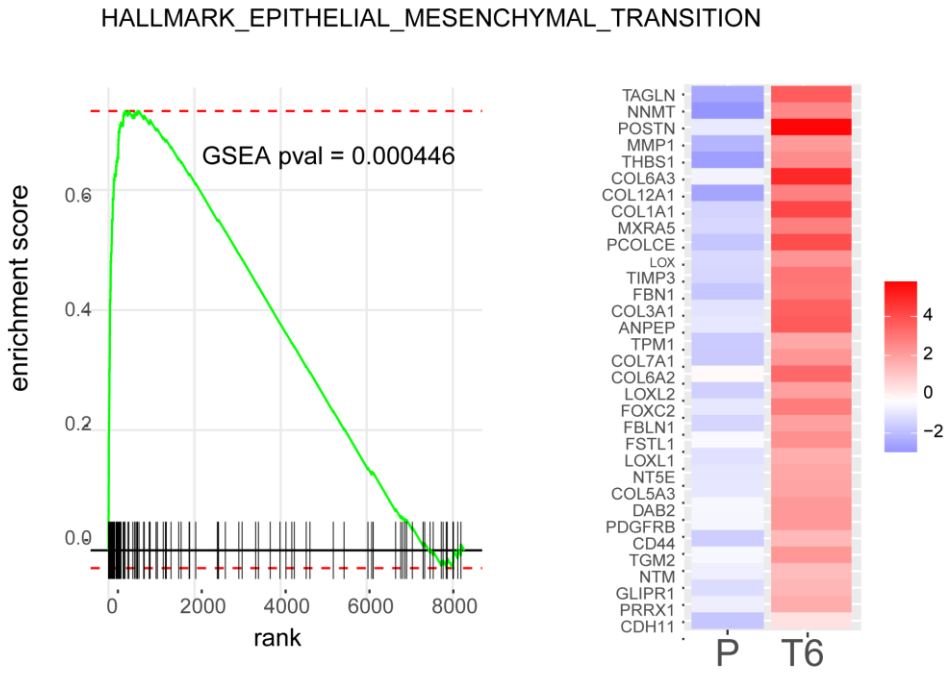
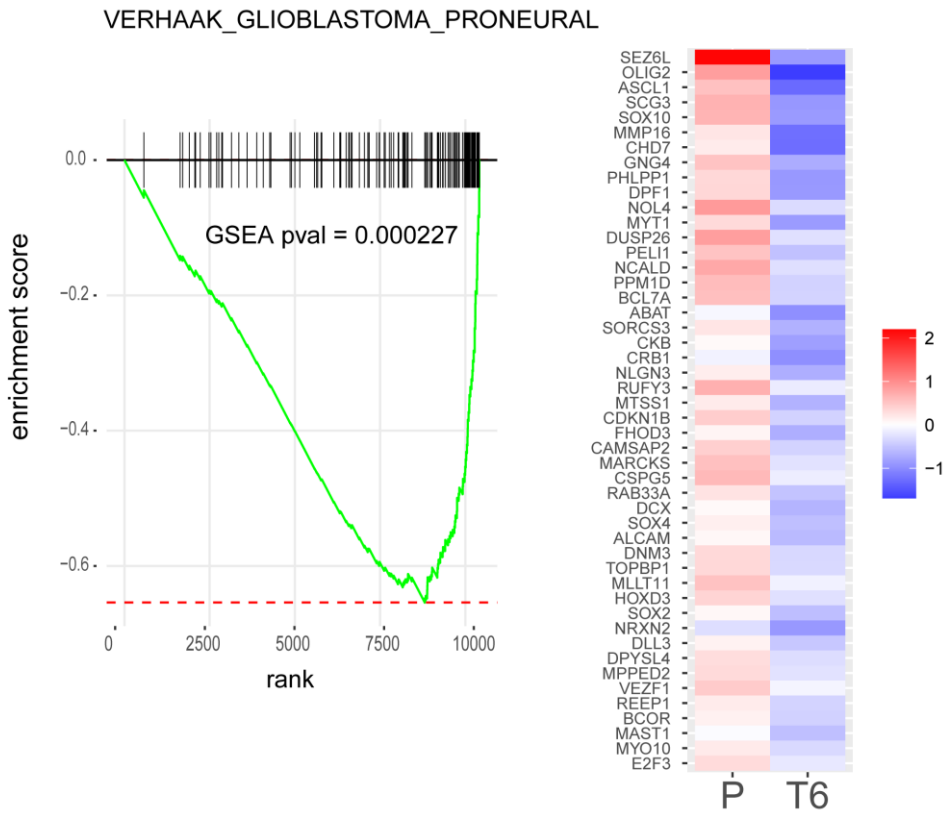


Figure 7-17 Gene set enrichment (GSEA) heatmap of the normalised enrichment scores (NES) from RNA-sequencing of the *MEK1/2* resistant clones compared to BIOMEDE-169 parental. The figure shows (A) up-regulated signatures and (B) down-regulated signatures. GSEA signatures are shown in rows and NES values for each clone are shown in columns. Up-regulated (NES ≥ 1.4) and depleted (NES ≥1.4) GSEA signatures seen in GSEA from full-proteome are coloured in orange.

A


HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION
GSEA pval = 0.000446
enrichment score
0.6
0.4
0.2
0.0
0
2000
4000
6000
8000
rank
TAGLN
NNMT
POSTN
MMP1
THBS1
COL6A3
COL12A1
COL1A1
MXRA5
PCOLCE
LOX
TIMP3
FBN1
COL3A1
ANPEP
TPM1
COL7A1
COL6A2
LOXL2
FOXC2
FBLN1
FSTL1
LOXL1
NT5E
COL5A3
DAB2
PDGFRB
CD44
TGM2
NTM
GLIPR1
PRRX1
CDH11
4
2
0
−2
P
T6

B


VERHAAK_GLIOBLASTOMA_PRONEURAL
GSEA pval = 0.000227
enrichment score
0.0
−0.2
−0.4
−0.6
0
2500
5000
7500
10000
rank
SEZ6L
OLIG2
ASCL1
SCG3
SOX10
MMP16
CHD7
GNG4
PHLPP1
DPF1
NOL4
MYT1
DUSP26
PELI1
NCALD
PPM1D
BCL7A
ABAT
SORCS3
CKB
CRB1
NLGN3
RUFY3
MTSS1
CDKN1B
FHOD3
CAMSAP2
MARCKS
CSPG5
RAB33A
DCX
SOX4
ALCAM
DNM3
TOPBP1
MLLT11
HOXD3
SOX2
NRXN2
DLL3
DPYSL4
MPPED2
VEZF1
REEP1
BCOR
MAST1
MYO10
E2F3
2
1
0
−1
P
T6

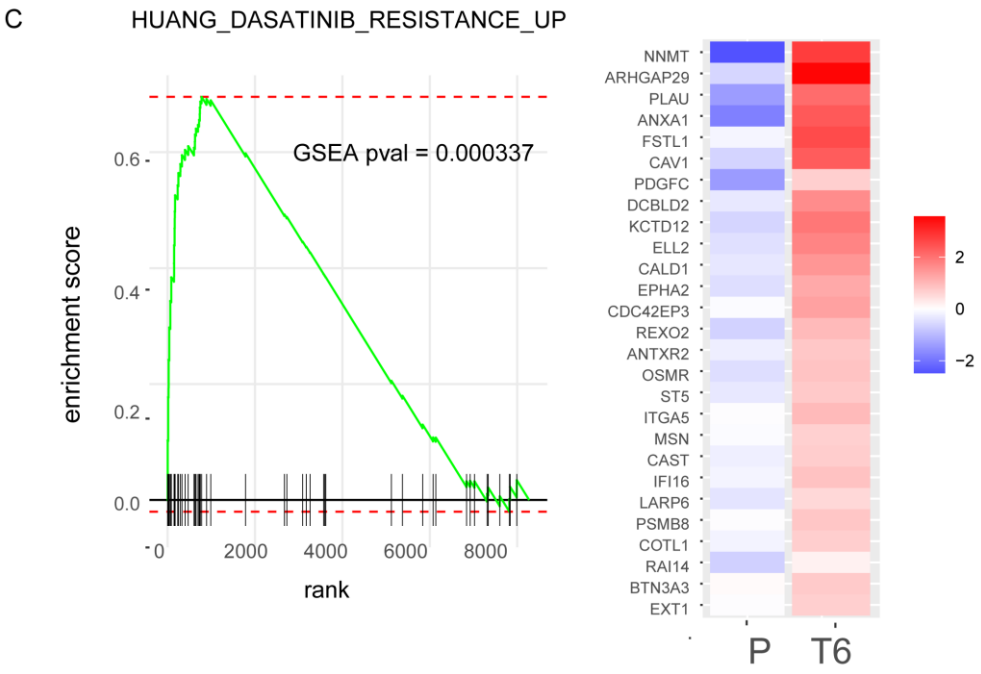


Figure 7-18 GSEA from proteome analysis of *MEK1*/2 resistant T6 clone compared to BIOMEDE-169 parental cells.GSEA plots are shown on the left and heatmap of the normalised enrichment scores (NES) for each signature of proteins up-regulated (red) or down-regulated (blue) are shown on the right for (A) HALLMARK EPITHELIAL MESENCHYMAL TRANSITION, (B) VERHAAK GLIOBLASTOMA PRONEUAL and (C) HUANG DASATINIB RESISTANCE UP. The curves show the enrichment score on the y-axis and the rank list metric on the x-axis. GSEA p-value are reported for each signature. The vertical black lines corresponding to the associated signatures and are ordered by their statistical tests of differential analysis between T6 and BIOMEDE-169 parental for each of the GSEA signatures. FastGSEA package was used to generate these figures.

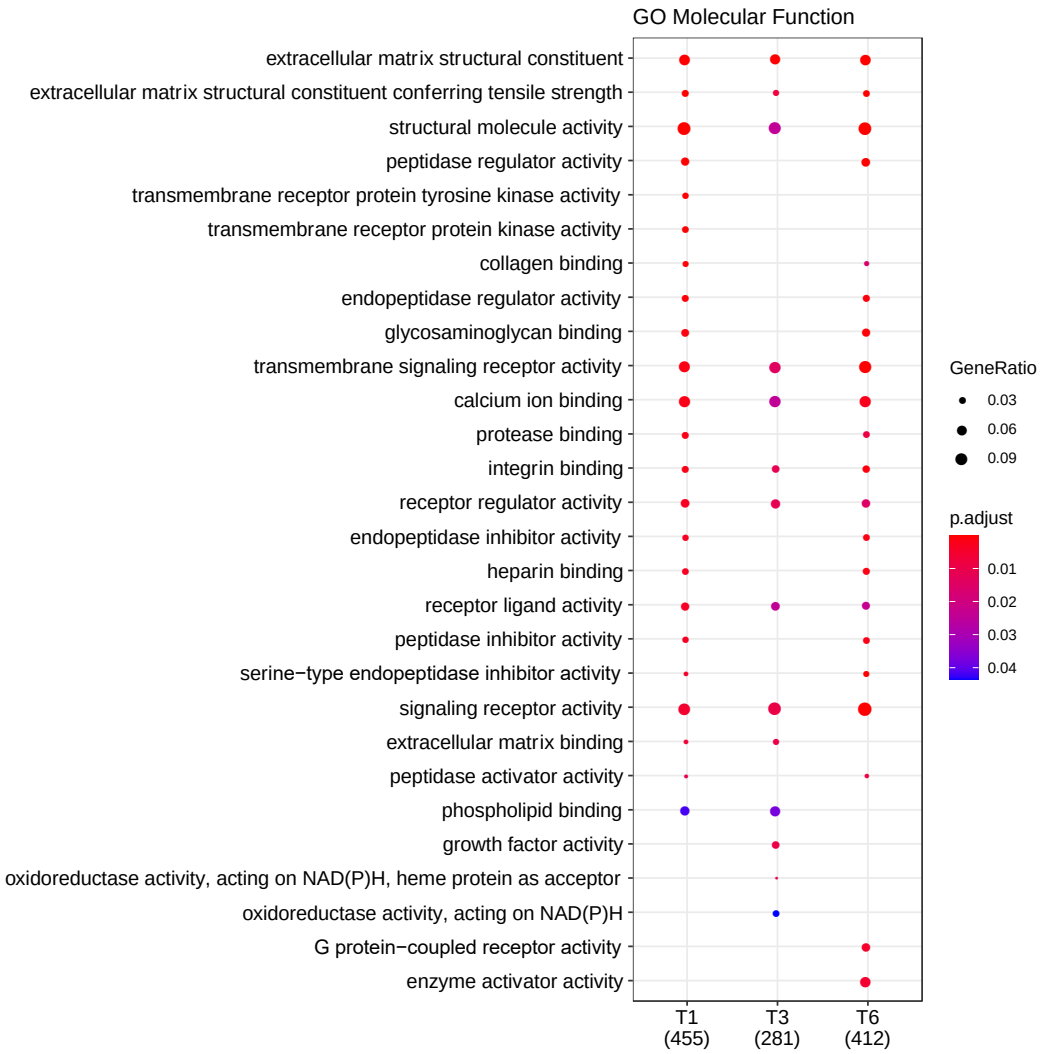


Figure 7-19 Dot plot of gene ontology (GO) analysis from showing shared molecular functions in the *MEK1*/2 resistant clones. The most significant enriched molecular functions are shown in rows and clones in columns. The circle size is equivalent of the number of genes associated to the pathway and coloured according to the adjusted p-value. The package clusterProfiler was used to generate this figure.

Having observed that a dasatinib-sensitive signature was identified in the resistant clones, I then sought to compare the GI50 values for trametinib and dasatinib in the BIOMEDE cultures obtained from the drug screen performed in Chapter 6. This revealed that cultures differentially sensitive to trametinib (GI50 0.016-0.05 µM) were conversely insensitive to the kinase inhibitor dasatinib (GI50 5.6-10 µM) and *vice versa* **{Figure 7-20}**. To identify specific gene expression signatures associated with these differential sensitivities, GSEA was applied to the RNA-seq data from the BIOMEDE cells. Many of the signatures identified in the A.T1, B.T3 and C.T6 resistant clones were present in the inherently trametinib-resistant and dasatinib-sensitive

BIOMEDE cultures **{Figure 7-21}**. These included the signatures associated with mesenchymal transition (VERHAAK GLIOBLASTOMA MESENCHYMAL, ONDER CDH1 TARGETS 2 UP, JECHLINGER EPITHELIAL TO MESENCHYMAL TRANSITION UP) **{Figure 7-22A}**. By contrast, the VERHAAK GLIOBLASTOMA PRONEURAL signature was identified in trametinib-sensitive and dasatinib-resistant cultures **{Figure 7-22B}**, and the HUANG DASATINIB RESISTANCE UP was identified in the inherently trametinib-resistant and dasatinib-sensitive cultures **{Figure 7-22C}**.

In addition, an enrichment of inflammatory/interferon-related gene signatures (HECKER IFNB1 TARGETS, SANA TNF SIGNALLING UP, MOSERLE IFNA RESPONSE) was observed in the inherently trametinib-resistant and dasatinib-sensitive cells. Associated genes with IFN signatures included an up-regulation of interferon IFN-inducible genes (IFI6, IFI44L, IFIT3, IFIT2, IFIT1, IFI27), CCL2, MX1, and MX2 amongst others **{Figure 7-22D}**.

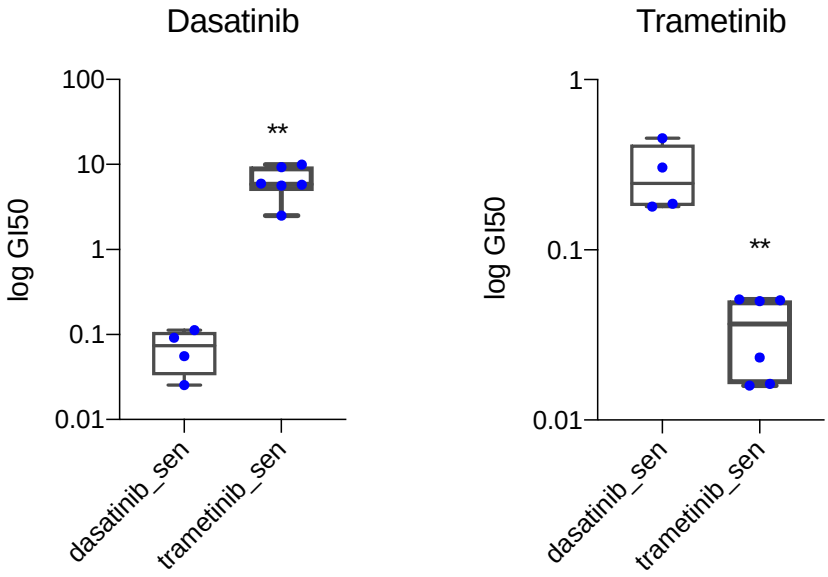
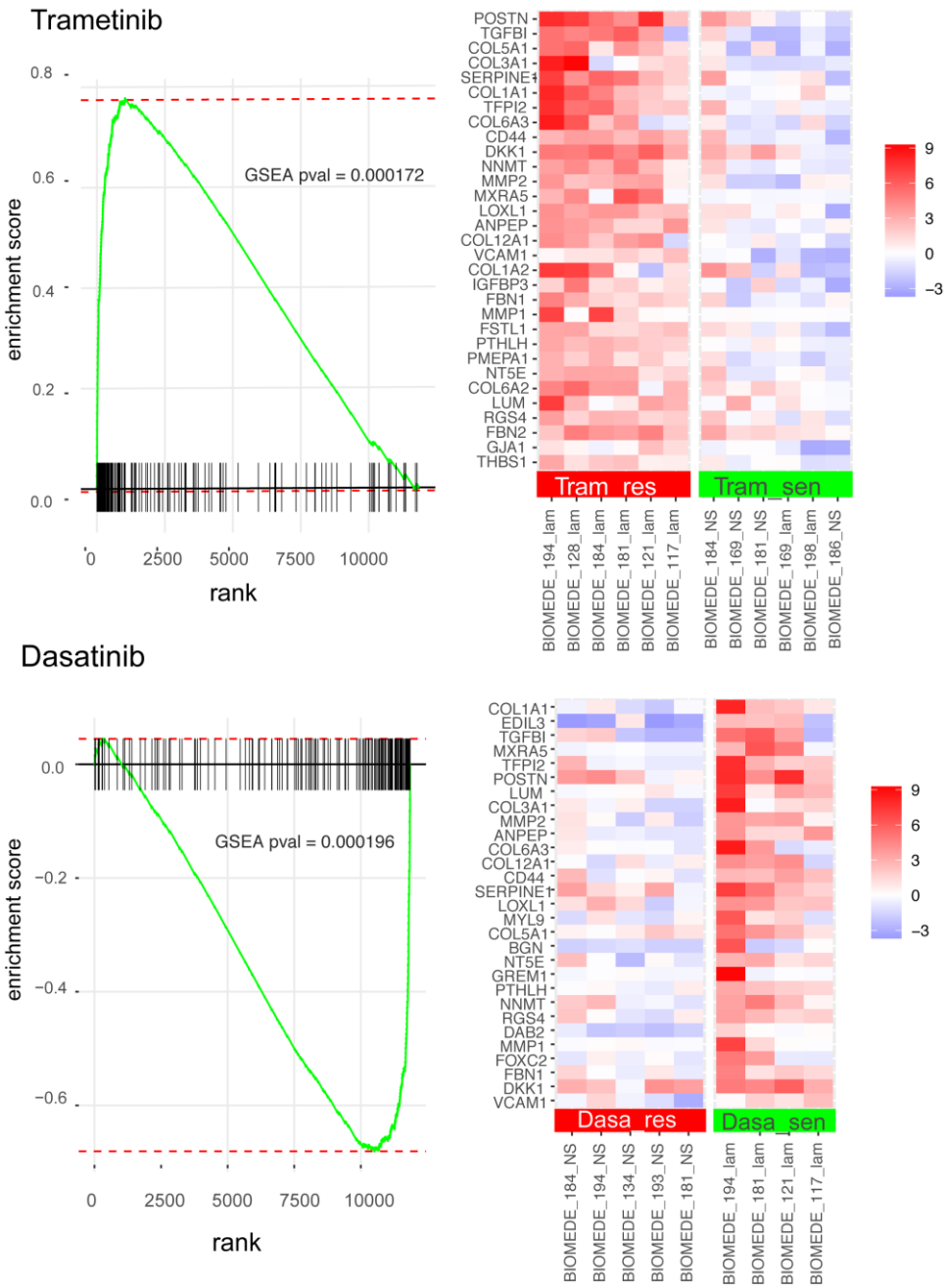


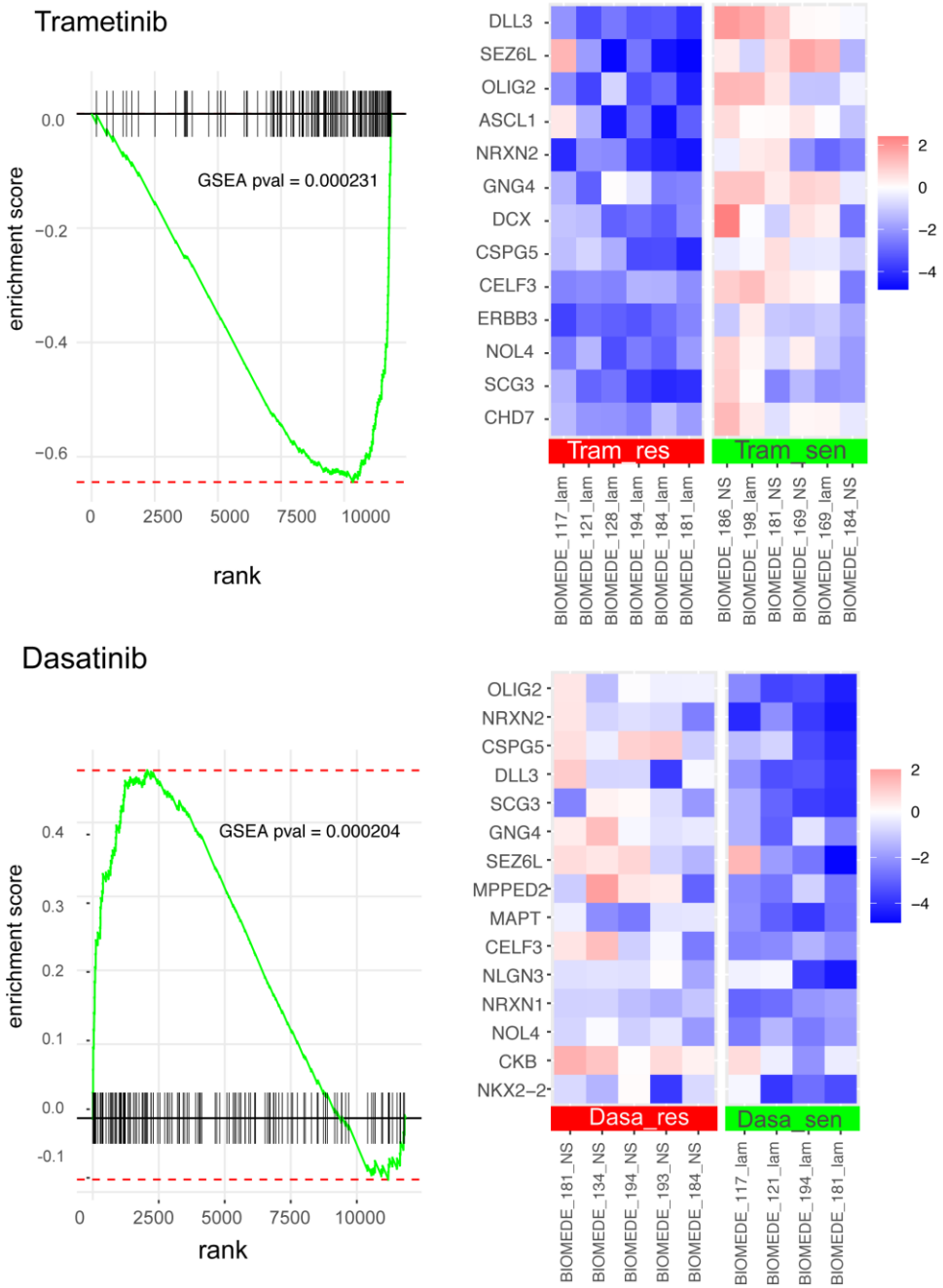
Figure 7-20 Box-plot showing log 10 of GI50 (y-axis) for dasatinib and trametinib across the biomede cultures (x-axis) where samples cultures are divided by sensitivity to dasatinib and trametinib. The whiskers go down to the smallest value and up to the largest, all points the line in the middle represents the mean.

Figure 7-21 Gene set enrichment analysis (GSEA) heatmap of the normalised enchiment scores (NES) for inherently dasatinib and trametinib resistant BIOMEDE-UK cultures. The top 25 signatures up-regulated and down-regulated are ploted in raws sorted by dasatinib resistant and trametinib resistant cultures NES values in columns. GSEA signatures from RNA-seq that were found up-regulated with NES ≥ 1.4 (within the top 25 signatures) and down-regulated NES ≤1.4 (within bottom 25 signatures) in all three *MEK1/2* resistant clones (A.T1, B.T3 and C.T6) are coloured in orange.

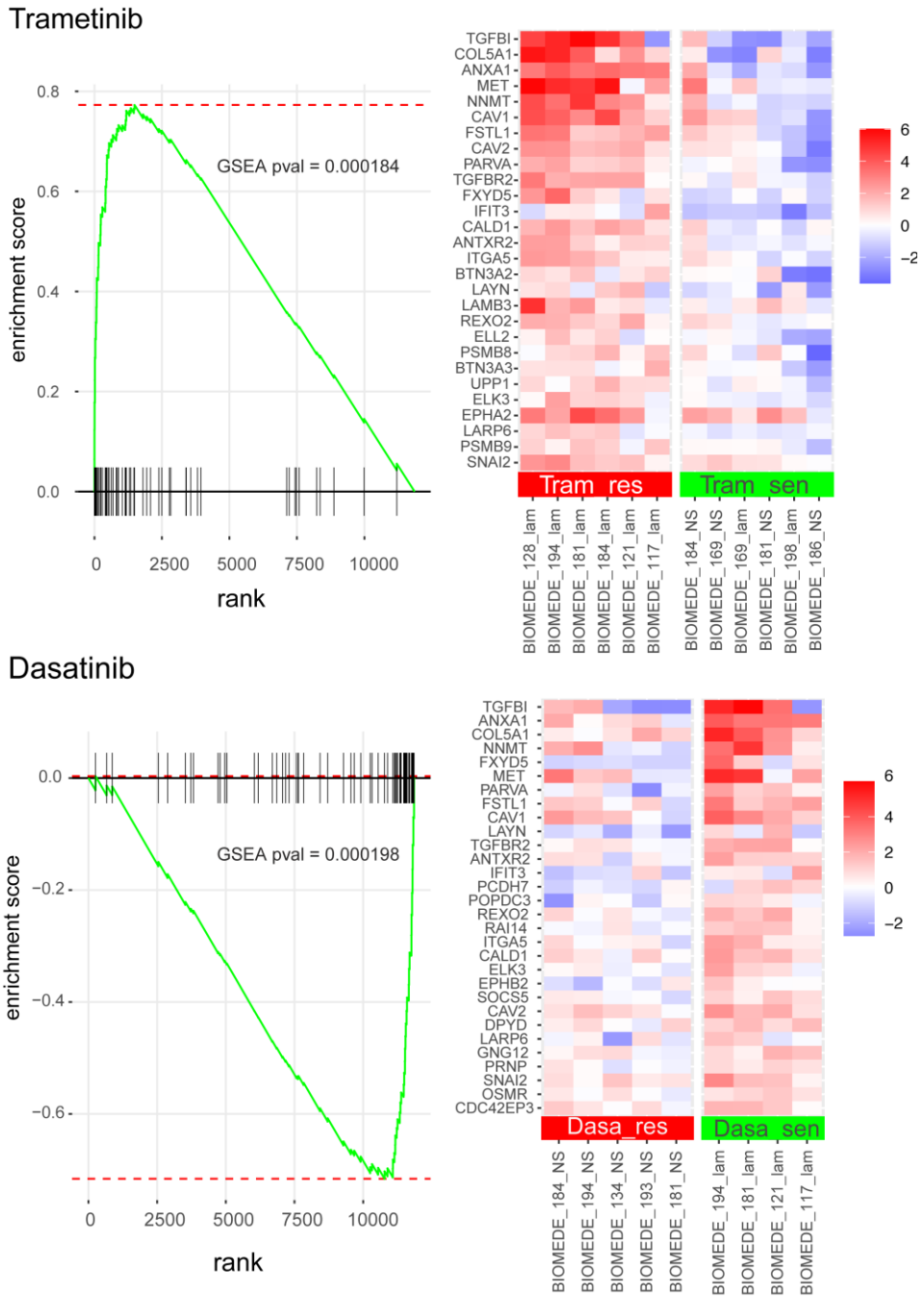
A HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION



B VERHAAK_GLIOBLASTOMA_PRONEURAL



C HUANG_DASATINIB_RESISTANCE_UP



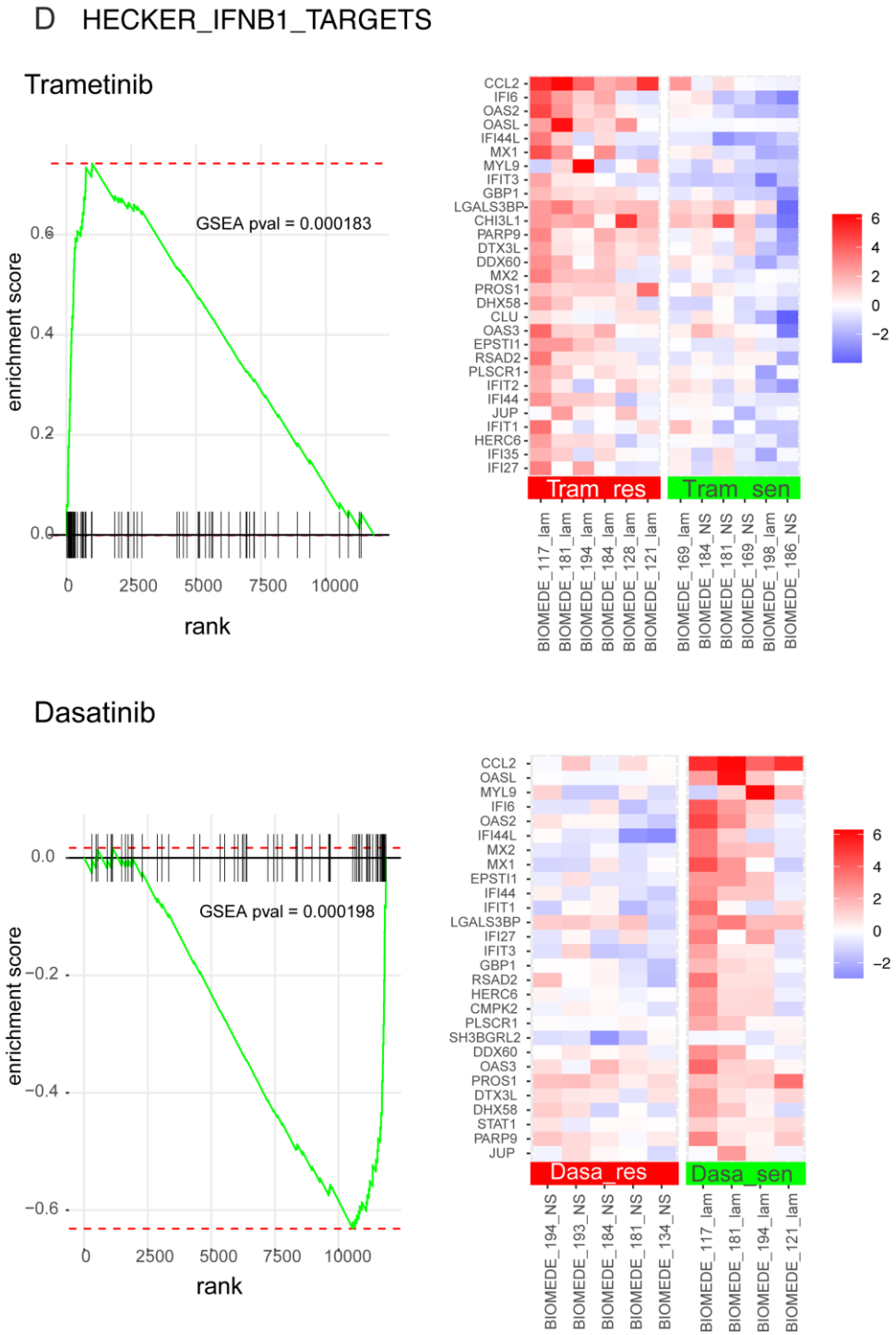


Figure 7-22 Gene set enrichment analysis from RNA-sequencing analysis from inherently trametinib-resistance compared to trametinib-sensitive cultures and inherently dasatinib-resistance compared to dasatinib-sensitive BIOMEDE-UK cultures. GSEA plots are shown on the left and heatmap of the normalised enrichment scores (NES) for each signature of proteins up-regulated (red) or down-regulated (blue) are shown on the right for (A) HALLMARK EPITHELIAL MESENCHYMAL TRANSITION, (B) VERHAAK GLIOBLASTOMA PRONEUAL (C) HUANG DASATINIB RESISTANCE UP and (D) HECKER_IFNB1 TARGETS. The curves show the enrichment score on the y-axis and the rank list metric on the x-axis. GSEA p-value are reported for each signature. The vertical black lines corresponding to the associated signatures and are ordered by their statistical tests of differential analysis between resistant and sensitive cultures for each of the GSEA signatures. FastGSEA package was used to generate these figures.

7.2.4 *MEK1/2* resistant clones are sensitive to dasatinib and to dual trametinib/dasatinib treatment

In order to explore rational alternate treatment approaches in the trametinib-resistant cells, these clones were treated with upstream (dasatinib, multi-RTK) and downstream (ulixertinib, ERK) MAPK pathway inhibitors. No significant differential sensitivity to ulixertinib was observed between the BIOMEDE-169 parental line (GI50 0.8241 µM) compared to A.T1 and B.T3 (GI50 2.271 and 1.060 µM, p=0.2385 and p=0.7884, one-way ANOVA, multiple comparison Dunnett's test, respectively). However C.T6 was significant less sensitive to the ERK inhibitor, with a GI50 ~10 µM (p<0.0001 one-way ANOVA, multiple comparison Dunnett's test) **{Figure 7-23A}.** Conversely, resistant clones were found to be more sensitive to the upstream kinase inhibitor dasatinib A.T1, B.T3 and C.T6 (GI50 0.0930, 0.0400 and 0.0367 µM, respectively) compared to BIOMEDE-169 parental cells (GI50 6.369 µM) (p<0.0001 one-way multiple comparison ANOVA Dunnett's test) **{Figure 7-23B}**.

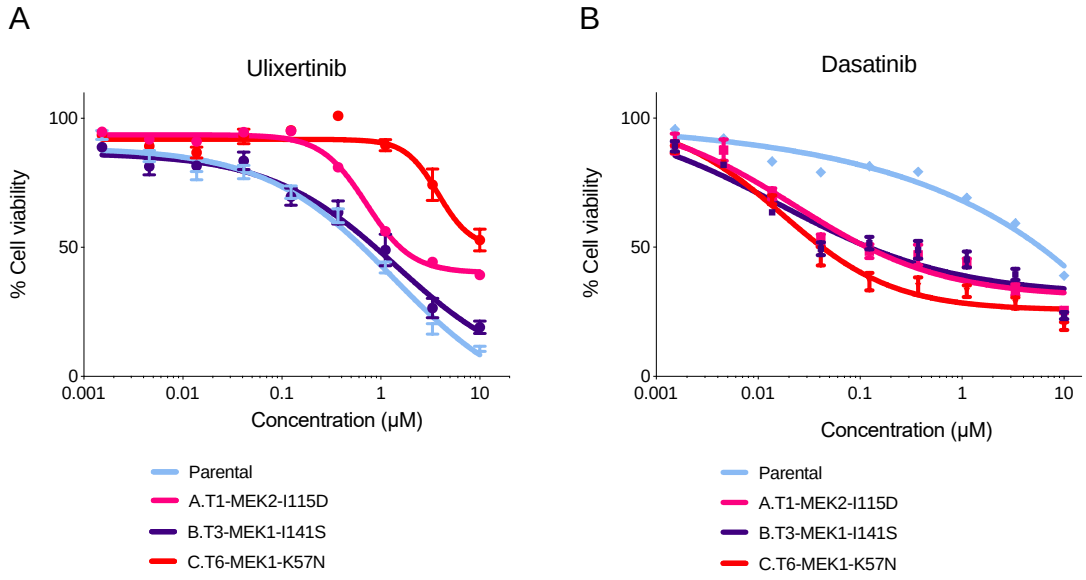


Figure 7-23 Ulixertinib and dasatinib drug response BIOMEDE-169 parental cells, and *MEK1/2* resistant clones. The effect on cell viability (surviving fraction on y-axes) of (A) ulixertinib and (B) dasatinib (x-axes) was assessed in BIOMEDE-169 parental cells, A.T1, B.T3 and C.T6 cells. The errors bars representing the standard error of the mean.

To further investigate the effect on cell viability, three drug combinations were tested by using a 6x10 dose matrix. Combined trametinib and dasatinib, as well as trametinib and ulixertinib, were found to be synergistic or additive at lower concentrations of trametinib in the parental cells (0.0045-0.0137 µM) compared to the resistant clones

(0.12-1.10 µM) **{Figure 7-24}** and **{Figure 7-25}**. The most synergistic Bliss area score from trametinib and dasatinib combination was greatest in BIOMEDE-169 parental cells (19.9) followed by A.T1 (16.16), C.T6 (8.69) and B.T3 (7.67). Whilst C.T6 was the least sensitive culture to ulixertinib as a single agent, in combination with trametinib it was found to have the highest most synergistic Bliss area (45.63) followed by the BIOMEDE-169 parental cells (24.59), B.T3 (22.66) and A.T1 (13.92). The third combination evaluated, ulixertinib and dasatinib, was the least effective at inhibiting cell viability (most synergistic Bliss area scores 3.63-14.35) **{Figure 7-26}**.

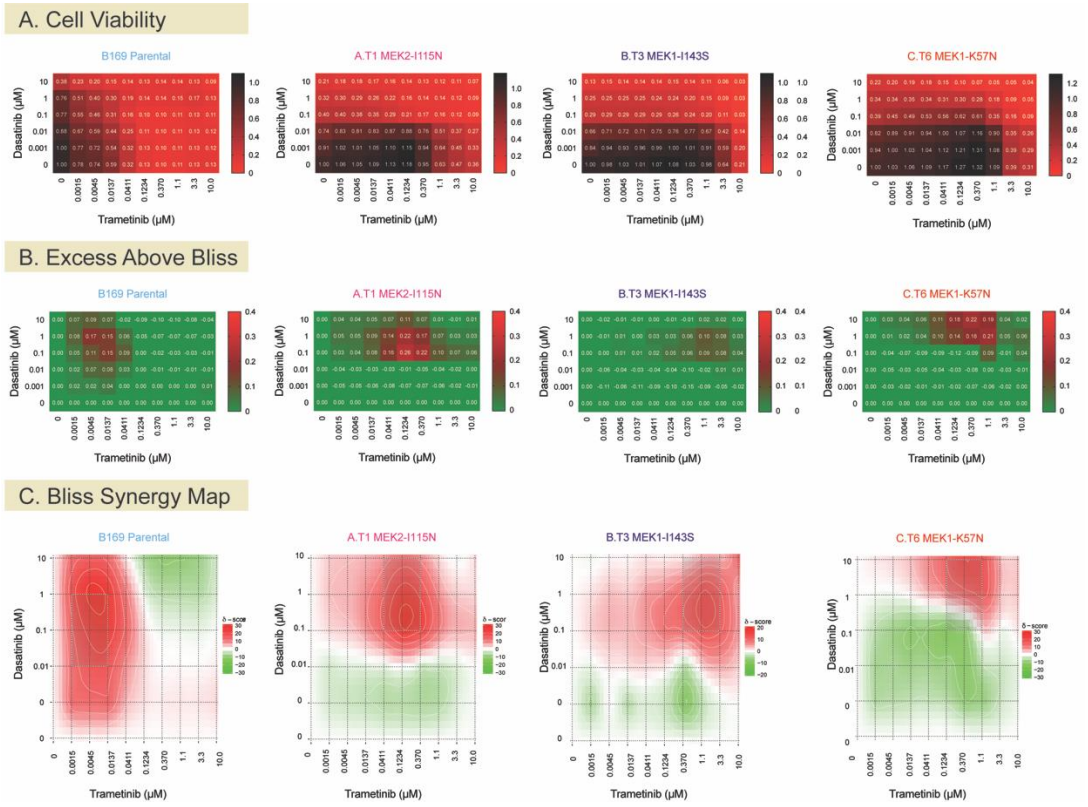


Figure 7-24 Dasatinib and Trametinib matrix combination in BIOMEDE-169 parental cells, and *MEK1/2* resistant clones. 0-10 µM range dose was used for dasatinib and trametinib. (A) Cell proliferation was determined by Cell Titer-glo. (B) Excess above Bliss was calculated and (C) Bliss synergy scores were conducted by using synergy.finder.

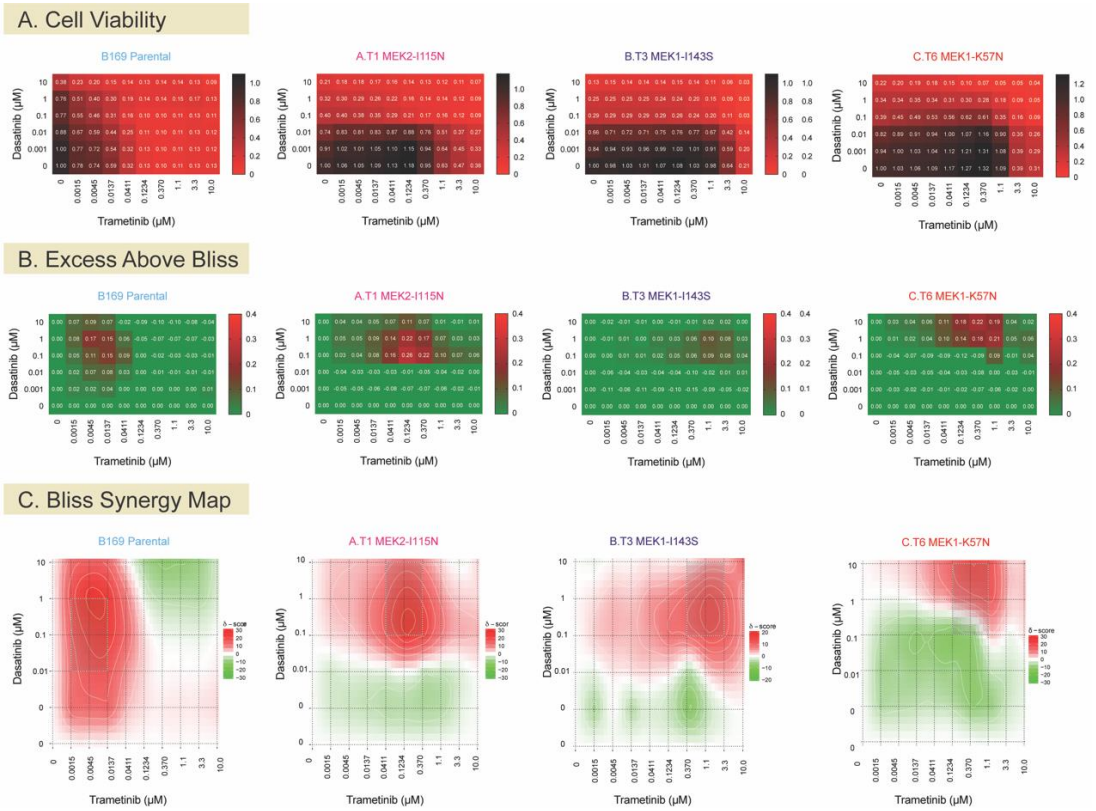


Figure 7-25 Ulixertinib and Trametinib matrix combination in BIOMEDE-169 parental cells, and *MEK1/2* resistant clones. 0-10 μM range dose was used for ulixertinib and trametinib. (A) Cell proliferation was determined by Cell Titer-glo. (B) Excess above Bliss was calculated and (C) Bliss synergy scores were conducted by using synergy.finder .

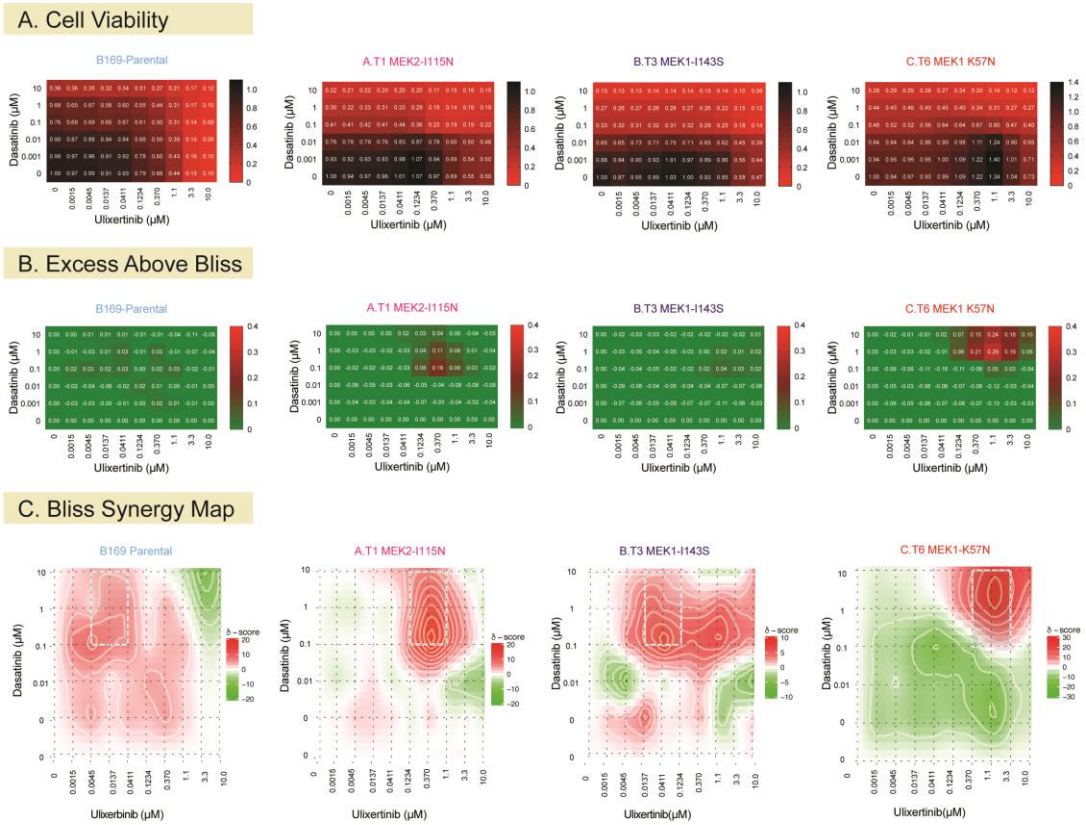


Figure 7-26 Dasatinib and Ulixertinib matrix combination in BIOMEDE-169 parental cells, and *MEK1/2* resistant clones. 0-10 μM range dose was used for dasatinib and ulixertinib. (A) Cell proliferation was determined by Cell Titer-glo. (B) Excess above Bliss was calculated and (C) Bliss synergy scores were conducted by using synergy.finder.

7.3 Discussion

In the setting of a co-clinical trial of prospectively-established DIPG patient-derived *in vitro* models, I show the identification of biomarkers of response to trametinib, a selective reversible inhibitor of MEK1/2 that binds to the allosteric pocket of MEK. Specifically, these involve multiple nodes of the MAPK signalling pathway, and include *PIK3R1*-N564D, *NF1*-I1824S and *BRAF*-G469V. These results suggest a possible rationale for the use of trametinib in DIPG, such as the ongoing phase II clinical trial called TRAM-01 (NCT03363217), which is exploring the use of trametinib in paediatric gliomas harbouring MAPK alterations independently of the tumour entity [41].

PIK3R1-N564D is an oncogenic hotspot mutation known to promote cell survival *in vitro* and oncogenesis *in vivo* [490]. The *PIK3R1*-N564D hotspot mutation is found in various cancers, most frequently in glioma, acute lymphoblastic T cell leukaemia, endometrial and colorectal cancers [378, 491, 492]. The mutation lies within the regulatory subunit of PI3-kinase resulting in loss-of-function, predicted to destabilise protein interaction which may impact tumour suppressive function (FATHMM pathogenic score of 0.99) [489, 493]. In line with the results of this chapter, Cheung and colleagues identified that *PIK3R1* oncogenic mutations activate the MAPK pathway and exhibited sensitivity to MAPK inhibitors [494]. *NF1*-I1824S lies in the helix domain and the change replaces isoleucine with serine at codon 1824. The isoleucine residue is moderately conserved and there is a large physicochemical difference between isoleucine and serine. This variant is a rare missense change, not present in population databases such as ExAC, with uncertain impact on protein function [495]. To the best of my knowledge this variant has only been reported in one individual with neurofibromatosis type 1 [496]. The efficacy of MEK inhibitors has previously been shown in NF1-deficient glioblastoma cell lines [497], in addition to clinical benefit in refractory neurofibromatosis-associated glioma harbouring *NF1* mutation [498-500]. By stochastic selection, an imbalance of variant allele frequency between 3D and 2D cultures was observed in BIOMEDE-181 (*PIK3R1*-N564D) and BIOMEDE-184 (*NF1*-I1824S). Trametinib sensitivity was only identified in the mutant cultures which would support the hypothesis that these mutations were responsible of trametinib efficacy *in vitro.*

BRAF-G469V is a class II *BRAF* mutation within the protein kinase domain, resulting in increased kinase activity and downstream MEK and ERK activation [472]. *BRAF* class II mutations have constitutively activated BRAF dimers independent of RAS activation [472]. Supporting the results of this chapter, *BRAF*-G469V has been shown to confer sensitivity to trametinib in melanoma and lung cancer [473, 501]. With this evidence, the BIOMEDE-169 patient was treated with trametinib at progression, under a compassionate use programme, after 15 weeks of being treated with everolimus, radiotherapy and several steroid pulses. The patient was treated with trametinib for 12 weeks, however soon after progressed and succumbed to the disease. Although an MRI before progression was not conducted to assess response to the MEK inhibitor, it was considered that during the first 9-10 weeks of trametinib, the patient had stable disease before progressing. Notably, I did not observe any efficacy of trametinib as a single agent in the BIOMEDE-169 orthotopic xenograft *in vivo*. One issue with designing such experiments in models with a long tumourigenic latency (167 days) is choosing the treatment window. In this context, different therapeutic strategies will be further evaluated with drug intervals (e.g. 2 weeks of treatment and 1 week of drug holiday) and treatment will continue until end of survival. Another explanation however could be the emergence of resistance to the single MEK inhibitor as previously described in melanoma or colorectal cancers [482-485].

To better understand this process in the context of DIPG, I generated BIOMEDE-169 trametinib-resistant clones *in vitro*. *MEK1/2* mutations were identified in the resistant clones resulting in MAPK pathway up-regulation by an increased expression of pMEK1/2, pAKT or pERK1/2. By using ddPCR, I demonstrated that *MEK1/2* mutations were acquired over time and the VAF increased with the length of trametinib exposure. Furthermore, I provide evidence that the resistance mechanism was irreversible, as cells on “drug holiday” for two months remained insensitive to trametinib. MEK1 and MEK2 exhibit 85% peptide sequence homology [502]. *MEK1*-K57N lies on the helix-A domain within the N-terminal negative regulatory region and is associated with high levels of RAF-independent activation of ERK signalling [476, 503]. Interestingly, *MEK1*-K57N has been attributed to cause resistance to BRAF and MEK inhibitors *in vitro* and in melanoma patients [504, 505]. To the best of my knowledge neither *MEK1*-I141S nor *MEK2*-I115N have been previously detected, and both mutations lie within the protein kinase domain. Notably, *MEK1*-I111N, the equivalent of *MEK2*-I115N, has been demonstrated to confer resistance *in vitro* to allosteric MEK inhibitors [476, 504].

Using RNA-seq and phospho-/total proteomics, a drift from a proneural to mesenchymal phenotype was observed in the three *MEK1/2*-mutant resistant clones. Mesenchymal transition (MT) was induced by up-regulation of important mesenchymal proteins involved in invasion and migration such as CD44, POSTN, collagen-family proteins, CDH11 and FBN1 amongst others. MT has been broadly reported as a hallmark of metastasis and resistance to multitherapy in cancer. In particular, glioma initiating clones displaying drug resistance and radio-resistance have been previously linked to proneural-mesenchymal transition [506]. *In vitro* models derived from BIOMEDE patients that were inherently insensitive to trametinib displayed a more mesenchymal phenotype compared to trametinib-sensitive cultures, which were more proneural. Furthermore, increased expression of biomarkers of response to the multi-kinase inhibitor dasatinib were found in the clones as well as in the inherently trametinib-resistant cultures. This suggested that trametinib-resistant cultures could be sensitive to dasatinib, which I was able to confirm, in both the inherently resistant cultures and selected clones. The combination of trametinib with dasatinib and the downstream ERK inhibitor ulixertinib synergistically inhibited cell proliferation in the parental cells and the *MEK1/2*-mutant clones. Supporting our results, Jing and colleagues found that a proportion of *KRAS*-mutant cell lines resistant to trametinib had a mesenchymal signature associated with increased gene expression of *FN1, S100A4, VIM, ACTA2 and CSH2* [507]. Interestingly, dasatinib has been found to overcome EMT-associated resistance to erlotinib in non-small cell lung cancers [508]. In addition, a study has found that dasatinib sensitises *KRAS*-mutant cancers to trametinib both *in vivo* and *in vitro* [509].

It remains to be explored whether these combinations will show efficacy *in vivo* and could be further translated in the event of resistance to MEK inhibitors in ongoing clinical trials. Or more importantly, these results support the use of an intermittent multitherapy regimen to control population dynamics and potentially prevent emergence of treatment resistance to begin with. The data shown in this chapter, highlights the MAPK pathway as a therapeutic target in DIPG, and show the importance of parallel resistance modelling and rational combinatorial treatments likely to be required for meaningful clinical translation.

CHAPTER 8 : Discussion

Remarkable efforts over the last decade have been made to elucidate the genomic landscape of childhood cancer [12, 22, 24, 73, 76, 187, 193, 339, 340, 342, 368]. These studies have led to the identification of the underlying genetic alterations in paediatric tumours revealing diagnostic, prognostic and predictive biomarkers to molecularly targeted therapies. The rapid expansion in understanding the biology of these tumours have led to world-wide routine genomic testing for many children with cancer [254, 265, 267, 510]. However, paediatric cancers are a very heterogenous group of tumours with distinct mutation signatures and genetic make-up that require further pre-clinical testing in order to identify biomarkers of response to novel treatments. To this end, in this thesis I have developed and validated multiple assays to identify genetic alterations in paediatric patients which have been directly translated into the clinic and are now part of the National Health Service diagnostic service. In addition, in an international phase II clinical trial in DIPG called BIOMEDE, I was able to establish patient-derived *in vitro* models identifying biomarkers of response to bespoke compounds as well as undertaking parallel resistance modelling to further inform novel treatment strategies at tumour relapse.

In Chapter 3, I describe the design and implementation of targeted sequencing assays in a clinically accredited laboratory for children with solid tumours. The capture sequencing panels were able to detect accurately genetic alterations including point mutations, small insertion and deletions at VAF > 5%, as well as structural variants, amplifications and homozygous deep deletions. The design and validation of the first version of the paediatric panel led to a manuscript where I am a first author [354]. The capture panel covers a total of 78 genes for the first version, and 91 for an expanded version. The genes were selected in collaboration with international experts in paediatric oncology. This technology was first applied locally in children with cancer from the Royal Marsden, and later expanded to the rest of the UK. The clinical application of the first 233 patients, described in Chapter 4, has also been published in a manuscript on which I am joint first author [383]. The sequencing success rate from FFPE samples was 82%, and at least one alteration was detected in 70% of sequenced samples with an actionability alteration of 51% of the patients. The paediatric panel sequencing assay is now offered to every child from the UK with a solid tumour as part of routine diagnostic testing, with a turnaround time of 3-4 weeks

from sample dispatch to reporting. The panel has been reviewed by NICE and is recommended for the use on clinical biopsy material in children with solid tumours. The value of targeted gene sequencing as a practical and cost-effective clinical tool to enable improved diagnosis, prognostication and therapeutic stratification for children with cancer was demonstrated.

However, as technology and knowledge of the biology of paediatric cancers evolves, more comprehensive genetic characterisation is needed to better define the biology of these tumours and obtain the best treatment possible. In this context, the Stratified Medicine Programme Paediatrics (SMPaeds), is currently being validated in the UK to deliver multi-omics profiling in children and young adults with solid tumours at time of relapse or refractory cancers. This approach will support the international biomarker-driven trial ESMART. To do this, comprehensive analysis will be performed including methylation sequencing using EPIC BeadChip, whole exome sequencing (WES), RNA-seq and low coverage whole genome sequencing (lcWGS). These approaches complement panel sequencing, with methylation profiling in particular critical for accurate diagnostics of brain tumours [20] and increasingly relevant for sarcomas [206]. RNA-seq enables the identification of structural variants and expression profile assessment, by performing WES alterations in other genes not included in the panel would be identified, and lcWGS will is used to detect genome-wide copy number profile. Furthermore, where sufficient tissue is available, concurrent analysis via the National Health Service England WGS programme (100,000 Genomes Project) will be compared with SMPaeds genomic and clinical data. This approach will provide an unbiased assessment of the clinical utility and cost effectiveness of multiple different modalities to enable formal recommendations for implementation into routine molecular diagnostics. Similar strategies, (INFORM, MAPPYACTS, ITHER, ZERO, and Precision Oncology for Young People Program), are conducted by other countries to support national and international precision medicine programs for paediatric oncology (INFORM2, ESMART and Paediatric MATCH) [267, 269, 270].

Although precision medicine programmes and specific tumour type trials using novel treatment strategies are becoming more available, the lack of “actionable” alterations that can be matched to a drug at present remains unfortunately very low. Moreover, some of the current trials are not accessible world-wide, which creates an equality problem across patients from different geographical locations. Multiple international

platforms have emerged to rapidly implement the pre-clinical efforts undertaken into the clinic in a systematic manner including the European Innovative Therapies for Children with Cancer Paediatric Preclinical Proof of Concept Platform (ITTC-P4) and the US Pediatric Preclinical Testing consortium (PPTC). Another challenge that arises for the large amount of data generated from multiple platforms is variant interpretation, data integration and how to prioritise relevant targets for clinical decision-making. A consensus hierarchal algorithm strategy will need to be established to better define and assign molecular alterations to targeted therapies.

The role of genetic predisposition variants in childhood cancer is becoming increasingly apparent [13, 14, 253, 363]. Owing to ethical and consent constraints in the studies included in this thesis, we were not allowed to report germline variants because there was no direct link with a genetics clinic. However, this is extremely important and a recognised deficiency of these studies. Ethical approval for simultaneous germline variant detection is now in place for routine patient diagnostics and is being reported via an accredited genetics clinic at Great Ormond Street Hospital so that the appropriate genetic counselling can be provided.

Through the application of the fusion-panel in the HERBY clinical trial samples, I describe in Chapter 4 the identification of a novel internal tandem duplication in the *NTRK2* gene, predominantly in *H3F3A*-K27M thalamic glioma. Whilst this alteration has not yet been reported in cancer, activating gene alteration through tandem duplications are seen involving *FGFR1* in paediatric glioma [34] and *FLT3* in acute myeloid leukaemia [511]. This discovery provides a rationale to examine this type of genetic modification in pHGG, supporting the use of NTRK inhibitors for which there are existing clinical trials for children with solid tumours with TRK fusions. However, as this is a novel variant, further analysis exploring functional characterisation will be explored in the Jones lab. In order to decipher the role of *NTRK2* tandem duplication, gene editing by using CRISPR/Cas9 will be used to knock-out (KO) ITD-*NTRK2* in a positive-derived model and different isoforms of the ITD-*NTRK2* will be knock-in (KI). These engineered models will be used to determine pathogenicity, pathway modulation and sensitivity to therapeutic agents. For the latter, a high-throughput combinatorial screen will be undertaken in collaboration with Mimi Bandopadhayay at Dana-Farber Cancer Institute, Boston. In addition, this finding highlights the probability that other SVs may be present in these tumours which have up to now been missed, which long read sequencing and/or deeper RNA and WGS may help

identify. Currently our laboratory is performing a large meta-analysis from retrospective data where RNA-seq and/or WGS is available, integrating different pipelines of SV detection.

Whilst sequencing tissue samples of patients is crucial, liquid biopsies offer the possibility of a non-invasive source for tumour genotyping and disease monitoring. The preliminary findings described in Chapter 5 demonstrate that ctDNA can be obtained and tumour variants can be detected in pHGG and DIPG patients from CSF and less successfully from plasma. The development of deep sequencing approaches with a higher detection rate might increase the utilisation of plasma assessment of ctDNA in patients with brain tumours. In this context, in the Hubank lab at the ICR and RMH, they are currently developing a bespoke ctDNA panel for paediatric solid tumours to be incorporated as part of the diagnostics pipeline. They are using a novel technology that combines ultra-deep capture sequencing with unique molecular identifiers. By a large-scale validation study, they are comparing tumour and serial ctDNA findings in children with cancer to define the clinical utility of ctDNA analysis. In my opinion, the incorporation of ctDNA from liquid biopsies in the design of future clinical trials will be critical. This has already been applied in adult cancers to monitor patient response to treatment and importantly to exploit the emergence of resistance alterations to targeted therapies as a consequence of cancer evolution [279, 281, 282, 284] . Using deep-sequencing of ctDNA at different time-points of the patient disease will enable clinicians to track the evolution of sub-clonal populations of cancer cell, as well as identify acquired mutations that might inform the choice of appropriate therapeutics.

In a prospective biopsy-stratified clinical trial BIOMEDE, which is described in Chapter 6, the paediatric panel sequencing was used to identify actionable alterations to guide therapeutic of choice of DIPG patients at relapse. At least one molecular alteration was found in all the patients with 79% patients of the cases with Tier 2B and/or Tier 3 alterations. Treatment recommendations based on the molecular profiling were reported to the treating physician via the UK trial steering group. Currently, the clinical information of these patients has not been released to our laboratory, however, as the trial data becomes available, patients receiving treatment recommendations according to their molecular profile and their clinical outcome will be disclosed as part of a future publication. Additionally, although not discussed in the thesis, molecular profiling was still performed in patients that were not eligible for the trial and findings

still reported to the treating clinician. Notably, in one patient at the Royal Marsden Hospital who was ineligible for the trial, *HIST2H3A*-K27M and *ACVR1*-G328V mutations were identified. Pre-clinical data from the Jones lab has identified *in vivo* efficacy of combined everolimus and vandetanib in *ACVR1*-mutant DIPG, and Dr Fernando Carceller was able to treat this patient with the combination under a compassionate use programme. In this patient, the ctDNA panel mentioned above has been applied enabling the identification of molecular alterations at longitudinal plasma samples.

The development of patient-specific *in vitro* and *in vivo* models in parallel to the patient disease allow the testing of molecularly-driven hypotheses in a highly relevant biological context. Such pre-clinical initiatives enable the identification and confirmation of specific genetic vulnerabilities which may provide novel treatment strategies. This has been particularly challenging to do in DIPG as disease progression is very rapid and the generation of these models can take a long period of time. To expand the scope of personalised medicine, I established patient-derived models that recapitulated the original disease from the BIOMEDE trial. Detailed molecular profiling (methylation BeadArray, exome, RNAseq) was linked to drug screening in newly-established patient-derived models of DIPG *in vitro* and *in vivo* and is described in Chapter 6. Efficacy to the three compounds of the clinical trial (dasatinib, erlotinib and everolimus) was assessed *in vitro*, resulting to a wide range of differential sensitivity across the 2D and 3D models. It remains to be explored whether the drug screening conducted in the avatars could have predicted treatment efficacy in the original patient. This will be investigated once the clinical data is collected and will be incorporated into a manuscript.

The personalised drug screening evaluated in Chapter 6 was a feasibility study hence the low number of drugs tested per patient (n = 11). This is being continuously expanded in the Jones lab, currently with a library of 44 compounds, many of which are accessible in clinical trials and/or FDA approved. The drugs include a wide range of mTOR/PI3K inhibitors, upstream and downstream MAPK inhibitors, cyclin-dependent kinase (CDK) inhibitors, bromodomain and extra-terminal (BET) inhibitors, janus-kinase 2 (JAK2) inhibitors, deacetylases and demethylases inhibitors, MDM2 inhibitors, and PARP inhibitors amongst others. The drug screening is being undertaken in the cultures described in this chapter but also in the more recent models that were established from the BIOMEDE trial by other members of the lab.

Nevertheless, I was able to identify differential drug sensitivities linked to specific genetic vulnerabilities, with such agents currently in clinical trials for children with cancer and specifically with DIPG. These include the WEE1 kinase inhibitor AZD1775, and the MEK inhibitor trametinib.

Sensitive models to the WEE1 kinase inhibitor were found to harbour *TP53* frameshift or stop coding mutations associated with lack of gene expression compared to missense or wild-type *TP53* cultures. AZD1775 has been described to be a radiosensitiser in different tumour types inducing DNA damage [445, 447]. The combination of AZD1775 and irradiation will be further explored *in vitro* and *in vivo* in our DIPG models with distinct *TP53* statuses. AZD1775 is a targeted agent included the international trial ESMART, with one of the arms including it with the chemotherapy agent carboplatin for *TP53*-mutant tumours. Interestingly, in the session "Ordering genomic changes as actionable targets in paediatric cancers" at last year's AACR meeting in Atlanta, some of the results from the ESMART trial were presented by Gudrun Schleiermacher. During her talk there was some controversial discussion about the utility of *TP53* mutation status as biomarker of response to AZD1775. I would like to emphasise the importance of validating my findings in a larger cohort of samples in order to confirm if the type of *TP53* mutation will inform of sensitivity to AZD1775 and whether this is associated to a specific tumour type or is also observed across any childhood cancer.

MAPK pathway alterations were found to be biomarkers of response to trametinib, including the non-canonical *BRAF*-G469V mutation, the hotspot *PIK3R1*-N564D and *NF1*-I1824S. However, treatment of trametinib in the *BRAF*-G469V CDX model and the original patient at relapse failed to elicit a significant response. To better understand the mechanism of resistance to trametinib in DIPG, I used the *BRAF*-G469V *in vitro* model to generate resistance to trametinib through continuous drug exposure, as described in Chapter 7. To the best of my knowledge this is the first-time resistant cells have been generated from a patient-derived DIPG model. *MEK1*/2 mutations were identified in three independent clones derived from the same parental primary culture. Further analysis (RNA-seq and phospho-/total proteomics) identified a mesenchymal transition phenotype in the clones which was additionally detected in the inherently trametinib-resistant BIOMEDE primary cultures. By contrast, glioblastoma proneural biomarkers were enriched in the parental cells as well as in the trametinib-sensitive primary cells. Besides the proneural to mesenchymal shift,

biomarkers of response to dasatinib were increased in the trametinib-resistant clones and trametinib-resistant cultures. This led to the confirmation that there was an inverse sensitivity between trametinib and dasatinib cultures as well as inverse correlation of their expression profile. I identified a synergistic effect from the combination of dasatinib and trametinib as well as ulixertinib and trametinib in the resistant clones. This can be promising therapeutic combination for targeting *BRAF*-G469V DIPG tumours. There were two other models that also showed sensitivity to trametinib harbouring MAPK alterations involving *NF1* and *PIK3R1*, genes that are commonly mutated in DIPG. It remains to be explored whether the similar mechanism of resistance would emerge in these cultures, and if the combinations found to be effective will show a synergistic effect. If this hypothesis can be proved, the combinatorial treatment found could be further extended to DIPG tumours with other mechanisms of MAPK pathway activation.

The use of a barcoding library to assign a molecular barcode to each cell has the advantage of being able to identify evolutionary processes over time, in response to drug treatments for example. This can be used to identify resistant clones, *in vitro* and *in vivo*, and has the capacity to distinguish if the mechanism was due to pre-existing clonal expansion or acquired over time. Unfortunately, although this was a methodology I wanted to undertake, due to capacity and time limitations over the course of this PhD, I was not able to conduct such experiments. Nevertheless, there will be a new PhD student starting in our lab in October 2020 following on from the work presented in my thesis, in collaboration with Mimi Bandopadhayay at Dana Farber Cancer Institute, Boston. The use of a novel selectable CRISPR-Cas9 barcoding library (EvoSeq), encompassing 4,000,000 unique guides, will allow for the tracking of cell lineages under the challenge of specific targeted-agents including MEK inhibitors. The barcodes of the surviving cells from different replicates will be compared to the pre-treatment population of cells. If resistance is acquired the barcode distribution will be stochastic, however, if resistance emerge from pre-existing clones the barcode distribution will be very similar amongst the replicates. I am currently drafting a manuscript with the results of Chapter 6 and 7, which excitingly have been the basis of an awarded grant from Children with Cancer UK to the Jones lab. This will allow to expand my work including the implementation of the state-of-art methodology described above, CRISPR-cas9 barcoding library, to track resistance *in vitro* and *in vitro* to MAPK inhibitors.

To summarise, I describe the use of panel sequencing in a clinically accredited laboratory applied to childhood cancers as well as the use of molecular profiling and establishment of patient-derived cell models from several clinical trials in pHGG and DIPG. In future clinical trials, the combination of genomic sequencing at biopsy, longitudinal liquid biopsy samples to track biomarker of response and identify resistance alterations, as well as the establishment of “avatars” to enhance the utility of precision medicine will be key strategies towards more successful treatments. With the ultimate goal to find cures for children with cancer, the integration of such comprehensive studies will be only possible with the collaboration of multidisciplinary fields between researchers, pathologists, treating physicians, radiologists, industry partners.

BIBLIOGRAPHY

1. Stratton, M.R., P.J. Campbell, and P.A. Futreal, The cancer genome. Nature, 2009. 458(7239): p. 719-24.
2. Siegel, R., D. Naishadham, and A. Jemal, Cancer statistics, 2012. CA Cancer J Clin, 2012. 62(1): p. 10-29.
3. Kopp, L.M., et al., Late effects in adult survivors of pediatric cancer: a guide for the primary care physician. Am J Med, 2012. 125(7): p. 636-41.
4. Vinchon, M., et al., Morbidity and tumor-related mortality among adult survivors of pediatric brain tumors: a review. Childs Nerv Syst, 2011. 27(5): p. 697-704.
5. Pritchard-Jones, K. and R. Sullivan, Children with cancer: driving the global agenda. Lancet Oncol, 2013. 14(3): p. 189-91.
6. Downing, J.R., et al., The Pediatric Cancer Genome Project. Nat Genet, 2012. 44(6): p. 619-22.
7. Filbin, M. and M. Monje, Developmental origins and emerging therapeutic opportunities for childhood cancer. Nat Med, 2019. 25(3): p. 367-376.
8. Chen, X., A. Pappo, and M.A. Dyer, Pediatric solid tumor genomics and developmental pliancy. Oncogene, 2015. 34(41): p. 5207-15.
9. Baker, S.J., D.W. Ellison, and D.H. Gutmann, Pediatric gliomas as neurodevelopmental disorders. Glia, 2016. 64(6): p. 879-95.
10. Huether, R., et al., The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. Nat Commun, 2014. 5: p. 3630.
11. Ma, X., et al., Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. Nature, 2018. 555(7696): p. 371-376.
12. Gröbner, S.N., et al., The landscape of genomic alterations across childhood cancers. Nature, 2018. 555(7696): p. 321-327.
13. Zhang, J., et al., Germline Mutations in Predisposition Genes in Pediatric Cancer. New England Journal of Medicine, 2015. 373(24): p. 2336-2346.
14. Brodeur, G.M., et al., Pediatric Cancer Predisposition and Surveillance: An Overview, and a Tribute to Alfred G. Knudson Jr. Clin Cancer Res, 2017. 23(11): p. e1-e5.
15. Ramkissoon, S.H., et al., Clinical targeted exome-based sequencing in combination with genome-wide copy number profiling: precision medicine analysis of 203 pediatric brain tumors. Neuro Oncol, 2017. 19(7): p. 986-996.
16. Cole, B.L., et al., Year 1 in the Molecular Era of Pediatric Brain Tumor Diagnosis: Application of Universal Clinical Targeted Sequencing in an Unselected Cohort of Children. JCO Precision Oncology, 2018(2): p. 1-13.
17. Steliarova-Foucher, E., et al., International incidence of childhood cancer, 2001-10: a population-based registry study. Lancet Oncol, 2017. 18(6): p. 719-731.
18. Ostrom, Q.T., et al., CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2009-2013. Neuro Oncol, 2016. 18(suppl_5): p. v1-v75.
19. Smith, M.A., et al., Declining childhood and adolescent cancer mortality. Cancer, 2014. 120(16): p. 2497-506.
20. Capper, D., et al., DNA methylation-based classification of central nervous system tumours. Nature, 2018. 555(7697): p. 469-474.
21. Sturm, D., et al., New Brain Tumor Entities Emerge from Molecular Classification of CNS-PNETs. Cell, 2016. 164(5): p. 1060-72.
22. Mackay, A., et al., Integrated Molecular Meta-Analysis of 1,000 Pediatric High-Grade and Diffuse Intrinsic Pontine Glioma. Cancer Cell, 2017.

23. Taylor, M.D., et al., Molecular subgroups of medulloblastoma: the current consensus. Acta Neuropathol, 2012. 123(4): p. 465-72.
24. Pajtler, K.W., et al., Molecular Classification of Ependymal Tumors across All CNS Compartments, Histopathological Grades, and Age Groups. Cancer Cell, 2015. 27(5): p. 728-43.
25. Ostrom, Q.T., et al., Alex's Lemonade Stand Foundation Infant and Childhood Primary Brain and Central Nervous System Tumors Diagnosed in the United States in 2007-2011. Neuro Oncol, 2015. 16 Suppl 10: p. x1-x36.
26. Louis, D.N., et al., The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta Neuropathologica, 2016. 131(6): p. 803-820.
27. Bandopadhayay, P., 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. Pediatr Blood Cancer, 2014. 61(7): p. 1173-9.
28. Krishnatry, R., et al., Clinical and treatment factors determining long-term outcomes for adult survivors of childhood low-grade glioma: A population-based study. Cancer, 2016. 122(8): p. 1261-9.
29. Szudek, J., et al., Associations of clinical features in neurofibromatosis 1 (NF1). Genetic Epidemiology, 2000. 19(4): p. 429-439.
30. Adriaensen, M.E., et al., Prevalence of subependymal giant cell tumors in patients with tuberous sclerosis and a review of the literature. Eur J Neurol, 2009. 16(6): p. 691-6.
31. Jones, D.T.W., P. Bandopadhayay, and N. Jabado, The Power of Human Cancer Genetics as Revealed by Low-Grade Gliomas. Annu Rev Genet, 2019. 53: p. 483-503.
32. Filbin, M.G. and D. Sturm, Gliomas in Children. Semin Neurol, 2018. 38(1): p. 121-130.
33. Zhang, J., et al., Whole-genome sequencing identifies genetic alterations in pediatric low-grade gliomas. Nat Genet, 2013. 45(6): p. 602-12.
34. Jones, D.T., et al., Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. Nat Genet, 2013. 45(8): p. 927-32.
35. Jones, D.T., et al., Tandem duplication producing a novel oncogenic BRAF fusion gene defines the majority of pilocytic astrocytomas. Cancer Res, 2008. 68(21): p. 8673-7.
36. Rivera, B., et al., Germline and somatic FGFR1 abnormalities in dysembryoplastic neuroepithelial tumors. Acta Neuropathol, 2016. 131(6): p. 847-63.
37. Qaddoumi, I., et al., Genetic alterations in uncommon low-grade neuroepithelial tumors: BRAF, FGFR1, and MYB mutations occur at high frequency and align with morphology. Acta Neuropathol, 2016. 131(6): p. 833-45.
38. Ramkissoon, L.A., et al., Genomic analysis of diffuse pediatric low-grade gliomas identifies recurrent oncogenic truncating rearrangements in the transcription factor MYBL1. Proc Natl Acad Sci U S A, 2013. 110(20): p. 8188-93.
39. Wefers, A.K., et al., Isomorphic diffuse glioma is a morphologically and molecularly distinct tumour entity with recurrent gene fusions of MYBL1 or MYB and a benign disease course. Acta Neuropathol, 2020. 139(1): p. 193-209.
40. Bandopadhayay, P., et al., MYB-QKI rearrangements in angiocentric glioma drive tumorigenicity through a tripartite mechanism. Nat Genet, 2016. 48(3): p. 273-82.
41. Perreault, S., et al., A phase 2 study of trametinib for patients with pediatric glioma or plexiform neurofibroma with refractory tumor and activation of the MAPK/ERK pathway: TRAM-01. BMC Cancer, 2019. 19(1): p. 1250.

42. Banerjee, A., et al., A phase I trial of the MEK inhibitor selumetinib (AZD6244) in pediatric patients with recurrent or refractory low-grade glioma: a Pediatric Brain Tumor Consortium (PBTC) study. Neuro Oncol, 2017. 19(8): p. 1135-1144.
43. Dombi, E., et al., Activity of Selumetinib in Neurofibromatosis Type 1-Related Plexiform Neurofibromas. N Engl J Med, 2016. 375(26): p. 2550-2560.
44. van Tilburg, C.M., et al., Response in a child with a BRAF V600E mutated desmoplastic infantile astrocytoma upon retreatment with vemurafenib. Pediatr Blood Cancer, 2018. 65(3).
45. Lassaletta, A., et al., Therapeutic and Prognostic Implications of BRAF V600E in Pediatric Low-Grade Gliomas. J Clin Oncol, 2017. 35(25): p. 2934-2941.
46. Krueger, D.A., et al., Everolimus for subependymal giant-cell astrocytomas in tuberous sclerosis. N Engl J Med, 2010. 363(19): p. 1801-11.
47. Franz, D.N., et al., Everolimus for subependymal giant cell astrocytoma: 5-year final analysis. Ann Neurol, 2015. 78(6): p. 929-38.
48. Fangusaro, J., Pediatric high grade glioma: a review and update on tumor clinical characteristics and biology. Front Oncol, 2012. 2: p. 105.
49. Jones, C., L. Perryman, and D. Hargrave, Paediatric and adult malignant glioma: close relatives or distant cousins? Nat Rev Clin Oncol, 2012. 9(7): p. 400-13.
50. Hargrave, D., Paediatric high and low grade glioma: the impact of tumour biology on current and future therapy. British journal of neurosurgery, 2009. 23(4): p. 351-63.
51. Michaeli, O. and U. Tabori, Pediatric High Grade Gliomas in the Context of Cancer Predisposition Syndromes. J Korean Neurosurg Soc, 2018. 61(3): p. 319-332.
52. Davis, F., et al., Medical diagnostic radiation exposures and risk of gliomas. Radiat Res, 2011. 175(6): p. 790-6.
53. Walter, A.W., et al., Secondary brain tumors in children treated for acute lymphoblastic leukemia at St Jude Children's Research Hospital. J Clin Oncol, 1998. 16(12): p. 3761-7.
54. Wang, Y., et al., Radiation-induced glioblastoma with rhabdoid characteristics following treatment for medulloblastoma: A case report and review of the literature. Mol Clin Oncol, 2018. 9(4): p. 415-418.
55. Vanan, M.I. and D.D. Eisenstat, DIPG in Children - What Can We Learn from the Past? Front Oncol, 2015. 5: p. 237.
56. Reddy, A.T. and J.C. Wellons, 3rd, Pediatric high-grade gliomas. Cancer J, 2003. 9(2): p. 107-12.
57. Louis DN, O.H., Wiestler OD, Cavenee WK, WHO Classification of tumours of the central nervous system. 2007.
58. Anghileri, E., et al., Gliomatosis cerebri (GC) or GC-like? A picture to be reconsidered in neuro-oncology based on large retrospective analysis of GC series. Neurol Sci, 2020.
59. Panigrahy, A. and S. Bluml, Neuroimaging of pediatric brain tumors: from basic to advanced magnetic resonance imaging (MRI). J Child Neurol, 2009. 24(11): p. 1343-65.
60. Tisnado, J., et al., Conventional and Advanced Imaging of Diffuse Intrinsic Pontine Glioma. Journal of child neurology, 2016. 31(12): p. 1386-1393.
61. Ranjan, S. and K.E. Warren, Gliomatosis Cerebri: Current Understanding and Controversies. Frontiers in oncology, 2017. 7: p. 165-165.
62. Shin, Y.M., et al., Gliomatosis cerebri: comparison of MR and CT features. AJR Am J Roentgenol, 1993. 161(4): p. 859-62.
63. Hoffman, L.M., et al., Spatial genomic heterogeneity in diffuse intrinsic pontine and midline high-grade glioma: implications for diagnostic biopsy and targeted therapeutics. Acta neuropathologica communications, 2016. 4: p. 1-1.

64. Stupp, R., et al., Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N Engl J Med, 2005. 352(10): p. 987-96.
65. Warren, K.E., Diffuse intrinsic pontine glioma: poised for progress. Front Oncol, 2012. 2: p. 205.
66. Bouffet, E., et al., Radiotherapy followed by high dose busulfan and thiotepa: a prospective assessment of high dose chemotherapy in children with diffuse pontine gliomas. Cancer, 2000. 88(3): p. 685-92.
67. Macy, M.E., et al., A pediatric trial of radiation/cetuximab followed by irinotecan/cetuximab in newly diagnosed diffuse pontine gliomas and high-grade astrocytomas: A Pediatric Oncology Experimental Therapeutics Investigators' Consortium study. Pediatr Blood Cancer, 2017. 64(11).
68. Hummel, T.R., et al., A pilot study of bevacizumab-based therapy in patients with newly diagnosed high-grade gliomas and diffuse intrinsic pontine gliomas. J Neurooncol, 2016. 127(1): p. 53-61.
69. Grill, J., et al., Phase II, Open-Label, Randomized, Multicenter Trial (HERBY) of Bevacizumab in Pediatric Patients With Newly Diagnosed High-Grade Glioma. J Clin Oncol, 2018. 36(10): p. 951-958.
70. Qaddoumi, I., et al., Phase II Trial of Erlotinib during and after Radiotherapy in Children with Newly Diagnosed High-Grade Gliomas. Front Oncol, 2014. 4: p. 67.
71. Hoffman, L.M., et al., TR-14: A FEASIBILITY AND RANDOMIZED PHASE II STUDY OF VORINOSTAT, BEVACIZUMAB, OR TEMOZOLOMIDE DURING RADIATION FOLLOWED BY MAINTENANCE CHEMOTHERAPY IN NEWLY-DIAGNOSED PEDIATRIC HIGH-GRADE GLIOMA: CHILDREN'S ONCOLOGY GROUP STUDY ACNS0822. Neuro-Oncology, 2015. 17(Suppl 3): p. iii39-iii40.
72. Truffaux, N., et al., Preclinical evaluation of dasatinib alone and in combination with cabozantinib for the treatment of diffuse intrinsic pontine glioma. Neuro Oncol, 2015. 17(7): p. 953-64.
73. Schwartzentruber, J., et al., Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature, 2012. 482(7384): p. 226-31.
74. Wu, G., et al., Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. Nat Genet, 2012. 44(3): p. 251-3.
75. Wu, G., et al., The genomic landscape of diffuse intrinsic pontine glioma and pediatric non-brainstem high-grade glioma. Nat Genet, 2014. 46(5): p. 444-50.
76. Taylor, K.R., et al., Recurrent activating ACVR1 mutations in diffuse intrinsic pontine glioma. Nat Genet, 2014. 46(5): p. 457-461.
77. Bax, D.A., et al., A distinct spectrum of copy number aberrations in pediatric high-grade gliomas. Clin Cancer Res, 2010. 16(13): p. 3368-77.
78. Sturm, D., et al., Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. Cancer Cell, 2012. 22(4): p. 425-37.
79. Verhaak, R.G., et al., Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell, 2010. 17(1): p. 98-110.
80. Ceccarelli, M., et al., Molecular Profiling Reveals Biologically Discrete Subsets and Pathways of Progression in Diffuse Glioma. Cell, 2016. 164(3): p. 550-63.
81. Brennan, C.W., et al., The somatic genomic landscape of glioblastoma. Cell, 2013. 155(2): p. 462-477.
82. Killela, P.J., et al., 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 U S A, 2013. 110(15): p. 6021-6.
83. Heaphy, C.M., et al., Altered telomeres in tumors with ATRX and DAXX mutations. Science, 2011. 333(6041): p. 425.

84. Fontebasso, A.M., et al., Recurrent somatic mutations in ACVR1 in pediatric midline high-grade astrocytoma. Nat Genet, 2014. 46(5): p. 462-6.
85. Sturm, D., et al., Paediatric and adult glioblastoma: multiform (epi)genomic culprits emerge. Nat Rev Cancer, 2014. 14(2): p. 92-107.
86. Shlien, A., et al., Combined hereditary and somatic mutations of replication error repair genes result in rapid onset of ultra-hypermutated cancers. Nat Genet, 2015. 47(3): p. 257-62.
87. Mackay, A., et al., Molecular, Pathological, Radiological, and Immune Profiling of Non-brainstem Pediatric High-Grade Glioma from the HERBY Phase II Randomized Trial. Cancer Cell, 2018. 33(5): p. 829-842 e5.
88. Korshunov, A., et al., H3-/IDH-wild type pediatric glioblastoma is comprised of molecularly and prognostically distinct subtypes with associated oncogenic drivers. Acta Neuropathol, 2017. 134(3): p. 507-516.
89. Clarke, M., et al., Infant high grade gliomas comprise multiple subgroups characterized by novel targetable gene fusions and favorable outcomes. Cancer Discovery, 2020: p. CD-19-1030.
90. Wu, G., et al., Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. Nat Genet, 2012. 44(3): p. 251-3.
91. Behjati, S., et al., Distinct H3F3A and H3F3B driver mutations define chondroblastoma and giant cell tumor of bone. Nat Genet, 2013. 45(12): p. 1479-82.
92. Ryall, S., et al., H3 K27M mutations are extremely rare in posterior fossa group A ependymoma. Childs Nerv Syst, 2017. 33(7): p. 1047-1051.
93. Gessi, M., et al., Evidence of H3 K27M mutations in posterior fossa ependymomas. Acta Neuropathol, 2016. 132(4): p. 635-7.
94. Castel, D., et al., Histone H3F3A and HIST1H3B K27M mutations define two subgroups of diffuse intrinsic pontine gliomas with different prognosis and phenotypes. Acta Neuropathol, 2015. 130(6): p. 815-27.
95. Castel, D., et al., Transcriptomic and epigenetic profiling of 'diffuse midline gliomas, H3 K27M-mutant' discriminate two subgroups based on the type of histone H3 mutated and not supratentorial or infratentorial location. Acta Neuropathol Commun, 2018. 6(1): p. 117.
96. Nagaraja, S., et al., Histone Variant and Cell Context Determine H3K27M Reprogramming of the Enhancer Landscape and Oncogenic State. Mol Cell, 2019. 76(6): p. 965-980.e12.
97. Filbin, M.G., et al., Developmental and oncogenic programs in H3K27M gliomas dissected by single-cell RNA-seq. Science, 2018. 360(6386): p. 331-335.
98. Jones, C., et al., Pediatric high-grade glioma: biologically and clinically in need of new thinking. Neuro Oncol, 2016.
99. Funato, K., et al., Use of human embryonic stem cells to model pediatric gliomas with H3.3K27M histone mutation. Science, 2014. 346(6216): p. 1529-33.
100. Mohammad, F., et al., EZH2 is a potential therapeutic target for H3K27M-mutant pediatric gliomas. Nat Med, 2017. 23(4): p. 483-492.
101. Pathania, M., et al., H3.3(K27M) Cooperates with Trp53 Loss and PDGFRA Gain in Mouse Embryonic Neural Progenitor Cells to Induce Invasive High-Grade Gliomas. Cancer Cell, 2017. 32(5): p. 684-700.e9.
102. Nikbakht, H., et al., Spatial and temporal homogeneity of driver mutations in diffuse intrinsic pontine glioma. Nat Commun, 2016. 7: p. 11185.
103. Vinci, M., et al., Functional diversity and cooperativity between subclonal populations of pediatric glioblastoma and diffuse intrinsic pontine glioma cells. Nature Medicine, 2018. 24(8): p. 1204-1215.
104. Baroni, L.V., et al., Indolent course of brainstem tumors with K27M-H3.3 mutation. Pediatr Blood Cancer, 2019: p. e28102.

105. Larson, J.D., et al., Histone H3.3 K27M Accelerates Spontaneous Brainstem Glioma and Drives Restricted Changes in Bivalent Gene Expression. Cancer Cell, 2019. 35(1): p. 140-155.e7.
106. Silveira, A.B., et al., H3.3 K27M depletion increases differentiation and extends latency of diffuse intrinsic pontine glioma growth in vivo. Acta Neuropathol, 2019. 137(4): p. 637-655.
107. Jessa, S., et al., Stalled developmental programs at the root of pediatric brain tumors. Nat Genet, 2019. 51(12): p. 1702-1713.
108. Jones, C. and S.J. Baker, Unique genetic and epigenetic mechanisms driving paediatric diffuse high-grade glioma. Nat Rev Cancer, 2014. 14(10).
109. Lewis, P.W., et al., Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. Science, 2013. 340(6134): p. 857-61.
110. Chan, K.M., et al., The histone H3.3K27M mutation in pediatric glioma reprograms H3K27 methylation and gene expression. Genes Dev, 2013. 27(9): p. 985-90.
111. Bender, S., et al., Reduced H3K27me3 and DNA hypomethylation are major drivers of gene expression in K27M mutant pediatric high-grade gliomas. Cancer Cell, 2013. 24(5): p. 660-72.
112. Bjerke, L., et al., Histone H3.3. mutations drive pediatric glioblastoma through upregulation of MYCN. Cancer Discov, 2013. 3(5): p. 512-9.
113. Arvanitis, C.D., G.B. Ferraro, and R.K. Jain, The blood–brain barrier and blood–tumour barrier in brain tumours and metastases. Nature Reviews Cancer, 2020. 20(1): p. 26-41.
114. Warren, K.E., Beyond the Blood:Brain Barrier: The Importance of Central Nervous System (CNS) Pharmacokinetics for the Treatment of CNS Tumors, Including Diffuse Intrinsic Pontine Glioma. Front Oncol, 2018. 8: p. 239.
115. Veringa, S.J., et al., In vitro drug response and efflux transporters associated with drug resistance in pediatric high grade glioma and diffuse intrinsic pontine glioma. PLoS One, 2013. 8(4): p. e61512.
116. McCully, C.M., et al., Model for concomitant microdialysis sampling of the pons and cerebral cortex in rhesus macaques (Macaca mulatta). Comp Med, 2013. 63(4): p. 355-60.
117. Subashi, E., et al., Tumor location, but not H3.3K27M, significantly influences the blood-brain-barrier permeability in a genetic mouse model of pediatric high-grade glioma. J Neurooncol, 2016. 126(2): p. 243-51.
118. Shergalis, A., et al., Current Challenges and Opportunities in Treating Glioblastoma. Pharmacological Reviews, 2018. 70(3): p. 412.
119. Vogelbaum, M.A. and M.K. Aghi, Convection-enhanced delivery for the treatment of glioblastoma. Neuro Oncol, 2015. 17 Suppl 2: p. ii3-ii8.
120. Souweidane, M.M., et al., Convection-enhanced delivery for diffuse intrinsic pontine glioma: a single-centre, dose-escalation, phase 1 trial. Lancet Oncol, 2018. 19(8): p. 1040-1050.
121. Chakraborty, S., et al., Superselective intraarterial cerebral infusion of cetuximab after osmotic blood/brain barrier disruption for recurrent malignant glioma: phase I study. J Neurooncol, 2016. 128(3): p. 405-15.
122. Szychot, E., et al., DIPG-16. APPLYING THE SIOPE DIPG REGISTRY SURVIVAL PREDICTION TOOL TO SELECT A SURVIVAL EXTENSION TARGET FOR A POWER CALCULATION FOR A NEW TRIAL OF CONVECTION ENHANCED DRUG DELIVERY (CED) OF CARBOPLATIN AND SODIUM VALPROATE IN DIFFUSE INTRINSIC PONTINE GLIOMA (DIPG). Neuro-Oncology, 2019. 21(Supplement_2): p. ii71-ii72.

123. Faltings, L., et al., Rechallenging Recurrent Glioblastoma with Intra-Arterial Bevacizumab with Blood Brain-Barrier Disruption Results in Radiographic Response. World Neurosurg, 2019. 131: p. 234-241.
124. Etame, A.B., et al., Focused ultrasound disruption of the blood-brain barrier: a new frontier for therapeutic delivery in molecular neurooncology. Neurosurgical focus, 2012. 32(1): p. E3-E3.
125. Harutyunyan, A.S., et al., H3K27M induces defective chromatin spread of PRC2-mediated repressive H3K27me2/me3 and is essential for glioma tumorigenesis. Nat Commun, 2019. 10(1): p. 1262.
126. Piunti, A., et al., Therapeutic targeting of polycomb and BET bromodomain proteins in diffuse intrinsic pontine gliomas. Nat Med, 2017. 23(4): p. 493-500.
127. Kasper, L.H. and S.J. Baker, Invited Review: Emerging functions of histone H3 mutations in paediatric diffuse high-grade gliomas. Neuropathol Appl Neurobiol, 2019.
128. Carvalho, D., et al., ALK2 inhibitors display beneficial effects in preclinical models of ACVR1 mutant diffuse intrinsic pontine glioma. Commun Biol, 2019. 2: p. 156.
129. Pal, S., et al., Dual HDAC and PI3K Inhibition Abrogates NFkappaB- and FOXM1-Mediated DNA Damage Response to Radiosensitize Pediatric High-Grade Gliomas. Cancer Res, 2018. 78(14): p. 4007-4021.
130. Miklja, Z., et al., Everolimus improves the efficacy of dasatinib in the treatment of PDGFRA-driven glioma. medRxiv, 2020: p. 2019.12.28.19015974.
131. Kumar, S.S., et al., BMI-1 is a potential therapeutic target in diffuse intrinsic pontine glioma. Oncotarget, 2017. 8(38): p. 62962-62975.
132. Majzner, R.G., et al., CAR T Cells Targeting B7-H3, a Pan-Cancer Antigen, Demonstrate Potent Preclinical Activity Against Pediatric Solid Tumors and Brain Tumors. Clin Cancer Res, 2019. 25(8): p. 2560-2574.
133. Mount, C.W., et al., Potent antitumor efficacy of anti-GD2 CAR T cells in H3-K27M(+) diffuse midline gliomas. Nat Med, 2018. 24(5): p. 572-579.
134. Wagner, L.M. and V.R. Adams, Targeting the PD-1 pathway in pediatric solid tumors and brain tumors. Onco Targets Ther, 2017. 10: p. 2097-2106.
135. Johanns, T.M., et al., Immunogenomics of Hypermutated Glioblastoma: A Patient with Germline POLE Deficiency Treated with Checkpoint Blockade Immunotherapy. Cancer Discov, 2016. 6(11): p. 1230-1236.
136. Pollack, I.F., et al., Antigen-specific immune responses and clinical outcome after vaccination with glioma-associated antigen peptides and polyinosinic-polycytidylic acid stabilized by lysine and carboxymethylcellulose in children with newly diagnosed malignant brainstem and nonbrainstem gliomas. J Clin Oncol, 2014. 32(19): p. 2050-8.
137. Pollack, I.F., et al., Antigen-specific immunoreactivity and clinical outcome following vaccination with glioma-associated antigen peptides in children with recurrent high-grade gliomas: results of a pilot study. J Neurooncol, 2016. 130(3): p. 517-527.
138. Bouffet, E., et al., Immune Checkpoint Inhibition for Hypermutant Glioblastoma Multiforme Resulting From Germline Biallelic Mismatch Repair Deficiency. J Clin Oncol, 2016. 34(19): p. 2206-11.
139. Martinez-Velez, N., et al., The oncolytic virus Delta-24-RGD elicits an antitumor effect in pediatric glioma and DIPG mouse models. Nat Commun, 2019. 10(1): p. 2235.
140. Martinez-Velez, N., et al., Delta-24-RGD combined with radiotherapy exerts a potent antitumor effect in diffuse intrinsic pontine glioma and pediatric high grade glioma models. Acta Neuropathol Commun, 2019. 7(1): p. 64.

141. Khanna, V., et al., Incidence and survival trends for medulloblastomas in the United States from 2001 to 2013. J Neurooncol, 2017. 135(3): p. 433-441.
142. Medulloblastoma. Nature Reviews Disease Primers, 2019. 5(1): p. 12.
143. Cavalli, F.M.G., et al., Intertumoral Heterogeneity within Medulloblastoma Subgroups. Cancer Cell, 2017. 31(6): p. 737-754.e6.
144. Schwalbe, E.C., et al., Novel molecular subgroups for clinical classification and outcome prediction in childhood medulloblastoma: a cohort study. Lancet Oncol, 2017. 18(7): p. 958-971.
145. Hovestadt, V., et al., Medulloblastomics revisited: biological and clinical insights from thousands of patients. Nat Rev Cancer, 2020. 20(1): p. 42-56.
146. Gibson, P., et al., Subtypes of medulloblastoma have distinct developmental origins. Nature, 2010. 468(7327): p. 1095-9.
147. Vladoiu, M.C., et al., Childhood cerebellar tumours mirror conserved fetal transcriptional programs. Nature, 2019. 572(7767): p. 67-73.
148. Hovestadt, V., et al., Resolving medulloblastoma cellular architecture by single-cell genomics. Nature, 2019. 572(7767): p. 74-79.
149. Waszak, S.M., et al., Spectrum and prevalence of genetic predisposition in medulloblastoma: a retrospective genetic study and prospective validation in a clinical trial cohort. Lancet Oncol, 2018. 19(6): p. 785-798.
150. Begemann, M., et al., Germline GPR161 Mutations Predispose to Pediatric Medulloblastoma. J Clin Oncol, 2020. 38(1): p. 43-50.
151. Waszak, S.M., et al., Germline Elongator mutations in Sonic Hedgehog medulloblastoma. Nature, 2020.
152. Suzuki, H., et al., Recurrent noncoding U1 snRNA mutations drive cryptic splicing in SHH medulloblastoma. Nature, 2019. 574(7780): p. 707-711.
153. Robinson, G.W., et al., Risk-adapted therapy for young children with medulloblastoma (SJYC07): therapeutic and molecular outcomes from a multicentre, phase 2 trial. Lancet Oncol, 2018. 19(6): p. 768-784.
154. Kool, M., et al., Genome sequencing of SHH medulloblastoma predicts genotype-related response to smoothened inhibition. Cancer cell, 2014. 25(3): p. 393-405.
155. Sharma, T., et al., Second-generation molecular subgrouping of medulloblastoma: an international meta-analysis of Group 3 and Group 4 subtypes. Acta neuropathologica, 2019. 138(2): p. 309-326.
156. Hilden, J.M., et al., Central nervous system atypical teratoid/rhabdoid tumor: results of therapy in children enrolled in a registry. J Clin Oncol, 2004. 22(14): p. 2877-84.
157. Dho, Y.S., et al., Investigation of the location of atypical teratoid/rhabdoid tumor. Childs Nerv Syst, 2015. 31(8): p. 1305-11.
158. Ginn, K.F. and A. Gajjar, Atypical teratoid rhabdoid tumor: current therapy and future directions. Front Oncol, 2012. 2: p. 114.
159. Erkek, S., et al., Comprehensive Analysis of Chromatin States in Atypical Teratoid/Rhabdoid Tumor Identifies Diverging Roles for SWI/SNF and Polycomb in Gene Regulation. Cancer Cell, 2019. 35(1): p. 95-110.e8.
160. Ho, B., et al., Molecular subgrouping of Atypical Teratoid / Rhabdoid Tumors (ATRT) - a reinvestigation and current consensus. Neuro Oncol, 2019.
161. Chun, H.E., et al., Identification and Analyses of Extra-Cranial and Cranial Rhabdoid Tumor Molecular Subgroups Reveal Tumors with Cytotoxic T Cell Infiltration. Cell Rep, 2019. 29(8): p. 2338-2354.e7.
162. Khatua, S., V. Ramaswamy, and E. Bouffet, Current therapy and the evolving molecular landscape of paediatric ependymoma. Eur J Cancer, 2017. 70: p. 34-41.
163. Kilday, J.P., et al., Pediatric ependymoma: biological perspectives. Mol Cancer Res, 2009. 7(6): p. 765-86.

164. McGuire, C.S., K.L. Sainani, and P.G. Fisher, Incidence patterns for ependymoma: a surveillance, epidemiology, and end results study. J Neurosurg, 2009. 110(4): p. 725-9.
165. Plotkin, S.R., et al., Spinal ependymomas in neurofibromatosis Type 2: a retrospective analysis of 55 patients. J Neurosurg Spine, 2011. 14(4): p. 543-7.
166. Merchant, T.E., et al., Conformal radiotherapy after surgery for paediatric ependymoma: a prospective study. Lancet Oncol, 2009. 10(3): p. 258-66.
167. Venkatramani, R., et al., Supratentorial ependymoma in children: to observe or to treat following gross total resection? Pediatr Blood Cancer, 2012. 58(3): p. 380-3.
168. Marinoff, A.E., et al., Rethinking childhood ependymoma: a retrospective, multi-center analysis reveals poor long-term overall survival. J Neurooncol, 2017. 135(1): p. 201-211.
169. Bayliss, J., et al., Lowered H3K27me3 and DNA hypomethylation define poorly prognostic pediatric posterior fossa ependymomas. Sci Transl Med, 2016. 8(366): p. 366ra161.
170. Pajtler, K.W., et al., Molecular heterogeneity and CXorf67 alterations in posterior fossa group A (PFA) ependymomas. Acta Neuropathol, 2018. 136(2): p. 211-226.
171. Jain, S.U., et al., PFA ependymoma-associated protein EZHIP inhibits PRC2 activity through a H3 K27M-like mechanism. Nat Commun, 2019. 10(1): p. 2146.
172. Hubner, J.M., et al., EZHIP / CXorf67 mimics K27M mutated oncohistones and functions as an intrinsic inhibitor of PRC2 function in aggressive posterior fossa ependymoma. Neuro Oncol, 2019.
173. Mack, S.C., et al., Therapeutic targeting of ependymoma as informed by oncogenic enhancer profiling. Nature, 2018. 553(7686): p. 101-105.
174. Ghasemi, D.R., et al., MYCN amplification drives an aggressive form of spinal ependymoma. Acta neuropathologica, 2019. 138(6): p. 1075-1089.
175. Pfister, S., et al., Novel genomic amplification targeting the microRNA cluster at 19q13.42 in a pediatric embryonal tumor with abundant neuropil and true rosettes. Acta Neuropathol, 2009. 117(4): p. 457-64.
176. Li, M., et al., Frequent amplification of a chr19q13.41 microRNA polycistron in aggressive primitive neuroectodermal brain tumors. Cancer Cell, 2009. 16(6): p. 533-46.
177. Korshunov, A., et al., LIN28A immunoreactivity is a potent diagnostic marker of embryonal tumor with multilayered rosettes (ETMR). Acta Neuropathol, 2012. 124(6): p. 875-81.
178. Kleinman, C.L., et al., Fusion of TTYH1 with the C19MC microRNA cluster drives expression of a brain-specific DNMT3B isoform in the embryonal brain tumor ETMR. Nat Genet, 2014. 46(1): p. 39-44.
179. Sin-Chan, P., et al., A C19MC-LIN28A-MYCN Oncogenic Circuit Driven by Hijacked Super-enhancers Is a Distinct Therapeutic Vulnerability in ETMRs: A Lethal Brain Tumor. Cancer Cell, 2019. 36(1): p. 51-67.e7.
180. Lambo, S., et al., The molecular landscape of ETMR at diagnosis and relapse. Nature, 2019. 576(7786): p. 274-280.
181. Louis, D.N., et al., The 2007 WHO classification of tumours of the central nervous system. Acta Neuropathol, 2007. 114(2): p. 97-109.
182. Pizer, B.L., et al., Analysis of patients with supratentorial primitive neuro-ectodermal tumours entered into the SIOP/UKCCSG PNET 3 study. Eur J Cancer, 2006. 42(8): p. 1120-8.
183. London, W.B., et al., Evidence for an age cutoff greater than 365 days for neuroblastoma risk group stratification in the Children's Oncology Group. J Clin Oncol, 2005. 23(27): p. 6459-65.

184. De Bernardi, B., et al., Excellent outcome with reduced treatment for infants with disseminated neuroblastoma without MYCN gene amplification. J Clin Oncol, 2009. 27(7): p. 1034-40.
185. Moreno, L., et al., Outcome of children with relapsed or refractory neuroblastoma: A meta-analysis of ITCC/SIOPEN European phase II clinical trials. Pediatr Blood Cancer, 2017. 64(1): p. 25-31.
186. Rubie, H., et al., Excellent outcome with reduced treatment in infants with nonmetastatic and unresectable neuroblastoma without MYCN amplification: results of the prospective INES 99.1. J Clin Oncol, 2011. 29(4): p. 449-55.
187. Ackermann, S., et al., A mechanistic classification of clinical phenotypes in neuroblastoma. Science, 2018. 362(6419): p. 1165.
188. Schwab, M., et al., Neuroblastoma: biology and molecular and chromosomal pathology. Lancet Oncol, 2003. 4(8): p. 472-80.
189. Mosse, Y.P., et al., Identification of ALK as a major familial neuroblastoma predisposition gene. Nature, 2008. 455(7215): p. 930-5.
190. Janoueix-Lerosey, I., et al., Somatic and germline activating mutations of the ALK kinase receptor in neuroblastoma. Nature, 2008. 455(7215): p. 967-70.
191. George, R.E., et al., Activating mutations in ALK provide a therapeutic target in neuroblastoma. Nature, 2008. 455(7215): p. 975-8.
192. Cheung, N.K.V., et al., Association of Age at Diagnosis and Genetic Mutations in Patients With Neuroblastoma. Jama-Journal of the American Medical Association, 2012. 307(10): p. 1062-1071.
193. Pugh, T.J., et al., The genetic landscape of high-risk neuroblastoma. Nat Genet, 2013. 45(3): p. 279-84.
194. Valentijn, L.J., et al., TERT rearrangements are frequent in neuroblastoma and identify aggressive tumors. Nat Genet, 2015. 47(12): p. 1411-4.
195. Peifer, M., et al., Telomerase activation by genomic rearrangements in high-risk neuroblastoma. Nature, 2015. 526(7575): p. 700-4.
196. Koche, R.P., et al., Extrachromosomal circular DNA drives oncogenic genome remodeling in neuroblastoma. Nat Genet, 2019.
197. Speleman, F., J.R. Park, and T.O. Henderson, Neuroblastoma: A Tough Nut to Crack. Am Soc Clin Oncol Educ Book, 2016. 35: p. e548-57.
198. Bresler, S.C., et al., Differential inhibitor sensitivity of anaplastic lymphoma kinase variants found in neuroblastoma. Sci Transl Med, 2011. 3(108): p. 108ra114.
199. Johnsen, J.I., et al., Molecular mechanisms and therapeutic targets in neuroblastoma. Pharmacol Res, 2018. 131: p. 164-176.
200. DuBois, S.G., et al., Phase I Study of the Aurora A Kinase Inhibitor Alisertib in Combination With Irinotecan and Temozolomide for Patients With Relapsed or Refractory Neuroblastoma: A NANT (New Approaches to Neuroblastoma Therapy) Trial. J Clin Oncol, 2016. 34(12): p. 1368-75.
201. Demetri, G.D., et al., Soft tissue sarcoma clinical practice guidelines in oncology. J Natl Compr Canc Netw, 2005. 3(2): p. 158-94.
202. Pappo, A.S. and U. Dirksen, Rhabdomyosarcoma, Ewing Sarcoma, and Other Round Cell Sarcomas. J Clin Oncol, 2018. 36(2): p. 168-179.
203. Slater, O. and J. Shipley, Clinical relevance of molecular genetics to paediatric sarcomas. J Clin Pathol, 2007. 60(11): p. 1187-94.
204. Chang, C.C. and V.B. Shidham, Molecular genetics of pediatric soft tissue tumors: clinical application. J Mol Diagn, 2003. 5(3): p. 143-54.
205. Mertens, F., C.R. Antonescu, and F. Mitelman, Gene fusions in soft tissue tumors: Recurrent and overlapping pathogenetic themes. Genes, chromosomes & cancer, 2016. 55(4): p. 291-310.

206. Koelsche, C., et al., Array-based DNA-methylation profiling in sarcomas with small blue round cell histology provides valuable diagnostic information. Modern Pathology, 2018. 31(8): p. 1246-1256.
207. Sandler, G., A. Yokoi, and A. Hayes-Jordan, An update in the management of pediatric sarcoma. Curr Opin Pediatr, 2019. 31(3): p. 368-377.
208. Hawkins, D.S., S.L. Spunt, and S.X. Skapek, Children's Oncology Group's 2013 blueprint for research: Soft tissue sarcomas. Pediatr Blood Cancer, 2013. 60(6): p. 1001-8.
209. Marshall, A.D. and G.C. Grosveld, Alveolar rhabdomyosarcoma - The molecular drivers of PAX3/7-FOXO1-induced tumorigenesis. Skelet Muscle, 2012. 2(1): p. 25.
210. Skapek, S.X., et al., Rhabdomyosarcoma. Nature Reviews Disease Primers, 2019. 5(1): p. 1.
211. McCarville, M.B., S.L. Spunt, and A.S. Pappo, Rhabdomyosarcoma in pediatric patients: the good, the bad, and the unusual. AJR Am J Roentgenol, 2001. 176(6): p. 1563-9.
212. Galili, N., et al., Fusion of a fork head domain gene to PAX3 in the solid tumour alveolar rhabdomyosarcoma. Nat Genet, 1993. 5(3): p. 230-5.
213. Shapiro, D.N., et al., Fusion of PAX3 to a member of the forkhead family of transcription factors in human alveolar rhabdomyosarcoma. Cancer Res, 1993. 53(21): p. 5108-12.
214. Shern, J.F., et al., Comprehensive genomic analysis of rhabdomyosarcoma reveals a landscape of alterations affecting a common genetic axis in fusion-positive and fusion-negative tumors. Cancer Discov, 2014. 4(2): p. 216-31.
215. Chen, X., et al., Targeting oxidative stress in embryonal rhabdomyosarcoma. Cancer Cell, 2013. 24(6): p. 710-24.
216. Kohsaka, S., et al., A recurrent neomorphic mutation in MYOD1 defines a clinically aggressive subset of embryonal rhabdomyosarcoma associated with PI3K-AKT pathway mutations. Nature genetics, 2014. 46(6): p. 595-600.
217. Agaram, N.P., et al., Recurrent MYOD1 mutations in pediatric and adult sclerosing and spindle cell rhabdomyosarcomas: evidence for a common pathogenesis. Genes Chromosomes Cancer, 2014. 53(9): p. 779-87.
218. Waxweiler, T.V., et al., Non-Rhabdomyosarcoma Soft Tissue Sarcomas in Children: A Surveillance, Epidemiology, and End Results Analysis Validating COG Risk Stratifications. Int J Radiat Oncol Biol Phys, 2015. 92(2): p. 339-48.
219. Clark, J., et al., Identification of novel genes, SYT and SSX, involved in the t(X;18)(p11.2;q11.2) translocation found in human synovial sarcoma. Nat Genet, 1994. 7(4): p. 502-8.
220. Knezevich, S.R., et al., A novel ETV6-NTRK3 gene fusion in congenital fibrosarcoma. Nat Genet, 1998. 18(2): p. 184-7.
221. Lawrence, B., et al., TPM3-ALK and TPM4-ALK oncogenes in inflammatory myofibroblastic tumors. Am J Pathol, 2000. 157(2): p. 377-84.
222. Bridge, J.A., et al., Fusion of the ALK gene to the clathrin heavy chain gene, CLTC, in inflammatory myofibroblastic tumor. Am J Pathol, 2001. 159(2): p. 411-5.
223. Antonescu, C.R., et al., Molecular diagnosis of clear cell sarcoma: detection of EWS-ATF1 and MITF-M transcripts and histopathological and ultrastructural analysis of 12 cases. J Mol Diagn, 2002. 4(1): p. 44-52.
224. Ladanyi, M. and W. Gerald, Fusion of the EWS and WT1 genes in the desmoplastic small round cell tumor. Cancer Res, 1994. 54(11): p. 2837-40.
225. Antonescu, C.R., et al., Prognostic impact of P53 status, TLS-CHOP fusion transcript structure, and histological grade in myxoid liposarcoma: a molecular and clinicopathologic study of 82 cases. Clin Cancer Res, 2001. 7(12): p. 3977-87.

226. Geller, D.S. and R. Gorlick, Osteosarcoma: a review of diagnosis, management, and treatment strategies. Clin Adv Hematol Oncol, 2010. 8(10): p. 705-18.
227. Nik-Zainal, S., et al., Mutational processes molding the genomes of 21 breast cancers. Cell, 2012. 149(5): p. 979-93.
228. Chen, X., et al., Recurrent somatic structural variations contribute to tumorigenesis in pediatric osteosarcoma. Cell Rep, 2014. 7(1): p. 104-12.
229. Ribi, S., et al., TP53 intron 1 hotspot rearrangements are specific to sporadic osteosarcoma and can cause Li-Fraumeni syndrome. Oncotarget, 2015. 6(10): p. 7727-7740.
230. Delattre, O., et al., Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. Nature, 1992. 359(6391): p. 162-5.
231. Solomon, D.A., et al., Mutational inactivation of STAG2 causes aneuploidy in human cancer. Science, 2011. 333(6045): p. 1039-43.
232. Tirode, F., et al., Genomic landscape of Ewing sarcoma defines an aggressive subtype with co-association of STAG2 and TP53 mutations. Cancer Discov, 2014. 4(11): p. 1342-53.
233. Baldauf, M.C., et al., Are EWSR1-NFATc2-positive sarcomas really Ewing sarcomas? Mod Pathol, 2018. 31(6): p. 997-999.
234. Grunewald, T.G.P., et al., Ewing sarcoma. Nat Rev Dis Primers, 2018. 4(1): p. 5.
235. Antonescu, C.R., et al., Sarcomas With CIC-rearrangements Are a Distinct Pathologic Entity With Aggressive Outcome: A Clinicopathologic and Molecular Study of 115 Cases. Am J Surg Pathol, 2017. 41(7): p. 941-949.
236. Gambarotti, M., et al., CIC-DUX4 fusion-positive round-cell sarcomas of soft tissue and bone: a single-institution morphological and molecular analysis of seven cases. Histopathology, 2016. 69(4): p. 624-34.
237. Kao, Y.C., et al., BCOR-CCNB3 Fusion Positive Sarcomas: A Clinicopathologic and Molecular Analysis of 36 Cases With Comparison to Morphologic Spectrum and Clinical Behavior of Other Round Cell Sarcomas. Am J Surg Pathol, 2018. 42(5): p. 604-615.
238. Watson, S., et al., Transcriptomic definition of molecular subgroups of small round cell sarcomas. J Pathol, 2018. 245(1): p. 29-40.
239. Szuhai, K., et al., The NFATc2 gene is involved in a novel cloned translocation in a Ewing sarcoma variant that couples its function in immunology to oncology. Clin Cancer Res, 2009. 15(7): p. 2259-68.
240. Grovas, A., et al., The National Cancer Data Base report on patterns of childhood cancers in the United States. Cancer, 1997. 80(12): p. 2321-32.
241. Gadd, S., et al., A Children's Oncology Group and TARGET initiative exploring the genetic landscape of Wilms tumor. Nat Genet, 2017. 49(10): p. 1487-1494.
242. Ruteshouser, E.C., S.M. Robinson, and V. Huff, Wilms tumor genetics: mutations in WT1, WTX, and CTNNB1 account for only about one-third of tumors. Genes, chromosomes & cancer, 2008. 47(6): p. 461-470.
243. Torrezan, G.T., et al., Recurrent somatic mutation in DROSHA induces microRNA profile changes in Wilms tumour. Nature Communications, 2014. 5(1): p. 4039.
244. Rao, Q., et al., Renal cell carcinoma in children and young adults: clinicopathological, immunohistochemical, and VHL gene analysis of 46 cases with follow-up. Int J Surg Pathol, 2011. 19(2): p. 170-9.
245. Caliò, A., et al., MiT Family Translocation Renal Cell Carcinoma: from the Early Descriptions to the Current Knowledge. Cancers, 2019. 11(8): p. 1110.
246. Roy, A., et al., Recurrent internal tandem duplications of BCOR in clear cell sarcoma of the kidney. Nature Communications, 2015. 6(1): p. 8891.

247. Astolfi, A., et al., Whole transcriptome sequencing identifies BCOR internal tandem duplication as a common feature of clear cell sarcoma of the kidney. Oncotarget, 2015. 6(38): p. 40934-40939.
248. Dimaras, H., et al., Retinoblastoma. Nature Reviews Disease Primers, 2015. 1(1): p. 15021.
249. Dunn, J.M., et al., Identification of germline and somatic mutations affecting the retinoblastoma gene. Science, 1988. 241(4874): p. 1797-800.
250. Zhang, J., et al., A novel retinoblastoma therapy from genomic and epigenetic analyses. Nature, 2012. 481(7381): p. 329-334.
251. Kooi, I.E., et al., Somatic genomic alterations in retinoblastoma beyond RB1 are rare and limited to copy number changes. Scientific Reports, 2016. 6(1): p. 25264.
252. Rushlow, D.E., et al., Characterisation of retinoblastomas without RB1 mutations: genomic, gene expression, and clinical studies. Lancet Oncol, 2013. 14(4): p. 327-34.
253. Kuhlen, M., et al., Family-based germline sequencing in children with cancer. Oncogene, 2019. 38(9): p. 1367-1380.
254. Mody, R.J., et al., Integrative Clinical Sequencing in the Management of Refractory or Relapsed Cancer in Youth. Jama, 2015. 314(9): p. 913-25.
255. McBride, K.A., et al., Li-Fraumeni syndrome: cancer risk assessment and clinical management. Nat Rev Clin Oncol, 2014. 11(5): p. 260-71.
256. Ripperger, T., et al., Childhood cancer predisposition syndromes-A concise review and recommendations by the Cancer Predisposition Working Group of the Society for Pediatric Oncology and Hematology. Am J Med Genet A, 2017. 173(4): p. 1017-1037.
257. Moorcraft, S.Y., D. Gonzalez, and B.A. Walker, Understanding next generation sequencing in oncology: A guide for oncologists. Crit Rev Oncol Hematol, 2015. 96(3): p. 463-74.
258. Goodwin, S., J.D. McPherson, and W.R. McCombie, Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet, 2016. 17(6): p. 333-351.
259. Kamps, R., et al., Next-Generation Sequencing in Oncology: Genetic Diagnosis, Risk Prediction and Cancer Classification. Int J Mol Sci, 2017. 18(2).
260. Mattocks, C.J., et al., A standardized framework for the validation and verification of clinical molecular genetic tests. Eur J Hum Genet, 2010. 18(12): p. 1276-88.
261. Hardwick, S.A., I.W. Deveson, and T.R. Mercer, Reference standards for next-generation sequencing. Nat Rev Genet, 2017. 18(8): p. 473-484.
262. Jennings, L.J., et al., Guidelines for Validation of Next-Generation Sequencing–Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. The Journal of Molecular Diagnostics, 2017. 19(3): p. 341-365.
263. Lih, C.J., et al., Analytical Validation of the Next-Generation Sequencing Assay for a Nationwide Signal-Finding Clinical Trial: Molecular Analysis for Therapy Choice Clinical Trial. J Mol Diagn, 2017. 19(2): p. 313-327.
264. Roy, S., et al., Standards and Guidelines for Validating Next-Generation Sequencing Bioinformatics Pipelines: A Joint Recommendation of the Association for Molecular Pathology and the College of American Pathologists. J Mol Diagn, 2018. 20(1): p. 4-27.
265. Harris, M.H., et al., Multicenter Feasibility Study of Tumor Molecular Profiling to Inform Therapeutic Decisions in Advanced Pediatric Solid Tumors: The Individualized Cancer Therapy (iCat) Study. JAMA Oncol, 2016. 2(5): p. 608-615.
266. Oberg, J.A., et al., Implementation of next generation sequencing into pediatric hematology-oncology practice: moving beyond actionable alterations. Genome Med, 2016. 8(1): p. 133.

267. Worst, B.C., et al., Next-generation personalised medicine for high-risk paediatric cancer patients - The INFORM pilot study. Eur J Cancer, 2016. 65: p. 91-101.
268. Rusch, M., et al., Clinical cancer genomic profiling by three-platform sequencing of whole genome, whole exome and transcriptome. Nat Commun, 2018. 9(1): p. 3962.
269. Harttrampf, A.C., et al., Molecular Screening for Cancer Treatment Optimization (MOSCATO-01) in Pediatric Patients: A Single-Institutional Prospective Molecular Stratification Trial. Clin Cancer Res, 2017. 23(20): p. 6101-6112.
270. Allen, C.E., et al., Target and Agent Prioritization for the Children's Oncology Group-National Cancer Institute Pediatric MATCH Trial. J Natl Cancer Inst, 2017. 109(5).
271. Moreno, L., et al., Early phase clinical trials of anticancer agents in children and adolescents — an ITCC perspective. Nature Reviews Clinical Oncology, 2017. 14(8): p. 497-507.
272. Hamisch, C., et al., Update on the diagnostic value and safety of stereotactic biopsy for pediatric brainstem tumors: a systematic review and meta-analysis of 735 cases. J Neurosurg Pediatr, 2017. 20(3): p. 261-268.
273. Gupta, N., et al., Prospective feasibility and safety assessment of surgical biopsy for patients with newly diagnosed diffuse intrinsic pontine glioma. Neuro Oncol, 2018. 20(11): p. 1547-1555.
274. Yates, L.R., et al., Subclonal diversification of primary breast cancer revealed by multiregion sequencing. Nat Med, 2015. 21(7): p. 751-9.
275. Gerlinger, M., et al., Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. The New England journal of medicine, 2012. 366(10): p. 883-892.
276. Parikh, A.R., et al., Liquid versus tissue biopsy for detecting acquired resistance and tumor heterogeneity in gastrointestinal cancers. Nat Med, 2019. 25(9): p. 1415-1421.
277. Cohen, J.D., et al., Combined circulating tumor DNA and protein biomarker-based liquid biopsy for the earlier detection of pancreatic cancers. Proc Natl Acad Sci U S A, 2017. 114(38): p. 10202-10207.
278. Cohen, J.D., et al., Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science, 2018. 359(6378): p. 926-930.
279. Dawson, S.J., et al., Analysis of circulating tumor DNA to monitor metastatic breast cancer. N Engl J Med, 2013. 368(13): p. 1199-209.
280. Mok, T., et al., Detection and Dynamic Changes of EGFR Mutations from Circulating Tumor DNA as a Predictor of Survival Outcomes in NSCLC Patients Treated with First-line Intercalated Erlotinib and Chemotherapy. Clin Cancer Res, 2015. 21(14): p. 3196-203.
281. Murtaza, M., et al., Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. Nature, 2013. 497(7447): p. 108-12.
282. Garcia-Murillas, I., et al., Assessment of Molecular Relapse Detection in Early-Stage Breast Cancer. JAMA Oncol, 2019.
283. Abbosh, C., et al., Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. Nature, 2017. 545(7655): p. 446-451.
284. Pearson, A., et al., Inactivating NF1 Mutations Are Enriched in Advanced Breast Cancer and Contribute to Endocrine Therapy Resistance. Clin Cancer Res, 2019.
285. De Rubis, G., S. Rajeev Krishnan, and M. Bebawy, Liquid Biopsies in Cancer Diagnosis, Monitoring, and Prognosis. Trends Pharmacol Sci, 2019. 40(3): p. 172-186.

286. Wan, J.C.M., et al., Liquid biopsies come of age: towards implementation of circulating tumour DNA. Nat Rev Cancer, 2017. 17(4): p. 223-238.
287. Andersson, D., et al., Circulating cell-free tumor DNA analysis in pediatric cancers. Mol Aspects Med, 2019: p. 100819.
288. Bettegowda, C., et al., Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med, 2014. 6(224): p. 224ra24.
289. Pan, W., et al., Brain tumor mutations detected in cerebral spinal fluid. Clin Chem, 2015. 61(3): p. 514-22.
290. Wang, Y., et al., Detection of tumor-derived DNA in cerebrospinal fluid of patients with primary tumors of the brain and spinal cord. Proc Natl Acad Sci U S A, 2015. 112(31): p. 9704-9.
291. De Mattos-Arruda, L., et al., Cerebrospinal fluid-derived circulating tumour DNA better represents the genomic alterations of brain tumours than plasma. 2015. 6: p. 8839.
292. Pentsova, E.I., et al., Evaluating Cancer of the Central Nervous System Through Next-Generation Sequencing of Cerebrospinal Fluid. J Clin Oncol, 2016. 34(20): p. 2404-15.
293. Bronkhorst, A.J., V. Ungerer, and S. Holdenrieder, The emerging role of cell-free DNA as a molecular marker for cancer management. Biomolecular Detection and Quantification, 2019. 17: p. 100087.
294. Diehl, F., et al., Circulating mutant DNA to assess tumor dynamics. Nat Med, 2008. 14(9): p. 985-90.
295. Sacher, A.G., et al., Prospective Validation of Rapid Plasma Genotyping for the Detection of EGFR and KRAS Mutations in Advanced Lung Cancer. JAMA Oncol, 2016. 2(8): p. 1014-22.
296. Gevensleben, H., et al., Noninvasive detection of HER2 amplification with plasma DNA digital PCR. Clin Cancer Res, 2013. 19(12): p. 3276-84.
297. Forshew, T., et al., Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. Sci Transl Med, 2012. 4(136): p. 136ra68.
298. Plagnol, V., et al., Analytical validation of a next generation sequencing liquid biopsy assay for high sensitivity broad molecular profiling. PLoS One, 2018. 13(3): p. e0193802.
299. Newman, A.M., et al., An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. Nat Med, 2014. 20(5): p. 548-54.
300. Newman, A.M., et al., Integrated digital error suppression for improved detection of circulating tumor DNA. Nat Biotechnol, 2016. 34(5): p. 547-555.
301. Lanman, R.B., et al., Analytical and Clinical Validation of a Digital Sequencing Panel for Quantitative, Highly Accurate Evaluation of Cell-Free Circulating Tumor DNA. PloS one, 2015. 10(10): p. e0140712-e0140712.
302. Odegaard, J.I., et al., Validation of a Plasma-Based Comprehensive Cancer Genotyping Assay Utilizing Orthogonal Tissue- and Plasma-Based Methodologies. Clin Cancer Res, 2018. 24(15): p. 3539-3549.
303. Cohen, J.D., et al., Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science (New York, N.Y.), 2018. 359(6378): p. 926-930.
304. Combaret, V., et al., Detection of tumor ALK status in neuroblastoma patients using peripheral blood. Cancer Med, 2015. 4(4): p. 540-50.
305. Chicard, M., et al., Genomic copy number profiling using circulating free tumor DNA highlights heterogeneity in neuroblastoma. Clin Cancer Res, 2016.
306. Van Roy, N., et al., Shallow Whole Genome Sequencing on Circulating Cell-Free DNA Allows Reliable Noninvasive Copy-Number Profiling in Neuroblastoma Patients. Clin Cancer Res, 2017. 23(20): p. 6305-6314.

307. Krumbholz, M., et al., Genomic EWSR1 Fusion Sequence as Highly Sensitive and Dynamic Plasma Tumor Marker in Ewing Sarcoma. Clin Cancer Res, 2016. 22(17): p. 4356-65.
308. Shukla, N.N., et al., Plasma DNA-based molecular diagnosis, prognostication, and monitoring of patients with EWSR1 fusion-positive sarcomas. JCO Precis Oncol, 2017. 2017.
309. Eguchi-Ishimae, M., et al., Early detection of the PAX3-FOXO1 fusion gene in circulating tumor-derived DNA in a case of alveolar rhabdomyosarcoma. Genes Chromosomes Cancer, 2019. 58(8): p. 521-529.
310. Berry, J.L., et al., Genomic cfDNA Analysis of Aqueous Humor in Retinoblastoma Predicts Eye Salvage: The Surrogate Tumor Biopsy for Retinoblastoma. Mol Cancer Res, 2018. 16(11): p. 1701-1712.
311. Gerrish, A., et al., Non-invasive diagnosis of retinoblastoma using cell-free DNA from aqueous humour. Br J Ophthalmol, 2019.
312. Jimenez, I., et al., Circulating tumor DNA analysis enables molecular characterization of pediatric renal tumors at diagnosis. Int J Cancer, 2019. 144(1): p. 68-79.
313. Ueno-Yokohata, H., et al., Preoperative diagnosis of clear cell sarcoma of the kidney by detection of BCOR internal tandem duplication in circulating tumor DNA. Genes Chromosomes Cancer, 2018. 57(10): p. 525-529.
314. Paret, C., et al., Personalized therapy: CNS HGNET-BCOR responsiveness to arsenic trioxide combined with radiotherapy. Oncotarget, 2017. 8(69): p. 114210-114225.
315. Huang, T.Y., et al., Detection of Histone H3 mutations in cerebrospinal fluid-derived tumor DNA from children with diffuse midline glioma. Acta neuropathologica communications, 2017. 5(1): p. 28-28.
316. Stallard, S., et al., CSF H3F3A K27M circulating tumor DNA copy number quantifies tumor growth and in vitro treatment response. Acta Neuropathol Commun, 2018. 6(1): p. 80.
317. Panditharatna, E., et al., Clinically Relevant and Minimally Invasive Tumor Surveillance of Pediatric Diffuse Midline Gliomas Using Patient-Derived Liquid Biopsy. Clin Cancer Res, 2018. 24(23): p. 5850-5859.
318. Mueller, S., et al., A pilot precision medicine trial for children with diffuse intrinsic pontine glioma-PNOC003: A report from the Pacific Pediatric Neuro-Oncology Consortium. Int J Cancer, 2019. 145(7): p. 1889-1901.
319. Heritier, S., et al., Circulating cell-free BRAF(V600E) as a biomarker in children with Langerhans cell histiocytosis. Br J Haematol, 2017. 178(3): p. 457-467.
320. Klega, K., et al., Detection of Somatic Structural Variants Enables Quantification and Characterization of Circulating Tumor DNA in Children With Solid Tumors. JCO precision oncology, 2018. 2018: p. 10.1200/PO.17.00285.
321. Rodriguez, D., et al., Evaluation of the Implementation of the Response Assessment in Neuro-Oncology Criteria in the HERBY Trial of Pediatric Patients with Newly Diagnosed High-Grade Gliomas. AJNR Am J Neuroradiol, 2019. 40(3): p. 568-575.
322. KAPA HyperPlus/SeqCap EZ workflow: Improving Data Quality and Turnaround Times for Targeted Next-Generation Sequencing of FFPE DNA. 2016; Available from: https://www.kapabiosystems.com/assets/KAPA_Generating-high-quality-capture-libraries-from-FFPE-samples_Note.pdf.
323. Barretina, J., et al., The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature, 2012. 483(7391): p. 603-7.
324. Hinson, A.R.P., et al., Human Rhabdomyosarcoma Cell Lines for Rhabdomyosarcoma Research: Utility and Pitfalls. Front Oncol, 2013. 3.

325. Chen, Y., et al., Oncogenic mutations of ALK kinase in neuroblastoma. Nature, 2008. 455(7215): p. 971-4.
326. McDermott, U., et al., Genomic alterations of anaplastic lymphoma kinase may sensitize tumors to anaplastic lymphoma kinase inhibitors. Cancer Res, 2008. 68(9): p. 3389-95.
327. Eleveld, T.F., et al., Relapsed neuroblastomas show frequent RAS-MAPK pathway mutations. Nat Genet, 2015. 47(8): p. 864-71.
328. Ellison, D.W., et al., beta-Catenin status predicts a favorable outcome in childhood medulloblastoma: the United Kingdom Children's Cancer Study Group Brain Tumour Committee. J Clin Oncol, 2005. 23(31): p. 7951-7.
329. Hill, R.M., et al., Combined MYC and P53 defects emerge at medulloblastoma relapse and define rapidly progressive, therapeutically targetable disease. Cancer Cell, 2015. 27(1): p. 72-84.
330. Lamont, J.M., et al., Combined histopathological and molecular cytogenetic stratification of medulloblastoma patients. Clin Cancer Res, 2004. 10(16): p. 5482-93.
331. Thway, K., et al., Evaluation of the optimal provision of formalin-fixed, paraffin-embedded material for reverse transcription-PCR in soft-tissue tumour diagnosis. J Clin Pathol, 2016.
332. Thway, K., et al., Utility of sarcoma-specific fusion gene analysis in paraffin-embedded material for routine diagnosis at a specialist centre. J Clin Pathol, 2010. 63(6): p. 508-12.
333. Guan, P. and W.K. Sung, Structural variation detection using next-generation sequencing data: A comparative technical review. Methods, 2016. 102: p. 36-49.
334. Li, Y., et al., Tumor DNA in cerebral spinal fluid reflects clinical course in a patient with melanoma leptomeningeal brain metastases. J Neurooncol, 2016. 128(1): p. 93-100.
335. Rodriguez Gutierrez, D., et al., Radiological evaluation of newly diagnosed non-brainstem pediatric high-grade glioma in the HERBY phase II trial. Clin Cancer Res, 2020.
336. Ianevski, A., A.K. Giri, and T. Aittokallio, SynergyFinder 2.0: visual analytics of multi-drug combination synergies. Nucleic Acids Res, 2020.
337. Ianevski, A., et al., SynergyFinder: a web application for analyzing drug combination dose-response matrix data. Bioinformatics, 2017. 33(15): p. 2413-2415.
338. Kilkenny, C., et al., Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. PLoS Biol, 2010. 8(6): p. e1000412.
339. Seki, M., et al., Integrated genetic and epigenetic analysis defines novel molecular subgroups in rhabdomyosarcoma. 2015. 6: p. 7557.
340. Shukla, N., et al., Oncogene mutation profiling of pediatric solid tumors reveals significant subsets of embryonal rhabdomyosarcoma and neuroblastoma with mutated genes in growth signaling pathways. Clin Cancer Res, 2012. 18(3): p. 748-57.
341. Hosoi, H., et al., Recurrent somatic structural variations contribute to tumorigenesis in pediatric osteosarcoma. Nat Commun, 2014. 7(1): p. 104-12.
342. Jones, D.T., et al., Dissecting the genomic complexity underlying medulloblastoma. Nature, 2012. 488(7409): p. 100-5.
343. Pugh, T.J., et al., Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. Nature, 2012. 488(7409): p. 106-10.
344. Robinson, G., et al., Novel mutations target distinct subgroups of medulloblastoma. Nature, 2012. 488(7409): p. 43-8.
345. Alexandrov, L.B., et al., Signatures of mutational processes in human cancer. Nature, 2013. 500(7463): p. 415-21.
346. Hertwig, F., M. Peifer, and M. Fischer, Telomere maintenance is pivotal for high-risk neuroblastoma. Cell Cycle, 2016. 15(3): p. 311-2.

347. Salto-Tellez, M. and D. Gonzalez de Castro, Next-generation sequencing: a change of paradigm in molecular diagnostic validation. J Pathol, 2014. 234(1): p. 5-10.
348. Hagemann, I.S., C.E. Cottrell, and C.M. Lockwood, Design of targeted, capture-based, next generation sequencing tests for precision cancer therapy. Cancer Genet, 2013. 206(12): p. 420-31.
349. Cheng, D.T., et al., Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): A Hybridization Capture-Based Next-Generation Sequencing Clinical Assay for Solid Tumor Molecular Oncology. J Mol Diagn, 2015. 17(3): p. 251-64.
350. Frampton, G.M., et al., Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. Nat Biotechnol, 2013. 31(11): p. 1023-31.
351. Cottrell, C.E., et al., Validation of a next-generation sequencing assay for clinical molecular oncology. J Mol Diagn, 2014. 16(1): p. 89-105.
352. Pritchard, C.C., et al., Validation and implementation of targeted capture and sequencing for the detection of actionable mutation, copy number variation, and gene rearrangement in clinical cancer specimens. J Mol Diagn, 2014. 16(1): p. 56-67.
353. Harris, M.H., et al., Multicenter Feasibility Study of Tumor Molecular Profiling to Inform Therapeutic Decisions in Advanced Pediatric Solid Tumors: The Individualized Cancer Therapy (iCat) Study. JAMA Oncol, 2016.
354. Izquierdo, E., et al., Development of a targeted sequencing approach to identify prognostic, predictive and diagnostic markers in paediatric solid tumours. Oncotarget, 2017. 8(67): p. 112036-112050.
355. Burford, A., et al., The ten-year evolutionary trajectory of a highly recurrent paediatric high grade neuroepithelial tumour with MN1:BEND2 fusion. Sci Rep, 2018. 8(1): p. 1032.
356. Richards, S., et al., Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. 2015. 17(5): p. 405-24.
357. Johnson, B.E., et al., Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. Science, 2014. 343(6167): p. 189-93.
358. Hawkins, C., et al., BRAF-KIAA1549 fusion predicts better clinical outcome in pediatric low-grade astrocytoma. Clin Cancer Res, 2011. 17(14): p. 4790-8.
359. Middleton, G., et al., The National Lung Matrix Trial: translating the biology of stratification in advanced non-small-cell lung cancer. Ann Oncol, 2015. 26(12): p. 2464-9.
360. Do, K., G. O'Sullivan Coyne, and A.P. Chen, An overview of the NCI precision medicine trials-NCI MATCH and MPACT. Chin Clin Oncol, 2015. 4(3): p. 31.
361. Chapman, P.B., et al., Improved survival with vemurafenib in melanoma with BRAF V600E mutation. N Engl J Med, 2011. 364(26): p. 2507-16.
362. Lynch, T.J., et al., Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med, 2004. 350(21): p. 2129-39.
363. Grobner, S.N., et al., The landscape of genomic alterations across childhood cancers. Nature, 2018. 555(7696): p. 321-327.
364. Mu, W., et al., Detection of structural variation using target captured next-generation sequencing data for genetic diagnostic testing. Genetics in Medicine, 2019. 21(7): p. 1603-1610.
365. Parker, M., et al., C11orf95-RELA fusions drive oncogenic NF-kappaB signalling in ependymoma. Nature, 2014. 506(7489): p. 451-5.

366. Doebele, R.C., et al., Entrectinib in patients with advanced or metastatic NTRK fusion-positive solid tumours: integrated analysis of three phase 1-2 trials. Lancet Oncol, 2019.
367. Bender, S., et al., Recurrent MET fusion genes represent a drug target in pediatric glioblastoma. Nature Medicine, 2016. 22(11): p. 1314-1320.
368. Sturm, D., et al., New Brain Tumor Entities Emerge from Molecular Classification of CNS- PNETs. Cell, 2016. 164(5): p. 1060-1072.
369. Pfaff, E., et al., Brainstem biopsy in pediatric diffuse intrinsic pontine glioma in the era of precision medicine: the INFORM study experience. European Journal of Cancer, 2019. 114: p. 27-35.
370. Guerreiro Stucklin, A.S., et al., Alterations in ALK/ROS1/NTRK/MET drive a group of infantile hemispheric gliomas. Nat Commun, 2019. 10(1): p. 4343.
371. Recurrent MET fusion genes represent a drug target in pediatric glioblastoma. Nat Med, 2016. 22(11): p. 1314-1320.
372. Johnson, A., et al., Comprehensive Genomic Profiling of 282 Pediatric Low- and High-Grade Gliomas Reveals Genomic Drivers, Tumor Mutational Burden, and Hypermutation Signatures. The oncologist, 2017. 22(12): p. 1478-1490.
373. Aghajan, Y., et al., Novel PPP1CB-ALK fusion protein in a high-grade glioma of infancy. BMJ Case Rep, 2016. 2016.
374. Maruggi, M., et al., A novel KIF5B-ALK fusion in a child with an atypical central nervous system inflammatory myofibroblastic tumour. BMJ Case Rep, 2018. 2018.
375. Geyer, J.R., et al., Survival of infants with malignant astrocytomas. A Report from the Childrens Cancer Group. Cancer, 1995. 75(4): p. 1045-50.
376. Grundy, R.G., et al., Primary postoperative chemotherapy without radiotherapy for treatment of brain tumours other than ependymoma in children under 3 years: results of the first UKCCSG/SIOP CNS 9204 trial. Eur J Cancer, 2010. 46(1): p. 120-33.
377. Chakravarty, D., et al., OncoKB: A Precision Oncology Knowledge Base. JCO Precis Oncol, 2017. 2017.
378. Forbes, S.A., et al., COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. Nucleic Acids Res, 2011. 39(Database issue): p. D945-50.
379. Gatz, S.A., et al., Chemotherapy responsiveness in a patient with multiply relapsed ameloblastic fibro-odontosarcoma of the maxilla. Pediatr Blood Cancer, 2015. 62(11): p. 2029-32.
380. Campbell, B.B., et al., Comprehensive Analysis of Hypermutation in Human Cancer. Cell, 2017. 171(5): p. 1042-1056.e10.
381. Van Allen, E.M., et al., The genetic landscape of clinical resistance to RAF inhibition in metastatic melanoma. Cancer discovery, 2014. 4(1): p. 94-109.
382. Whittaker, S.R., et al., A genome-scale RNA interference screen implicates NF1 loss in resistance to RAF inhibition. Cancer discovery, 2013. 3(3): p. 350-362.
383. George, S.L., et al., A tailored molecular profiling programme for children with cancer to identify clinically actionable genetic alterations. Eur J Cancer, 2019. 121: p. 224-235.
384. Cohen, B., et al., Pediatric Oncology Provider Views on Performing a Biopsy of Solid Tumors in Children with Relapsed or Refractory Disease for the Purpose of Genomic Profiling. Ann Surg Oncol, 2016. 23(Suppl 5): p. 990-997.
385. Mateo, J., et al., A First-Time-in-Human Study of GSK2636771, a Phosphoinositide 3 Kinase Beta-Selective Inhibitor, in Patients with Advanced Solid Tumors. Clin Cancer Res, 2017. 23(19): p. 5981-5992.
386. Andre, F., et al., Alpelisib for PIK3CA-Mutated, Hormone Receptor-Positive Advanced Breast Cancer. N Engl J Med, 2019. 380(20): p. 1929-1940.

387. Voss, M.H., et al., PTEN Expression, Not Mutation Status in TSC1, TSC2, or mTOR, Correlates with the Outcome on Everolimus in Patients with Renal Cell Carcinoma Treated on the Randomized RECORD-3 Trial. Clin Cancer Res, 2019. 25(2): p. 506-514.
388. Kwiatkowski, D.J., et al., Mutations in TSC1, TSC2, and MTOR Are Associated with Response to Rapalogs in Patients with Metastatic Renal Cell Carcinoma. Clin Cancer Res, 2016. 22(10): p. 2445-2452.
389. Franz, D.N., et al., Efficacy and safety of everolimus for subependymal giant cell astrocytomas associated with tuberous sclerosis complex (EXIST-1): a multicentre, randomised, placebo-controlled phase 3 trial. Lancet, 2013. 381(9861): p. 125-32.
390. Zhang, J., et al., Combined BRAF(V600E) and MEK blockade for BRAF(V600E)-mutant gliomas. J Neurooncol, 2017. 131(3): p. 495-505.
391. Perez, E. and D. Capper, DNA-Methylation-based Classification of Paediatric Brain Tumours. Neuropathol Appl Neurobiol, 2020.
392. Bao, Z.S., et al., RNA-seq of 272 gliomas revealed a novel, recurrent PTPRZ1-MET fusion transcript in secondary glioblastomas. Genome Res, 2014. 24(11): p. 1765-73.
393. Roskoski, R., Jr., Properties of FDA-approved small molecule protein kinase inhibitors: A 2020 update. Pharmacol Res, 2019. 152: p. 104609.
394. Koschmann, C., et al., Clinically Integrated Sequencing Alters Therapy in Children and Young Adults With High-Risk Glial Brain Tumors. JCO Precision Oncology, 2018(2): p. 1-34.
395. Akshulakov, S.K., et al., Current Trends for Improving Safety of Stereotactic Brain Biopsies: Advanced Optical Methods for Vessel Avoidance and Tumor Detection. Front Oncol, 2019. 9: p. 947.
396. Hayashi, M., et al., Highly personalized detection of minimal Ewing sarcoma disease burden from plasma tumor DNA. Cancer, 2016. 122(19): p. 3015-23.
397. Pan, C., et al., Molecular profiling of tumors of the brainstem by sequencing of CSF-derived circulating tumor DNA. Acta Neuropathol, 2019. 137(2): p. 297-306.
398. Miller, A.M., et al., Tracking tumour evolution in glioma through liquid biopsies of cerebrospinal fluid. Nature, 2019. 565(7741): p. 654-658.
399. Martinez-Ricarte, F., et al., Molecular Diagnosis of Diffuse Gliomas through Sequencing of Cell-Free Circulating Tumor DNA from Cerebrospinal Fluid. Clin Cancer Res, 2018. 24(12): p. 2812-2819.
400. Piccioni, D.E., et al., Analysis of cell-free circulating tumor DNA in 419 patients with glioblastoma and other primary brain tumors. CNS Oncol, 2019. 8(2): p. Cns34.
401. Varlet, P., et al., WHO grade has no prognostic value in the pediatric high-grade glioma included in the HERBY trial. Neuro Oncol, 2019.
402. Thompson, J.C., et al., Detection of Therapeutically Targetable Driver and Resistance Mutations in Lung Cancer Patients by Next-Generation Sequencing of Cell-Free Circulating Tumor DNA. Clin Cancer Res, 2016. 22(23): p. 5772-5782.
403. Mouliere, F., et al., Enhanced detection of circulating tumor DNA by fragment size analysis. Science translational medicine, 2018. 10(466): p. eaat4921.
404. Christensen, E., et al., Early Detection of Metastatic Relapse and Monitoring of Therapeutic Efficacy by Ultra-Deep Sequencing of Plasma Cell-Free DNA in Patients With Urothelial Bladder Carcinoma. J Clin Oncol, 2019. 37(18): p. 1547-1557.
405. Hirotsu, Y., et al., Dual-molecular barcode sequencing detects rare variants in tumor and cell free DNA in plasma. Scientific Reports, 2020. 10(1): p. 3391.
406. Roujeau, T., et al., Stereotactic biopsy of diffuse pontine lesions in children. Journal of neurosurgery, 2007. 107(1 Suppl): p. 1-4.

407. Cage, T.A., et al., A systematic review of treatment outcomes in pediatric patients with intracranial ependymomas. J Neurosurg Pediatr, 2013. 11(6): p. 673-81.
408. Northcott, P.A., S.M. Pfister, and D.T. Jones, Next-generation (epi)genetic drivers of childhood brain tumours and the outlook for targeted therapies. Lancet Oncol, 2015. 16(6): p. e293-302.
409. Gao, H., et al., High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response. Nat Med, 2015. 21(11): p. 1318-25.
410. Vlachogiannis, G., et al., Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. Science, 2018. 359(6378): p. 920-926.
411. Campbell, K.M., et al., Oral Cavity Squamous Cell Carcinoma Xenografts Retain Complex Genotypes and Intertumor Molecular Heterogeneity. Cell Rep, 2018. 24(8): p. 2167-2178.
412. Stebbing, J., et al., Patient-derived xenografts for individualized care in advanced sarcoma. Cancer, 2014. 120(13): p. 2006-15.
413. Pauli, C., et al., Personalized In Vitro and In Vivo Cancer Models to Guide Precision Medicine. Cancer Discov, 2017. 7(5): p. 462-477.
414. Vargas, R., et al., Case study: patient-derived clear cell adenocarcinoma xenograft model longitudinally predicts treatment response. npj Precision Oncology, 2018. 2(1): p. 14.
415. Lee, J.K., et al., Pharmacogenomic landscape of patient-derived tumor cells informs precision oncology therapy. Nat Genet, 2018. 50(10): p. 1399-1411.
416. Stewart, E., et al., Orthotopic patient-derived xenografts of paediatric solid tumours. Nature, 2017. 549(7670): p. 96-100.
417. Brabetz, S., et al., A biobank of patient-derived pediatric brain tumor models. Nat Med, 2018. 24(11): p. 1752-1761.
418. Rokita, J.L., et al., Genomic Profiling of Childhood Tumor Patient-Derived Xenograft Models to Enable Rational Clinical Trial Design. Cell Rep, 2019. 29(6): p. 1675-1689.e9.
419. Loh, A.H.P., et al., Combinatorial screening using orthotopic patient derived xenograft-expanded early phase cultures of osteosarcoma identify novel therapeutic drug combinations. Cancer Lett, 2019. 442: p. 262-270.
420. Lin, G.L., et al., Therapeutic strategies for diffuse midline glioma from high-throughput combination drug screening. Sci Transl Med, 2019. 11(519).
421. Hall, M.D., et al., First clinical experience with DRD2/3 antagonist ONC201 in H3 K27M-mutant pediatric diffuse intrinsic pontine glioma: a case report. J Neurosurg Pediatr, 2019: p. 1-7.
422. Geoerger, B., et al., Innovative Therapies for Children with Cancer pediatric phase I study of erlotinib in brainstem glioma and relapsing/refractory brain tumors. Neuro-oncology, 2011. 13(1): p. 109-18.
423. Gilbertson, R.J., et al., ERBB1 is amplified and overexpressed in high-grade diffusely infiltrative pediatric brain stem glioma. Clinical cancer research : an official journal of the American Association for Cancer Research, 2003. 9(10 Pt 1): p. 3620-4.
424. Geoerger, B., et al., Target-driven exploratory study of imatinib mesylate in children with solid malignancies by the Innovative Therapies for Children with Cancer (ITCC) European Consortium. Eur J Cancer, 2009. 45(13): p. 2342-51.
425. Puget, S., et al., Mesenchymal transition and PDGFRA amplification/mutation are key distinct oncogenic events in pediatric diffuse intrinsic pontine gliomas. PLoS One, 2012. 7(2): p. e30313.
426. Zarghooni, M., et al., Whole-genome profiling of pediatric diffuse intrinsic pontine gliomas highlights platelet-derived growth factor receptor alpha and poly (ADP-ribose) polymerase as potential therapeutic targets. Journal of clinical oncology

: official journal of the American Society of Clinical Oncology, 2010. 28(8): p. 1337-44.
427. Truffaux, N., et al., Preclinical evaluation of dasatinib alone and in combination with cabozantinib for the treatment of diffuse intrinsic pontine glioma. Neuro Oncol, 2014.
428. Broniscer, A., et al., Phase I trial, pharmacokinetics, and pharmacodynamics of vandetanib and dasatinib in children with newly diagnosed diffuse intrinsic pontine glioma. Clin Cancer Res, 2013. 19(11): p. 3050-8.
429. Chinnaiyan, P., et al., RTOG 0913: a phase 1 study of daily everolimus (RAD001) in combination with radiation therapy and temozolomide in patients with newly diagnosed glioblastoma. Int J Radiat Oncol Biol Phys, 2013. 86(5): p. 880-4.
430. Yang, W., et al., Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res, 2013. 41(Database issue): p. D955-61.
431. Chi, A.S., et al., Pediatric and adult H3 K27M-mutant diffuse midline glioma treated with the selective DRD2 antagonist ONC201. J Neurooncol, 2019. 145(1): p. 97-105.
432. Tauziède-Espariat, A., et al., An integrative radiological, histopathological and molecular analysis of pediatric pontine histone-wildtype glioma with MYCN amplification (HGG-MYCN). Acta Neuropathologica Communications, 2019. 7(1): p. 87.
433. Castel, D., et al., Histone H3 wild-type DIPG/DMG overexpressing EZHIP extend the spectrum diffuse midline gliomas with PRC2 inhibition beyond H3-K27M mutation. Acta Neuropathologica, 2020.
434. Turner, K.M., et al., Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. Nature, 2017. 543(7643): p. 122-125.
435. Sanborn, J.Z., et al., Double minute chromosomes in glioblastoma multiforme are revealed by precise reconstruction of oncogenic amplicons. Cancer research, 2013. 73(19): p. 6036-6045.
436. Nathanson, D.A., et al., Targeted Therapy Resistance Mediated by Dynamic Regulation of Extrachromosomal Mutant EGFR DNA. Science, 2014. 343(6166): p. 72.
437. Zheng, S., et al., A survey of intragenic breakpoints in glioblastoma identifies a distinct subset associated with poor survival. Genes & development, 2013. 27(13): p. 1462-1472.
438. Nikolaev, S., et al., Extrachromosomal driver mutations in glioblastoma and low-grade glioma. Nat Commun, 2014. 5: p. 5690.
439. deCarvalho, A.C., et al., Discordant inheritance of chromosomal and extrachromosomal DNA elements contributes to dynamic disease evolution in glioblastoma. Nat Genet, 2018. 50(5): p. 708-717.
440. Von Hoff, D.D., et al., Elimination of extrachromosomally amplified MYC genes from human tumor cells reduces their tumorigenicity. Proc Natl Acad Sci U S A, 1992. 89(17): p. 8165-9.
441. Caretti, V., et al., WEE1 kinase inhibition enhances the radiation response of diffuse intrinsic pontine gliomas. Mol Cancer Ther, 2013. 12(2): p. 141-50.
442. Mueller, S., et al., Targeting Wee1 for the treatment of pediatric high-grade gliomas. Neuro Oncol, 2014. 16(3): p. 352-60.
443. Sanai, N., et al., Phase 0 Trial of AZD1775 in First-Recurrence Glioblastoma Patients. Clin Cancer Res, 2018. 24(16): p. 3820-3828.
444. Hirai, H., et al., Small-molecule inhibition of Wee1 kinase by MK-1775 selectively sensitizes p53-deficient tumor cells to DNA-damaging agents. Mol Cancer Ther, 2009. 8(11): p. 2992-3000.

445. Webster, P.J., et al., AZD1775 induces toxicity through double-stranded DNA breaks independently of chemotherapeutic agents in p53-mutated colorectal cancer cells. Cell cycle (Georgetown, Tex.), 2017. 16(22): p. 2176-2182.
446. Ku, B.M., et al., Mutational status of TP53 defines the efficacy of Wee1 inhibitor AZD1775 in KRAS-mutant non-small cell lung cancer. Oncotarget, 2017. 8(40): p. 67526-67537.
447. Diab, A., et al., Multiple Defects Sensitize p53-Deficient Head and Neck Cancer Cells to the WEE1 Kinase Inhibition. Mol Cancer Res, 2019. 17(5): p. 1115-1128.
448. McGranahan, N. and C. Swanton, Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. Cancer Cell, 2015. 27(1): p. 15-26.
449. Hata, A.N., et al., Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. Nat Med, 2016. 22(3): p. 262-9.
450. Shi, H., et al., Acquired resistance and clonal evolution in melanoma during BRAF inhibitor therapy. Cancer Discov, 2014. 4(1): p. 80-93.
451. Bhang, H.E., et al., Studying clonal dynamics in response to cancer therapy using high-complexity barcoding. Nat Med, 2015. 21(5): p. 440-8.
452. Turke, A.B., et al., Preexistence and clonal selection of MET amplification in EGFR mutant NSCLC. Cancer cell, 2010. 17(1): p. 77-88.
453. Ramirez, M., et al., Diverse drug-resistance mechanisms can emerge from drug-tolerant cancer persister cells. Nature Communications, 2016. 7(1): p. 10690.
454. Drilon, A., et al., Efficacy of Larotrectinib in TRK Fusion-Positive Cancers in Adults and Children. N Engl J Med, 2018. 378(8): p. 731-739.
455. Arcila, M.E., et al., Rebiopsy of lung cancer patients with acquired resistance to EGFR inhibitors and enhanced detection of the T790M mutation using a locked nucleic acid-based assay. Clin Cancer Res, 2011. 17(5): p. 1169-80.
456. Chandrasekhar, C., P.S. Kumar, and P.V.G.K. Sarma, Novel mutations in the kinase domain of BCR-ABL gene causing imatinib resistance in chronic myeloid leukemia patients. Scientific Reports, 2019. 9(1): p. 2412.
457. Bresler, S.C., et al., ALK mutations confer differential oncogenic activation and sensitivity to ALK inhibition therapy in neuroblastoma. Cancer Cell, 2014. 26(5): p. 682-94.
458. Nissan, M.H., et al., Loss of NF1 in cutaneous melanoma is associated with RAS activation and MEK dependence. Cancer Res, 2014. 74(8): p. 2340-50.
459. Hrustanovic, G., et al., RAS-MAPK dependence underlies a rational polytherapy strategy in EML4-ALK-positive lung cancer. Nat Med, 2015. 21(9): p. 1038-47.
460. Engelman, J.A., et al., MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. Science, 2007. 316(5827): p. 1039-43.
461. Cocco, E., et al., Resistance to TRK inhibition mediated by convergent MAPK pathway activation. Nat Med, 2019. 25(9): p. 1422-1427.
462. Boumahdi, S. and F.J. de Sauvage, The great escape: tumour cell plasticity in resistance to targeted therapy. Nature Reviews Drug Discovery, 2020. 19(1): p. 39-56.
463. Suda, K., et al., CD44 Facilitates Epithelial-to-Mesenchymal Transition Phenotypic Change at Acquisition of Resistance to EGFR Kinase Inhibitors in Lung Cancer. Molecular Cancer Therapeutics, 2018. 17(10): p. 2257.
464. Mu, P., et al., SOX2 promotes lineage plasticity and antiandrogen resistance in TP53- and RB1-deficient prostate cancer. Science, 2017. 355(6320): p. 84-88.
465. Biehs, B., et al., A cell identity switch allows residual BCC to survive Hedgehog pathway inhibition. Nature, 2018. 562(7727): p. 429-433.

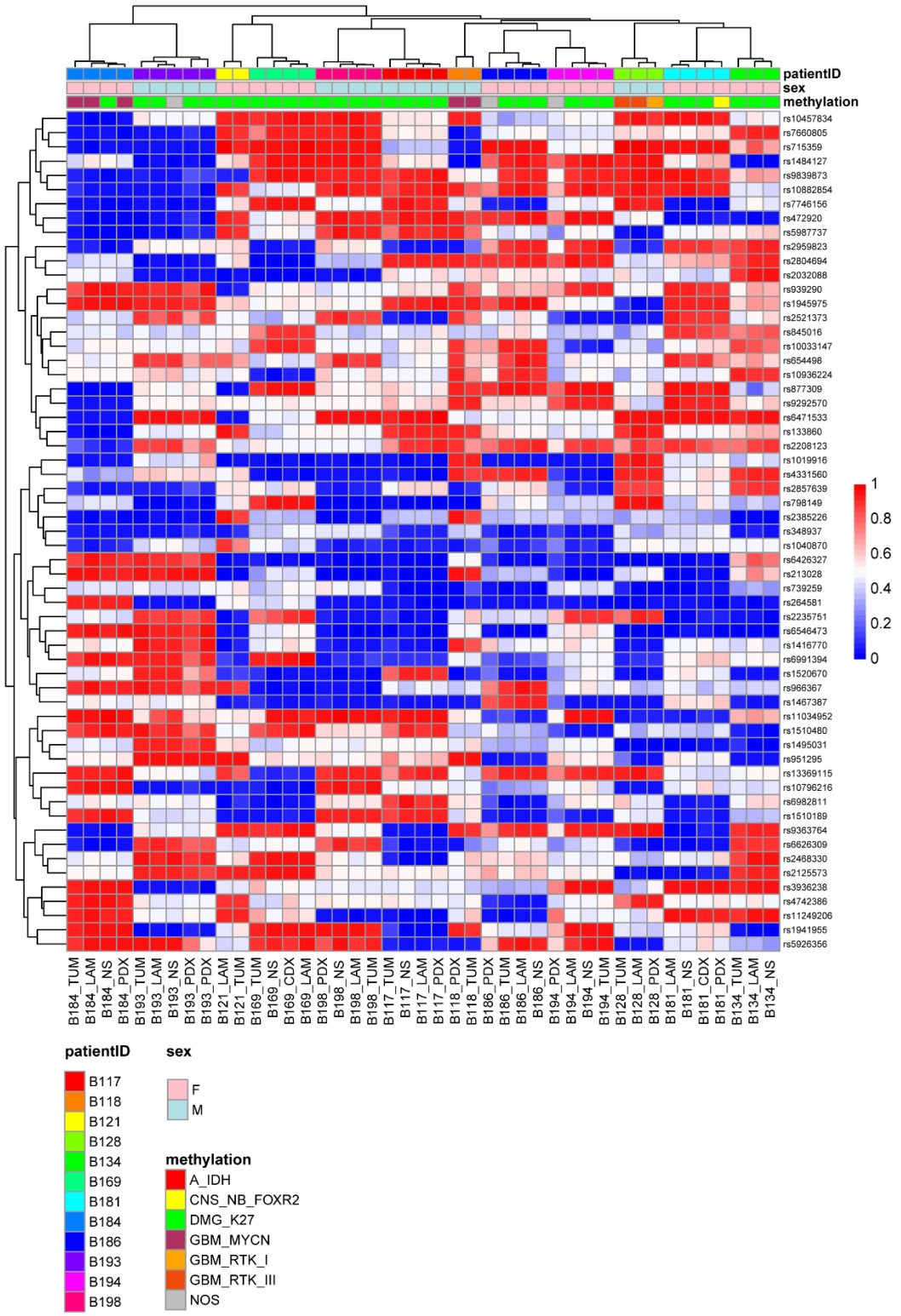
466. Greaves, M., Evolutionary Determinants of Cancer. Cancer Discovery, 2015. 5(8): p. 806.
467. Turajlic, S., et al., Resolving genetic heterogeneity in cancer. Nat Rev Genet, 2019. 20(7): p. 404-416.
468. Cortes, J.E., et al., Ponatinib in refractory Philadelphia chromosome-positive leukemias. N Engl J Med, 2012. 367(22): p. 2075-88.
469. Park, S.R., et al., Safety and feasibility of targeted agent combinations in solid tumours. Nature Reviews Clinical Oncology, 2013. 10(3): p. 154-168.
470. Avruch, J., et al., Ras activation of the Raf kinase: tyrosine kinase recruitment of the MAP kinase cascade. Recent Prog Horm Res, 2001. 56: p. 127-55.
471. Zaman, A., W. Wu, and T.G. Bivona, Targeting Oncogenic BRAF: Past, Present, and Future. Cancers, 2019. 11(8): p. 1197.
472. Yao, Z., et al., BRAF Mutants Evade ERK-Dependent Feedback by Different Mechanisms that Determine Their Sensitivity to Pharmacologic Inhibition. Cancer Cell, 2015. 28(3): p. 370-83.
473. Yao, Z., et al., Tumours with class 3 BRAF mutants are sensitive to the inhibition of activated RAS. Nature, 2017. 548(7666): p. 234-238.
474. Flaherty, K.T., et al., Inhibition of mutated, activated BRAF in metastatic melanoma. N Engl J Med, 2010. 363(9): p. 809-19.
475. Flaherty, K.T., et al., Improved survival with MEK inhibition in BRAF-mutated melanoma. N Engl J Med, 2012. 367(2): p. 107-14.
476. Emery, C.M., et al., MEK1 mutations confer resistance to MEK and B-RAF inhibition. Proceedings of the National Academy of Sciences, 2009. 106(48): p. 20411.
477. Wagle, N., et al., Dissecting therapeutic resistance to RAF inhibition in melanoma by tumor genomic profiling. J Clin Oncol, 2011. 29(22): p. 3085-96.
478. Rizos, H., et al., BRAF inhibitor resistance mechanisms in metastatic melanoma: spectrum and clinical impact. Clin Cancer Res, 2014. 20(7): p. 1965-77.
479. Johnson, D.B., et al., Acquired BRAF inhibitor resistance: A multicenter meta-analysis of the spectrum and frequencies, clinical behaviour, and phenotypic associations of resistance mechanisms. Eur J Cancer, 2015. 51(18): p. 2792-9.
480. Flaherty, K.T., et al., Combined BRAF and MEK inhibition in melanoma with BRAF V600 mutations. N Engl J Med, 2012. 367(18): p. 1694-703.
481. Long, G.V., et al., Combined BRAF and MEK inhibition versus BRAF inhibition alone in melanoma. N Engl J Med, 2014. 371(20): p. 1877-88.
482. Villanueva, J., et al., Concurrent MEK2 mutation and BRAF amplification confer resistance to BRAF and MEK inhibitors in melanoma. Cell Rep, 2013. 4(6): p. 1090-9.
483. Long, G.V., et al., Increased MAPK reactivation in early resistance to dabrafenib/trametinib combination therapy of BRAF-mutant metastatic melanoma. Nature Communications, 2014. 5(1): p. 5694.
484. Wagle, N., et al., MAP kinase pathway alterations in BRAF-mutant melanoma patients with acquired resistance to combined RAF/MEK inhibition. Cancer Discov, 2014. 4(1): p. 61-8.
485. Moriceau, G., et al., Tunable-combinatorial mechanisms of acquired resistance limit the efficacy of BRAF/MEK cotargeting but result in melanoma drug addiction. Cancer Cell, 2015. 27(2): p. 240-56.
486. Stone, T.J., et al., Comprehensive molecular characterisation of epilepsy-associated glioneuronal tumours. Acta Neuropathol, 2018. 135(1): p. 115-129.
487. Schindler, G., 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 Neuropathol, 2011. 121(3): p. 397-405.

488. Mistry, M., et al., BRAF mutation and CDKN2A deletion define a clinically distinct subgroup of childhood secondary high-grade glioma. J Clin Oncol, 2015. 33(9): p. 1015-22.
489. Jubb, H.C., et al., COSMIC-3D provides structural perspectives on cancer genetics for drug discovery. Nat Genet, 2018. 50(9): p. 1200-1202.
490. Jaiswal, B.S., et al., Somatic mutations in p85alpha promote tumorigenesis through class IA PI3K activation. Cancer cell, 2009. 16(6): p. 463-474.
491. Chang, M.T., et al., Identifying recurrent mutations in cancer reveals widespread lineage diversity and mutational specificity. Nat Biotechnol, 2016. 34(2): p. 155-63.
492. AACR Project GENIE: Powering Precision Medicine through an International Consortium. Cancer Discov, 2017. 7(8): p. 818-831.
493. Ross, R.L., et al., Identification of mutations in distinct regions of p85 alpha in urothelial cancer. PloS one, 2013. 8(12): p. e84411-e84411.
494. Cheung, L.W., et al., Naturally occurring neomorphic PIK3R1 mutations activate the MAPK pathway, dictating therapeutic response to MAPK pathway inhibitors. Cancer Cell, 2014. 26(4): p. 479-94.
495. Nykamp, K., et al., Sherloc: a comprehensive refinement of the ACMG-AMP variant classification criteria. Genet Med, 2017. 19(10): p. 1105-1117.
496. Brinckmann, A., et al., Detection of novel NF1 mutations and rapid mutation prescreening with Pyrosequencing. Electrophoresis, 2007. 28(23): p. 4295-301.
497. See, W.L., et al., Sensitivity of glioblastomas to clinically available MEK inhibitors is defined by neurofibromin 1 deficiency. Cancer Res, 2012. 72(13): p. 3350-9.
498. Ameratunga, M., et al., Prolonged disease control with MEK inhibitor in neurofibromatosis type I-associated glioblastoma. J Clin Pharm Ther, 2016. 41(3): p. 357-359.
499. Romo, C.G., et al., Trametinib for aggressive gliomas in adults with neurofibromatosis type 1. Journal of Clinical Oncology, 2019. 37(15_suppl): p. e13562-e13562.
500. Fangusaro, J., et al., Selumetinib in paediatric patients with BRAF-aberrant or neurofibromatosis type 1-associated recurrent, refractory, or progressive low-grade glioma: a multicentre, phase 2 trial. Lancet Oncol, 2019. 20(7): p. 1011-1022.
501. Noeparast, A., et al., Non-V600 BRAF mutations recurrently found in lung cancer predict sensitivity to the combination of Trametinib and Dabrafenib. Oncotarget, 2017. 8(36): p. 60094-60108.
502. Wu, P.-K. and J.-I. Park, MEK1/2 Inhibitors: Molecular Activity and Resistance Mechanisms. Seminars in Oncology, 2015. 42(6): p. 849-862.
503. Gao, Y., et al., Allele-Specific Mechanisms of Activation of MEK1 Mutants Determine Their Properties. Cancer Discov, 2018. 8(5): p. 648-661.
504. Emery, C.M., et al., BRAF-inhibitor Associated MEK Mutations Increase RAF-Dependent and -Independent Enzymatic Activity. Mol Cancer Res, 2017. 15(10): p. 1431-1444.
505. Catalanotti, F., et al., Phase II trial of MEK inhibitor selumetinib (AZD6244, ARRY-142886) in patients with BRAFV600E/K-mutated melanoma. Clinical cancer research : an official journal of the American Association for Cancer Research, 2013. 19(8): p. 2257-2264.
506. Segerman, A., et al., Clonal Variation in Drug and Radiation Response among Glioma-Initiating Cells Is Linked to Proneural-Mesenchymal Transition. Cell Rep, 2016. 17(11): p. 2994-3009.
507. Jing, J., et al., Comprehensive predictive biomarker analysis for MEK inhibitor GSK1120212. Mol Cancer Ther, 2012. 11(3): p. 720-9.

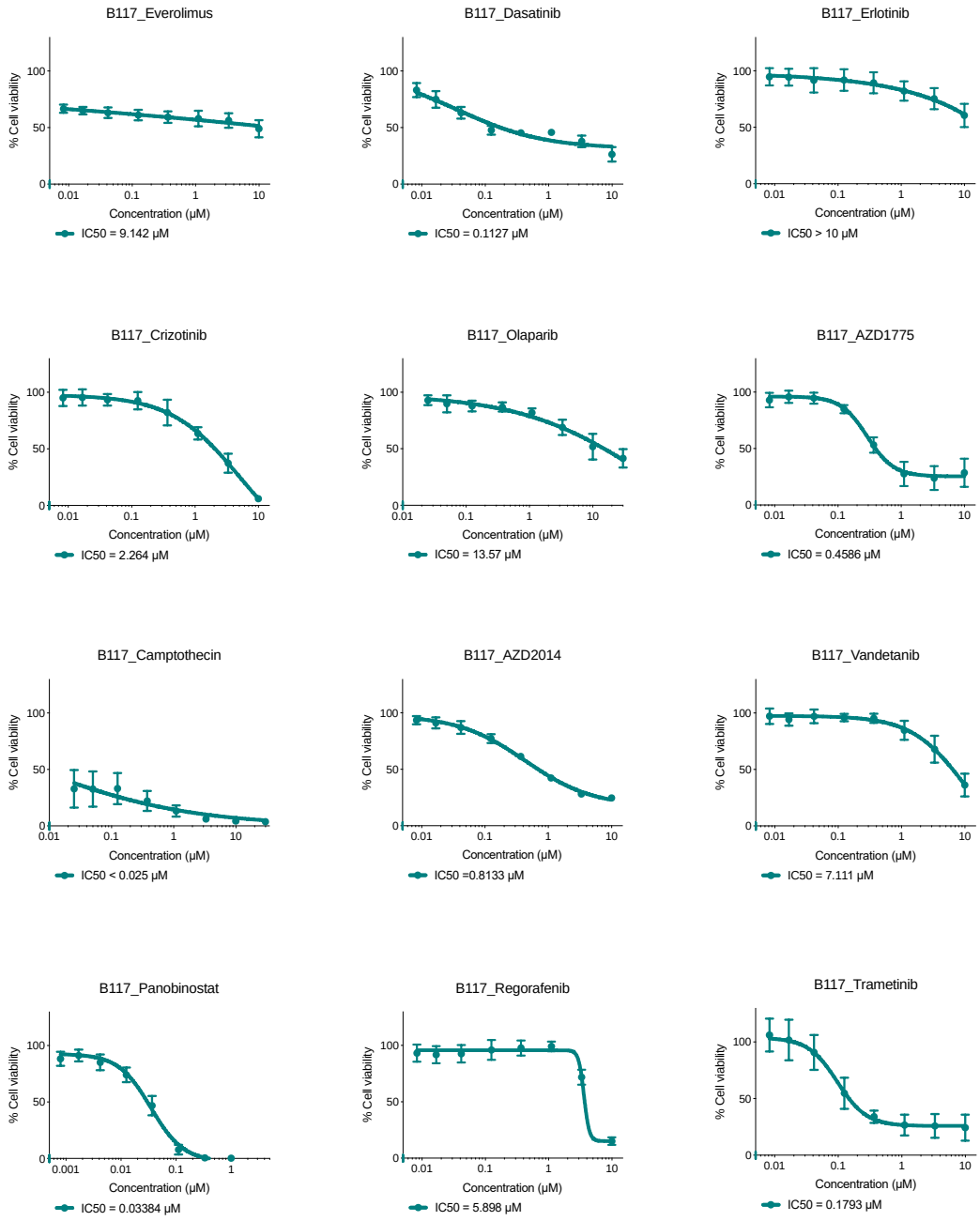
508. Wilson, C., et al., Overcoming EMT-associated resistance to anti-cancer drugs via Src/FAK pathway inhibition. Oncotarget, 2014. 5(17): p. 7328-41.
509. Rao, G., et al., Dasatinib sensitises KRAS-mutant cancer cells to mitogen-activated protein kinase kinase inhibitor via inhibition of TAZ activity. Eur J Cancer, 2018. 99: p. 37-48.
510. Parsons, D.W., et al., Diagnostic Yield of Clinical Tumor and Germline Whole-Exome Sequencing for Children With Solid Tumors. JAMA Oncol, 2016. 2(5): p. 616-624.
511. Lacayo, N.J., et al., Gene expression profiles at diagnosis in de novo childhood AML patients identify FLT3 mutations with good clinical outcomes. Blood, 2004. 104(9): p. 2646-54.

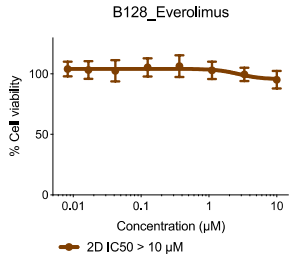
APPENDIX I

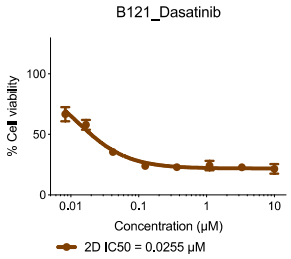
Appendix Figure 1 Heatmap of SNP genotype from methylation array of the BIOMEDE-UK models. 65 known commonly variable SNP measured with Minfi are represented from the methylation arrays for tumour, in vitro cultures (2D, LAM and 3D NS) and in vivo (PDX and CDX). Models derived from the same patient clustered together.

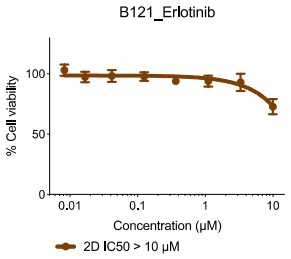


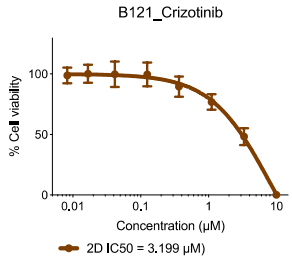
Appendix Figure 2 Drug response curves of the DIPG cultures derived from BIOMEDE-UK patients. Each model derived from one patient is represented in one page (n=11). Surviving fraction is shown on y-axes and every drug is represented in each curve (drug range 0-10 µM x-axes) across the models in 2D and/or 3D. The errors bars represent the standard deviation of the mean, IC50 values for each drug are plotted below the graph curve.

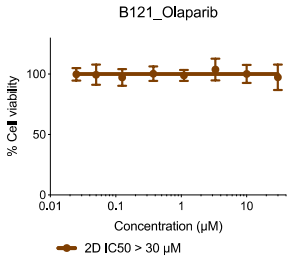


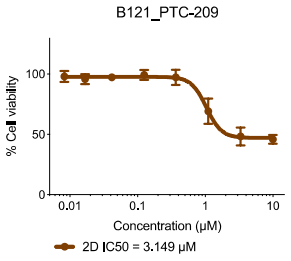

B128_Everolimus
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 > 10 μM

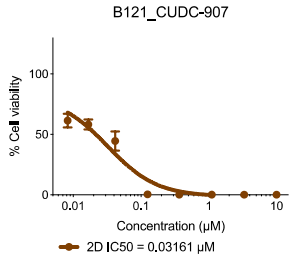

B121_Dasatinib
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 = 0.0255 μM

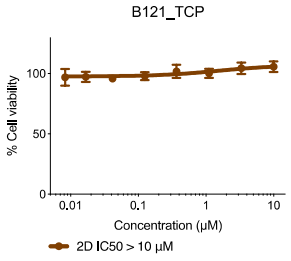

B121_Erlotinib
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 > 10 μM

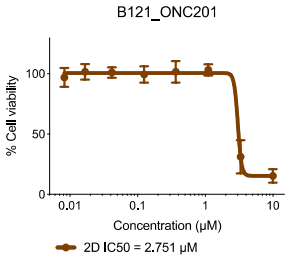

B121_Crizotinib
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 = 3.199 μM)

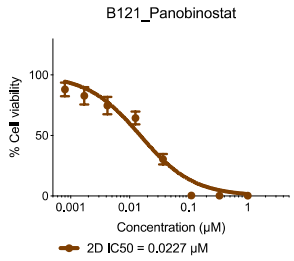

B121_Olaparib
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 > 30 μM

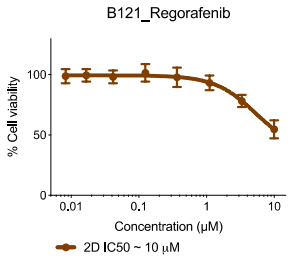

B121_PTC-209
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 = 3.149 μM

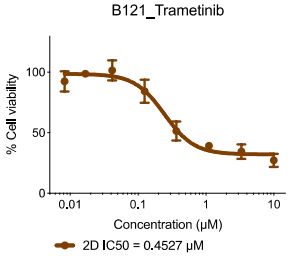

B121_CUDC-907
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 = 0.03161 μM

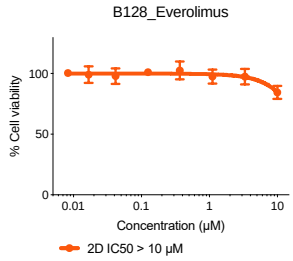

B121_TCP
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 > 10 μM

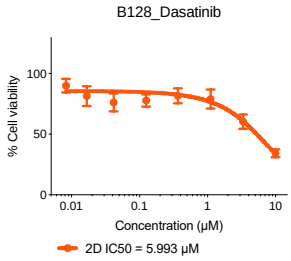

B121_ONC201
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 = 2.751 μM

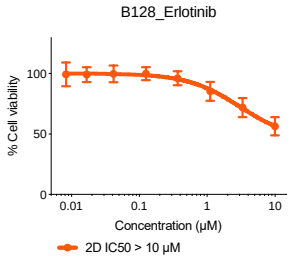

B121_Panobinostat
% Cell viability
100
50
0
0.001
0.01
0.1
1
Concentration (μM)
2D IC50 = 0.0227 μM

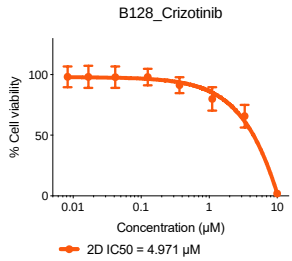

B121_Regorafenib
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 ~ 10 μM

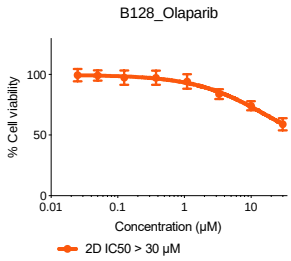

B121_Trametinib
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 = 0.4527 μM

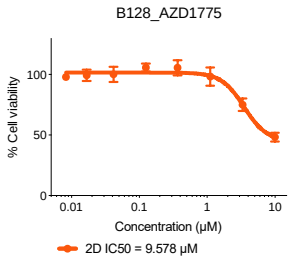

B128_Everolimus
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 > 10 μM

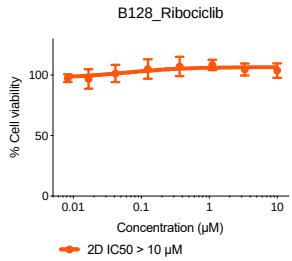

B128_Dasatinib
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 = 5.993 μM

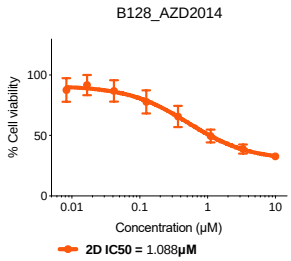

B128_Erlotinib
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 > 10 μM

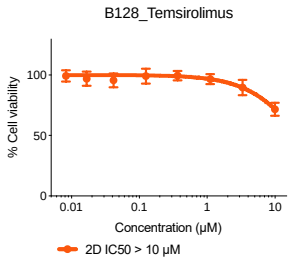

B128_Crizotinib
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 = 4.971 μM

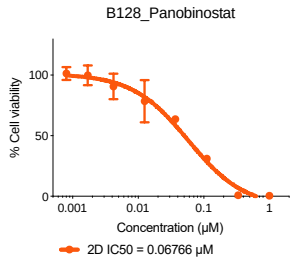

B128_Olaparib
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 > 30 μM

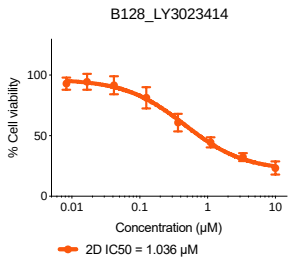

B128_AZD1775
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 = 9.578 μM

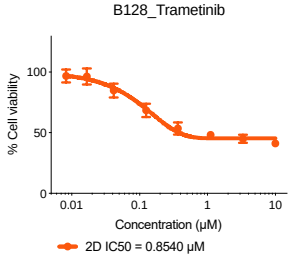

B128_Ribociclib
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 > 10 μM

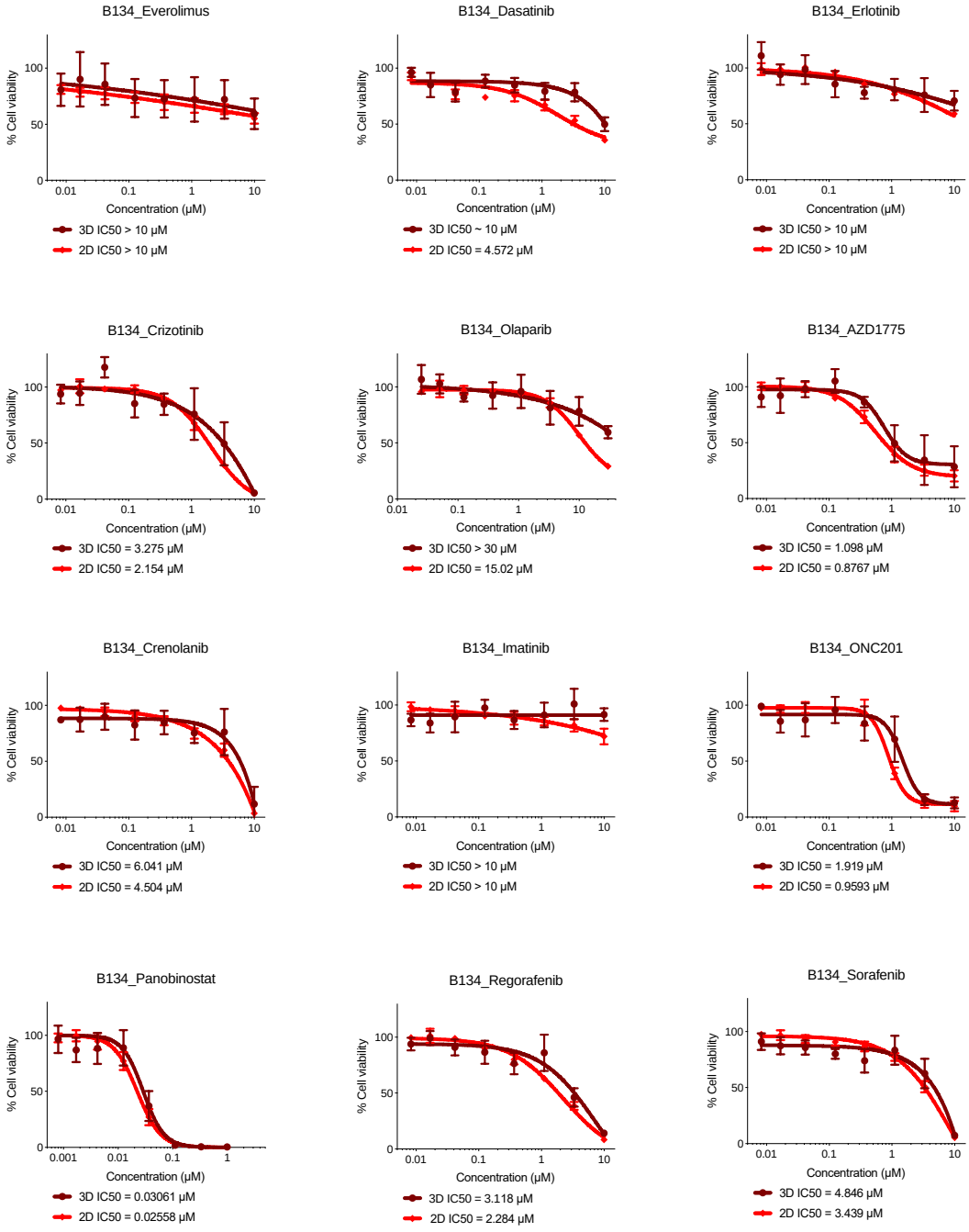

B128_AZD2014
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 = 1.088μM

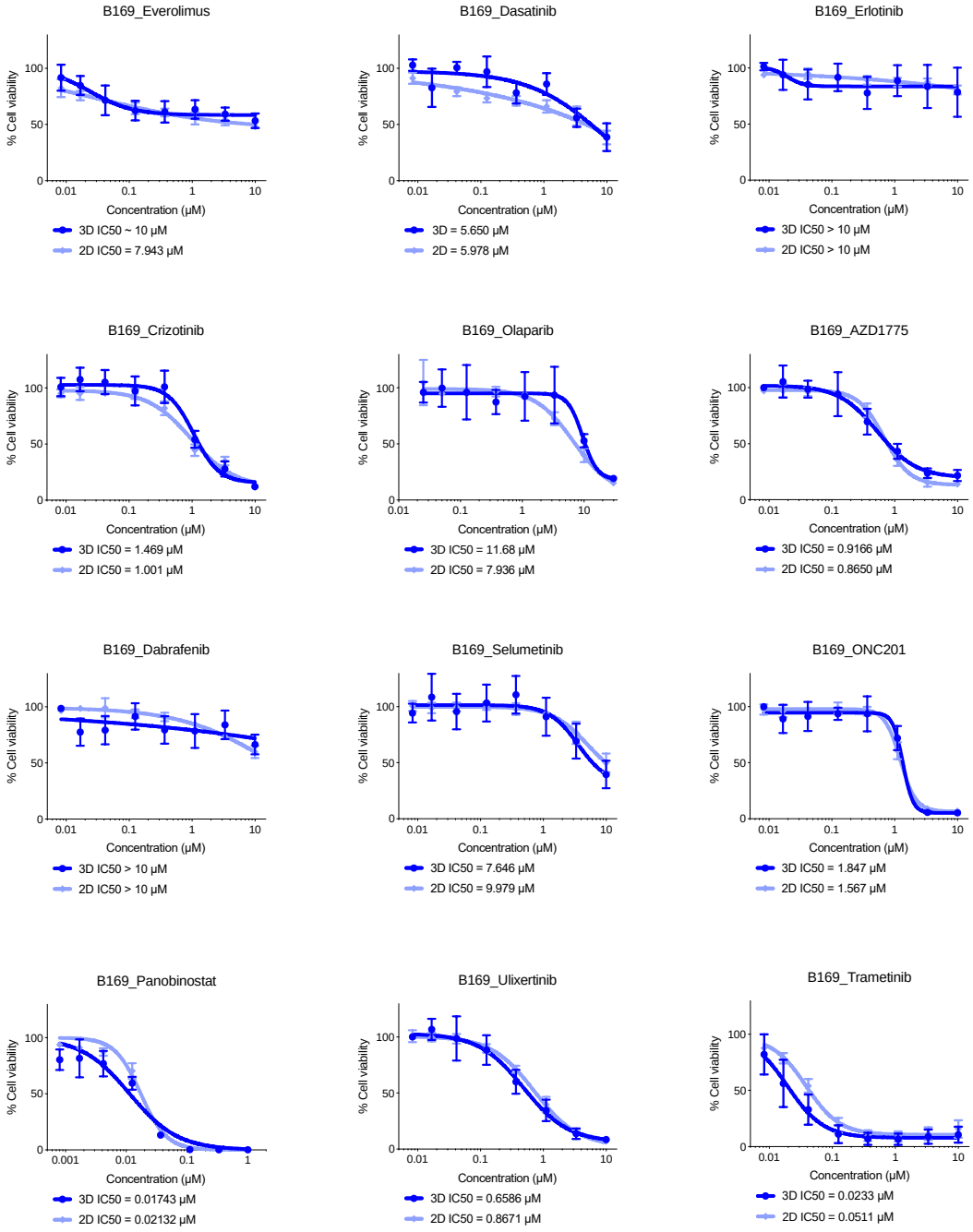

B128_Temsirolimus
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 > 10 μM

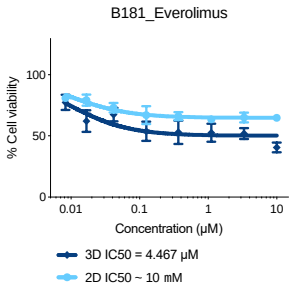

B128_Panobinostat
% Cell viability
100
50
0
0.001
0.01
0.1
1
Concentration (μM)
2D IC50 = 0.06766 μM

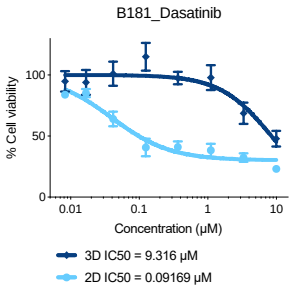

B128_LY3023414
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 = 1.036 μM

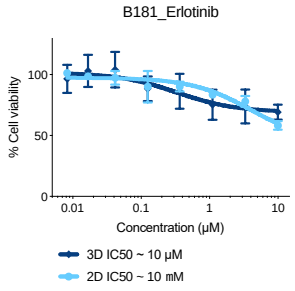

B128_Trametinib
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 = 0.8540 μM

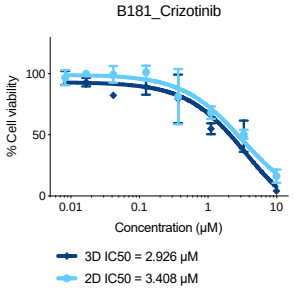

B134_Everolimus
% Cell viability
Concentration (μM)
3D IC50 > 10 μM
2D IC50 > 10 μM
B134_Dasatinib
% Cell viability
Concentration (μM)
3D IC50 ~ 10 μM
2D IC50 = 4.572 μM
B134_Erlotinib
% Cell viability
Concentration (μM)
3D IC50 > 10 μM
2D IC50 > 10 μM
B134_Crizotinib
% Cell viability
Concentration (μM)
3D IC50 = 3.275 μM
2D IC50 = 2.154 μM
B134_Olaparib
% Cell viability
Concentration (μM)
3D IC50 > 30 μM
2D IC50 = 15.02 μM
B134_AZD1775
% Cell viability
Concentration (μM)
3D IC50 = 1.098 μM
2D IC50 = 0.8767 μM
B134_Crenolanib
% Cell viability
Concentration (μM)
3D IC50 = 6.041 μM
2D IC50 = 4.504 μM
B134_Imatinib
% Cell viability
Concentration (μM)
3D IC50 > 10 μM
2D IC50 > 10 μM
B134_ONC201
% Cell viability
Concentration (μM)
3D IC50 = 1.919 μM
2D IC50 = 0.9593 μM
B134_Panobinostat
% Cell viability
Concentration (μM)
3D IC50 = 0.03061 μM
2D IC50 = 0.02558 μM
B134_Regorafenib
% Cell viability
Concentration (μM)
3D IC50 = 3.118 μM
2D IC50 = 2.284 μM
B134_Sorafenib
% Cell viability
Concentration (μM)
3D IC50 = 4.846 μM
2D IC50 = 3.439 μM

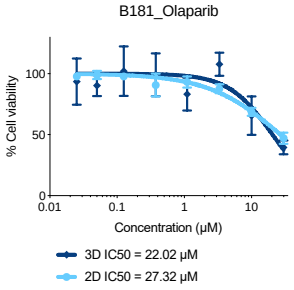

B169_Everolimus
% Cell viability
Concentration (μM)
3D IC50 ~ 10 μM
2D IC50 = 7.943 μM
B169_Dasatinib
% Cell viability
Concentration (μM)
3D = 5.650 μM
2D = 5.978 μM
B169_Erlotinib
% Cell viability
Concentration (μM)
3D IC50 > 10 μM
2D IC50 > 10 μM
B169_Crizotinib
% Cell viability
Concentration (μM)
3D IC50 = 1.469 μM
2D IC50 = 1.001 μM
B169_Olaparib
% Cell viability
Concentration (μM)
3D IC50 = 11.68 μM
2D IC50 = 7.936 μM
B169_AZD1775
% Cell viability
Concentration (μM)
3D IC50 = 0.9166 μM
2D IC50 = 0.8650 μM
B169_Dabrafenib
% Cell viability
Concentration (μM)
3D IC50 > 10 μM
2D IC50 > 10 μM
B169_Selumetinib
% Cell viability
Concentration (μM)
3D IC50 = 7.646 μM
2D IC50 = 9.979 μM
B169_ONC201
% Cell viability
Concentration (μM)
3D IC50 = 1.847 μM
2D IC50 = 1.567 μM
B169_Panobinostat
% Cell viability
Concentration (μM)
3D IC50 = 0.01743 μM
2D IC50 = 0.02132 μM
B169_Ulixertinib
% Cell viability
Concentration (μM)
3D IC50 = 0.6586 μM
2D IC50 = 0.8671 μM
B169_Trametinib
% Cell viability
Concentration (μM)
3D IC50 = 0.0233 μM
2D IC50 = 0.0511 μM

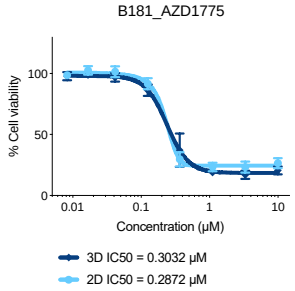

B181_Everolimus
% Cell viability
Concentration (μM)
3D IC50 = 4.467 μM
2D IC50 ~ 10 mM

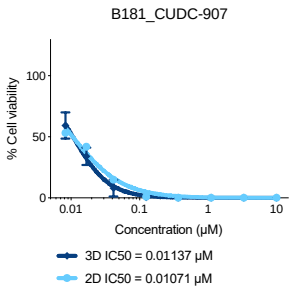

B181_Dasatinib
% Cell viability
Concentration (μM)
3D IC50 = 9.316 μM
2D IC50 = 0.09169 μM

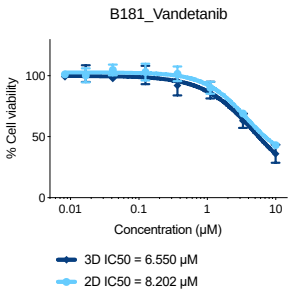

B181_Erlotinib
% Cell viability
Concentration (μM)
3D IC50 ~ 10 μM
2D IC50 ~ 10 mM

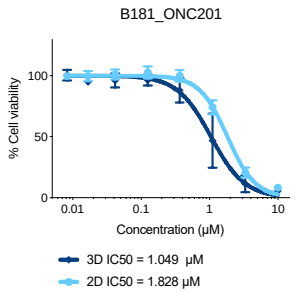

B181_Crizotinib
% Cell viability
Concentration (μM)
3D IC50 = 2.926 μM
2D IC50 = 3.408 μM

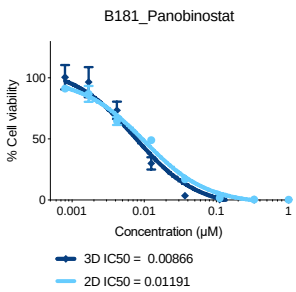

B181_Olaparib
% Cell viability
Concentration (μM)
3D IC50 = 22.02 μM
2D IC50 = 27.32 μM

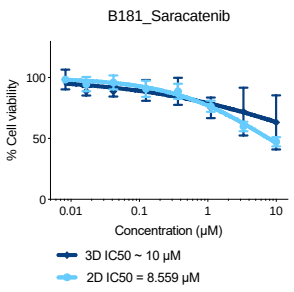

B181_AZD1775
% Cell viability
Concentration (μM)
3D IC50 = 0.3032 μM
2D IC50 = 0.2872 μM

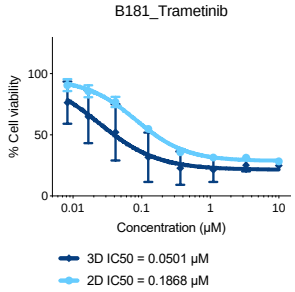

B181_CUDC-907
% Cell viability
Concentration (μM)
3D IC50 = 0.01137 μM
2D IC50 = 0.01071 μM

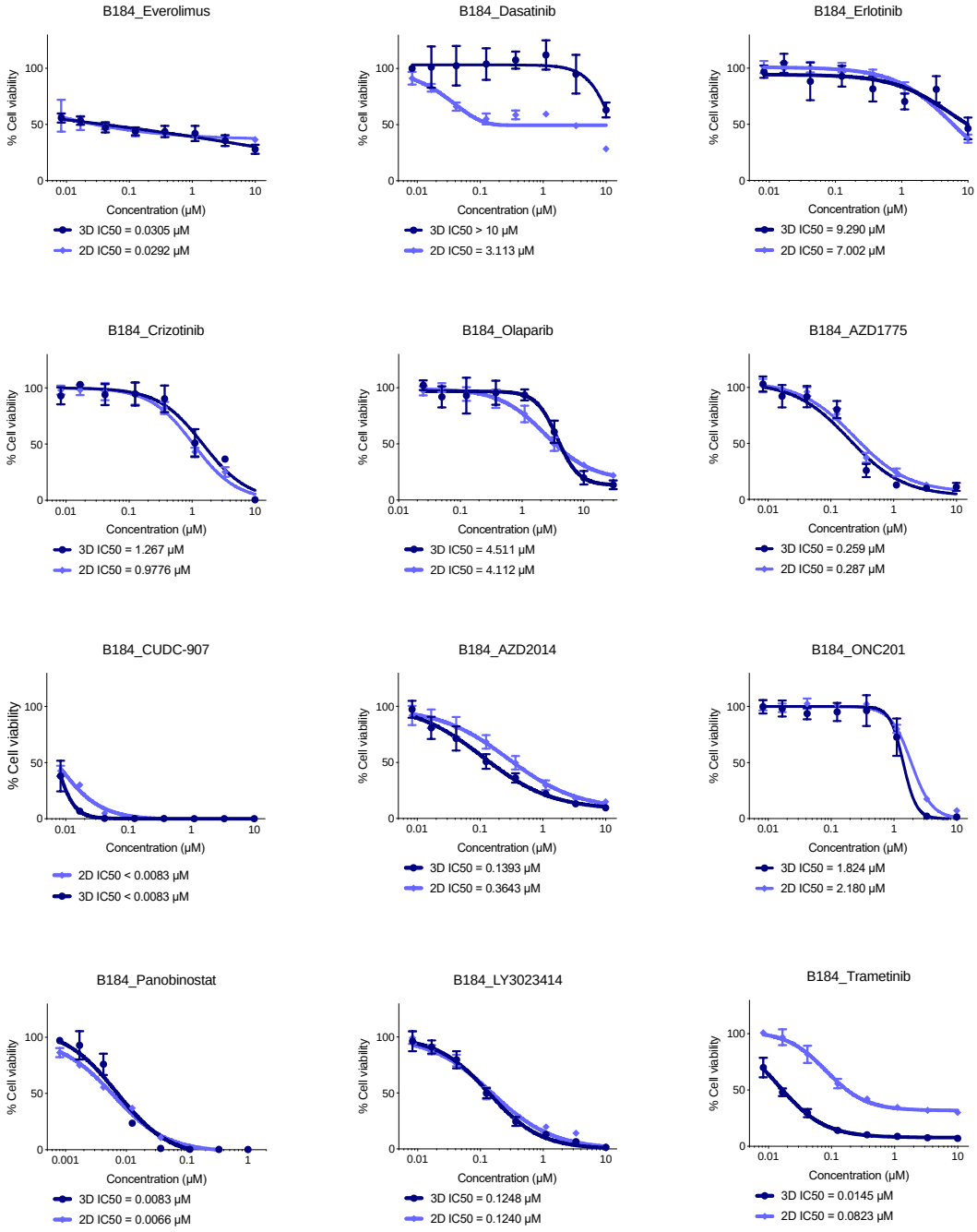

B181_Vandetanib
% Cell viability
Concentration (μM)
3D IC50 = 6.550 μM
2D IC50 = 8.202 μM

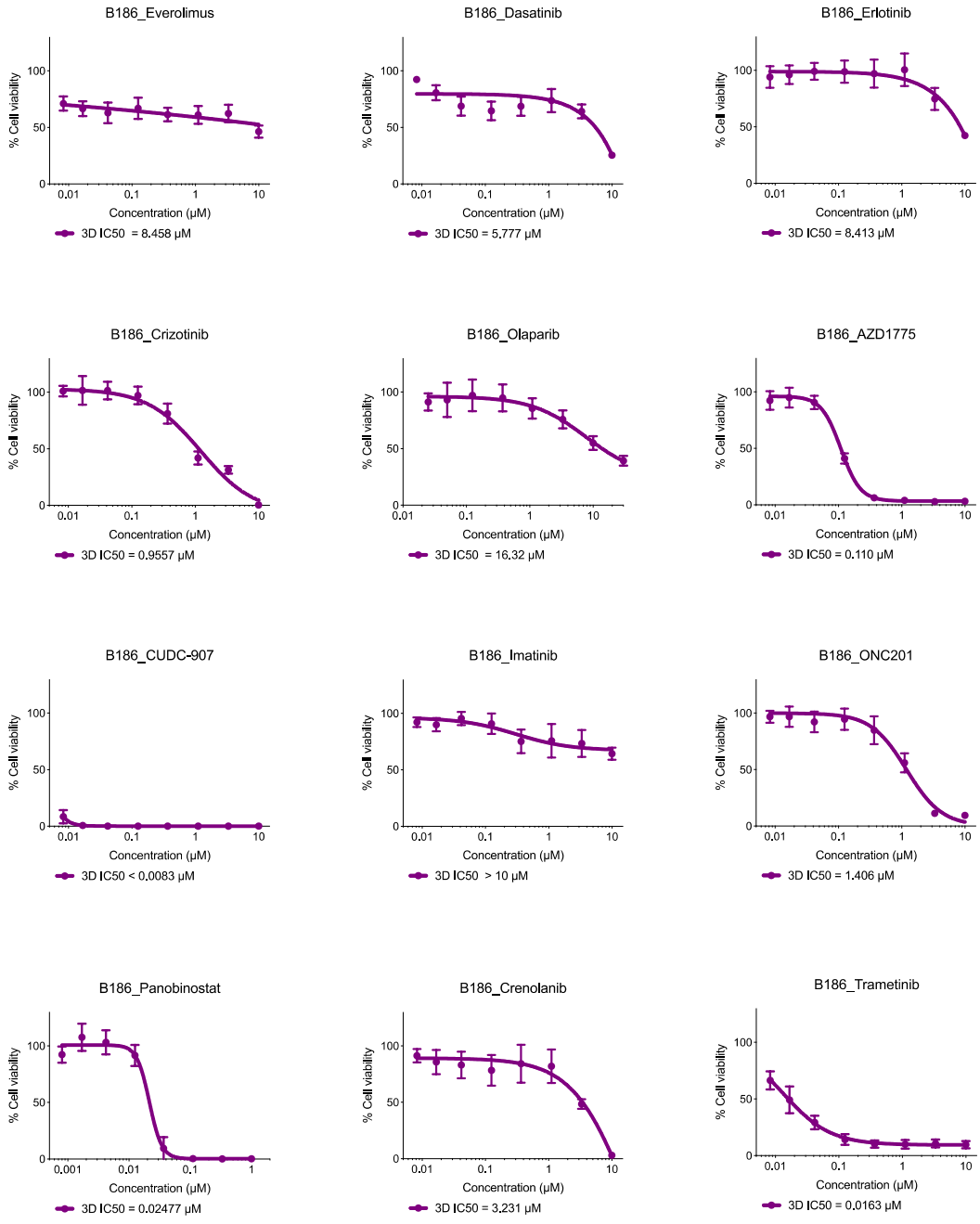

B181_ONC201
% Cell viability
Concentration (μM)
3D IC50 = 1.049 μM
2D IC50 = 1.828 μM

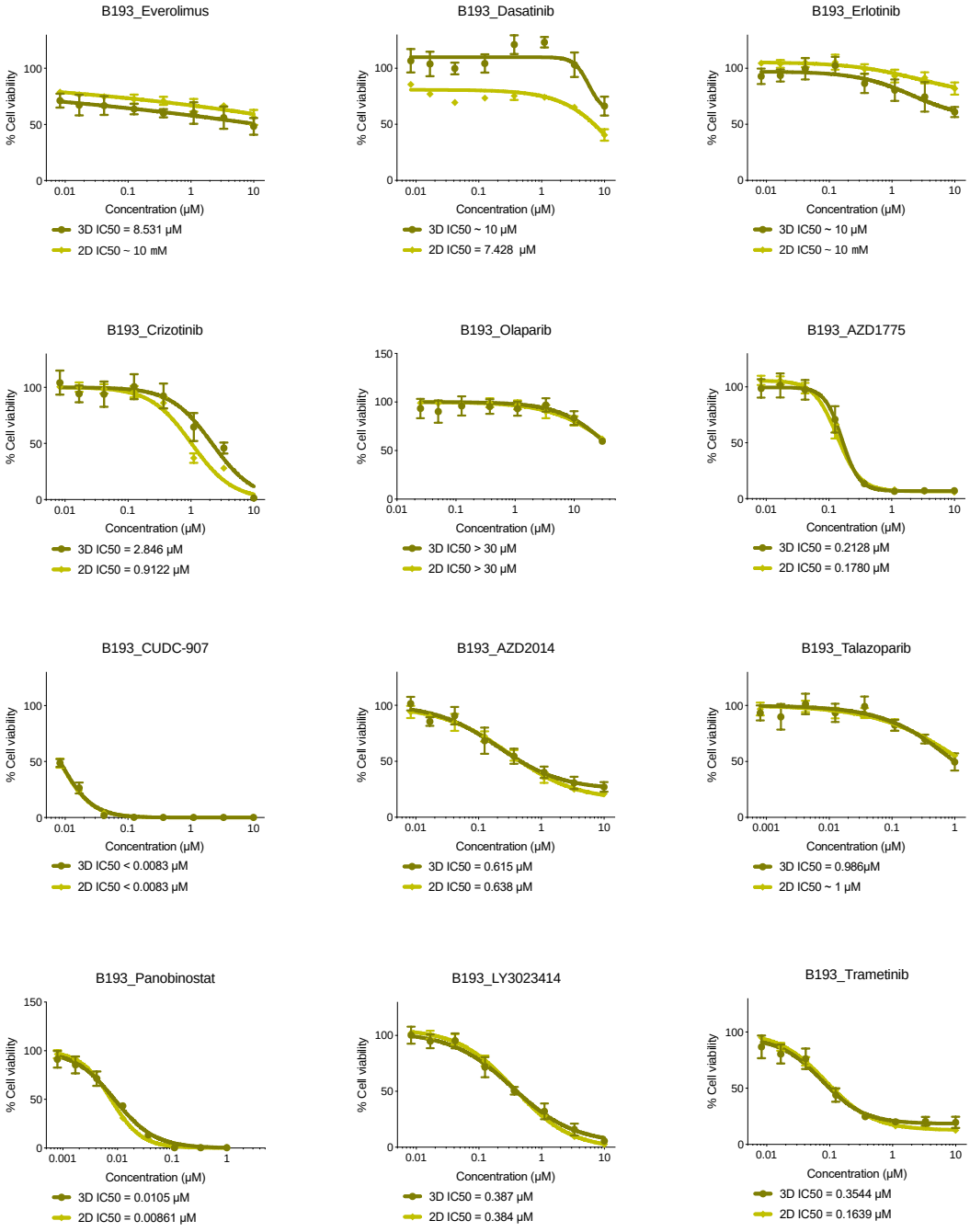

B181_Panobinostat
% Cell viability
Concentration (μM)
3D IC50 = 0.00866
2D IC50 = 0.01191

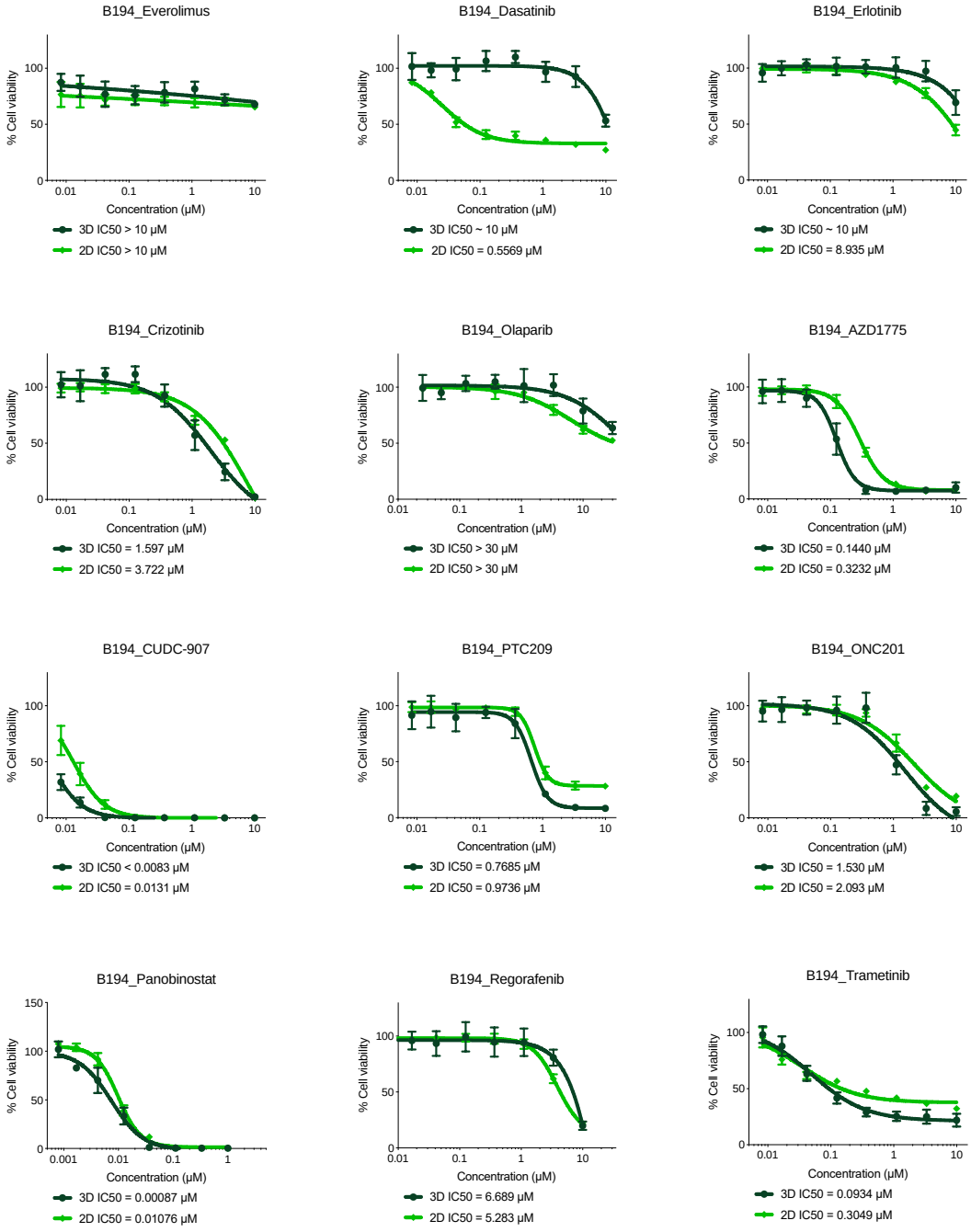

B181_Saracatenib
% Cell viability
Concentration (μM)
3D IC50 ~ 10 μM
2D IC50 = 8.559 μM

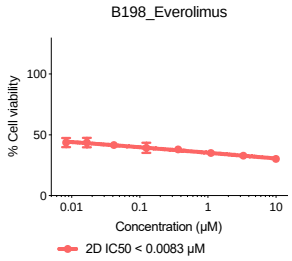

B181_Trametinib
% Cell viability
Concentration (μM)
3D IC50 = 0.0501 μM
2D IC50 = 0.1868 μM

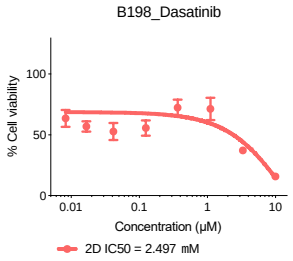

B184_Everolimus
% Cell viability
Concentration (μM)
3D IC50 = 0.0305 μM
2D IC50 = 0.0292 μM
B184_Dasatinib
% Cell viability
Concentration (μM)
3D IC50 > 10 μM
2D IC50 = 3.113 μM
B184_Erlotinib
% Cell viability
Concentration (μM)
3D IC50 = 9.290 μM
2D IC50 = 7.002 μM
B184_Crizotinib
% Cell viability
Concentration (μM)
3D IC50 = 1.267 μM
2D IC50 = 0.9776 μM
B184_Olaparib
% Cell viability
Concentration (μM)
3D IC50 = 4.511 μM
2D IC50 = 4.112 μM
B184_AZD1775
% Cell viability
Concentration (μM)
3D IC50 = 0.259 μM
2D IC50 = 0.287 μM
B184_CUDC-907
% Cell viability
Concentration (μM)
2D IC50 < 0.0083 μM
3D IC50 < 0.0083 μM
B184_AZD2014
% Cell viability
Concentration (μM)
3D IC50 = 0.1393 μM
2D IC50 = 0.3643 μM
B184_ONC201
% Cell viability
Concentration (μM)
3D IC50 = 1.824 μM
2D IC50 = 2.180 μM
B184_Panobinostat
% Cell viability
Concentration (μM)
3D IC50 = 0.0083 μM
2D IC50 = 0.0066 μM
B184_LY3023414
% Cell viability
Concentration (μM)
3D IC50 = 0.1248 μM
2D IC50 = 0.1240 μM
B184_Trametinib
% Cell viability
Concentration (μM)
3D IC50 = 0.0145 μM
2D IC50 = 0.0823 μM

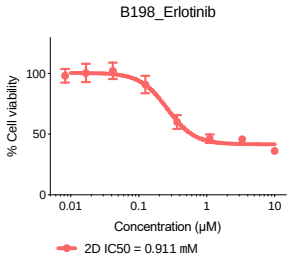

B186_Everolimus
% Cell viability
Concentration (μM)
3D IC50 = 8.458 μM
B186_Dasatinib
3D IC50 = 5.777 μM
B186_Erlotinib
3D IC50 = 8.413 μM
B186_Crizotinib
3D IC50 = 0.9557 μM
B186_Olaparib
3D IC50 = 16.32 μM
B186_AZD1775
3D IC50 = 0.110 μM
B186_CUDC-907
3D IC50 < 0.0083 μM
B186_Imatinib
3D IC50 > 10 μM
B186_ONC201
3D IC50 = 1.406 μM
B186_Panobinostat
3D IC50 = 0.02477 μM
B186_Crenolanib
3D IC50 = 3.231 μM
B186_Trametinib
3D IC50 = 0.0163 μM

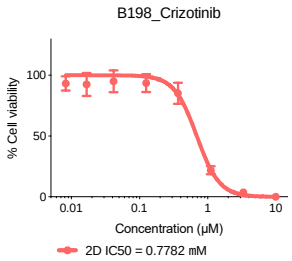

B193_Everolimus
% Cell viability
Concentration (μM)
3D IC50 = 8.531 μM
2D IC50 ~ 10 mM
B193_Dasatinib
% Cell viability
Concentration (μM)
3D IC50 ~ 10 μM
2D IC50 = 7.428 μM
B193_Erlotinib
% Cell viability
Concentration (μM)
3D IC50 ~ 10 μM
2D IC50 ~ 10 mM
B193_Crizotinib
% Cell viability
Concentration (μM)
3D IC50 = 2.846 μM
2D IC50 = 0.9122 μM
B193_Olaparib
% Cell viability
Concentration (μM)
3D IC50 > 30 μM
2D IC50 > 30 μM
B193_AZD1775
% Cell viability
Concentration (μM)
3D IC50 = 0.2128 μM
2D IC50 = 0.1780 μM
B193_CUDC-907
% Cell viability
Concentration (μM)
3D IC50 < 0.0083 μM
2D IC50 < 0.0083 μM
B193_AZD2014
% Cell viability
Concentration (μM)
3D IC50 = 0.615 μM
2D IC50 = 0.638 μM
B193_Talazoparib
% Cell viability
Concentration (μM)
3D IC50 = 0.986μM
2D IC50 ~ 1 μM
B193_Panobinostat
% Cell viability
Concentration (μM)
3D IC50 = 0.0105 μM
2D IC50 = 0.00861 μM
B193_LY3023414
% Cell viability
Concentration (μM)
3D IC50 = 0.387 μM
2D IC50 = 0.384 μM
B193_Trametinib
% Cell viability
Concentration (μM)
3D IC50 = 0.3544 μM
2D IC50 = 0.1639 μM

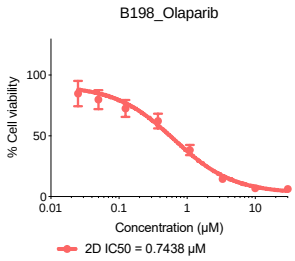

B194_Everolimus
% Cell viability
Concentration (μM)
3D IC50 > 10 μM
2D IC50 > 10 μM
B194_Dasatinib
% Cell viability
Concentration (μM)
3D IC50 ~ 10 μM
2D IC50 = 0.5569 μM
B194_Erlotinib
% Cell viability
Concentration (μM)
3D IC50 ~ 10 μM
2D IC50 = 8.935 μM
B194_Crizotinib
% Cell viability
Concentration (μM)
3D IC50 = 1.597 μM
2D IC50 = 3.722 μM
B194_Olaparib
% Cell viability
Concentration (μM)
3D IC50 > 30 μM
2D IC50 > 30 μM
B194_AZD1775
% Cell viability
Concentration (μM)
3D IC50 = 0.1440 μM
2D IC50 = 0.3232 μM
B194_CUDC-907
% Cell viability
Concentration (μM)
3D IC50 < 0.0083 μM
2D IC50 = 0.0131 μM
B194_PTC209
% Cell viability
Concentration (μM)
3D IC50 = 0.7685 μM
2D IC50 = 0.9736 μM
B194_ONC201
% Cell viability
Concentration (μM)
3D IC50 = 1.530 μM
2D IC50 = 2.093 μM
B194_Panobinostat
% Cell viability
Concentration (μM)
3D IC50 = 0.00087 μM
2D IC50 = 0.01076 μM
B194_Regorafenib
% Cell viability
Concentration (μM)
3D IC50 = 6.689 μM
2D IC50 = 5.283 μM
B194_Trametinib
% Cell viability
Concentration (μM)
3D IC50 = 0.0934 μM
2D IC50 = 0.3049 μM

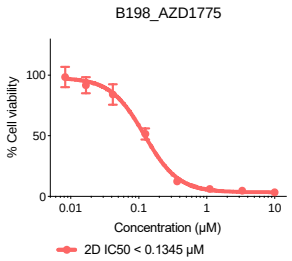

B198_Everolimus
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 < 0.0083 μM

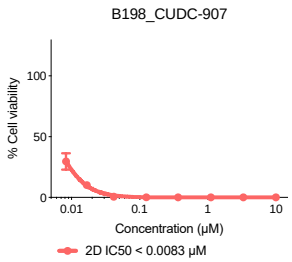

B198_Dasatinib
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 = 2.497 mM

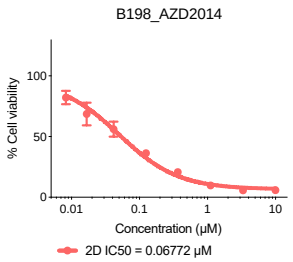

B198_Erlotinib
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 = 0.911 mM

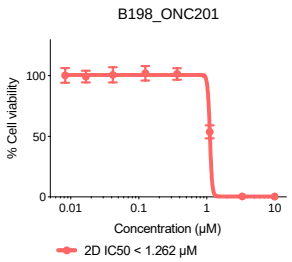

B198_Crizotinib
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 = 0.7782 mM

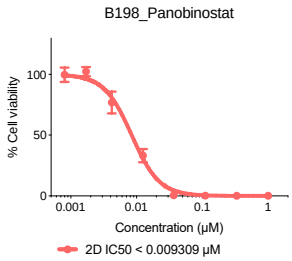

B198_Olaparib
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 = 0.7438 μM

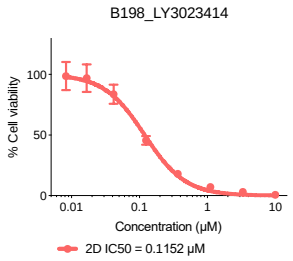

B198_AZD1775
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 < 0.1345 μM

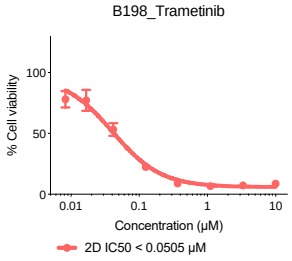

B198_CUDC-907
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 < 0.0083 μM


B198_AZD2014
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 = 0.06772 μM


B198_ONC201
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 < 1.262 μM

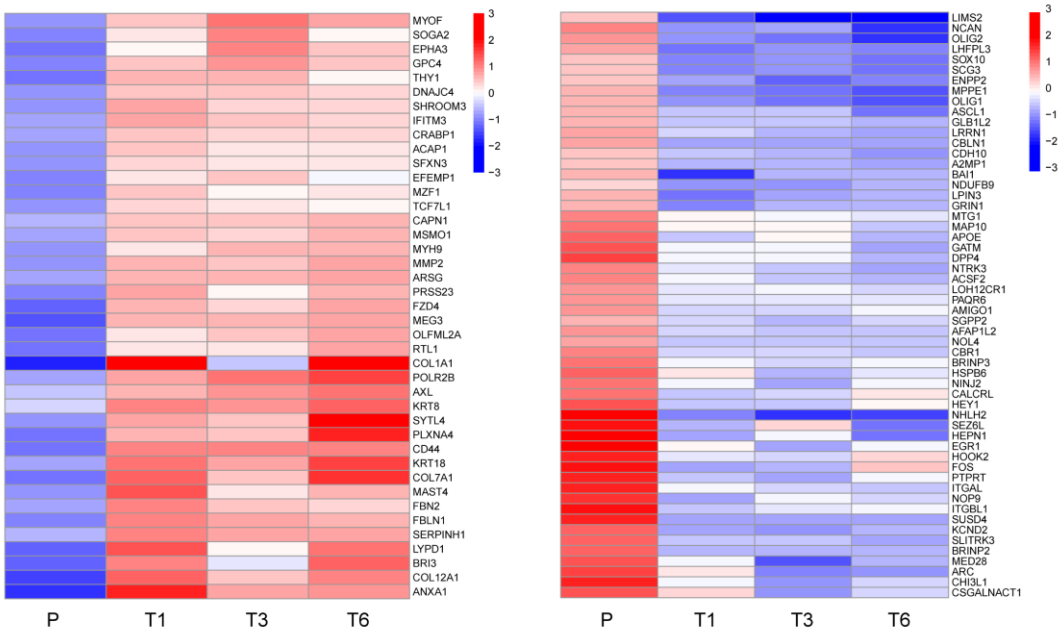

B198_Panobinostat
% Cell viability
100
50
0
0.001
0.01
0.1
1
Concentration (μM)
2D IC50 < 0.009309 μM


B198_LY3023414
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 = 0.1152 μM

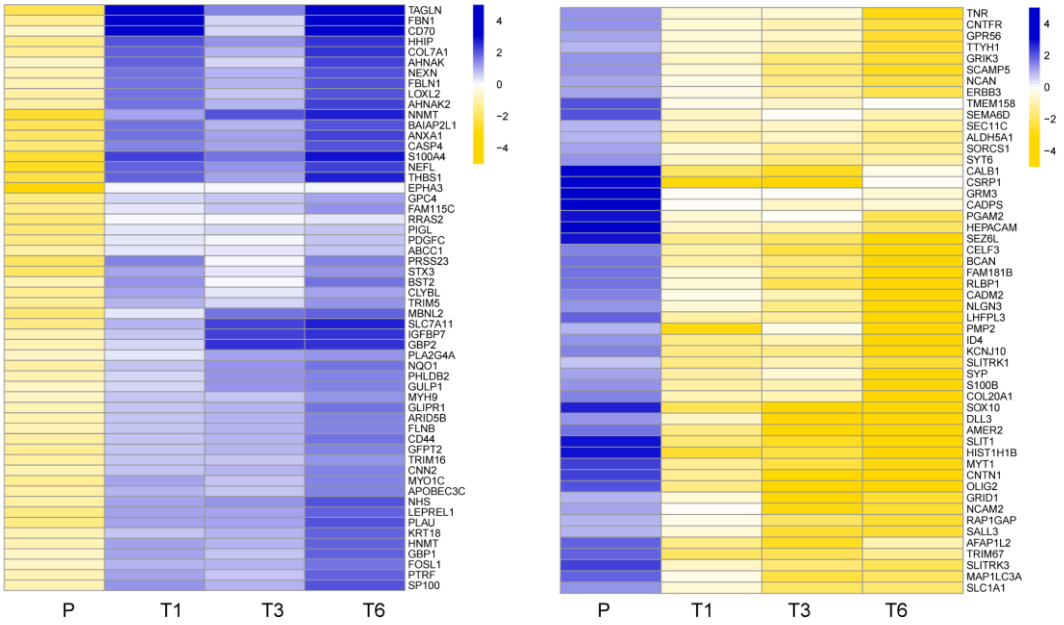

B198_Trametinib
% Cell viability
100
50
0
0.01
0.1
1
10
Concentration (μM)
2D IC50 < 0.0505 μM

Appendix Figure 3 Intersection of shared up-regulated and down-regulated targets by RNA-seq, Full proteome and phosphoproteomics. (A) Shared differentially expressed genes across the *MEK1*/2 clones compared to the BIOMEDE-169 are shown (left increased expression, righ decreased expression). (B) Shared differentially expressed proteins across the *MEK1*/2 clones compared to the BIOMEDE-169 are shown (left increased expression, righ decreased expression). (C) Shared differentially phosphorylated proteinsacrossed the *MEK1*/2 clones compared to the BIOMEDE-169 are shown (left increased phosphorylation, decreased phosphorylation). Intersection was done by selecting up-regulated and down-regulated genes / proteins and phosprorylation by the log-fold change being greated than 1 and less than -1 in any of the three clones for RNA, proteome and phosphoproteome and these three were then intersected and shown in the representative heatmps.

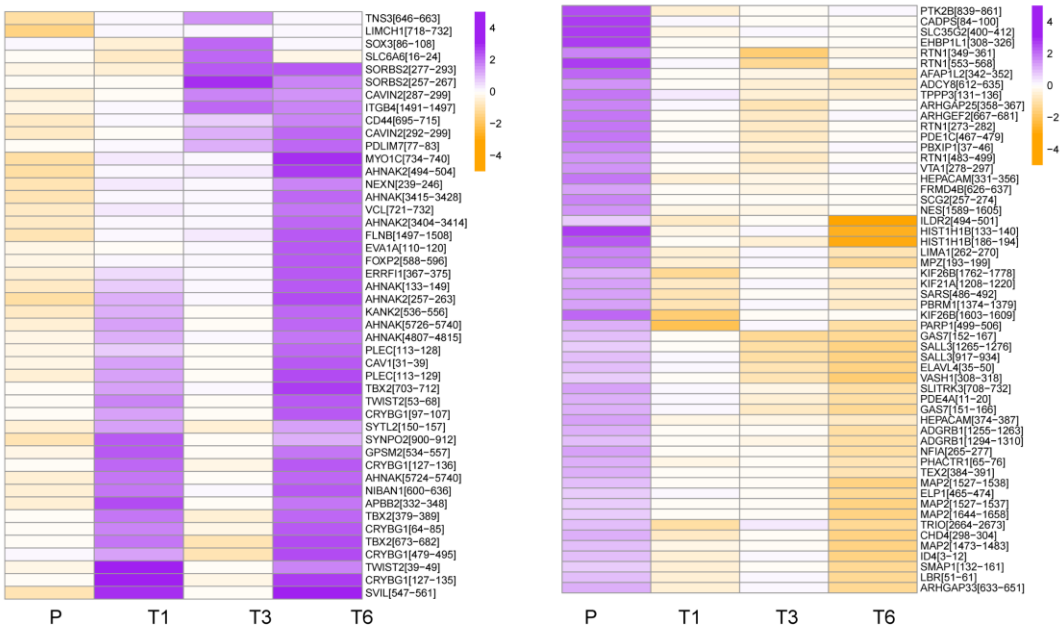
A



B



C


TNS3[646−663]
LIMCH1[718−732]
SOX3[86−108]
SLC6A6[16−24]
SORBS2[277−293]
SORBS2[257−267]
CAVIN2[287−299]
ITGB4[1491−1497]
CD44[695−715]
CAVIN2[292−299]
PDLIM7[77−83]
MYO1C[734−740]
AHNAK2[494−504]
NEXN[239−246]
AHNAK[3415−3428]
VCL[721−732]
AHNAK2[3404−3414]
FLNB[1497−1508]
EVA1A[110−120]
FOXP2[588−596]
ERRFI1[367−375]
AHNAK[133−149]
AHNAK2[257−263]
KANK2[536−556]
AHNAK[5726−5740]
AHNAK[4807−4815]
PLEC[113−128]
CAV1[31−39]
PLEC[113−129]
TBX2[703−712]
TWIST2[53−68]
CRYBG1[97−107]
SYTL2[150−157]
SYNPO2[900−912]
GPSM2[534−557]
CRYBG1[127−136]
AHNAK[5724−5740]
NIBAN1[600−636]
APBB2[332−348]
TBX2[379−389]
CRYBG1[64−85]
TBX2[673−682]
CRYBG1[479−495]
TWIST2[39−49]
CRYBG1[127−135]
SVIL[547−561]
4
2
0
−2
−4
P
T1
T3
T6
PTK2B[839−861]
CADPS[84−100]
SLC35G2[400−412]
EHBP1L1[308−326]
RTN1[349−361]
RTN1[553−568]
AFAP1L2[342−352]
ADCY8[612−635]
TPPP3[131−136]
ARHGAP25[358−367]
ARHGEF2[667−681]
RTN1[273−282]
PDE1C[467−479]
PBXIP1[37−46]
RTN1[483−499]
VTA1[278−297]
HEPACAM[331−356]
FRMD4B[626−637]
SCG2[257−274]
NES[1589−1605]
ILDR2[494−501]
HIST1H1B[133−140]
HIST1H1B[186−194]
LIMA1[262−270]
MPZ[193−199]
KIF26B[1762−1778]
KIF21A[1208−1220]
SARS[486−492]
PBRM1[1374−1379]
KIF26B[1603−1609]
PARP1[499−506]
GAS7[152−167]
SALL3[1265−1276]
SALL3[917−934]
ELAVL4[35−50]
VASH1[308−318]
SLITRK3[708−732]
PDE4A[11−20]
GAS7[151−166]
HEPACAM[374−387]
ADGRB1[1255−1263]
ADGRB1[1294−1310]
NFIA[265−277]
PHACTR1[65−76]
TEX2[384−391]
MAP2[1527−1538]
ELP1[465−474]
MAP2[1527−1537]
MAP2[1644−1658]
TRIO[2664−2673]
CHD4[298−304]
MAP2[1473−1483]
ID4[3−12]
SMAP1[132−161]
LBR[61−81]
ARHGAP33[633−651]
4
2
0
−2
−4
P
T1
T3
T6

Appendix Table 1 List of genes included in the Paeds-v1 and Paeds-v2 panels.

Basis criteria for gene selections are defined according to: Level 1 predictive biomarker, Level 2: prognostic biomarker, Level 3: diagnostic biomarker, Level 4: potentially targetable with inhibitors available or under development, Level 5: known germline or high-risk SNP, Level 6: research evidence only.

Actionability criteria is defined by: Tier 1: Recognised (FDA/EMA approved) predictive biomarker for response to drug in that indication (OncoKB Level 1).Tier 2B recognised standard of care predictive biomarker for response in another indication, Tier 3: open clinical trial for predictive biomarker for paediatric solid tumours and Tier 4: compelling biological evidence supports biomarker as being predictive of response to drug.

For more detailed information refer to [381]

Gene	v1	v2	Alteration	Basis criteria	Actionability criteria
ACVR1	YES	YES	Mutation	Level 2,3,4	4
AKT1	YES	YES	Mutation (EK17), amplification	Level 4	3
ALK	YES	YES	Mutation, amplification, translocation	Level 1	mutation 3, translocation 2A
AMER1	YES	YES	Mutation, deletion	Level 3	
APC	YES	YES	Mutation	Level 6	
ARID1A	YES	YES	Mutation	Level 6	4
ARID1B	YES	YES	Mutation	Level 6	
ASXL1	YES	YES	Mutation	Level 6	
ATM	YES	YES	Mutation/Loss	Level 4	3
ATRX	YES	YES	Mutation, deletion	Level 3,4	3
BARD1	YES	NO	SNP	Level 5	
BCOR	YES	YES	Mutation	Level 3,6	
BRAF	YES	YES	Mutation (V600E/K)	Level 1,3	2B (V600E, V600K)
CASC15	YES	NO	SNP	Level 5	
C19MC	NO	YES	Amplification	Level 6	
CCND1	NO	YES	Amplification	Level 4	4
CCND2	NO	YES	Amplification	Level 4	4
CCNE1	NO	YES	Amplification	Level 4	4
CDK12	NO	YES	Mutation	Level 6	
CDK4	YES	YES	Amplification	Level 1	3 (amplification)
CDK6	YES	YES	Amplification	Level 1	3
CDKN2A	YES	YES	Deletion	Level 4	3
CFL1	YES	NO	SNP	Level 5	
CDKN2B	YES	YES	Deletion	Level 4	3
CHEK2	YES	YES	Mutation	Level 4	3
CIC	NO	YES	Mutation	Level 6	

CREBBP	NO	YES	Mutation	Level 6	
CTNNB1	YES	YES	Mutation	Level 2,4	4
DAXX	NO	YES	Mutation	Level 6	
DDX1	NO	YES	Amplification	Level 6	
DDR2	YES	NO	Mutation	Level 1	3
DDX3X	YES	YES	Mutation	Level 6	
DICER1	YES	YES	Mutation	Level 6	
DROSHA	NO	YES	Mutation	Level 6	
EGFR	YES	YES	Mutation, ampflication	Level 1	2B (mutation), 4 (amplification)
ERBB2	YES	YES	Mutation, ampflication	Level 1	2B (amplification)
ERG	YES	NO	Translocation	Level 3	
ETV6	YES	NO	Translocation	Level 3	
EWSR1	YES	NO	Translocation	Level 3	3
EZH2	NO	YES	Mutation	Level 4	3
FBXW7	YES	YES	Mutation	Level 4	4
FGFR1	YES	YES	Mutation, amplification, translocation	Level 1	3
FGFR2	YES	YES_	Mutation, amplification, translocation	Level 1	3
FGFR3	YES	YES_	Mutation, amplification, translocation	Level 1	3
FGFR4	YES	YES_	Mutation, amplification, translocation	Level 4	3
FRS2	NO	YES	Amplification	Level 6	
FUBP1	NO	YES	Mutation	Level 6	
FUS	YES	NO	Translocation	Level 3	
GLI2	NO	YES	Amplification	Level 3	
H3F3A	YES	YES	Mutation (K27M, G34R/V)	Level 2,3,4	3
HIST1H3B	YES	YES	Mutation	Level 2,3,4	3
HIST1H3C	YES	YES	Mutation	Level 2,3,4	3
HIST2H3A/C	NO	YES	Mutation	Level 2,3,4	3
HRAS	YES	YES	Mutation	Level 4	3
IDH1	YES	YES	Mutation (R132X)	Level 2,3,4	2B (oncogenic)
IDH2	NO	YES	Mutation (R172X)	Level 4	2B (oncogenic)
IGF1R	NO	YES	Amplification	Level 4	4
IGF2	NO	YES	Amplification	Level 6	
IL3	YES	NO	SNP	Level 5	
IL6	YES	NO	SNP	Level 5	
KIT	YES	YES	Mutation, ampflication	Level 1	2B
KMT2D	YES	NO	Mutation	Level 6	
KRAS	YES	YES	Mutation	Level 4	3
LMO1	YES	NO	SNP	Level 5	

MAP2K1	YES	YES	Mutation	Level 4	4
MAP2K2	YES	YES	Mutation	Level 4	4
MDM2	YES	YES	Amplification	Level 4	4
MDM4	NO	YES	Amplification, mutation	Level 6	
MET	NO	YES	Mutation, amplification, translocation	Level 1	2B
MLH1	NO	YES	Mutation	Level 5	
MSH2	NO	YES	Mutation	Level 5	
MSH6	NO	YES	Mutation	Level 5	
MYC	NO	YES	Amplification	Level 2,4	4
MYCL	NO	YES	Amplification	Level 6	
MYCN	YES	YES	Amplification, Mutation (P44L)	Level 2,4	3
MYOD1	YES	YES	Mutation	Level 2	
NCOA2	YES	NO	Translocation	Level 3	
NF1	YES	YES	Mutation, deletion	Level 4,5	3
NF2	NO	YES	Mutation, deletion	Level 4,5	4
NRAS	YES	YES	Mutation	Level 4	3
OTX2	NO	YES	Mutation	Level 6	
PAX3	YES	NO	Translocation	Level 3	
PAX7	YES	NO	Translocation	Level 3	
PDGFRA	YES	YES	Mutation, ampflication	Level 1,2	2B
PHOX2B	YES	YES	Mutation, SNP	Level 5	
PIK3CA	YES	YES	Mutation	Level 1	3
PIK3R1	YES	YES	Mutation	Level 4	3
PMS2	NO	YES	Mutation	Level 5	
PPM1D	YES	YES	Mutation	Level 6	
PTCH1	YES	YES	Mutation	Level 4	4
PTEN	YES	YES	Mutation	Level 4	3
PTPN11	YES	YES	Mutation	Level 4	
PTPRD	YES	NO	Mutation	Level 6	
RB1	YES	YES	Mutation, deletion	Level 3,4	3
RELA	NO	YES	Translocation	Level 2,3,4	
RET	YES	NO	Translocation	Level 1	2B (fusions and oncogenic mutations)
ROS1	YES	NO	Translocation	Level 1	3 (fusions)
SETD2	YES	YES	Mutation	Level 6	
SMARCA2	NO	YES	Mutation, deletion	Level 6	
SMARCA4	YES	YES	Mutation, deletion	Level 4,5	3
SMARCB1	YES	YES	Mutation, deletion	Level 4,5	3
SMO	NO	YES	Mutation	Level 1	4

SS18	YES	NO	Translocation	Level 3	
SUFU	YES	YES	Translocation	Level 3	
TENM3	YES	NO	Mutation	Level 6	
TERT	NO	YES	Mutation, translocation, amplification	Level 2	
TP53	YES	YES	Mutation, deletion	Level 2,5	3
TSC1	YES	YES	Mutation	Level 1	2B
TSC2	NO	YES	Mutation	Level 1	2B
WT1	YES	YES	Mutation	Level 3	
YAP1	NO	YES	Amplification	Level 6	
ZHX2	YES	NO	Mutation	Level 6	

Appendix Table 2 List of the 330 genes included in the pHGG-panel.

ACVR1	*CCNE1*	*FANCA*	*HIST1H2BD*	*ID1*	*MDM2*	*PPM1D*	*SMARCB1*
ADCY1	*CCNJ*	*FANCG*	*HIST1H2BE*	*ID2*	*MDM4*	*PRKCZ*	*SMARCE1*
AHDC1	*CCNK*	*FANCM*	*HIST1H2BF*	*ID3*	*MET*	*PRKRA*	*SMO*
AKT1	*CDC42BPB*	*FBXO10*	*HIST1H2BG*	*IDH1*	*MGMT*	*PTCH1*	*SORBS2*
AKT2	*CDK12*	*FBXW7*	*HIST1H2BH*	*IDH2*	*MLH1*	*PTEN*	*SOX10*
AMER1	*CDK16*	*FEM1A*	*HIST1H2BI*	*IFIH1*	*MLH3*	*PTPN11*	*SRCAP*
APC	*CDK4*	*FGFR1*	*HIST1H2BJ*	*IGF1R*	*MSH2*	*PTPN12*	*STAG2*
APOBEC3H	*CDK6*	*FGFR2*	*HIST1H2BK*	*IGF2BP2*	*MSH5*	*PTPN23*	*STAG3*
ARID1A	*CDKN1B*	*FHOD1*	*HIST1H2BL*	*IGFBP4*	*MSH6*	*PTPRF*	*STMN4*
ARID1B	*CDKN1C*	*FLT1*	*HIST1H2BM*	*IGFBP7*	*MTOR*	*PTPRT*	*SUFU*
ASXL1	*CDKN2A*	*FUBP1*	*HIST1H2BN*	*INO80E*	*MYC*	*PUM1*	*SUV420H1*
ATF5	*CDKN2B*	*GAB2*	*HIST1H2BO*	*KAT6B*	*MYCN*	*PXDN*	*SUV420H2*
ATG12	*CDKN2C*	*GABRP*	*HIST1H3A*	*KDM3A*	*NF1*	*QKI*	*SUZ12*
ATM	*CENPB*	*GLI2*	*HIST1H3B*	*KDM3B*	*NF2*	*RAB11FIP1*	*SYVN1*
ATN1	*CHAF1A*	*GLI3*	*HIST1H3C*	*KDM4B*	*NFIB*	*RAB36*	*TBX3*
ATR	*CHD5*	*GNB2L1*	*HIST1H3D*	*KDM5B*	*NRAS*	*RAD50*	*TCF7L2*
ATRX	*CHD7*	*GOLPH3*	*HIST1H3E*	*KDM6B*	*NTRK1*	*RAD51C*	*TEAD2*
AURKB	*CHEK2*	*H3F3A*	*HIST1H3F*	*KDR*	*NTRK2*	*RALGAPB*	*TERT*
AXL	*CIC*	*HCN1*	*HIST1H3G*	*KIT*	*NTRK3*	*RASGRF2*	*TFR2*
BCOR	*CNPY4*	*HDAC1*	*HIST1H3H*	*KMT2B*	*PALB2*	*RB1*	*TLR7*
BCORL1	*CRAMP1L*	*HDAC6*	*HIST1H3I*	*KMT2C*	*PARP10*	*RERE*	*TOP3A*
BCR	*CSF1R*	*HECW1*	*HIST1H3J*	*KMT2D*	*PDGFA*	*RET*	*TP53*
BLM	*CSPG4*	*HIST1H1A*	*HIST1H4A*	*KRAS*	*PDGFB*	*ROS1*	*TP53BP2*
BMP1	*CTNNB1*	*HIST1H1B*	*HIST1H4B*	*LPL*	*PDGFRA*	*RPTOR*	*TSC2*
BMP2	*DAXX*	*HIST1H1C*	*HIST1H4C*	*MAFF*	*PIK3AP1*	*RXFP2*	*UBA1*
BMP2K	*DEPDC5*	*HIST1H1D*	*HIST1H4D*	*MAFK*	*PIK3C2A*	*SALL4*	*ULK2*
BMP3	*DICER1*	*HIST1H1E*	*HIST1H4E*	*MAP2K4*	*PIK3C2B*	*SCAF1*	*UNC80*
BMP8A	*DROSHA*	*HIST1H1T*	*HIST1H4F*	*MAP2K7*	*PIK3C2G*	*SETD1A*	*VIL1*
BPTF	*DSG1*	*HIST1H2AA*	*HIST1H4G*	*MAP3K1*	*PIK3CA*	*SETD1B*	*WDR20*
BRAF	*E2F1*	*HIST1H2AB*	*HIST1H4H*	*MAP3K13*	*PIK3CB*	*SETD2*	*WEE1*
BRCA1	*EGFR*	*HIST1H2AC*	*HIST1H4I*	*MAP3K15*	*PIK3R1*	*SF3A1*	*WNT11*
BRCA2	*EIF3A*	*HIST1H2AD*	*HIST1H4J*	*MAP3K4*	*PIK3R2*	*SF3A2*	*WNT2*
BRD4	*ELMO3*	*HIST1H2AE*	*HIST1H4K*	*MAP3K6*	*PIK3R5*	*SF3A3*	*WNT7B*
BRIP1	*ENC1*	*HIST1H2AG*	*HIST1H4L*	*MAP3K9*	*PIM1*	*SF3B1*	*WNT8A*
BUB3	*EPHB3*	*HIST1H2AI*	*HIST2H2AA3*	*MAPK7*	*PLAGL2*	*SF3B2*	*WNT9A*
CBL	*ERBB2*	*HIST1H2AJ*	*HIST2H2AA4*	*MAPKAPK2*	*PLCG1*	*SF3B3*	*ZKSCAN7*
CBX4	*ERBB3*	*HIST1H2AK*	*HIST2H2AB*	*MAPKAPK3*	*PML*	*SHKBP1*	
CC2D1A	*ERBB4*	*HIST1H2AL*	*HIST2H2AC*	*MARS*	*PMS1*	*SHROOM4*	
CCNA1	*ERCC3*	*HIST1H2AM*	*HIST2H3C*	*MAX*	*PMS2*	*SMAD7*	
CCNB3	*ESPL1*	*HIST1H2BA*	*HNRNPUL1*	*MCM2*	*POLK*	*SMARCA2*	
CCND1	*ETV6*	*HIST1H2BB*	*HRAS*	*MCM5*	*POLQ*	*SMARCA4*	
CCND2	*EZH2*	*HIST1H2BC*	*HUWE1*	*MCM8*	*POLR1B*	*SMARCA5*	

Appendix Table 3 List of primers and probes used for ddPCR

Assay	FW	RV	WT-probe	Dye	Mutant-probe	Dye
H3F3A-K27M	GGTAAAGCACCCAGGAAG	CAAGAGAGACTTTGTCCC	TC+GC+A+A+GA+GT+GC	HEX	TC+GC+A+**T**+GA+GTGC	FAM
BRAF-V600E	CATGAAGACCTCACAGTAAAAATAGGTGAT	TGGGACCCACTCCATCGA	CTAGCTACAGTGAAATC	VIC	TAGCTACAGAGAAATC	FAM
IDH1-R132G	CTTGTGAGTGGATGGGTAAAACCTA	CACATTATTGCCAACATGACTTACTTGAT	AAGCATGACGACCTATG	VIC	AAGCATGACCACCTATG	FAM
IDH1-R132H	CTTGTGAGTGGATGGGTAAAACCTA	CCAACATGACTTACTTGATCCCCATA	CATCATAGGTCGTCATGC	VIC	ATCATAGGTCATCATGC	FAM
TP53-R282W	GCTTTGAGGTGCGTGTTTGTG	CTTTCTTGCGGAGATTCTCTTCCT	TGCGCCGGTCTCT	VIC	TGCGCCAGTCTCT	FAM
PIK3CA-E542K	GGGAAAATGACAAAGAACAGCTCAA	GCACTTACCTGTGACTCCATAGAAA	CCTCTCTCTGAAATCA	VIC	CCTCTCTCTAAAATCA	FAM
PIK3CA-H1047R	GCAAGAGGCTTTGGAGTATTTCATG	GCTGTTTAATTGTGTGGAAGATCCAA	CCACCATGATGTGCATC	VIC	CACCATGACGTGCATC	FAM

Appendix Table 4 STR profiling of DIPG BIOMEDE-UK samples. The table includes the tissue (FF, fresh frozen), blood and in vivo primary cultures (2D and 3D).

ID	Tyoe	D21S11	TH01	TPOX	VWA	AMEL	CSF1PO	D16S539	D7S820	D13S317	D5S818
B117	FF	29, 30	6, 9.3	8	18	XY	12	11, 12	11, 12	11	11
B117	Blood	29, 30	6, 9.3	8	15, 18	XY	12	11, 12	11, 12	11	11
B117	2D	29, 30	6, 9.3	8	18	XY	12	11 12	11, 12	11	11
B117	3D	29,30	6,9.3	8	18	X,Y	12	11,12	11,12	11	11
B134	FF	28	9.3	8, 9	20	X	10, 12	8, 11	9, 11	9, 13	9, 11
B134	Blood	28, 29	9.3	8, 9	14, 20	X	10, 12	8, 11	9, 11	9, 13	9, 11
B134	2D	28, 29	9.3	8, 9	14, 20	X	10, 12	8, 11	9, 11	9, 13	9, 11
B134	3D	28	9.3	8, 9	20	X	10, 12	8, 11	9, 11	9, 13	9, 11

B121	FF	29, 30	7, 8	8, 11	15, 16	X	11, 12	9, 11	12	8, 12	11
B121	Blood	29, 30	7, 8	8, 11	15, 16	X	11, 12	9, 11	12	8, 12	11
B121	2D	29,30	7,8	8,11	15,16	X	11,12	9,11	12	8,12	11
B128	FF	28, 32.2	7	8	14, 17	XY	12	11, 12	9	8	11, 12
B128	Blood	28, 32.2	6, 7	8	14, 17	XY	12	11, 12	9	8	11, 12
B128	2D	28,30.2	7	8	14,17	X,Y	11,12	11,12	9	8,	11,12
B169	FF	29,30	7,9	8	16,18	X	10,14	9,11	10,11	11,14	11,12
B169	Blood	29,30	7,9	8	16,18	X	10,14	9,11	10,11	11,14	11,12
B169	2D	29,30	9	8	16,18	X	10,13	10,11	10,11	11	11,12
B169	3D	29,30	9	8	16,18	X	10,14	11	10,11	11	11,12
B181	Blood	29,32.2	9.3	8,11	15,16	X	12	11	8,9	8,9	12
B181	2D	29,32.2	9.3	8,11	15,16	X	12	11	8,9	8,9	12
B181	3D	29,32.2	9.3	8,11	15,16	X	12	11	8,9	8,9	12
B184	FF	28,31	7	8,11	15,19	X	10,11	9,13	7,11	9,11	12,13
B184	Blood	28,31	7	8,11	15,19	X	10,11	9,13	7,11	9,11	12,13
B186	FF	28,29	6	8	16,18	X	12	11,12	8,11	11	11,12
B186	Blood	28,29	6,7	8	16,18	X	12	11,12	8,11	11	11,12
B186	2D	28,29	6	8	16,18	X	12	11,12	8,11	11	11,12
B193	FF	30,35	6	11	15,16	X,Y	7,12	11,12	10	11,12	12,13
B193	Blood	30,35	6,9	11	15,16	X,Y	7,12	11,12	10	11,12	12,13
B193	2D	35	6	11	15,16	X,Y	7,12	11	9,10	12	12,13
B193	3D	35	6	11	15,17	X,Y	7,12	11,12	9,10	12,13	12,13
B194	Blood	27,30	7	8,11	18,19	X	9,12	11,13	8,10	8,12	12
B194	2D	27,30	7	8,11	18,19	X	9,12	11,13	8,10	8,12	12
B194	3D	27,30	7	8,11	18,19	X	9,12	11,13	8,10	12	12
B198	FF	29,31	7,9	9,11	16,17	X,Y	11,13	9,12	10,12	11	11,14
B198	Blood	29,31	7,9	9,11	16,17	X,Y	11,13	9,12	10,12	11	11,14
B198	3D	29,31	7,9	9,11	16,17	X,Y	11,13	9,12	10,12	11	11,14

APPENDIX II Manuscripts published

www.impactjournals.com/oncotarget/ Oncotarget, 2017, Vol. 8, (No. 67), pp: 112036-112050

Research Paper

Development of a targeted sequencing approach to identify prognostic, predictive and diagnostic markers in paediatric solid tumours

Elisa Izquierdo[1,2], Lina Yuan[1], Sally George[3,4], Michael Hubank[1], Chris Jones[2], Paula Proszek[1], Janet Shipley[5], Susanne A. Gatz[4,5], Caedyn Stinson[6], Andrew S. Moore[6,7,8], Steven C. Clifford[9], Debbie Hicks[9], Janet C. Lindsey[9], Rebecca M. Hill[9], Thomas S. Jacques[10,11], Jane Chalker[12], Khin Thway[13], Simon O'Connor[14], Lynley Marshall[4], Lucas Moreno[4,15], Andrew Pearson[4], Louis Chesler[3,4], Brian A. Walker[1,16,*] and David Gonzalez De Castro[1,17,*]

[1]Molecular Diagnostics Department, The Institute of Cancer Research and Clinical Genomics, The Royal Marsden NHS Foundation, London, United Kingdom

[2]Glioma Team, Division of Molecular Pathology and Cancer Therapeutics, The Institute of Cancer Research, London, United Kingdom

[3]Paediatric Tumour Biology, Division of Molecular Pathology, The Institute of Cancer Research, London, United Kingdom

[4]Paediatric Drug Development Team, Children and Young People's Unit, Royal Marsden NHS Foundation Trust, London, United Kingdom

[5]Sarcoma Molecular Pathology Team, Divisions of Molecular Pathology and Cancer Therapeutics, The Institute of Cancer Research, London, United Kingdom

[6]The University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, Australia

[7]Oncology Service, Children's Health Queensland Hospital and Health Service, Brisbane, Australia

[8]UQ Child Health Research Centre, The University of Queensland, Brisbane, Australia

[9]Wolfson Childhood Cancer Research Centre, Northern Institute for Cancer Research, Newcastle University, Newcastle Upon Tyne, United Kingdom

[10]Department of Histology, Great Ormond Street Hospital for Children NHS Foundation Trust, London, United Kingdom

[11]Developmental Biology and Cancer Programme, UCL GOS Institute of Child Health, London, United Kingdom

[12]Haematology, Cellular and Molecular Diagnostics Service, UCL GOS Institute of Child Health, London, United Kingdom

[13]Sarcoma Unit, Royal Marsden NHS Foundation Trust, London, United Kingdom

[14]Haemato-Oncology Unit, Royal Marsden NHS Foundation Trust, London, United Kingdom

[15]HNJ-CNIO Clinical Research Unit and Hospital Universitario Niño Jesus, Madrid, Spain

[16]Myeloma Institute, University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA

[17]Centre for Cancer Research and Cell Biology, Queen's University Belfast, Belfast, United Kingdom

[*]These authors have contributed equally to this work

Correspondence to: *David Gonzalez De Castro,* ***email:*** *D.GonzalezdeCastro@qub.ac.uk*

Keywords: *childhood cancer; targeted sequencing; molecular diagnostics; panel validation; targeted therapies*

Received: *July 28, 2017* ***Accepted:*** *November 16, 2017* ***Published:*** *December 06, 2017*

ABSTRACT

The implementation of personalised medicine in childhood cancers has been limited by a lack of clinically validated multi-target sequencing approaches specific for paediatric solid tumours. In order to support innovative clinical trials in high-risk patients with unmet need, we have developed a clinically relevant targeted sequencing panel spanning 311 kb and comprising 78 genes involved in childhood cancers. A total of 132 samples were used for the validation of the panel, including Horizon Discovery

cell blends (n=4), cell lines (n=15), formalin-fixed paraffin embedded (FFPE, n=83) and fresh frozen tissue (FF, n=30) patient samples. Cell blends containing known single nucleotide variants (SNVs, n=528) and small insertion-deletions (indels n=108) were used to define panel sensitivities of ≥98% for SNVs and ≥83% for indels [95% CI] and panel specificity of ≥98% [95% CI] for SNVs. FFPE samples performed comparably to FF samples (n=15 paired). Of 95 well-characterised genetic abnormalities in 33 clinical specimens and 13 cell lines (including SNVs, indels, amplifications, rearrangements and chromosome losses), 94 (98.9%) were detected by our approach. We have validated a robust and practical methodology to guide clinical management of children with solid tumours based on their molecular profiles. Our work demonstrates the value of targeted gene sequencing in the development of precision medicine strategies in paediatric oncology.

INTRODUCTION

Cancer remains the leading cause of death due to disease in children aged >1 year [1]. Cure rates for paediatric solid tumours have not substantially improved in the past decade with patients having recurrent disease performing particularly badly, reflecting the limitations of current approaches that employ intensive chemotherapy, surgery and radiation [2–4]. In adults, the stratification of patients by genetic profiling using high throughput sequencing has supported adaptive clinical trials [5, 6], and there is an urgent need to translate such opportunities to the treatment of childhood disease.

The genomic landscape of paediatric cancer is becoming increasingly well-defined leading to the conclusion that childhood cancers have in general fewer somatic mutations than adults, but that mutations in epigenetic regulators occur at a higher incidence [7–17]. Key recent findings include recurrent mutations in the genes encoding histones 3.3 and 3.1 (*H3F3A* and *HIST1H3B*) as well as the activin A receptor type I (*ACVR1*) that are unique to paediatric high-grade glioma (pHGG) and diffuse intrinsic pontine glioma (DIPG) [18–20]. Similarly, *ATRX* mutations, *TERT* rearrangements and *MYCN* amplification define mutually exclusive molecular subgroups of neuroblastoma, all of which are associated with poor prognosis [21–23]. The newly proposed molecular-based medulloblastoma sub-classification defines subgroups, each of which potentially requires a tailored therapeutic strategy [7, 11, 24].

Despite our improved knowledge of somatic alterations in paediatric cancers, precision medicine remains unavailable for the majority of patients. For example, a small number of early-phase paediatric trials are recruiting children whose tumours harbour genetic alterations including *ALK* genomic alterations (mutations, amplifications or translocations) that can be treated with ALK inhibitors and *BRAF* V600 mutant tumours that can be treated with BRAF or MEK inhibitors.

In addition, there is now an extensive list of recurrent genetic alterations with potential diagnostic, prognostic or predictive value, and sequential testing of single genes using standard methods has become unfeasible due to lack of available material and high costs. High-throughput sequencing (also known as next generation sequencing or NGS) offers a solution to these issues. In particular, panel-based NGS assays which simultaneously sequence a targeted set of genes with recurrent alterations, associated with known clinical or biological implications are cheaper, less challenging in terms of interpretation and more suited to clinical diagnostics than current approaches [25]. Despite this, development and validation of high throughput gene panel sequencing is challenging. Typically, DNA is only available from formalin-fixed, paraffin-embedded (FFPE) samples, which yields relatively poor quality DNA. DNA extraction and library construction to clinical laboratory standards requires optimisation, and it is necessary to construct a standardised informatics pipeline that identifies and interprets actionable mutations. Appropriate and rapid clinical reporting of identified variants and incorporation of the results into the electronic patient records also need to be considered if molecular stratification of childhood cancer is to be successfully translated to the clinic [26]. There are several examples of validation and implementation of targeted sequencing in adult cancer [27–30]. In the past two years, several approaches using high-throughput sequencing have been applied for clinical decision-making in children with solid tumours [31–34], however a clinically validated panel specifically targeting recurrent alterations in childhood cancers using archival FFPE specimens would significantly assist the development of molecular stratification strategies in paediatric oncology.

Here we describe the development and validation, within an accredited clinical pathology laboratory (CPA UK), of a paediatric solid tumour sequencing panel for use with either routine FFPE or fresh frozen (FF) samples. As part of the validation, we established overall performance, sensitivity, specificity, repeatability, reproducibility, accuracy and limit of detection, following guidelines previously described for validation of genetic tests [35].

RESULTS

Selection of panel content

The panel design covers a total of 78 genes (Table 1), either recurrently altered in paediatric cancers or clinically actionable in adult cancers and with potential application in childhood solid tumours. The genes were selected in wide-collaboration with national experts in paediatric oncology patient care covering all areas of paediatric solid tumours (glioma, medulloblastoma, bone sarcomas, soft tissue sarcomas, renal tumours and neuroblastoma among others). Targets were chosen by consensus based on most clinically relevant aberrations including: i) predictive biomarker (level 1), prognostic biomarker (level 2), diagnostic biomarker (3) potentially targetable biomarkers with inhibitors available or under development (level 4), known germline or high risk single nucleotide polymorphism (level 5) or unclear significance, research only (level 6). Factors influencing the choice of targets included: childhood tumour type where alterations have been reported, molecules targeting these genes and clinical trials available for children with solid tumours (Supplementary Table 1A). A library of customized biotinylated DNA probes was designed to capture a total of ~311(kilobase) kb for the detection of single nucleotide variants (SNVs), short insertion-deletions (indels), copy number variations and structural rearrangements (Supplementary Table 1B). Exons were padded with 5 base pairs (bp) of intronic sequence to increase exon depth and for detection of splice-site variants.

Panel validation

Research use of sequence capture assays has become common, but basing clinical care on gene panel sequencing results requires confident calling of both variant and non-variant sequence, and a full understanding of the performance of the assay. Implementation in the clinic therefore requires robust validation in an accredited laboratory.

To validate the paediatric gene panel, we followed the standardised framework for clinical assay validation set out by Mattocks *et al.* [35]. We determined overall performance of the panel across the target regions, measuring precision, sensitivity and specificity. As a standard, we used a set of four Horizon cell blends previously characterized by NGS and droplet digital PCR (ddPCR) (Supplementary Table 2A and 2B) and 15 paediatric cell lines with known variants. For further validation, we performed capture and sequencing on 83 FFPE and 30 FF clinical samples (Supplementary Table 3).

Overall performance

Overall, the panel performed well, with over 96% of 901 regions of interest achieving specification. Only 24 (2.7%) regions were classified as underperforming across the four cell blends and five FFPE samples, with read depth lower than 2 x standard deviation (SD) of the mean based on $\log_2$ (Supplementary Table 4A and 4B). 22 of 24 underperforming regions were located within highly GC-enriched regions, which are known to be refractory to efficient hybridization and/or amplification (Figure 1 and Supplementary Table 4C).

Quality and coverage metrics were generated across all samples (Supplementary Tables 5 and 6). The average total number of reads was $8.8x10^6$ (SD=$3.1x10^6$) for FFPE and $7.9x10^6$ (SD=$3x10^6$) for high molecular weight (HMW) samples (FF and cell lines). The percentage mapped (96.1±3.9 for FFPE vs 97.3±2.5 for high molecular weight samples) and percentage of bases from unique reads on target (45.9±3 for FFPE vs 42.7±2.4 for HMW) was very similar for both FFPE and HMW samples. Duplicates were higher in FFPE samples (60.2% for FFPE vs 36.1% for HMW). The overall mean depth was 698 ± 365 for FFPE vs 899 ± 347 for HMW (Table 2).

Limit of detection

To determine the limit of detection, SNVs present in the cell blends at known variant allele frequency (VAF) were used. The pipeline detected all 61 SNVs including 33 SNVs with an expected VAF of 4-5%. 15/17 expected indels were detected. Of the two indels not detected, one was 18 bp in length at an expected VAF of 4.2%, whilst the other was 2 bp at 5% VAF (Figure 2 and Supplementary Table 7A). We therefore established a minimum threshold of 5% VAF in the analysis pipeline, which allows for detection of a heterozygous mutation when >10% neoplastic cells are present in the tumour sample.

Assessment of precision

To measure precision, we took advantage of natural variants present as intrinsic "background" SNVs and indels in the captured regions from the four cell blends. Precision was assessed by comparing the alterations expected with those detected to obtain within run-precision (repeatability), and between run-precision data (intermediate precision). Variants ≤ 5% in all four blends and within poor performing regions were excluded leaving a total of 528 SNVs (132 variants in 4 blends) and 108 indels (27 indels in 4 blends) for analysis. All of the 528 SNVs and 90 out of 108 (83%) indels were detected (Supplementary Table 8).

Repeatability

Pairwise correlation of VAF between runs was $r^2 \geq 0.994$ [95%CI:0.991-0.996] for SNVs and $r^2 \geq 0.785$ [95%CI:0.652-0.919] for indels (Supplementary Figures 1 and 2) indicating that the panel accurately reproduces data from independently prepared samples on the same run.

Table 1: Gene panel list including 78 genes recurrently altered in paediatric cancers or clinically actionable

ACVR1	CTNNB1	IL3	PPM1D
AKT1	*DDR2*	*IL6*	*PTCH1*
ALK	*DDX3X*	*KIT*	*PTEN*
AMER1	*DICER1*	*KMT2D*	*PTPN11*
APC	*EGFR*	*KRAS*	*PTPRD*
ARID1A	*ERBB2*	*LMO1*	*RB1*
ARID1B	*ERG*	*MAP2K1*	*RET*
ASXL1	*ETV6*	*MAP2K2*	*ROS1*
ATM	*EWSR1*	*MDM2*	*SETD2*
ATRX	*FBXW7*	*MYCN*	*SMARCA4*
BARD1	*FGFR1*	*MYOD1*	*SMARCB1*
BCOR	*FGFR2*	*NCOA2*	*SS18*
BRAF	*FGFR3*	*NF1*	*SUFU*
CASC15	*FGFR4*	*NRAS*	*TENM3*
CDK4	*FUS*	*PAX3*	*TP53*
CDK6	*H3F3A*	*PAX7*	*TSC1*
CDKN2A	*HIST1H3B*	*PDGFRA*	*WT1*
CDKN2B	*HIST1H3C*	*PHOX2B*	*ZHX2*
CFL1	*HRAS*	*PIK3CA*	
CHEK2	*IDH1*	*PIK3R1*	

Intermediate precision

Pairwise correlation was $r^2 \geq 0.995$ [95%CI:0.993-0.997] and overall correlation was $r^2 = 0.996$ [95%CI:0.995-0.997] for SNV detection. For indels pairwise correlation was $r^2 \geq 0.827$ [95%CI: 0.716-0.937] and overall correlation was $r^2 \geq 0.875$ [95%CI:0.829-0.921] (Table 3 and Supplementary Figures 3 and 4) indicating that the panel accurately reproduces data from repeat samples on different runs.

Assessment of sensitivity and specificity

To determine sensitivity we used the same background 528 SNVs and 108 indels, together with the known cancer-specific variants (61 SNVs and 17 indels) from the four cell blends. SNVs and indels were called and their presence was compared to the list of variants expected in the capture regions from the cell blends (Supplementary Tables 7A and 8). All the SNVs were detected, resulting in a sensitivity of ≥98% [95%CI:0.98-1]. From the 108 background indels, 18 were not detected, as were 2 of the cancer-specific indels, obtaining a sensitivity of ≥83% [95%CI:0.761-0.897]. True Positive (TP) of all SNVs = 589; False-Negative (FN) of all SNVs = 0. TPs of all indels = 105; FNs of all indels = 20. The undetected indels were manually checked on Integrative Genomics Viewer (IGV). We observed that 12 of 20 were located +4 bp upstream of the exon (our bed file covers ±5 bp), four had poor coverage, two fell in highly repetitive regions and one was a long indel (18bp).

To determine specificity, we used the cancer-specific data from the four cell blends harbouring a total of 61 true positive and 87 true negative SNVs (Supplementary Table 7B). There were insufficient true negatives (n=3) to determine specificity for indels. SNVs were called and their presence was compared to the list of variants expected in the capture regions from the cell blends. The specificity of cancer-specific SNVs was ≥98% [95%CI:0.946-1]. Positive-Predictive Value (PPV) was ≥98% [95%CI:0.926-1] and the Negative-Predictive Value (NPV) was ≥98% [95%CI:0.946-1].

The range of VAF for the SNVs detected by our pipeline, including the background and the cancer specific variants (528 + 61 = 589), was 23% at ≥ 50% VAF (134/589), 35% at 50-20% of VAF (207/589) and 42% at < 20% (248/589). The range of VAF for the indels detected by our pipeline including the background and cancer specific indels (90 + 15 = 105) was 0% at > 50% of VAF (0/105), 31% at 50-20% of VAF (33/105) and 69%

<20% (72/105). The range of VAF for the cancer specific variants detected by the NGS panel was in line with the manufacturer's specifications.

Performance and variant detection comparison in paired FF-FFPE clinical samples

To assess the performance of the panel on real clinical material we compared 15 paired clinical DNA samples isolated from both FF and FFPE samples. For the FFPE samples, we obtained an average of 93.4% ± 5.42% and 80.3% ± 20.3% of targeted positions covered at depths of ≥ 100x and ≥ 250x respectively. The overall mean depth for FFPE was 785 ± 333. Overall percentage of bases from unique reads on target for FFPE was 47.6% ± 2.3%. For FF samples, we obtained an average of 96.6% ± 0.6% and 90.9% ± 1.9% of targeted positions covered at depths of ≥ 100x and ≥ 250x respectively. The overall mean depth for FF was 977 ± 142. Overall percentage of bases from unique reads on target for FF was 44% ± 2.2%. As expected, duplicates were substantially lower in FF samples (54.5% for FFPE vs 29.9% for FF). Insert size for the library pre-capture DNA was 285 bp ± 24 for FFPE and 326 bp ± 24 for FF (Table 4).

VAFs found in the paired FF-FFPE samples were compared, obtaining an overall correlation of r^2 = 0.983 (95%CI: 0.984-0.985; p<0.0001) (Figure 3 and Supplementary Figure 5). A total of 42.3% (5562/13146) variants were detected in FF but not in FFPE, of which 78.1% (4346/5562) had VAF below 5%, with 17.6% (982/5562) having VAF between 5-10%. Less than 5% variants missed in FFPE samples were present in FF at VAF above 10%.

Conversely, a total of 8.2% (1084/13146) variants were detected in FFPE but not in FF, of which 50.8% (551/1084) had VAF below 5%, with 33.2% (360/1084) having VAF between 5-10%, and the remaining 16.0% (173/1084) were present in FFPE only at VAF above 10%.

Detection of known variants in paediatric samples

To assess the ability of the panel to detect known variants in clinical samples, we performed a variant analysis of 41 paediatric samples with 90 known genetic abnormalities (30 alterations in 13 cell lines and 60 alterations in 14 FFPE and 14 FF samples with known genetic alterations identified by routine testing): 50 SNVs, including mutations in *TP53, ALK, CTNNB1, DDX3X, SMARCA4*, one duplication (*BRAF* p.Thr599dup), 7 indels including *DDX3X and TP53*, 13 amplifications including *MYCN* and *CDK4*, and 19 chromosome/gene losses, for example chr 9q loss including loss of *PTCH1* and *TSC1*. 100% of the variants interrogated by the panel were successfully detected (Tables 5 and 6 and Supplementary Table 9).

Detection of rearrangements

Five sarcoma FFPE samples were included in the analysis where translocations had previously been detected by RT-qPCR involving *EWSR1*. Rearrangements in *EWSR1* were detected in four out of the five FFPE samples (80%) leading to fusion genes of *EWSR1* with partners *ATF1* (detected in two samples), *FLI1* and *CREB1* (Supplementary Figure 6). The fusion not detected was *EWSR1-NR4A3*. This is too small a sample to confirm validation of the panel for detection of translocations at this stage and further work is in progress.

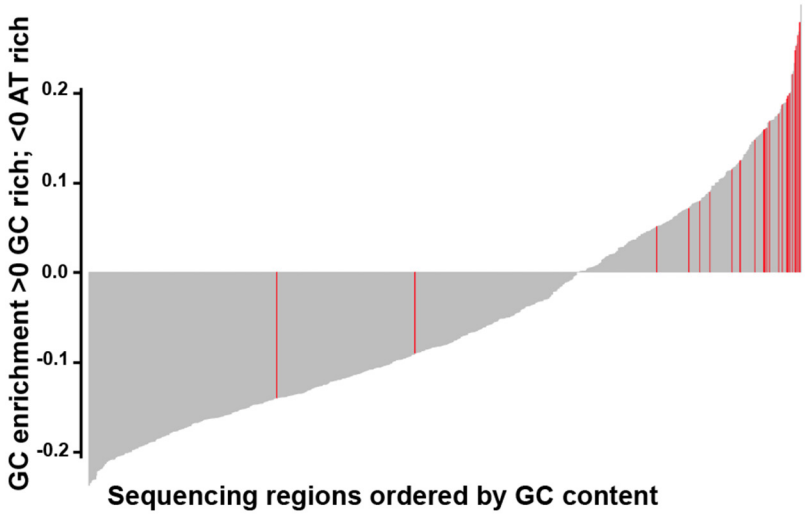


Figure 1: Bar plot showing GC content in the 901 regions captured by the panel. The plot is ordered from low to high GC-content of each region captured. Red bars highlight the underperforming regions (24/901), mainly located within GC-enriched regions.

Table 2: Average quality metrics across all samples. Data expressed as means ± standard deviation

	Total reads	Percentage of reads mapped	Percentage of duplicates	Percentage of unique on target	Mean depth
FFPE (n=83)	$8.8x10^6$±$3.1x10^6$	96.1±3.9	60.2±13.7	45.9±3	698±365
FF and cell lines (n=49)	$7.9x10^6$ ±$3x10^6$	97.3±2.5	36.1±9.7	42.7±2.4	899±347

DISCUSSION

Targeted therapies are already the standard of care for several molecular subgroups of adult cancers. *EGFR* mutations or *ALK* rearrangements in lung cancer, *BRAF* V600E mutations in metastatic melanoma and breast cancer patients harbouring *HER2* amplifications are examples of therapeutic biomarkers routinely used in the adult population [36–38]. The implementation of personalised medicine in paediatric oncology has remained challenging partly due to the low incidence of childhood cancer, accessibility of drugs and regulatory hurdles [39]. Nevertheless, the understanding of genetics in childhood cancer over the last decade has improved thanks to large sequencing initiatives across the world [31–33]. The updated World Health Organization Classification (WHO) classification of brain tumours based on molecular features is a clear example of the huge impact of applying molecular profiling to guide diagnosis and treatment with the potential to improve outcomes in childhood cancers [40].

We have developed an NGS targeted sequencing based diagnostic test to accurately detect clinically relevant genomic alterations across 78 cancer genes in routine FFPE as well as FF paediatric samples. The overall performance of our assay was excellent; from the 901 regions captured only 24 (<3%) failed the quality control metrics mainly as a result of being located in GC-rich regions, and should be noted for future panel design. VAF for known SNVs and indels were very similar in within-run and between-run replicates, demonstrating that the assay is repeatable and reproducible. SNVs were detected at a wide range of VAFs simulating the heterogeneity expected in cancer samples including 33 SNVs with an expected VAF of 4-5%. The detection of variants at low VAF is crucial, especially in samples with a low neoplastic cell content. Sensitivity was ≥98% for SNVs and ≥83% for indels and specificity ≥98% for SNVs. False-negative calls were mostly indels at low VAF (≤5%) and predominantly occurred at splice sites. Variants were analysed in exons and the surrounding 5 bp, but were not reported by our pipeline if they occurred outside ±2 bp of the coding region. This could be solved expanding the sequence covered by bed file at intron:exon boundaries, but the relevance of these variants remains unclear. There is currently no consensus as to the most

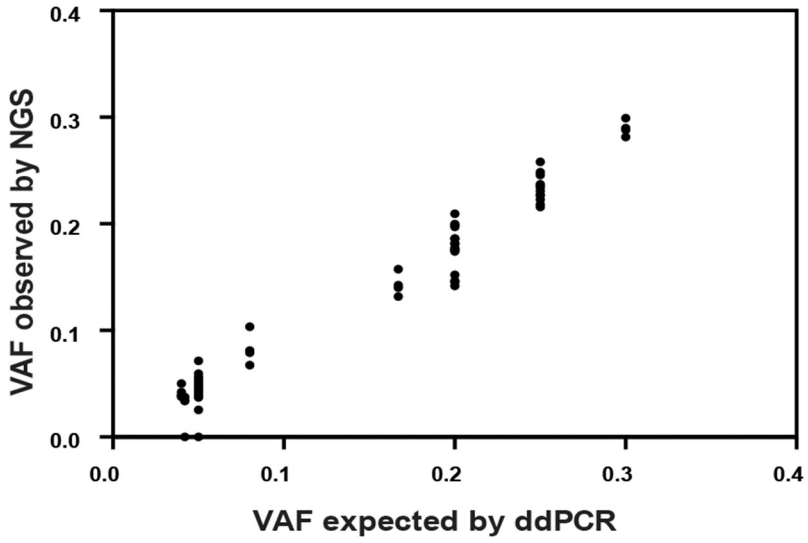


Figure 2: Comparison of known variant allele frequencies by droplet digital PCR (x axis) against variant allele frequency obtained by NGS (y axis) for all cancer-specific variants (61 single nucleotide variants, SNVs and 17 insertion-deletions, indels). Overall correlation was r^2=0.969 [95% CI: 0.975-0.990; p<0.0001].

Table 3: Pairwise correlation of (A) single nucleotide variants (SNVs) and (B) insertion-deletions (indels) for each of the 4 cell blends with identical background variants between the two runs

A Samples	Correlation	Standard Error	Lower 95%CI	Upper 95%CI
Tru-Q1-HD728-T	0.995	0.001	0.993	0.997
Tru-Q2-HD729-T	0.996	0.001	0.995	0.997
Tru-Q3-HD730-T	0.996	0.001	0.995	0.998
Tru-Q4-HD731-T	0.997	0.001	0.995	0.998
B Samples	**Correlation**	**Standard Error**	**Lower 95%CI**	**Upper 95%CI**
Tru-Q1-HD728-T	0.912	0.03	0.853	0.971
Tru-Q2-HD729-T	0.905	0.032	0.842	0.969
Tru-Q3-HD730-T	0.858	0.047	0.766	0.951
Tru-Q4-HD731-T	0.827	0.057	0.716	0.937

appropriate minimum region of interest to cover at splice sites for clinical reporting and in many cases the biological meaning of these mutations are unknown. The latest guidelines recommend calling likely disrupted gene function in nonsense, frameshift, canonical ±1 or ±2 splice sites, initiation codon, and single exon or multi-exon deletion, all of which would be covered with our current pipeline [41].

We also compared the performance of paired FFPE-FF specimens obtaining comparable quality metrics between both tissue types, as well as a high overall correlation of VAF. This is particularly important as most clinical samples routinely available are derived from FFPE tissue where nucleic acid quality is generally compromised and chemically challenged, leading to DNA degradation and potential deamination or oxidation artefacts. The discrepancies of the variants observed between FFPE and FF were mainly at low VAF, below or at the lower limit of detection of our approach. The discrepancies of the variants above 10% could be explained by variation in neoplastic cell content between FF and FFPE and intra-tumour heterogeneity leading to sub-clonal alterations. Three of the samples with more striking differences were brain tumours which are well known as highly heterogeneous tumours [42, 43].

We verified the accuracy of our NGS approach in cell lines and clinical specimens (FFPE and FF) containing known genetic abnormalities previously characterized by other methodologies and obtained a high concordance (r^2 = 0.983). The FFPE and FF samples used for the validation were a cohort of specimens from several hospitals across the world. We obtained reproducible and accurate results

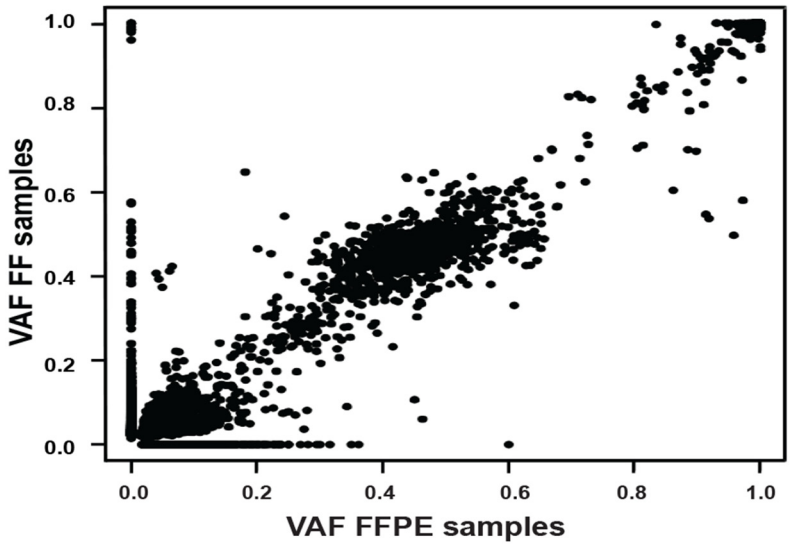


Figure 3: Overall correlation of variant allele frequency (VAFs) found between the 15 formalin-fixed paraffin embedded (x axis) and fresh frozen (y axis) paired samples.

Table 4: Comparison of quality metrics between formalin-fixed paraffin embedded (FFPE) and fresh frozen (FF) matched samples (n=15)

	Total reads	Percentage of reads mapped	Percentage of duplicates	Percentage of unique on target	Mean depth
FFPE (n=15)	$8.7x10^7 \pm 3.4x10^6$	95.5±2.2	54.5.2±9.2	47.6±2.3	785±333
FF (n=15)	$7.7x10^6 \pm 1.2x10^6$	98.6±0.7	29.9±6.9	44.2±2.2	977±142

Data expressed as means ± standard deviation.

Table 5: Known variants in paediatric cancer cell lines were compared against capture sequencing from the Cancer Cell Line Encyclopaedia and other published data

Cell line ID	Gene	Alteration	Detected	Allele frequency expected	Allele frequency observed
Be(2)C	TP53	p.C135F	YES	no data available	100%
Be(2)C	MYCN	AMPLIFICATION	YES	not applicable	not applicable
CCA	KRAS	p.Q61L	YES	no data available	29%
IMR32	ATM	p.V2716A	YES	59%	59%
IMR32	MYCN	AMPLIFICATION	YES	not applicable	not applicable
KELLY	ALK	p.F1174L	YES	39%	32%
KELLY	MAP2K1	p.A390T	YES	48%	47%
KELLY	TP53	p.P177T	YES	93%	99%
KELLY	MYCN	AMPLIFICATION	YES	not applicable	not applicable
LAN1	ALK	p.F1174L	YES	no data available	47%
LAN1	TP53	p.C182*	YES	no data available	99%
LAN1	MYCN	AMPLIFICATION	YES	not applicable	not applicable
LAN5	ALK	p.R1275Q	YES	no data available	50%
LAN5	MYCN	AMPLIFICATION	YES	not applicable	not applicable
NBLS	NF1	splice_acceptor_ variant c.6705-1G>T	YES	no data available	42%
RD	ATM†	p.D273N	YES	17%	2%
RD	NF1	p.E977*	YES	56%	59%
RD	NRAS	p.Q61H	YES	68%	61%
RD	TP53	p.R248W	YES	100%	100%
RH30	CDK4	AMPLIFICATION	YES	not applicable	not applicable
RH41	APC	p.M526L	YES	60%	59%
RH41	TP53	p.P152fs	YES	100%	100%
RMS559	FGFR4	p.V582L	YES	no data available	76%
SKNAS	NRAS	p.Q61L	YES	45%	46%
SKNAS	RB1	p.L477P	YES	47%	31%
SKNAS	TP53	DEL exons 10,11	YES	not applicable	not applicable
SKNSH	NRAS	p.Q61L	YES	15%	23%
SKNSH	SMARCA4	p.R973T	YES	32%	45%
SKNSH	CHEK2	p.T410fs	YES	59%	44%
SKNSH	ALK	p.F1174L	YES	no data available	36%

†*ATM* mutation in this cell line is subclonal and variation in AF is expected with on-going passages.

Table 6: Known variants in paediatric FFPE (n=14) and FF (n=14) samples were compared against other platforms such as RNA seq, 450k array, Sanger Sequencing and *FISH*

Genes with alterations detected by other methodologies	Alteration	Tumour type	Total cases expected	% of cases detected
DDX3X	SNV and indel	Medulloblastoma	6	100
PTCH1	SNV and indel	Medulloblastoma	5	100
TP53	SNV and indel	Medulloblastoma	3	100
MYCN	SNV	Medulloblastoma	2	100
MYCN	Amplification	Neuroblastoma (n=3) Medulloblastoma (n=4)	7	100
CTNNB1	SNV	Medulloblastoma	5	100
H3F3A	SNV	Glioma	3	100
SMARCA4	SNV	Medulloblastoma	3	100
BRAF	SNV	Glioma	2	100
ALK	SNV	Neuroblastoma	1	100
HIST1H3B	SNV	Glioma	1	100
AKT1	SNV	Medulloblastoma	1	100
ACVR1	SNV	Medulloblastoma	1	100
PIK3CA	SNV	Medulloblastoma	1	100
MLL2	SNV	Medulloblastoma	1	100
chr 9q - (PTCH1, TSC1)	loss	Medulloblastoma	5	100
chr 10- (PTEN, SUFU, FGFR2)	loss	Medulloblastoma	4	100
chr 6- (HIST1H3B, HIST1H3C, ROS1, ARID1B)	loss	Medulloblastoma	2	100
chr 9- (CDKN2A/B, PTHC1, TSC1)	loss	Medulloblastoma	2	100
chr12- (MLL2, CDK4)	loss	Medulloblastoma	1	100
ATM LOH	loss	Medulloblastoma	1	100
chr 3p- (CTNNB1, STED2)	loss	Medulloblastoma	1	100
chr17- (TP53, NF1, HER2, PPM1D)	loss	Medulloblastoma	1	100
chr17p- (NF1, TP53)	loss	Medulloblastoma	1	100
Total			60	

from different quality samples processed in different pathology laboratories, demonstrating the value of this approach for the development of national and international clinical trials in paediatric oncology.

Our data shows that this NGS approach can detect structural variants, including amplifications, deletions and chromosomal rearrangements. These types of variants are not generally detected with commercial amplicon-based NGS panels, despite being of critical importance for the clinical management and diagnosis of paediatric patients (e.g. *MYCN* amplification in neuroblastoma, *EWSR1* in Ewing's sarcoma). Only one out of five chromosomal rearrangements involving *EWSR1* was not identified by the assay which could be due to the lack of coverage at the intronic genomic location of the breakpoint. As expected, this is one of the limitations of the methodology, as capturing intronic regions commonly involved in translocations poses challenges associated to the presence of repetitive sequence elements. This can be partially overcome by including capture baits for the breakpoint regions of the most common partner genes involved in the translocations.

In summary, we have developed a robust clinical test that can detect SNVs, small indels, copy number variation and with high reproducibility and repeatability in routine clinical FFPE samples from a variety of centres.

Our approach has been incorporated into a pilot molecular profiling study for paediatric patients at the Royal Marsden Hospital (London, UK) and this has now been extended across the UK as the METEOR programme, an interim step towards the UK's more advanced paediatric molecular profiling programme, Stratified Medicine-Paediatrics (SM-Paeds) which is about to be rolled out throughout the UK. The NGS panel will form a key part of the SM-Paeds programme, which is underpinning UK patient eligibility screening for several clinical trials including the highly innovative international ITCC basket trial, called ESMART (NCT02813135), where patients are enrolled according to molecular alterations found in their tumours on biopsy at relapse. This is the first time that genomic results are incorporated into the patient's record in paediatric cancer in the UK within a clinically relevant timeframe of 3-5 weeks from DNA extraction to report generation. Our data shows that this NGS assay can be an accurate and a practical platform for molecular stratification and identification of actionable targets required to accelerate personalised medicine clinical trials in childhood solid tumours.

MATERIALS AND METHODS

Validation samples

A representative selection of common, poor risk paediatric tumours was used for the validation comprising 132 samples: i) Four cell blends with validated variants (Tru-Q1-4 HorizonDiscovery, Cambridge, UK), ii) 15 paediatric cell lines iii) 83 FFPE clinical samples and iv) 30 FF clinical samples (Supplementary Table 3).

Local institutional review board approval was obtained for the project in addition to separate approvals from the contributing tumour banks (The Children's Cancer and Leukaemia Group Tumour Bank and the Queensland Children's Tumour Bank).

Sample preparation

Assessment from haematoxylin and eosin (H&E) stained slides was performed by experienced pathologists to mark the region of the section containing tumour and to estimate neoplastic cell content, defined as the percentage of neoplastic cells out of total nucleated cells in the marked area. Tumour cellularity, reflecting the density of tumour nuclei, was also estimated. Macro-dissection of the marked area was performed when a distinct area of neoplastic cells from normal cells was observed in a large area and the overall tumour content without macro-dissection would have been <60%. 24 out of the 83 FFPE samples underwent macro-dissection to enrich the tumour content. DNA from blood and cell lines, FF and FFPE samples was extracted using the QIAamp DNA blood mini kit, the QIAamp DNA mini kit and the QIAamp DNA FFPE tissue kit (Qiagen, Hilden, Germany), respectively. For specimens where DNA was extracted at local centres, methods are provided in Supplementary Materials. DNA was quantified using Qubit dsDNA High Sensitivity Assay Kit with the Qubit 2.0 fluorometer, (Invitrogen, Carlsbad, CA). Analysis by TapeStation 2200 using the genomic DNA ScreenTape assay (Agilent Technologies, Santa Clara, CA) was performed to determine the degree of fragmentation of genomic DNA prior to library preparation. Based on optimization studies, samples yielding DNA with median fragment length > 1000 bp were processed using 200 ng DNA. Samples with DNA < 1000 bp were processed using 400 ng if there was sufficient DNA.

Gene panel capture and sequencing

Library preparation was performed using the KAPA Hyper and HyperPlus Kit (Kapa Biosystems, Wilmington, MA, USA) and SeqCap EZ adapters (Roche, NimbleGen, Madison WI, USA), following the manufacturer's protocol, including dual-SPRI size selection of the libraries (250-450 bp). In samples prepared using the KAPA Hyper Kit (n=39), DNA was sheared with the Covaris M220 (Covaris, Woburn, MA) using supplier protocols. KAPA HyperPlus employs enzymatic fragmentation and was used in 93 samples. Optimization of the process indicated that the change from enzymatic fragmentation resulted in a substantial improvement in library complexity and unique coverage depth compared to sonication [44]. Following fragmentation DNA was end-repaired, A-tailed and indexed adapters ligated. To optimise enrichment and reduce off-target capture, pooled, multiplexed, amplified pre-capture libraries (6 to 10 cycles according to the DNA input) were hybridized twice overnight (up to 13 samples per hybridization, consecutive days) using 1 μg of the pooled library DNA to a custom design of DNA baits complementary to the genomic regions of interest (NimbleGen SeqCap EZ library, Roche, Madison, WI, USA). A 5 cycle PCR was performed between hybridizations to enrich the captured product. After hybridisation, unbound capture baits were washed away and the remaining hybridised DNA was PCR amplified (12 cycles). PCR products were purified using AMPure XP beads (Beckman Coulter, Danvers, MA, USA) and quantified using the KAPA Quantification q-PCR Kit (KAPA Biosystems, Wilmington, MA, USA). Sequencing was performed on a MiSeq (Illumina, San Diego, CA, USA) with 75 bp paired-end reads and v3 chemistry according to the manufacturer's instructions. For samples where germline matched control was available (n=23), pools from tumour and control DNA libraries were multiplexed separately for hybridization and combined prior to sequencing at a ratio of 4:1, increasing the relative number of reads derived from tumour DNA.

Data analysis

Primary analysis was performed using MiSeq Reporter Software (v2.5.1; Illumina), generating nucleotide sequences and base quality scores in Fastq format. Resulting sequences were aligned against the human reference sequence build GRCh37/Hg19 to generate binary alignment (BAM) and variant call files (vcf). Secondary analysis was performed in-house using Molecular Diagnostics Information Management System to generate QC, variant annotation, data visualisation and a clinical report. In the Molecular Diagnostics Information Management System, reads were deduplicated using Picard (http://broadinstitute.github.io/picard/), and metrics generated for each panel region. Oncotator (v1.5.3.0) (https://portals.broadinstitute.org/oncotator) was used to annotate point mutations and indels using a minimum variant allele frequency (VAF) of 5% and a minimum number of 10 variant reads. Manta (https://github.com/Illumina/manta) was used for the detection of structural variants. Variants were annotated for gene names, nature of variant (e.g. missense), PolyPhen-2 predictions, and cancer-specific annotations from the variant databases including COSMIC, Tumorscape, and published MutSig results. Copy number variation (CNV) was assessed using the ratio of GC-normalized depth of region of interest (ROI) in tumour against GC-normalized read depth of ROI in either matched germline DNA (when available) or the male cell line G147A (Promega, Madison, WI USA). Any ratio below 0.65 fold was defined as a potential deletion whereas a ratio above 2.4 was flagged as a potential amplification. All potential mutations, structural variants and CNVs were visualised using IGV and two individuals were required to review the mutation report independently. Variant calls from samples with previously known SNVs and indels were checked manually on IGV.

Cell blends

The four cell blends contained 163 SNVs and 34 indels common to all four blends (background variants) (Supplementary Table 2A). Additionally, there were 61 SNVs and 17 indels, cancer variants, which were unique between blends, introduced at known VAF, and verified by ddPCR (Supplementary Table 2B). The four cell blends were used to assess overall performance, repeatability, intermediate precision, sensitivity and limit of detection. Specificity was determined using 87 true negative SNV sites (wild type) where another blend harboured a mutation at the corresponding position. The cell blends were processed and sequenced in two different runs by two independent users.

Overall performance

Four cell blends and five FFPE samples were used to measure performance across the capture design. The log mean depth across the panel was compared to the log depth of each region captured for each gene. Regions were classified as underperforming if the depth was lower than 2 x SD of the mean based on $\log_2$ [$\log_2$(ROI)>mean($\log_2$(ROI))-2xSD($\log_2$(ROI))]. GC content and mappability scores were compared against each region captured by the panel. Quality and coverage metrics were calculated across all the samples including i) total reads, ii) percentage of reads mapped to the reference sequence, iii) percentage of duplicates, iv) percentage of bases from unique reads de-duplicated on target, v) mean depth of targeted positions and vi) percentage of targeted positions with ≥50x, ≥100x and ≥250x coverage.

Limit of detection

To assess the limit of detection and determine a reliable cut off for the analysis we used the unique cancer-specific set of variants from the four cell blends introduced at range of VAFs from 4% to 30%, defined by ddPCR.

Precision

Repeatability (or within-run precision) was determined by comparing the cell blend background variant data across the 4 different samples in the same run for variant detection and VAF. Intra-run pairwise correlation was calculated for two runs where the cell blends were prepared and sequenced by different users generating two sets of repeatability data.

Intermediate precision (or between-run precision) was determined by comparing the cell blend background variant data between two runs for variant detection and VAF. Between-run pairwise correlation was calculated from two different runs prepared by different users and sequenced on different MiSeq instruments.

Sensitivity and specificity

The sensitivity of the panel was determined by separately comparing the cell blend background variants and the cancer-specific variants introduced at known VAF. Specificity was determined using the cell blend cancer-specific set of data with known variants and known true negative sites. Variants were classified according to the different ranges of frequencies of the variants present in the DNA blends. We also determined Positive-Predictive Value and Negative-Predictive Value.

Correlation between NGS targeted panel and other methodologies

13 paediatric cancer cell lines were tested harbouring a total of 30 known SNVs, deletions and amplifications previously identified by the Cancer Cell Line Encyclopaedia using Target Enrichment Sequencing (Agilent Technologies, Santa Clara, CA) and other

published data [45–50]. Furthermore 33 samples (FF=14, FFPE=19) had a total of 65 known genetic alterations including i) SNVs detected by Sanger Sequencing (*H3F3A, TP53, CTNNB1, HIST1H3B, ALK, BRAF*) [51–53] and RNA-Seq ii) copy number changes by FISH (*MYCN*) [54] and 450k array and iii) rearrangements by Real-Time Quantitative PCR involving *ESWR1* as previously described [55, 56] (Refer to Supplementary Materials).

Fresh frozen vs FFPE samples

15 paired FF and FFPE paediatric samples were compared for quality control metrics, coverage and the distribution of library inserts sizes between FFPE and FF paired samples. In addition, we correlated the VAF of the total variants found in the paired samples.

Abbreviations

FFPE (formalin-fixed paraffin embedded); FF (fresh frozen); SNVs (single nucleotide variants); Indels (insertion-deletions); DIPG (diffuse intrinsic pontine glioma); NGS (next-generation sequencing); kb (kilobase); bp (base pairs); ddPCR (droplet digital PCR); SD (standard deviation); CI (confidential interval); HMW (high molecular weight); VAF (variant allele frequency); TP (true positive); FN (false negative); PPV (positive-predictive Value); NPV (negative-predictive Value); WHO (world health organization); H&E (haematoxylin and eosin); BAM (binary alignment); vcf (variant call files); CNV (copy number variation); ROI (region of interest); IGV (integrative genomics viewer).

Author contributions

E Izquierdo designed and performed experiments, analysed-interpreted data, and wrote the manuscript; L Yuan analysed-interpreted the data and constructed analytical and visualisation tools; S George, L Chesler helped with the design, provided samples and gave advice; P Proszek performed experiments; C Jones, J Shipley, SA Gatz, L Marshall, C Stinson, AS Moore, SC Clifford, D Hicks, J Lindsey, R Hill, TS Jacques and J Chalker provided samples and/or gave advice; A Pearson, L Moreno and D Gonzalez de Castro conceived and designed the study; B Walker and D Gonzalez de Castro, supervised the research, interpreted data and reviewed the manuscript; D Gonzalez de Castro, K Thway and SO Connor carried out histopathological analysis; L Moreno reviewed the manuscript; M Hubank reviewed and edited the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We are enormously grateful to the Christopher's Smile charity (grant numbers CSM 002 and CSM 003) for their support and enthusiasm to provide children with more effective and less toxic targeted drugs through molecular profiling. We also thank the contribution of Children's Cancer and Leukaemia Group (CCLG).

CONFLICTS OF INTEREST

The authors declared no conflicts of interest with the submitted paper.

FINANCIAL SUPPORT

This work was supported by Christopher's Smile charity (CSM 002 and CSM 003) and the the NIHR Biomedical Research Centre at the Royal Marsden and the Institute of Cancer Research in London (Grant ref. A113).

REFERENCES

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. CA Cancer J Clin. 2012; 62:10-29.
2. London WB, Castel V, Monclair T, Ambros PF, Pearson AD, Cohn SL, Berthold F, Nakagawara A, Ladenstein RL, Iehara T, Matthay KK. Clinical and biologic features predictive of survival after relapse of neuroblastoma: a report from the International Neuroblastoma Risk Group project. J Clin Oncol. 2011; 29:3286-92.
3. Oberlin O, Rey A, Lyden E, Bisogno G, Stevens MC, Meyer WH, Carli M, Anderson JR. Prognostic factors in metastatic rhabdomyosarcomas: results of a pooled analysis from United States and European cooperative groups. J Clin Oncol. 2008; 26:2384-9.
4. Smith MA, Altekruse SF, Adamson PC, Reaman GH, Seibel NL. Declining childhood and adolescent cancer mortality. Cancer. 2014; 120:2497-506.
5. Middleton G, Crack LR, Popat S, Swanton C, Hollingsworth SJ, Buller R, Walker I, Carr TH, Wherton D, Billingham LJ. The National Lung Matrix Trial: translating the biology of stratification in advanced non-small-cell lung cancer. Ann Oncol. 2015; 26:2464-9.
6. Conley BA, Doroshow JH. Molecular analysis for therapy choice: NCI MATCH. Semin Oncol. 2014; 41:297-9.
7. Pugh TJ, Weeraratne SD, Archer TC, Pomeranz Krummel DA, Auclair D, Bochicchio J, Carneiro MO, Carter SL, Cibulskis K, Erlich RL, Greulich H, Lawrence MS, Lennon NJ, et al. Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. Nature. 2012; 488:106-10.
8. Shukla N, Ameur N, Yilmaz I, Nafa K, Lau CY, Marchetti A, Borsu L, Barr FG, Ladanyi M. Oncogene mutation profiling of pediatric solid tumors reveals significant subsets of embryonal rhabdomyosarcoma and neuroblastoma with mutated genes in growth signaling pathways. Clin Cancer Res. 2012; 18:748-57.

9. Jones DT, Hutter B, Jager N, Korshunov A, Kool M, Warnatz HJ, Zichner T, Lambert SR, Ryzhova M, Quang DA, Fontebasso AM, Stutz AM, Hutter S, et al. Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. Nat Genet. 2013; 45:927-32.
10. Pugh TJ, Morozova O, Attiyeh EF, Asgharzadeh S, Wei JS, Auclair D, Carter SL, Cibulskis K, Hanna M, Kiezun A, Kim J, Lawrence MS, Lichenstein L, et al. The genetic landscape of high-risk neuroblastoma. Nat Genet. 2013; 45:279-84.
11. Robinson G, Parker M, Kranenburg TA, Lu C, Chen X, Ding L, Phoenix TN, Hedlund E, Wei L, Zhu X, Chalhoub N, Baker SJ, Huether R, et al. Novel mutations target distinct subgroups of medulloblastoma. Nature. 2012; 488:43-8.
12. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Borresen-Dale AL, Boyault S, Burkhardt B, Butler AP, et al. Signatures of mutational processes in human cancer. Nature. 2013; 500:415-21.
13. Wu G, Diaz AK, Paugh BS, Rankin SL, Ju B, Li Y, Zhu X, Qu C, Chen X, Zhang J, Easton J, Edmonson M, Ma X, et al. The genomic landscape of diffuse intrinsic pontine glioma and pediatric non-brainstem high-grade glioma. Nat Genet. 2014; 46:444-50.
14. Hosoi H, Tanaka Y, Okita H, Miyachi M, Souzaki R, Taguchi T, Koh K, Hanada R, Kato K, Nomura Y, Akiyama M, Oka A, Igarashi T, et al. Recurrent somatic structural variations contribute to tumorigenesis in pediatric osteosarcoma. Nat Commun. 2014; 7:104-12.
15. Huether R, Dong L, Chen X, Wu G, Parker M, Wei L, Ma J, Edmonson MN, Hedlund EK, Rusch MC, Shurtleff SA, Mulder HL, Boggs K, et al. The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. Nat Commun. 2014; 5:3630.
16. Shern JF, Chen L, Chmielecki J, Wei JS, Patidar R, Rosenberg M, Ambrogio L, Auclair D, Wang J, Song YK, Tolman C, Hurd L, Liao H, et al. Comprehensive genomic analysis of rhabdomyosarcoma reveals a landscape of alterations affecting a common genetic axis in fusion-positive and fusion-negative tumors. Cancer Discov. 2014; 4:216-31.
17. Seki M, Nishimura R, Yoshida K, Shimamura T, Shiraishi Y, Sato Y, Kato M, Chiba K, Tanaka H, Hoshino N, Nagae G, Shiozawa Y, Okuno Y. Integrated genetic and epigenetic analysis defines novel molecular subgroups in rhabdomyosarcoma. Nat Commun. 2015; 6:7557.
18. Castel D, Philippe C, Calmon R, Le Dret L, Truffaux N, Boddaert N, Pages M, Taylor KR, Saulnier P, Lacroix L, Mackay A, Jones C, Sainte-Rose C, et al. Histone H3F3A and HIST1H3B K27M mutations define two subgroups of diffuse intrinsic pontine gliomas with different prognosis and phenotypes. Acta Neuropathol. 2015; 130:815-27.
19. Schwartzentruber J, Korshunov A, Liu XY, Jones DT, Pfaff E, Jacob K, Sturm D, Fontebasso AM, Quang DA, Tonjes M, Hovestadt V, Albrecht S, Kool M, et al. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature. 2012; 482:226-31.
20. Wu G, Broniscer A, McEachron TA, Lu C, Paugh BS, Becksfort J, Qu C, Ding L, Huether R, Parker M, Zhang J, Gajjar A, Dyer MA, et al. Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. Nat Genet. 2012; 44:251-3.
21. Peifer M, Hertwig F, Roels F, Dreidax D, Gartlgruber M, Menon R, Kramer A, Roncaioli JL, Sand F, Heuckmann JM, Ikram F, Schmidt R, Ackermann S, et al. Telomerase activation by genomic rearrangements in high-risk neuroblastoma. Nature. 2015; 526:700-4.
22. Valentijn LJ, Koster J, Zwijnenburg DA, Hasselt NE, van Sluis P, Volckmann R, van Noesel MM, George RE, Tytgat GA, Molenaar JJ, Versteeg R. TERT rearrangements are frequent in neuroblastoma and identify aggressive tumors. Nat Genet. 2015; 47:1411-4.
23. Hertwig F, Peifer M, Fischer M. Telomere maintenance is pivotal for high-risk neuroblastoma. Cell Cycle. 2016; 15:311-2.
24. Jones DT, Jager N, Kool M, Zichner T, Hutter B, Sultan M, Cho YJ, Pugh TJ, Hovestadt V, Stutz AM, Rausch T, Warnatz HJ, Ryzhova M, et al. Dissecting the genomic complexity underlying medulloblastoma. Nature. 2012; 488:100-5.
25. Salto-Tellez M, Gonzalez de Castro D. Next-generation sequencing: a change of paradigm in molecular diagnostic validation. J Pathol. 2014; 234:5-10.
26. Hagemann IS, Cottrell CE, Lockwood CM. Design of targeted, capture-based, next generation sequencing tests for precision cancer therapy. Cancer Genet. 2013; 206:420-31.
27. Cheng DT, Mitchell TN, Zehir A, Shah RH, Benayed R, Syed A, Chandramohan R, Liu ZY, Won HH, Scott SN, Brannon AR, O'Reilly C, Sadowska J, et al. Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): a hybridization capture-based next-generation sequencing clinical assay for solid tumor molecular oncology. J Mol Diagn. 2015; 17:251-64.
28. Frampton GM, Fichtenholtz A, Otto GA, Wang K, Downing SR, He J, Schnall-Levin M, White J, Sanford EM, An P, Sun J, Juhn F, Brennan K, et al. Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. Nat Biotechnol. 2013; 31:1023-31.
29. Cottrell CE, Al-Kateb H, Bredemeyer AJ, Duncavage EJ, Spencer DH, Abel HJ, Lockwood CM, Hagemann IS, O'Guin SM, Burcea LC, Sawyer CS, Oschwald DM, Stratman JL, et al. Validation of a next-generation sequencing assay for clinical molecular oncology. J Mol Diagn. 2014; 16:89-105.
30. Pritchard CC, Salipante SJ, Koehler K, Smith C, Scroggins S, Wood B, Wu D, Lee MK, Dintzis S, Adey A, Liu Y, Eaton KD, Martins R, et al. Validation and implementation

of targeted capture and sequencing for the detection of actionable mutation, copy number variation, and gene rearrangement in clinical cancer specimens. The J Mol Diagn. 2014; 16:56-67.

31. Mody RJ, Wu YM, Lonigro RJ, Cao X, Roychowdhury S, Vats P, Frank KM, Prensner JR, Asangani I, Palanisamy N, Dillman JR, Rabah RM, Kunju LP, et al. Integrative clinical sequencing in the management of refractory or relapsed cancer in youth. JAMA. 2015; 314:913-25.
32. Worst BC, van Tilburg CM, Balasubramanian GP, Fiesel P, Witt R, Freitag A, Boudalil M, Previti C, Wolf S, Schmidt S, Chotewutmontri S, Bewerunge-Hudler M, Schick M, et al. Next-generation personalised medicine for high-risk paediatric cancer patients - the INFORM pilot study. Eur J Cancer. 2016; 65:91-101.
33. Harris MH, DuBois SG, Glade Bender JL, Kim A, Crompton BD, Parker E, Dumont IP, Hong AL, Guo D, Church A, Stegmaier K, Roberts CW, Shusterman S, et al. Multicenter feasibility study of tumor molecular profiling to inform therapeutic decisions in advanced pediatric solid tumors: the Individualized Cancer Therapy (iCat) study. JAMA Oncol. 2016.
34. Allen CE, Laetsch TW, Mody R, Irwin MS, Lim MS, Adamson PC, Seibel NL, Parsons DW, Cho YJ, Janeway K. Target and agent prioritization for the Children's Oncology Group-National Cancer Institute pediatric MATCH trial. J Natl Cancer Inst. 2017.
35. Mattocks CJ, Morris MA, Matthijs G, Swinnen E, Corveleyn A, Dequeker E, Muller CR, Pratt V, Wallace A. A standardized framework for the validation and verification of clinical molecular genetic tests. Eur J Hum Genet. 2010; 18:1276-88.
36. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, Harris PL, Haserlat SM, Supko JG, Haluska FG, Louis DN, Christiani DC, Settleman J, Haber DA. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med. 2004; 350:2129-39.
37. Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, Larkin J, Dummer R, Garbe C, Testori A, Maio M, Hogg D, Lorigan P, Lebbe C, et al. Improved survival with vemurafenib in melanoma with BRAF V600E mutation. N Engl J Med. 2011; 364:2507-16.
38. Kwak EL, Bang YJ, Camidge DR, Shaw AT, Solomon B, Maki RG, Ou SH, Dezube BJ, Janne PA, Costa DB, Varella-Garcia M, Kim WH, Lynch TJ, et al. Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. N Engl J Med. 2010; 363:1693-703.
39. Vassal G, Rousseau R, Blanc P, Moreno L, Bode G, Schwoch S, Schrappe M, Skolnik J, Bergman L, Bradley-Garelik MB, Saha V, Pearson A, Zwierzina H. Creating a unique, multi-stakeholder Paediatric Oncology Platform to improve drug development for children and adolescents with cancer. Eur J Cancer. 2015; 51:218-24.
40. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, Ohgaki H, Wiestler OD, Kleihues P, Ellison DW. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta Neuropathol. 2016; 131:803-20.
41. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. 2015; 17:405-24.
42. Nikbakht H, Panditharatna E, Mikael LG, Li R, Gayden T, Osmond M, Ho CY, Kambhampati M, Hwang EI, Faury D, Siu A, Papillon-Cavanagh S, Bechet D, et al. Spatial and temporal homogeneity of driver mutations in diffuse intrinsic pontine glioma. Nat Commun. 2016; 7:11185.
43. Johnson BE, Mazor T, Hong C, Barnes M, Aihara K, McLean CY, Fouse SD, Yamamoto S, Ueda H, Tatsuno K, Asthana S, Jalbert LE, Nelson SJ, et al. Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. Science. 2014; 343:189-93.
44. Appel M, Walker B, Jones T, Izquierdo Delgado E, Raterman D, Rosenbaum H, Burgess D. KAPA HyperPlus/SeqCap EZ workflow: Improving Data Quality and Turnaround Times for Targeted Next-Generation Sequencing of FFPE DNA. Kapa Biosystems. 2016.
45. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, Wilson CJ, Lehar J, Kryukov GV, Sonkin D, Reddy A, Liu M, Murray L, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012; 483:603-7.
46. Hinson AR, Jones R, Crose LE, Belyea BC, Barr FG, Linardic CM. Human rhabdomyosarcoma cell lines for rhabdomyosarcoma research: utility and pitfalls. Front Oncol. 2013; 3:183.
47. Chen Y, Takita J, Choi YL, Kato M, Ohira M, Sanada M, Wang L, Soda M, Kikuchi A, Igarashi T, Nakagawara A, Hayashi Y, Mano H, Ogawa S. Oncogenic mutations of ALK kinase in neuroblastoma. Nature. 2008; 455:971-4.
48. George RE, Sanda T, Hanna M, Frohling S, Luther W 2nd, Zhang J, Ahn Y, Zhou W, London WB, McGrady P, Xue L, Zozulya S, Gregor VE, et al. Activating mutations in ALK provide a therapeutic target in neuroblastoma. Nature. 2008; 455:975-8.
49. McDermott U, Iafrate AJ, Gray NS, Shioda T, Classon M, Maheswaran S, Zhou W, Choi HG, Smith SL, Dowell L, Ulkus LE, Kuhlmann G, Greninger P, et al. Genomic alterations of anaplastic lymphoma kinase may sensitize tumors to anaplastic lymphoma kinase inhibitors. Cancer Res. 2008; 68:3389-95.
50. Eleveld TF, Oldridge DA, Bernard V, Koster J, Daage LC, Diskin SJ, Schild L, Bentahar NB, Bellini A, Chicard M, Lapouble E, Combaret V, Legoix-Ne P, et al. Relapsed

neuroblastomas show frequent RAS-MAPK pathway mutations. Nat Genet. 2015; 47:864-71.

51. Ellison DW, Onilude OE, Lindsey JC, Lusher ME, Weston CL, Taylor RE, Pearson AD, Clifford SC. beta-Catenin status predicts a favorable outcome in childhood medulloblastoma: the United Kingdom Children's Cancer Study Group Brain Tumour Committee. J Clin Oncol. 2005; 23:7951-7.

52. Hill RM, Kuijper S, Lindsey JC, Petrie K, Schwalbe EC, Barker K, Boult JK, Williamson D, Ahmad Z, Hallsworth A, Ryan SL, Poon E, Robinson SP, et al. Combined MYC and P53 defects emerge at medulloblastoma relapse and define rapidly progressive, therapeutically targetable disease. Cancer Cell. 2015; 27:72-84.

53. Bjerke L, Mackay A, Nandhabalan M, Burford A, Jury A, Popov S, Bax DA, Carvalho D, Taylor KR, Vinci M, Bajrami I, McGonnell IM, Lord CJ, et al. Histone H3.3. mutations drive pediatric glioblastoma through upregulation of MYCN. Cancer Discov. 2013; 3:512-9.

54. Lamont JM, McManamy CS, Pearson AD, Clifford SC, Ellison DW. Combined histopathological and molecular cytogenetic stratification of medulloblastoma patients. Clin Cancer Res. 2004; 10:5482-93.

55. Thway K, Wren D, Lee J, Thompson L, Fisher C, Gonzalez D. Evaluation of the optimal provision of formalin-fixed, paraffin-embedded material for reverse transcription-PCR in soft-tissue tumour diagnosis. J Clin Pathol. 2017; 70:20-4.

56. Thway K, Rockcliffe S, Gonzalez D, Swansbury J, Min T, Thompson L, Fisher C. Utility of sarcoma-specific fusion gene analysis in paraffin-embedded material for routine diagnosis at a specialist centre. J Clin Pathol. 2010; 63:508-12.

European Journal of Cancer 121 (2019) 224–235

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Original Research

A tailored molecular profiling programme for children with cancer to identify clinically actionable genetic alterations

Sally L. George [a,b,*,1], Elisa Izquierdo [c,d,1], James Campbell [e], Eleni Koutroumanidou [c], Paula Proszek [c], Sabri Jamal [c], Deborah Hughes [c], Lina Yuan [c], Lynley V. Marshall [a,b], Fernando Carceller [a,b], Julia C. Chisholm [a,b], Sucheta Vaidya [a,b], Henry Mandeville [b], Paola Angelini [b], Ajla Wasti [b], Tomas Bexelius [b], Khin Thway [f], Susanne A. Gatz [b,g,h], Matthew Clarke [d], Bissan Al-Lazikani [e], Giuseppe Barone [i], John Anderson [i,j], Deborah A. Tweddle [k], David Gonzalez [c,l], Brian A. Walker [c,m], Jack Barton [j], Sarita Depani [i], Jessica Eze [i,n], Saira W. Ahmed [i,n], Lucas Moreno [b,o,p], Andrew Pearson [b], Janet Shipley [g], Chris Jones [d], Darren Hargrave [h,i], Thomas S. Jacques [i,n], Michael Hubank [c], Louis Chesler [a,b]

[a] *Paediatric Tumour Biology, Division of Clinical Studies, The Institute of Cancer Research, London, UK*
[b] *Children and Young People's Unit, Royal Marsden NHS Foundation Trust, London, UK*
[c] *Molecular Diagnostics Department, The Institute of Cancer Research and Clinical Genomics, The Royal Marsden NHS Foundation, London, UK*
[d] *Glioma Team, Division of Molecular Pathology and Cancer Therapeutics, The Institute of Cancer Research, London, UK*
[e] *Bioinformatics Core Facility, The Institute of Cancer Research, London, UK*
[f] *Pathology Department, Royal Marsden NHS Foundation Trust, London, UK*
[g] *Sarcoma Molecular Pathology Team, Divisions of Molecular Pathology and Cancer Therapeutics, The Institute of Cancer Research, London, UK*
[h] *Cancer Research UK Clinical Trials Unit, Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, UK*
[i] *Department of Haematology and Oncology, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK*
[j] *Developmental Biology and Cancer Programme, UCL GOS Institute of Child Health, London, UK*
[k] *Northern Institute for Cancer Research, Newcastle University, Newcastle, UK*
[l] *Centre for Cancer Research and Cell Biology, Queens University Belfast, Belfast, UK*
[m] *Myeloma Center, University of Arkansas for Medical Sciences, Little Rock, AR, USA*
[n] *Department of Histology, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK*

* *Corresponding author*. Paediatric Tumour Biology, Division of Clinical Studies, The Institute of Cancer Research, London, UK.
E-mail address: sally.george@icr.ac.uk (S.L. George).
[1] These authors contributed equally.

https://doi.org/10.1016/j.ejca.2019.07.027

[o] HNJ-CNIO Clinical Research Unit, Hospital Universitario Nino Jesus, Madrid, Spain
[p] Paediatric Oncology & Haematology, Vall d'Hebron University Hospital, Barcelona, Spain

Received 18 April 2019; received in revised form 27 June 2019; accepted 23 July 2019
Available online 19 September 2019

KEYWORDS
Paediatric oncology;
Clinical targeted sequencing;
Personalised medicine;
Circulating tumour DNA

Abstract ***Background:*** For children with cancer, the clinical integration of precision medicine to enable predictive biomarker–based therapeutic stratification is urgently needed.
Methods: We have developed a hybrid-capture next-generation sequencing (NGS) panel, specifically designed to detect genetic alterations in paediatric solid tumours, which gives reliable results from as little as 50 ng of DNA extracted from formalin-fixed paraffin-embedded (FFPE) tissue. In this study, we offered an NGS panel, with clinical reporting via a molecular tumour board for children with solid tumours. Furthermore, for a cohort of 12 patients, we used a circulating tumour DNA (ctDNA)–specific panel to sequence ctDNA from matched plasma samples and compared plasma and tumour findings.
Results: A total of 255 samples were submitted from 223 patients for the NGS panel. Using FFPE tissue, 82% of all submitted samples passed quality control for clinical reporting. At least one genetic alteration was detected in 70% of sequenced samples. The overall detection rate of clinically actionable alterations, defined by modified OncoKB criteria, for all sequenced samples was 51%. A total of 8 patients were sequenced at different stages of treatment. In 6 of these, there were differences in the genetic alterations detected between time points. Sequencing of matched ctDNA in a cohort of extracranial paediatric solid tumours also identified a high detection rate of somatic alterations in plasma.
Conclusion: We demonstrate that tailored clinical molecular profiling of both tumour DNA and plasma-derived ctDNA is feasible for children with solid tumours. Furthermore, we show that a targeted NGS panel–based approach can identify actionable genetic alterations in a high proportion of patients.

1. Introduction

In adult malignancies, precision medicine initiatives enabling standardised, high-throughput molecular profiling and predictive biomarker–based stratification have been implemented to maximise clinical efficacy of targeted therapeutics [1–7]. Similar initiatives are urgently needed for childhood cancer, which remains the primary cause of death in children after infancy [8].

In children, comprehensive molecular profiling programmes have incorporated whole-exome sequencing (WES) and RNA sequencing (RNA-seq) and, in some cases, copy number analysis, whole-genome sequencing (WGS), microarray or methylation arrays. Such initiatives have detected potentially actionable findings in 46–60.9% of patients [9–11]. However, logistical and financial practicalities limit large-scale implementation of this approach in most health-care settings. Targeted next-generation sequencing (NGS) panels are typically more cost-effective and can be tailored to the study population and standardised according to regulatory requirements. Therefore, this may present a more suitable alternative for implementation into health-care systems.

Generic adult cancer gene panels have been used in children [12,13]; however, the spectrum of mutations differs between adult and paediatric tumours. For example, recurrent H3 mutations are a hallmark of paediatric high-grade glioma [14,15], and rearrangements upstream to the *TERT* promoter are frequent in neuroblastoma [16]. These differences necessitate a tailored approach to determine common and actionable events; hence, we have developed and clinically validated a paediatric-specific solid tumour NGS panel for use in precision medicine [17].

In children with relapsed/refractory cancer, access to adequate biopsy material remains challenging [18,19]. Therefore, our strategy has been to optimise the paediatric panel for use on formalin-fixed paraffin-embedded (FFPE) tissue if frozen tissue is unavailable and, in parallel, begin evaluating more-easily accessible sources of tumour DNA, such as plasma.

Plasma-derived circulating tumour DNA (ctDNA) has been shown to be an alternative to repeat biopsy in common adult malignancies [20–23]. ctDNA analysis is minimally invasive, amenable to serial sampling and may also give more representative information

regarding tumour heterogeneity [24,25]. Limited studies in children with cancer have detected somatic mutations in small volumes of plasma [26–30].

Here, we report the development of version 2 of our paediatric solid tumour–specific NGS panel and the national implementation of clinical NGS panel sequencing. We report on assay performance and the clinical relevance of the findings. In parallel, we evaluate the feasibility of performing targeted sequencing of ctDNA in a clinical laboratory setting using a ctDNA-specific NGS panel.

2. Materials and methods

2.1. Patients

A Royal Marsden Hospital (RM) pilot study for patients aged ≤24 years with solid tumours treated at our Children and Young People's Unit commenced in March 2016 and was subsequently expanded nationally for children aged ≤16 years. Ethical approval was obtained from the National Research Ethics Service (reference: 15/LO/07) and the Biological Studies Steering Group of the Children's Cancer and Leukaemia Group (reference: 2015 BS 09). Participants and/or guardians gave informed consent. Patients were eligible to enrol at any time including diagnosis and relapse/progression. Blood was taken for germline DNA analysis, and archival tissue was retrieved from the most recent surgery, or if indicated, a repeat biopsy could be requested at the treating clinician's discretion.

2.2. Sample preparation and sequencing

Sample preparation, DNA extraction, library preparation and sequencing were performed according to established protocols [17,31]. Two different panels were used: version 1 (v1, 78 genes, 311 kb) and version 2 (v2, 91 genes, 473 kb) (Table S1). The custom hybridisation panel is capable of detecting single-nucleotide variants (SNVs), small insertions and deletions (indels), copy number variations (CNVs) and structural variants for which we capture the region where the breakpoint occurs, for instance, 50 kb upstream to the *TERT* promoter [16]. Sequencing output files were processed as previously reported [31]. Only somatic variants, detected after subtraction of germline findings, were reported.

Samples were analysed initially using MiSeq Reporter version 2.5 (http://emea.support.illumina.com/sequencing/sequencing_software/miseq_reporter/downloads.html). Analysis was later executed using an in-house developed pipeline Molecular Diagnostic Information Management System version 3.0 (MDIMSv3) using the following bioinformatic software and versions: demultiplexing was performed using bcl2fastq 2.17.1.14, reads were aligned using BWA *0.7.12,* structural variants were identified using Manta 0.29.6, SNVs and indels were called with GATK 3.5.0 and variants were annotated with Oncotator version 1.5.1.0. CNVs were assessed as previously described [17].

2.3. Gene panel capture version 2, design and validation

Integral to the study design was the ability to update and adapt the regions included on the panel according to clinical need and target prioritisation. For v2, genes were ranked by consensus expert opinion according to set selection criteria (Table S1). The panel was validated using four cell blends (Tru-Q1-4 Horizon Discovery, Cambridge, United Kingdom [UK]) and 10 FFPE samples with known variants (SNVs = 554, indels = 79). Quality and coverage metrics were calculated across all the samples including (i) total reads, (ii) percentage of reads mapped to the reference sequence, (iii) percentage of duplicates, (iv) percentage of bases from unique reads deduplicated on target and (v) mean depth. Sensitivity, specificity and accuracy were determined by comparing the cell blends and FFPE samples with known variants and known true negatives.

2.4. Molecular tumour board

A monthly molecular tumour board (MTB) was established for discussion of findings, and the interpreted results were then reported to the treating clinician. The MTB core members included paediatric/adolescent oncologists, experts in early clinical trials, molecular pathologists, bioinformaticians and paediatric tumour biologists, from the RM, Great Ormond Street Hospital and The Institute of Cancer Research, London. OncoKB was used as a basis to define tiers of actionability [32]. In addition, COSMIC [33]-defined mutations/SNVs, genetic amplifications, gains or losses, for which a paediatric clinical trial was currently recruiting, were also considered, as well as alterations where compelling preclinical paediatric data existed for that target (Table S1). Heterozygous gene loss and missense mutations outside of defined hotspot regions were defined as not actionable.

2.5. ctDNA extraction and analysis

A total of 12 plasma samples were identified for sequencing where the corresponding tumour samples contained at least one genetic alteration present on the ctDNA panel. The plasma ctDNA sequencing results were not reported back to the MTB.

About 5 to 10 mL of blood was collected into cell-free DNA blood collection tubes (Streck, La Vista, United States of America) and centrifuged twice at 1600 g. ctDNA extraction and sequencing using a commercially available hybrid-capture panel (Avenio ctDNA expanded kit, Roche) was performed according to the manufacturer's instructions.

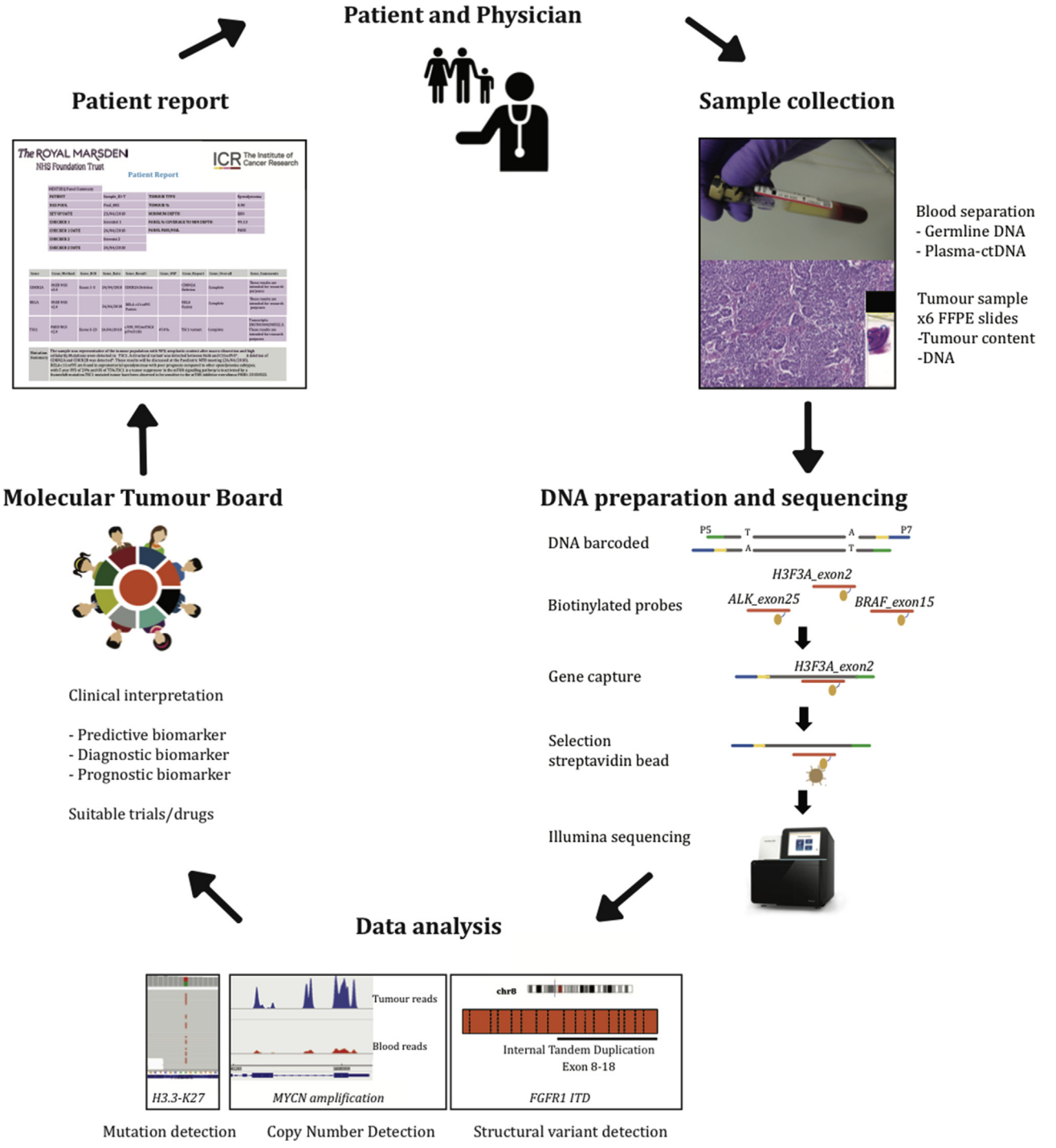


Fig. 1. Study overview. After obtaining informed consent, tumour and blood samples were collected. DNA was extracted, and sequence libraries were prepared using the capture-based paediatric solid tumour panel. After sequencing, samples underwent an in-house data analysis pipeline that detects mutations, structural variants and copy number changes. Genomic alterations were manually reviewed by two independent scientists and then discussed in a molecular tumour board before a clinical report was issued. FFPE, formalin-fixed paraffin-embedded.

3. Results

3.1. Version 2 of the paediatric solid tumour panel

v1 of the panel was validated as previously reported [17]. v2 was also validated to Good Laboratory and Clinical Practice standards and performed well, comparable with v1, obtaining a similar number of reads and percentage of unique on-target reads (Figure S1A-C). The polymerase chain reaction (PCR) duplicate percentage was improved (v1 = 55.3% and v2 = 20.3%) (Figure S1D). The sensitivity for detection of SNVs was ≥99% and ≥90% for indels at ≥5% variant allele frequency (VAF) (Table S2). The specificity for SNVs was ≥98% at ≥5%

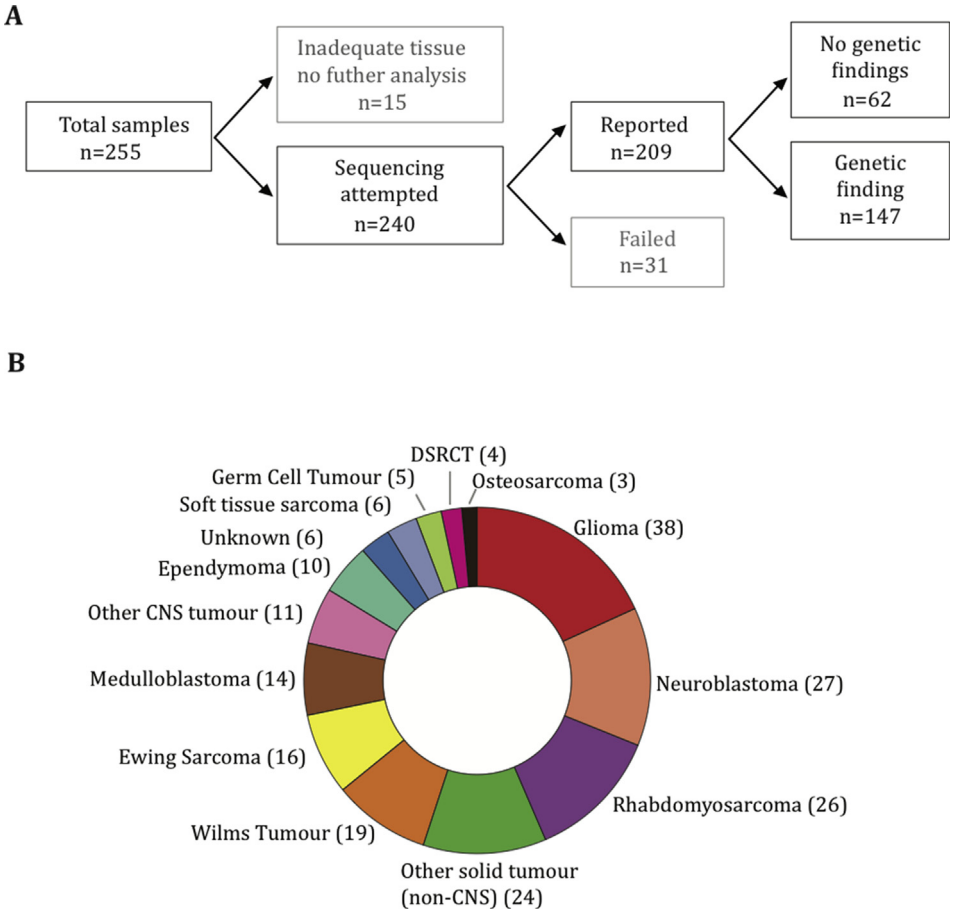


Fig. 2. Tumour samples submitted for sequencing. Summary of sample flow and the total number of samples successfully sequenced (A). Distribution of tumour types among reported cases (B). DSRCT, desmoplastic small round cell tumour; CNS, central nervous system.

allele frequency. The correlation (r^2) of VAF for SNVs and indels between droplet digital polymerase chain reaction (ddPCR) and v2 was 0.9527 (Figure S1E) and between v1 and v2 was 0.9301 (Figure S1F).

3.2. Patient samples and overall performance

An overview of the study is given in Fig. 1. A total of 255 samples were submitted from 223 patients. Although patients were eligible to enrol at any time, 90% of evaluable patients had at least one episode of progression/relapse before study enrolment. FFPE tissue from the most recent surgery was requested for all but 3 patients where fresh frozen tissue was used.

Adequate coverage for clinical reporting of results was obtained in 82% of submitted samples (Fig. 2A). Reasons for sample rejection or failure were as follows: tumour content less than 10%, DNA less than 20 ng and/or excessive DNA fragmentation. The median depth of coverage for all reported samples was 495 (interquartile range: 264–868). The most common cancers sequenced were glioma (38), neuroblastoma (27) and rhabdomyosarcoma (26) (Fig. 2B).

3.3. Genetic findings

At least one genetic alteration was detected in 70% (145/209) of samples at an allele frequency $\geq$ 5%. The somatic genetic alterations detected, grouped according to underlying diagnosis, are summarised in Fig. 3, Table S3 and Fig S2. In keeping with other studies [34], the most frequently mutated gene was *TP53* in 36/209 (17%); in addition high frequencies of alterations in genes known to be recurrently altered in paediatric malignancies such as *ATRX*, *CDKN2A*, *CTNNB1* in 12/209 (5.7%), *MYCN* in 11/209 (5.2%) and *H3F3A, PIK3CA* in 10/209 (4.3%) were detected.

3.4. Clinical actionability

Potentially targetable alterations, defined by OncoKB tiers of actionability in addition to predictive biomarkers for currently recruiting paediatric clinical trials, were detected in 51% of sequenced samples (Fig. 4A). Of the 107 tumour samples classified as potentially actionable, 42 (39%) had greater than one actionable alteration detected. For each tumour sample, only the alteration for which there was the highest tier of evidence for actionability was included.

Glioma was the tumour type with more defined actionable alterations found, followed by osteosarcoma and rhabdomyosarcoma (Fig. 4B). No tier 1 alterations (US Food and Drug Administration [FDA]–recognised biomarker predictive of response to an FDA-approved drug) were detected, indicative of the lack of regulatory approvals for paediatric indications. Only one patient had a tier 2A alteration: a patient with an inflammatory myofibroblastic tumour, harbouring an *ALK:SQSTM1* translocation. The patient had a complete surgical resection and did not require systemic therapy.

As a feasibility study, follow-up data were not routinely collected for all patients. Of the 57 patients with a tier 2B or 3 alteration and available follow-up data, only four (7%) received targeted therapies:

Three patients with $BRAF_{V600E}$ mutations were treated with dabrafenib/trametinib combination therapy: patient 1 had a pleomorphic xanthoastrocytoma and was commenced on dabrafenib/trametinib after third disease progression. The patient remains on treatment with stable disease after 9 months. Patient 2 had glioblastoma multiforme and was commenced on dabrafenib/trametinib after disease progression. The patient had stable disease for 13 months before further progression. Patient 3 had multiply relapsed metastatic ameloblastic fibroodontosarcoma [35]; by day 28 of treatment, there had been a partial response but asymptomatic cardiac toxicity, required discontinuation of both drugs. On normalisation of the shortening and ejection fractions, the patient was recommenced on single-agent dabrafenib and had sustained partial response for 15 further months. A patient with multiply relapsed metastatic germinoma and *PDGFRA*/*KIT* amplification was given dasatinib, but progressed on treatment.

One patient with high-grade glioma (patient ID 045-T) had a total of 49 somatic mutations (Table S3) (in ~0.18 Mb) consistent with a hypermutator phenotype, associated with mismatch repair deficiency and predictive of potential sensitivity to immune checkpoint blockade [36]. However, the patient was not fit for clinical trial enrolment by the time the sequencing results were available.

Other patients had findings that informed prognosis: a mutation in *CTNNB1* was found in a patient originally diagnosed with supratentorial primitive neuroectodermal tumour (PNET), biologically more in keeping with a WNT-activated medulloblastoma. Other examples included an *MYOD1* mutation in a patient with embryonal rhabdomyosarcoma, associated with distinct clinical features and poor prognosis [37], and a *RELA-c11orf95* fusion in a patient with supratentorial ependymoma, associated with high-risk disease [38].

3.5. Analysis of paired samples

For eight patients, paired samples were sequenced at different stages of treatment (Fig. 5). In six of these, there were differences between the variants detected at different time points. Mutations in *PTEN, NF1* and *TP53* were observed in a patient with high-grade glioma (patient 2) after dabrafenib/trametinib treatment but not in the pre-treatment sample. The patient subsequently received everolimus but progressed after 3 months on treatment. The acquisition of *NF1* mutations as a resistance mechanism after BRAF inhibition is consistent with findings in $BRAF_{V600E}$-mutant melanoma [39,40]. in another child with glioma sequenced at diagnosis and progression, the tumour harboured shared alterations in *H3F3A* and *TP53*, whereas *PTEN* was only present at diagnosis and *PIK3CA* at progression. In a patient with Wilms tumour, a potentially targetable *TSC2* mutation was found in the 3rd relapse sample, which was not present in the previous sample.

3.6. ctDNA analysis

ctDNA was sequenced in a cohort of 12 patients with extracranial tumours, in whom the tumour panel had detected a genetic alteration that was also covered by a commercially available ctDNA sequencing panel. In 3 patients, in whom ctDNA and FFPE were sequenced from the same time point, there was a direct concordance between findings. However, in 5 patients, from whom plasma was collected after at least one subsequent relapse, variants were detected in the plasma that were not detected in FFPE samples (Table 1). For example, in a patient with neuroblastoma, an *ALK* F1174L mutation was detected in both tumour and plasma; however, an additional *ALK* hotspot mutation was also detected in the plasma that was not present in the tumour sample. In addition, of note, in 2 cases, variants detected in plasma at relapse were only identified at very low levels in diagnostic tumour samples, below the predefined limit of detection for clinical reporting.

4. Discussion

Comprehensive molecular profiling strategies have been shown to be feasible in children with cancer [9–11] and show encouraging results. However, wide-scale implementation is impractical in most health-care settings, and even if resources were unlimited, it is also restricted by the availability of biopsy material. We show that using as little as 50 ng of DNA, this assay is an accurate, reproducible and practical platform for molecular stratification and identification of actionable targets, required to accelerate precision medicine clinical trials in childhood tumours.

We are aware that although capture-based panel sequencing is an excellent tool, it has limitations. With our targeted panel approach, only a small portion of the genome is sequenced, and therefore, it is not always possible to distinguish between focal gains or deletions

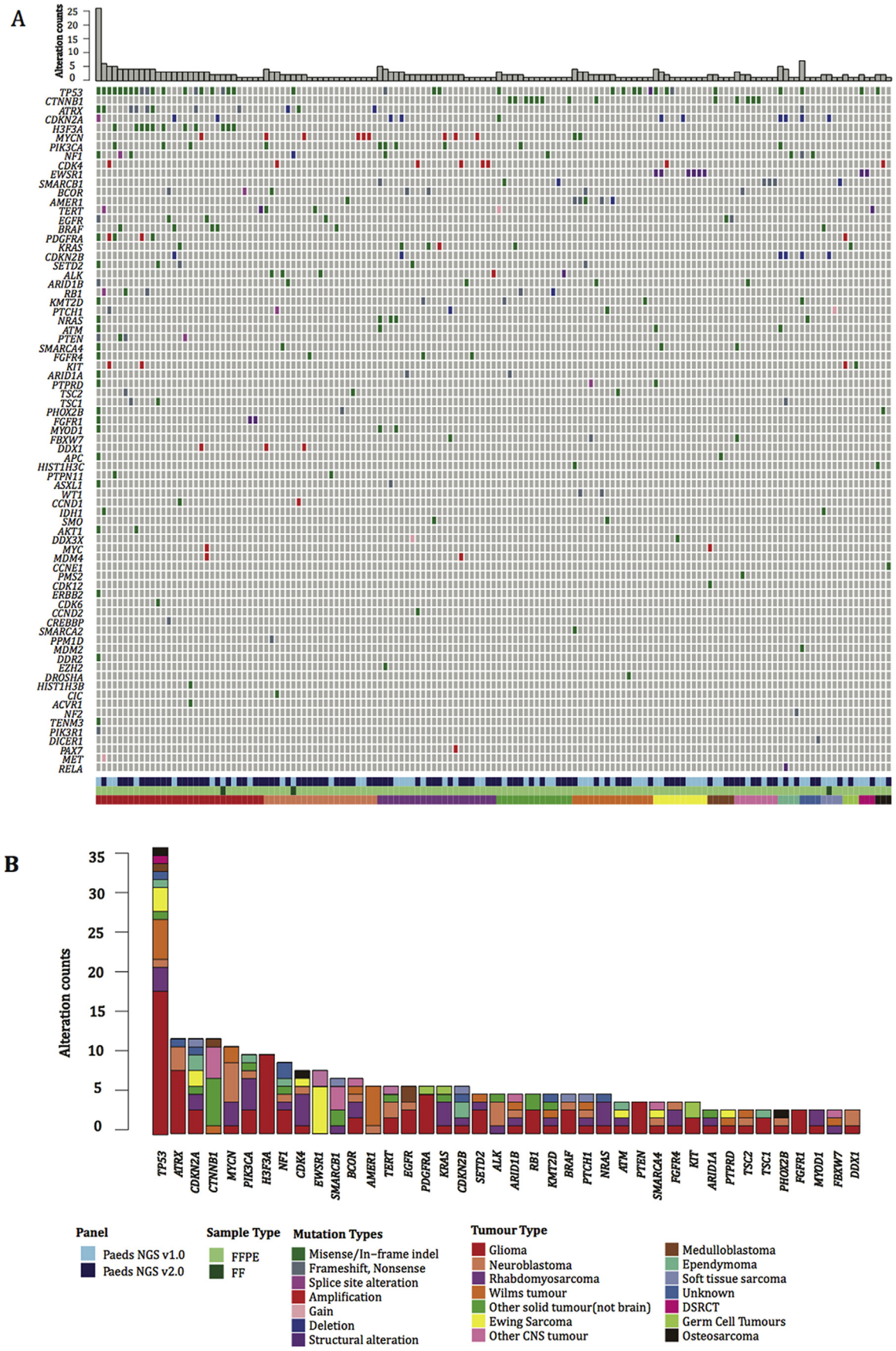

A
Alteration counts
25
20
15
10
5
0
TP53
CTNNB1
ATRX
CDKN2A
H3F3A
MYCN
PIK3CA
NF1
CDK4
EWSR1
SMARCB1
BCOR
AMER1
TERT
EGFR
BRAF
PDGFRA
KRAS
CDKN2B
SETD2
ALK
ARID1B
RB1
KMT2D
PTCH1
NRAS
ATM
PTEN
SMARCA4
FGFR4
KIT
ARID1A
PTPRD
TSC2
TSC1
PHOX2B
FGFR1
MYOD1
FBXW7
DDX1
APC
HIST1H3C
PTPN11
ASXL1
WT1
CCND1
IDH1
SMO
AKT1
DDX3X
MYC
MDM4
CCNE1
PMS2
CDK12
ERBB2
CDK6
CCND2
CREBBP
SMARCA2
PPM1D
MDM2
DDR2
EZH2
DROSHA
HIST1H3B
CIC
ACVR1
NF2
TENM3
PIK3R1
DICER1
PAX7
MET
RELA
B
Alteration counts
35
30
25
20
15
10
5
0
TP53
ATRX
CDKN2A
CTNNB1
MYCN
PIK3CA
H3F3A
NF1
CDK4
EWSR1
SMARCB1
BCOR
AMER1
TERT
EGFR
PDGFRA
KRAS
CDKN2B
SETD2
ALK
ARID1B
RB1
KMT2D
BRAF
PTCH1
NRAS
ATM
PTEN
SMARCA4
FGFR4
KIT
ARID1A
PTPRD
TSC2
TSC1
PHOX2B
FGFR1
MYOD1
FBXW7
DDX1
Panel
Paeds NGS v1.0
Paeds NGS v2.0
Sample Type
FFPE
FF
Mutation Types
Misense/In-frame indel
Frameshift, Nonsense
Splice site alteration
Amplification
Gain
Deletion
Structural alteration
Tumour Type
Glioma
Neuroblastoma
Rhabdomyosarcoma
Wilms tumour
Other solid tumour(not brain)
Ewing Sarcoma
Other CNS tumour
Medulloblastoma
Ependymoma
Soft tissue sarcoma
Unknown
DSRCT
Germ Cell Tumours
Osteosarcoma

Table 1
Results of ctDNA panel sequencing of matched plasma samples and comparison with tumour panel sequencing for genes covered by both panels, ordered by the time elapsed between samples.

Diagnosis	Days between samples	Treatment position with FFPE sample	Treatment position with blood sample	Isolated ctDNA (ng)	Gene	Amino acid change	AF FFPE DNA	AF ctDNA	Sequencing depth ctDNA	Sequencing depth tumour
Neuroblastoma	5	5th relapse	5th relapse	18.54	*TP53*	C135F	74.0%	20.30%	13348	393
Wilms tumour	19	Post induction	Post induction	32.22	*TP53*	G245D	77.0%	7.44%	5498	402
Ewing sarcoma	84	2nd relapse	2nd relapse	50	*TP53*	C176Y	87.0%	49.90%	3453	70
Neuroblastoma	214	Diagnosis	2nd relapse	7.5	*ALK*	R1275Q	N/D	3.11%[c]	2954	528
					ALK	F1174L	17.0%	3.88%	2242	354
					APC	R499*	0.24%[a]	0.31%	2580	412
Ewing sarcoma	315	Diagnosis	Relapse	34.02	*TP53*	R273C	48.0%[b]	N/D	3557	314
					TP53	R337C	N/D	31.40%[c]	5237	391
					CDKN2A	R80*	3.0%[a]	25.53%	2064	899
ACC	427	3rd progression	VGPR to 4th-line therapy	51.96	*CTNNB1*	S33Pro	33.00%[b]	N/D	5194	777
RMS	444	Diagnosis	2nd relapse	18.6	*TP53*	V173M	F	11.43%	2782	17
					PIK3CA	E542K	15.0%[b]	N/D	2166	167
					PIK3CA	E545K	17.0%	0.56%[a]	2065	180
Osteosarcoma	514	Diagnosis	2nd relapse	33.96	*TP53*	R248T	78.0%	11.08%	6334	91
					TP53	Y220C	N/D	0.29%[c]	5510	542
Neuroblastoma	738	Post induction	1st relapse	168.6	*TP53*	R249S	N/D	0.05%[c]	14825	193
					ALK	D1091N	8.0%	0.03%	22632	308
RMS	954	Diagnosis	2nd relapse	29.52	*KRAS*	G12C	92.0%	0.09%[a]	3233	1453
Wilms tumour	1211	Diagnosis	3rd relapse	50.76	*TP53*	R273C	100.0%	23.96%	3961	74
Wilms tumour	1322	Post induction	3rd relapse	19.86	*TP53*	R181C	86.0%	3.72%	2525	141
					TP53	C176Y	N/D	3.03%[c]	2439	174

FFPE, formalin-fixed paraffin-embedded; ctDNA, circulating tumour DNA; RMS, rhabdomyosarcoma; ACC, adrenocortical carcinoma; VGPR, very good partial response, postinduction, surgical resection after routine induction chemotherapy, AF, allele fraction; F, failed coverage; N/D, not detected.
[a] Below limit of detection.
[b] Detected in tumour only.
[c] Detected in plasma only.

and larger chromosomal gains or losses. Therefore, in version 3 of the panel, we are incorporating a new assay to determine this, which includes probes located across the chromosomes. In addition, as novel gene discoveries and/or targeted inhibitors become available, a wider approach is required for certain indications including a more extensive method for detection of structural variants/translocations. Capture NGS panels are able to detect translocations in DNA with the ability to determine the single-nucleotide breakpoint, so long as those breakpoints occur in or close to a targeted region. We used MANTA to detect spanning pair reads and split reads, thereby identifying fusion gene partners. However, detection of fusion genes is inevitably restricted. We are therefore currently validating a panel using anchored multiplex PCR-based enrichment to detect fusions from RNA, removing the need to sequence long and complex intronic regions. Furthermore, methylation profiling is particularly relevant for precise diagnostic classification of central nervous system (CNS) tumours, many of which harbour few if any recurrent somatic alterations.

Therefore, in the Stratified Medicine Paediatrics (SMPaeds) national molecular profiling study for children with relapsed and refractory cancers, we will retain the practical advantages of panel sequencing and run this alongside other more comprehensive profiling modalities including WES, RNA-seq, low-coverage WGS and methylation to support biomarker-driven clinical trials in the UK, such as eSMART [41]. Furthermore, where sufficient tissue is available, concurrent analysis via the National Health Service England WGS programme will be compared with SMPaeds genomic and

Fig. 3. Overview of sequencing results. Oncoprint represents somatic mutations and gains, amplification and deletions detected in genes that are covered by the targeted panel. Samples are grouped in columns with genes displayed along rows. Samples are arranged according to the tumour type and genes sorted by frequency. Panel version, sample type, molecular annotations and diagnosis are provided as bars according to the included key (A). Bar plot of most recurrent altered genes, sorted by frequency and colour coded according to the tumour type (B). FFPE, formalin-fixed paraffin-embedded; DSRCT, desmoplastic small round cell tumour; CNS, central nervous system; FF, fresh frozen.

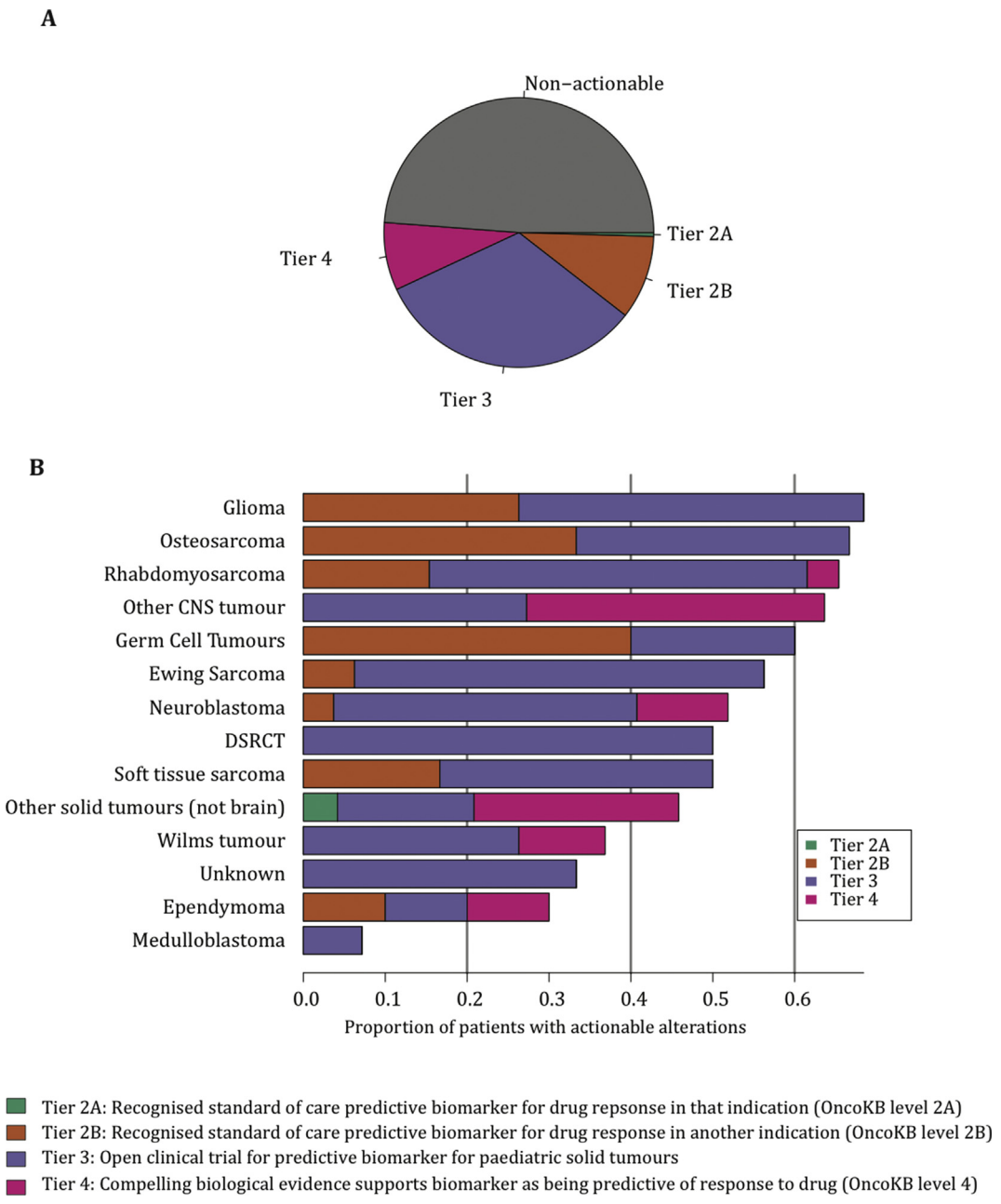


Fig. 4. Clinical actionability. Somatic alterations were defined according to OncoKB levels of evidence. Actionability tiers are described in the key. Distribution of actionability tiers for the entire sequenced cohort (A). Distribution of actionability tiers across common tumours, colour coded according to the tumour type (B). DSRCT, desmoplastic small round cell tumour; CNS, central nervous system.

clinical data. This approach will provide an unbiased assessment of the clinical utility and cost-effectiveness of multiple different modalities to enable formal recommendations for implementation into routine molecular diagnostics.

Despite the high detection rate of potentially actionable alterations, few patients received treatment with targeted agents. The reasons for this were multifactorial and include the following: lack of available clinical trials, difficulties accessing novel drugs on a compassionate-use basis and/or clinical deterioration of the patient. In addition, although many patients had relapsed/refractory disease, a considerable proportion of patients were still on either first-line therapy or proven standard relapse therapies at the time of sequencing. A number of patients were also enrolled in available phase I/II trials that did not require biomarker screening.

This was a pilot study, requiring retrieval of archival tissue, batching of samples for sequencing and infrequent MTBs. However, for the prospective SMPaeds study, which mandates biopsy at relapse for molecular

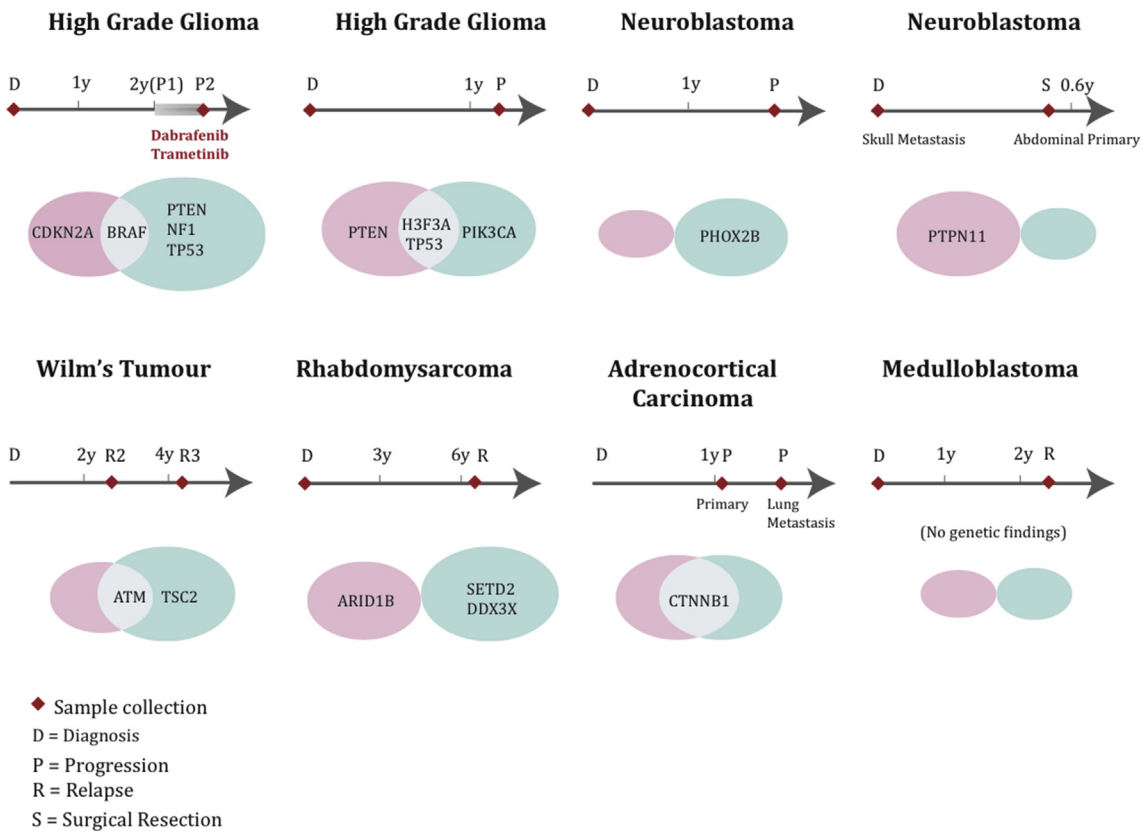


Fig. 5. Comparison of results from paired samples, sequenced at different time points. Venn diagrams compare the genetic findings in eight patients. Shared alterations are illustrated by the intersection of the two ovals. Alterations detected at only the 1st time point are represented in the pink oval, and alterations identified at the 2nd time point only are represented in the green oval. The size of the oval represents the number of variants identified in each patient.

preselection for clinical trials, samples will be processed in a clinically relevant time frame, which after clinical feedback is currently 3–4 weeks, with the final goal of returning data in two weeks. For children with primary solid tumours (who are not enrolled in SMPaeds), as a result of this study, NGS panel sequencing on the paediatric solid tumour panel v2 is now offered in the UK as part of routine National Health Service diagnostic testing with a turnaround time of 4 weeks from sample dispatch to reporting. Owing to ethical and consent constraints, we were not permitted to report germline findings in the present study. However, given the obvious clinical importance of predisposing mutations in paediatric cancer, we have now obtained suitable consents to report germline mutations via an accredited genetics clinic at Great Ormond Street Hospital.

The sequencing of paired tumour samples at different times demonstrates the importance of tumour heterogeneity and evolution, adding to the mounting literature in support of the clinical importance of biopsy at relapse for children with cancer [19,42]. Notably, many tumour mutations emerging at the time of relapse (*PTEN, NF1, PIK3CA* and *TSC2*) are recognised predictive biomarkers of a targeted therapeutic response.

Although sequencing tissue samples of patients is crucial, liquid biopsies offer the possibility of a non-invasive source for tumour genotyping and disease monitoring. Our preliminary findings from a small number of children demonstrate that high-depth sequencing of ctDNA can identify actionable somatic variants. We also identified some discrepancies between tumour and plasma, most likely a reflection of tumour heterogeneity and evolution. However, large-scale validation studies comparing tumour and serial ctDNA findings in children with cancer are needed to define the clinical utility of ctDNA analysis, for which a bespoke ctDNA panel for paediatric solid tumours is currently being developed to be incorporated as part of the diagnostics pipeline.

In summary, we demonstrate the value of targeted gene sequencing as a practical and cost-effective clinical tool to enable improved diagnosis, prognostication and therapeutic stratification for children with cancer.

Acknowledgements

This work was supported by Christopher's Smile, the National Institute of Health Research (NIHR) Royal

Marsden Biomedical Research Centre (BRC), Children With Cancer UK (CWC UK) Cancer Research UK (CRUK), Abbie's Fund, the Rosetree Trust and the KiCa Fund, managed by the King Baudouin Foundation. Roche provided support for Panel development. T.S.J. is funded by The Brain Tumour Charity, CWC UK, GOSH Children's Charity (GOSH CC), CRUK, the Olivia Hodson Cancer Fund and the NIHR GOSH BRC. J.A. and D.H. are funded by the GOSH CC and NIHR GOSH BRC. L.V.M. is funded by the Oak Foundation. The authors thank all participants and the CCLG Tissue Bank for access to samples and contributing CCLG Centres, including members of the ECMC Paediatric network. The CCLG Tissue Bank is funded by Cancer Research UK and CCLG.

Conflict of interest statement

There are no known conflicts of interest associated with this publication, and there has been no significant financial support for this work that could have influenced its outcome.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejca.2019.07.027.

References

[1] Fiore RN, Goodman KW. Precision medicine ethics: selected issues and developments in next-generation sequencing, clinical oncology, and ethics. Curr Opin Oncol 2016;28(1):83–7.
[2] Tuff-Lacey A, et al. A collaborative approach to enabling stratified cancer medicine in the UK. Drug Discov Today 2015;20(12):1414–8.
[3] Middleton G, et al. The National Lung Matrix Trial: translating the biology of stratification in advanced non-small-cell lung cancer. Ann Oncol 2015;26(12):2464–9.
[4] Do K, O'Sullivan Coyne G, Chen AP. An overview of the NCI precision medicine trials-NCI MATCH and MPACT. Chin Clin Oncol 2015;4(3):31.
[5] Kim G, et al. FDA approval summary: olaparib monotherapy in patients with deleterious germline BRCA-mutated advanced ovarian cancer treated with three or more lines of chemotherapy. Clin Cancer Res 2015;21(19):4257–61.
[6] Loong HH, et al. Crizotinib in the management of advanced-stage non-small-cell lung cancer. Future Oncol 2015;11(5):735–45.
[7] Stagno F, et al. Imatinib mesylate in chronic myeloid leukemia: frontline treatment and long-term outcomes. Expert Rev Anticancer Ther 2016;16(3):273–8.
[8] Siegel R, Naishadham D, Jemal A. Cancer statistics. CA Cancer J Clin 2012;62(1):10–29. 2012.
[9] Mody RJ, et al. Integrative clinical sequencing in the management of refractory or relapsed cancer in youth. J Am Med Assoc 2015;314(9):913–25.
[10] Worst BC, et al. Next-generation personalised medicine for high-risk paediatric cancer patients - the INFORM pilot study. Eur J Cancer 2016;65:91–101.
[11] Harttrampf AC, et al. Molecular screening for cancer treatment optimization (MOSCATO-01) in pediatric patients: a single-institutional prospective molecular stratification trial. Clin Cancer Res 2017;23(20):6101–12.
[12] Harris MH, et al. Multicenter feasibility study of tumor molecular profiling to inform therapeutic decisions in advanced pediatric solid tumors: the individualized cancer therapy (iCat) study. JAMA Oncol 2016;2(5):608–15.
[13] Ortiz MV, et al. Integrating genomics into clinical pediatric oncology using the molecular tumor board at the memorial sloan kettering cancer center. Pediatr Blood Cancer 2016;63(8):1368–74.
[14] Castel D, et al. Histone H3F3A and HIST1H3B K27M mutations define two subgroups of diffuse intrinsic pontine gliomas with different prognosis and phenotypes. Acta Neuropathol 2015;130(6):815–27.
[15] Schwartzentruber J, et al. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature 2012;482(7384):226–31.
[16] Peifer M, et al. Telomerase activation by genomic rearrangements in high-risk neuroblastoma. Nature 2015;526(7575):700–4.
[17] Izquierdo E, et al. Development of a targeted sequencing approach to identify prognostic, predictive and diagnostic markers in paediatric solid tumours. Oncotarget 2017;8(67):112036–50.
[18] Padovan-Merhar OM, et al. Enrichment of targetable mutations in the relapsed neuroblastoma genome. PLoS Genet 2016;12(12):e1006501.
[19] Cohen B, et al. Pediatric oncology provider views on performing a biopsy of solid tumors in children with relapsed or refractory disease for the purpose of genomic profiling. Ann Surg Oncol 2016;23(Suppl 5):990–7.
[20] Thompson JC, et al. Detection of therapeutically targetable driver and resistance mutations in lung cancer patients by next-generation sequencing of cell-free circulating tumor DNA. Clin Cancer Res 2016;22(23):5772–82.
[21] Rothe F, et al. Plasma circulating tumor DNA as an alternative to metastatic biopsies for mutational analysis in breast cancer. Ann Oncol 2014;25(10):1959–65.
[22] Xu S, et al. Circulating tumor DNA identified by targeted sequencing in advanced-stage non-small cell lung cancer patients. Cancer Lett 2016;370(2):324–31.
[23] Bettegowda C, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med 2014;6(224):224ra24.
[24] De Mattos-Arruda L, Caldas C. Cell-free circulating tumour DNA as a liquid biopsy in breast cancer. Mol Oncol 2016;10(3):464–74.
[25] Alix-Panabieres C, Pantel K. Clinical applications of circulating tumor cells and circulating tumor DNA as liquid biopsy. Cancer Discov 2016;6(5):479–91.
[26] Kurihara S, et al. Circulating free DNA as non-invasive diagnostic biomarker for childhood solid tumors. J Pediatr Surg 2015;50(12):2094–7.
[27] Combaret V, et al. Detection of tumor ALK status in neuroblastoma patients using peripheral blood. Cancer Med 2015;4(4):540–50.
[28] Panditharatna E, et al. Clinically relevant and minimally invasive tumor surveillance of pediatric diffuse midline gliomas using patient-derived liquid biopsy. Clin Cancer Res 2018;24(23):5850–9.
[29] Jimenez I, et al. Circulating tumor DNA analysis enables molecular characterization of pediatric renal tumors at diagnosis. Int J Cancer 2019;144(1):68–79.
[30] Chicard M, et al. Whole-exome sequencing of cell-free DNA reveals temporo-spatial heterogeneity and identifies treatment-resistant clones in neuroblastoma. Clin Cancer Res 2018;24(4):939–49.
[31] Allin DM, et al. Circulating tumour DNA is a potential biomarker for disease progression and response to targeted therapy in advanced thyroid cancer. Eur J Cancer 2018;103:165–75.
[32] Chakravarty D, et al. OncoKB: a precision oncology knowledge basevol. 2017. JCO Precis Oncol; 2017.

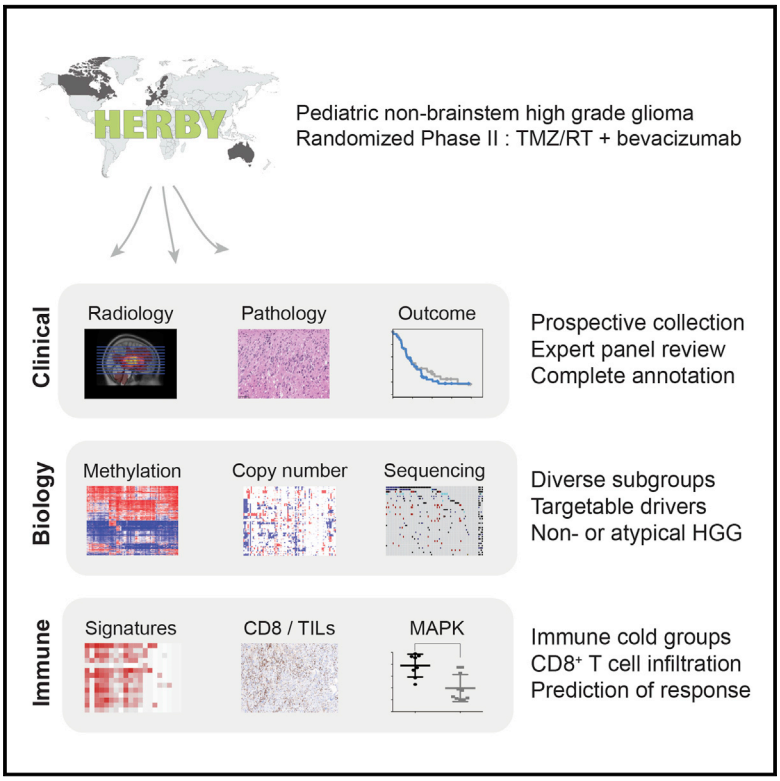
[33] Forbes SA, et al. COSMIC: somatic cancer genetics at high-resolution. Nucleic Acids Res 2017;45(D1):D777–83.
[34] Grobner SN, et al. The landscape of genomic alterations across childhood cancers. Nature 2018;555(7696):321–7.
[35] Gatz SA, et al. Chemotherapy responsiveness in a patient with multiply relapsed ameloblastic fibro-odontosarcoma of the maxilla. Pediatr Blood Cancer 2015;62(11):2029–32.
[36] Lee L, Gupta M, Sahasranaman S. Immune Checkpoint inhibitors: an introduction to the next-generation cancer immunotherapy. J Clin Pharmacol 2016;56(2):157–69.
[37] Kohsaka S, et al. A recurrent neomorphic mutation in MYOD1 defines a clinically aggressive subset of embryonal rhabdomyosarcoma associated with PI3K-AKT pathway mutations. Nat Genet 2014;46(6):595–600.
[38] Pajtler KW, et al. Molecular classification of ependymal tumors across all CNS compartments, histopathological grades, and age groups. Cancer Cell 2015;27(5):728–43.
[39] Whittaker SR, et al. A genome-scale RNA interference screen implicates NF1 loss in resistance to RAF inhibition. Cancer Discov 2013;3(3):350–62.
[40] Van Allen EM, et al. The genetic landscape of clinical resistance to RAF inhibition in metastatic melanoma. Cancer Discov 2014;4(1):94–109.
[41] European Proof-of Concept Therapeutic Stratification Trial of Molecular Anomalies in Relapsed or Refractory Tumours. https://clinicaltrials.gov/ct2/show/NCT02813135?term=esmart&rank=5.
[42] Eleveld TF, et al. Relapsed neuroblastomas show frequent RAS-MAPK pathway mutations. Nat Genet 2015;47(8):864–71.

Article

Cancer Cell

Molecular, Pathological, Radiological, and Immune Profiling of Non-brainstem Pediatric High-Grade Glioma from the HERBY Phase II Randomized Trial

Graphical Abstract



Highlights

- The HERBY trial tested the use of bevacizumab in pediatric non-brainstem HGG
- Parallel translational biology studies highlighted the diversity of the trial cohort
- Elevated levels of CD8$^+$ T cells were seen in PXA-like and hypermutant tumors
- MAPK-associated immune signatures predicted response to bevacizumab

Authors

Alan Mackay, Anna Burford, Valeria Molinari, ..., Tim Jaspan, Pascale Varlet, Chris Jones

Correspondence

chris.jones@icr.ac.uk

In Brief

In a pediatric high-grade non-brainstem glioma cohort, Mackay et al. show that hypermutator tumors and those resembling pleomorphic xanthoastrocytoma are highly infiltrated by CD8$^+$ lymphocytes and benefit from the addition of bevacizumab, whereas the histone H3 subgroups are immune cold and have a poor outcome.

Mackay et al., 2018, Cancer Cell *33*, 829–842
May 14, 2018
https://doi.org/10.1016/j.ccell.2018.04.004

CellPress

Cancer Cell
Article

Molecular, Pathological, Radiological, and Immune Profiling of Non-brainstem Pediatric High-Grade Glioma from the HERBY Phase II Randomized Trial

Alan Mackay,[1,2] Anna Burford,[1,2] Valeria Molinari,[1,2] David T.W. Jones,[3,4] Elisa Izquierdo,[1,2] Jurriaan Brouwer-Visser,[6] Felice Giangaspero,[7,8] Christine Haberler,[9,10] Torsten Pietsch,[11] Thomas S. Jacques,[12,13] Dominique Figarella-Branger,[14] Daniel Rodriguez,[15] Paul S. Morgan,[15] Pichai Raman,[16,17] Angela J. Waanders,[16,18] Adam C. Resnick,[16,17,18]

(Author list continued on next page)

[1]Division of Molecular Pathology, The Institute of Cancer Research, 15 Cotswold Road, Sutton, London, Surrey SM2 5NG, UK
[2]Division of Cancer Therapeutics, The Institute of Cancer Research, 15 Cotswold Road, Sutton, London, Surrey SM2 5NG, UK
[3]Hopp Children's Cancer Center at the NCT Heidelberg (KiTZ), Heidelberg, Germany
[4]Division of Paediatric Neuro-oncology, German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany
[5]Department of Pediatric Hematology and Oncology, Heidelberg University Hospital, Heidelberg, Germany
[6]Roche Innovation Center, New York, NY, USA
[7]Department of Radiology, Oncology and Anatomic-Pathology Sciences, Sapienza University, Rome, Italy
[8]IRCCS Neuromed, Pozzilli, Italy
[9]Institute of Neurology, Medical University of Vienna, Vienna, Austria
[10]Comprehensive Cancer Center, Medical University of Vienna, Vienna, Austria
[11]DGNN Brain Tumor Reference Center, Institute of Neuropathology, University of Bonn Medical Center, Bonn, Germany
[12]UCL Great Ormond Street Institute of Child Health, London, UK
[13]Department of Histopathology, Great Ormond Street Hospital for Children, London, UK

(Affiliations continued on next page)

SUMMARY

The HERBY trial was a phase II open-label, randomized, multicenter trial evaluating bevacizumab (BEV) in addition to temozolomide/radiotherapy in patients with newly diagnosed non-brainstem high-grade glioma (HGG) between the ages of 3 and 18 years. We carried out comprehensive molecular analysis integrated with pathology, radiology, and immune profiling. In *post-hoc* subgroup analysis, hypermutator tumors (mismatch repair deficiency and somatic *POLE/POLD1* mutations) and those biologically resembling pleomorphic xanthoastrocytoma ([PXA]-like, driven by *BRAF*_V600E or *NF1* mutation) had significantly more CD8$^+$ tumor-infiltrating lymphocytes, and longer survival with the addition of BEV. Histone H3 subgroups (hemispheric G34R/V and midline K27M) had a worse outcome and were immune cold. Future clinical trials will need to take into account the diversity represented by the term "HGG" in the pediatric population.

INTRODUCTION

High-grade gliomas (HGGs) in children, like their adult counterparts, continue to have a bleak prognosis, with a median overall survival (OS) of 9–15 months (Cohen et al., 2011; Jones et al., 2016; Ostrom et al., 2015). Recent integrated molecular-profiling initiatives have shown that pediatric HGGs (pHGGs) are biologically distinct from their adult counterparts, with subgroups of the disease marked by recurrent mutations in genes encoding histone H3 variants having different age of incidence, anatomical

Significance

We validate in the prospective clinical trial setting the biological and clinical diversity of pediatric high-grade glioma previously described in large retrospective series, underpinned by detailed pathological and radiological analysis. Although adding bevacizumab (BEV) to standard temozolomide/radiotherapy did not improve survival across the whole cohort, we identify disease subgroups with MAPK activation to harbor an enhanced CD8$^+$ T cell immune response, which may derive benefit from the addition of BEV. If confirmed in another study, this would represent a useful predictive biomarker for this regimen in these tumors, and points the way for therapeutic strategies for subgroups of children with high-grade glioma.

Maura Massimino,[19] Maria Luisa Garrè,[20] Helen Smith,[21] David Capper,[22,23,24] Stefan M. Pfister,[3,4,5] Thomas Würdinger,[25] Rachel Tam,[26] Josep Garcia,[21] Meghna Das Thakur,[26] Gilles Vassal,[27] Jacques Grill,[27] Tim Jaspan,[15] Pascale Varlet,[28] and Chris Jones[1,2,29,*]

[14]Department of Pathology and Neuropathology, La Timone Hospital, Aix Marseille University, Marseille, France
[15]Nottingham University Hospitals, Nottingham, UK
[16]The Center for Data Driven Discovery in Biomedicine (D^3b), Children's Hospital of Philadelphia, Philadelphia, PA, USA
[17]Division of Neurosurgery, Children's Hospital of Philadelphia, Philadelphia, PA, USA
[18]Division of Oncology, Children's Hospital of Philadelphia, Philadelphia, PA, USA
[19]Pediatric Oncology Unit, Fondazione IRCCS, Istituto Nazionale Tumori, Milan, Italy
[20]Centro di Neuro-Oncologia, Istituto Giannina Gaslini, Genoa, Italy
[21]F. Hoffmann-La Roche Ltd, Basel, Switzerland
[22]Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, Berlin, Germany
[23]Berlin Institute of Health, Institute of Neuropathology, Berlin, Germany
[24]Department of Neuropathology, University Hospital Heidelberg and Clinical Cooperation Unit Neuropathology, German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany
[25]Department of Neurosurgery, Brain Tumor Center Amsterdam, VU Medical Center, Amsterdam, the Netherlands
[26]Genentech, South San Francisco, CA, USA
[27]Pediatric and Adolescent Oncology and Unite Mixte de Recherche 8203 du Centre National de la Recherche Scientifique, Gustave Roussy, Paris-Saclay University, Villejuif, France
[28]Sainte-Anne Hospital, Paris-Descartes University, Paris, France
[29]Lead contact
*Correspondence: chris.jones@icr.ac.uk
https://doi.org/10.1016/j.ccell.2018.04.004

location, clinical outcome, and a range of biological parameters (Jones and Baker, 2014; Mackay et al., 2017; Paugh et al., 2010; Schwartzentruber et al., 2012; Sturm et al., 2012; Wu et al., 2012, 2014). Histone wild-type (WT) tumors have widely differing mutational burdens, ranging from infant cases (<3 years) driven by single gene fusion events through to patients with biallelic mismatch repair deficiency harboring some of the highest mutational rates in human cancer (Jones and Baker, 2014; Mackay et al., 2017; Shlien et al., 2015; Wu et al., 2014).

The rapid advances in our understanding of pHGGs have come predominantly from the accumulation of numerous disparate retrospective collections, a reflection of the rarity of the disease. Clinical trial cohorts with ancillary biomarker analyses have been relatively limited in their scope, and historically have focused on single-marker analyses. These include the Children's Oncology Group ACNS0126 (radiotherapy [RT]/temozolomide [TMZ]) (Cohen et al., 2011) and ACNS0423 (RT/TMZ followed by TMZ and lomustine) (Jakacki et al., 2016) studies, which report on the frequency and clinical correlations of O6-methylguanine-DNA methyltransferase (MGMT) expression (ACNS0126) (Jakacki et al., 2016; Pollack et al., 2006), *IDH1* mutation (ACNS0423) (Pollack et al., 2011), as well as phosphorylated Akt expression (Pollack et al., 2010a) and microsatellite instability (both) (Pollack et al., 2010b). The CCG-945 study ("8 in 1" chemotherapy) (Finlay et al., 1995) reported on the prognostic significance of p53 expression/mutation (Pollack et al., 2002), in addition to the presence/absence of 1p19q co-deletion (Pollack et al., 2003b).

This last study (Pollack et al., 2003b) also highlighted the critical importance of pathological review in the diagnosis of pHGG, and subsequent interpretation of clinical trial results (Gilles et al., 2008; Pollack et al., 2003a), particularly in midline locations (Eisenstat et al., 2015). It has subsequently become clear that numerous histological subtypes of HGG can harbor distinct genetic drivers and have considerably better clinical outcomes, such as *BRAF*_V600E mutations in epithelioid glioblastoma (GBM), anaplastic ganglioglioma, and anaplastic pleomorphic xanthoastrocytoma (PXA) (Hatae et al., 2016); in the latter two categories, this mutation is also found in lower-grade entities lacking obvious anaplasia. Additional histone WT cases of otherwise uncontroversial HGGs have been found to be biologically and clinically more similar to several types of low-grade glioma (LGG) and PXA (Korshunov et al., 2015), highlighting the importance of an integrated diagnosis combining molecular and histological features.

The HERBY trial (study BO25041; clinicaltrials.gov NCT01390948) was a phase II, open-label, randomized, multicenter, comparator study of the addition of the anti-angiogenic agent bevacizumab (BEV) to RT and TMZ in patients between the ages of 3 and 18 years with newly diagnosed non-brainstem HGG (Grill et al., 2018). Central confirmation of HGG diagnosis was mandatory before randomization, followed by consensus review by five independent expert neuropathologists. Real-time panel radiological assessment was also incorporated. An exploratory endpoint of the study was to establish a biospecimen repository for correlative molecular profiling. In addition to its role in tumor angiogenesis, vascular endothelial growth factor (VEGF) restricts immune cell activity, and BEV has been demonstrated to facilitate recruitment of T cells to infiltrate tumors (Wallin et al., 2016), as well as increase the ratio of $CD8^+CD3^+$ T cells in adult GBM specimens (Scholz et al., 2016). We therefore also sought to explore the immune profile of cases within the HERBY cohort.

RESULTS

The Translational Research Cohort Is Representative of the Whole Clinical Trial Population

The total HERBY cohort comprised 121 randomized patients at diagnosis (3–18 years) plus 3 infant cases (<3 years) at relapse. Of these, 113 patients consented to the translational research program (Table S1). Tumor tissue was collected from either

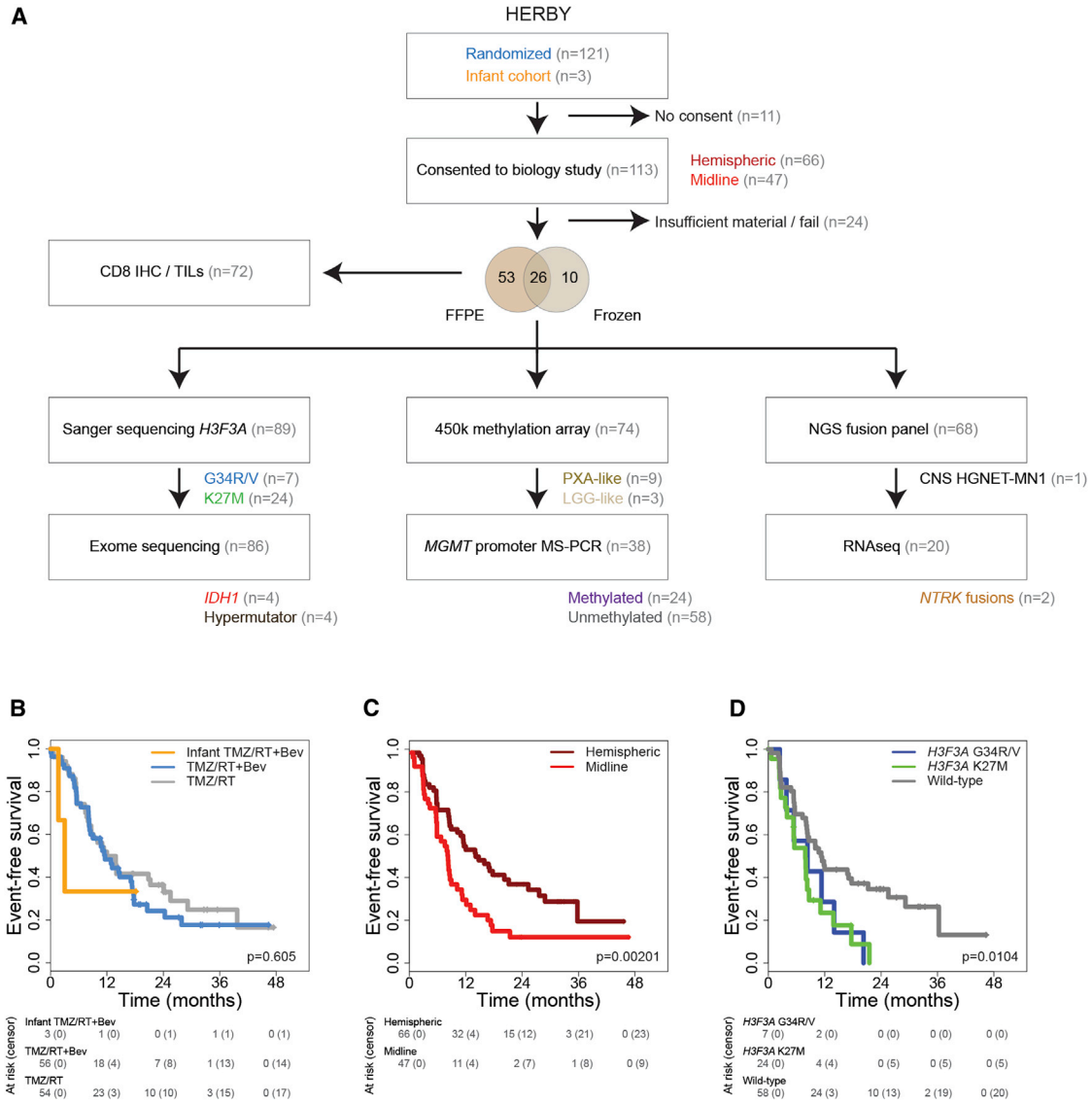


Figure 1. Sample Cohort

(A) Flow diagram indicating total HERBY trial cohort (n = 121 randomized plus 3 infants), those patients consenting to the biology study (n = 113) for whom sufficient formalin-fixed paraffin-embedded (FFPE) or frozen tumor was available (n = 89), and the respective molecular analyses undertaken.

(B–D) Kaplan-Meier plots of event-free survival of cases (y axis) separated by treatment arm (B), anatomical location (C), or *H3F3A* status (D), with time given in months (x axis) and the overall p value provided, calculated by the log rank test. Individual pairwise comparisons are provided in the text.

See also Table S1.

resection (n = 93) or biopsy (n = 20), although 24 cases failed to provide sufficient quantity or quality of sample for molecular analysis. For the remaining 89 cases, material was available in the form of either fresh-frozen material (n = 36), formalin-fixed paraffin-embedded pathology specimens (n = 79), or both (n = 26). These were subjected to Sanger sequencing for *H3F3A* (n = 89), exome sequencing (n = 86), Illumina 450k methylation BeadChip profiling (n = 74), CD8 immunohistochemistry (n = 70), methylation-specific PCR for *MGMT* promoter methylation (n = 36), a capture-based sequencing panel for common fusion gene detection (n = 68), and RNA sequencing (RNA-seq) (n = 20) (Figure 1A).

The translational research cohort, representing a subset (91%) of the randomized trial, displayed equivalent clinical characteristics to the full dataset (Grill et al., 2018), with no difference in the primary endpoint of 1-year event-free survival (EFS) with the addition of BEV to the standard therapy of TMZ and RT (median 12.0 versus 8.3 months, p = 0.372, log rank test), with an additional small (n = 3) infant cohort treated with BEV at relapse (Figure 1B). The cohort contained 66 (58%) hemispheric and 47 (42%) non-brainstem midline tumors, with the latter location conferring a significantly shorter EFS (median 8.0 versus 14.7 months, p = 0.00201, log rank test) (Figure 1C). Histone mutation status was a significant predictor of worse prognosis

compared with WT (median EFS = 11.3 months) for *H3F3A*_K27M (24/89, 27%; median EFS = 7.9 months; p = 0.0063, log rank test) and also trended toward shorter survival for *H3F3A*_G34R/V (7/89, 8%; median EFS = 8.3 months; p = 0.096, log rank test) (Figure 1D).

Integrated Molecular Analysis Defines the Major (Epi) genomic Alterations in pHGG

We used the Heidelberg brain tumor classifier on Illumina 450k methylation array data to assign a molecular subgroup to each of 74 samples for which such data were available (Table S2). After excluding low-scoring assignments (<0.2), we used a simplified system to classify tumors as either H3K27M (n = 18), H3G34R/V (n = 6), or IDH1 (n = 4) (integrating gene mutation data in low-scoring cases); as resembling PXA-like (n = 9) or other LGG-like (n = 3); and aggregating the remaining tumors (HGG-WT, n = 34) (Figure 2A). IDH1 tumors represented the oldest patients (median = 17.2 years, others = 11.2, p = 0.0107, t test), with LGG-like representing the youngest category (median = 5.7 years, p = 0.0098, t test) (Figure 2B). These two subgroups each had significantly better outcome in terms of EFS (p = 0.0281 and p = 0.0386, log rank test), although not OS (p = 0.0935 and p = 0.129, log rank test) (Figure 2C), when compared individually to the remaining tumors. The PXA-like showed a trend toward longer OS (p = 0.0867, log rank test) compared with the rest. When IDH1, PXA-like, and LGG-like tumors were excluded from the analysis, the significant differences between histone mutant and HGG-WT groups were absent (H3K27M, p = 0.257 EFS and p = 0.0746 OS; H3G34R/V, p = 0.552 EFS and p = 0.116 OS, log rank test). Twelve out of 78 (15%) samples harbored a methylated *MGMT* promoter, although this was largely restricted to the H3G34R/V (n = 3, p = 0.0249, Fishers exact test) and IDH1 (n = 3, p = 0.0062, Fisher's exact test) subgroups (Figure S1A), and was not significantly associated with survival (Figure S1B) in these uniformly TMZ-treated patients.

We used a 450k methylation array and exome-sequencing coverage to derive DNA copy-number profiles for 86 pHGG (Figure S2A), including focal amplifications/deletions, as well as whole-arm chromosomal gains/losses (Table S3). Taken with the somatic single-nucleotide variants (SNVs) and small insertion/deletions from whole-exome sequencing (Table S3), and candidate gene fusion events from capture-based panel sequencing (n = 68) and RNA-seq (n = 20) (Table S3), we derived an integrated map of genetic alterations across the translational research cohort (Figure 3A). The most common alteration was *TP53* mutation (39/82, 48%), followed by *ATRX* deletion/mutation (25/82, 30%), *PDGFRA* amplification/mutation (17/82, 21%), and *CDKN2A/B* deletion (15/82, 18%). Additional recurrent alterations in receptor tyrosine kinases (*EGFR, MET, ERBB3, IGF1R,* and *NTRK2*), phosphatidylinositol 3-kinase (PI3K)/mammalian target of rapamycin (*PTEN, PIK3CA, TSC2,* and *PIK3R1*), and MAP-kinase (*NF1, BRAF, PTPN11,* and *PTPN12*) pathways were common, as were amplifications/mutations in various genes associated with cell-cycle regulation (*RB1, CDK4, MDM2,* and *CCND2*). Taking a minimum variant allele frequency of 5% as a threshold, the median number of somatic mutations per sample was 15 (range = 0–337) (Figure 3B), with the exception of four cases for whom there were more than 100-fold more, and were excluded from gene-level counts.

pHGGs Comprise a Diverse Set of Biological and Clinicopathological Subgroups

Two cases were highlighted from the methylation subgrouping as potentially representing non-HGG entities. One case classifying as CNS neuroblastoma with FOXR2 activation (CNS NB-FOXR2, methylation classifier score = 0.617), was found to have no evidence of *FOXR2* alterations. A further case, a compact and necrotic tumor with perivascular radiating arrangements (Figure S2B), displayed a methylation classifier score strongly indicative of a high-grade neuroepithelial tumor with *MN1* alteration (CNS HGNET-MN1, methylation classifier score = 0.713) (Figure S2C). We identified a candidate alteration in this case fusing exon 1 of *MN1* (22q12.1) to exon 3 of *CARD6* (5p13.1) (Figure S2D), and thus appears most likely to fall into this categorization.

Three cases classified more closely to either pilocytic astrocytoma (n = 2) or desmoplastic infantile ganglioglioma (DIG) (n = 1) by 450k methylation profiling. The first two harbored mitogen-activated protein kinase (MAPK) dysregulation in the form of either *BRAF*_V600E or intragenic *FGFR1* duplication (Figure S2E). Histologically, after Pathology Committee re-review, no piloid features were seen, and anaplastic features were evident (Figure S2F). The DIG-like case was found in the infant cohort (2.7 years). None of these three patients died during the course of follow-up, and, although numbers are small, were all found in the right frontal and temporo-parietal lobes with central predominance (Figure S2G).

There were four cases with *IDH1* hotspot mutations (Figure S3A). Three were classified as WHO-grade III anaplastic astrocytoma (AA), *IDH1*_R132-positive by immunohistochemistry, with concurrent *TP53* and *ATRX* mutations. The remaining case was originally classified as a mixed oligo-astrocytoma, which, by virtue of the presence of *IDH1*_R132 and *TERT* promoter mutation (C228T), as well as copy-number loss of chromosomes 1p and 19q, would be described as an oligodendroglioma according to WHO 2016 (Figure S3B). Across the whole cohort, *IDH1* mutation conferred a significantly longer EFS (p = 0.0398, log rank test), although not OS (p = 0.110, log rank test) (Figure S3C), and were restricted to the frontal lobes (Figure S3D).

After excluding *IDH1* mutant cases, the remaining *H3F3A* and *BRAF* WT cases (n = 38) represented a heterogeneous mix of genomic profiles, with recurrent deletions/mutations in the common pHGG tumor suppressor genes *TP53* (n = 11), *ATRX* (n = 5), *CDKN2A/B* (n = 7), *NF1* (n = 8), *RB1* (n = 7), and *PTEN* (n = 5), but also with an enrichment of gene amplifications in *PDGFRA* (n = 5, with *KIT* and *KDR*, n = 4), *CDK4* (n = 7, with *MDM2*, n = 4), *EGFR* (n = 4), *MET* (n = 2), *CCND2* (n = 3), and *MYCN* (n = 3) (Figure S3E). The most common methylation subclass in these cases was designated GBM_RTK_MYCN (n = 6); however, it included only one of those with *MYCN* amplification, and with no other common amplifications or mutations. Seven cases harbored none of the recurrently altered genes previously described in pHGG, and clearly represent a subgroup that warrants further investigation. Together, these cases had bilateral hemispheric distribution with predominant deep right cerebral localization (Figure S3F).

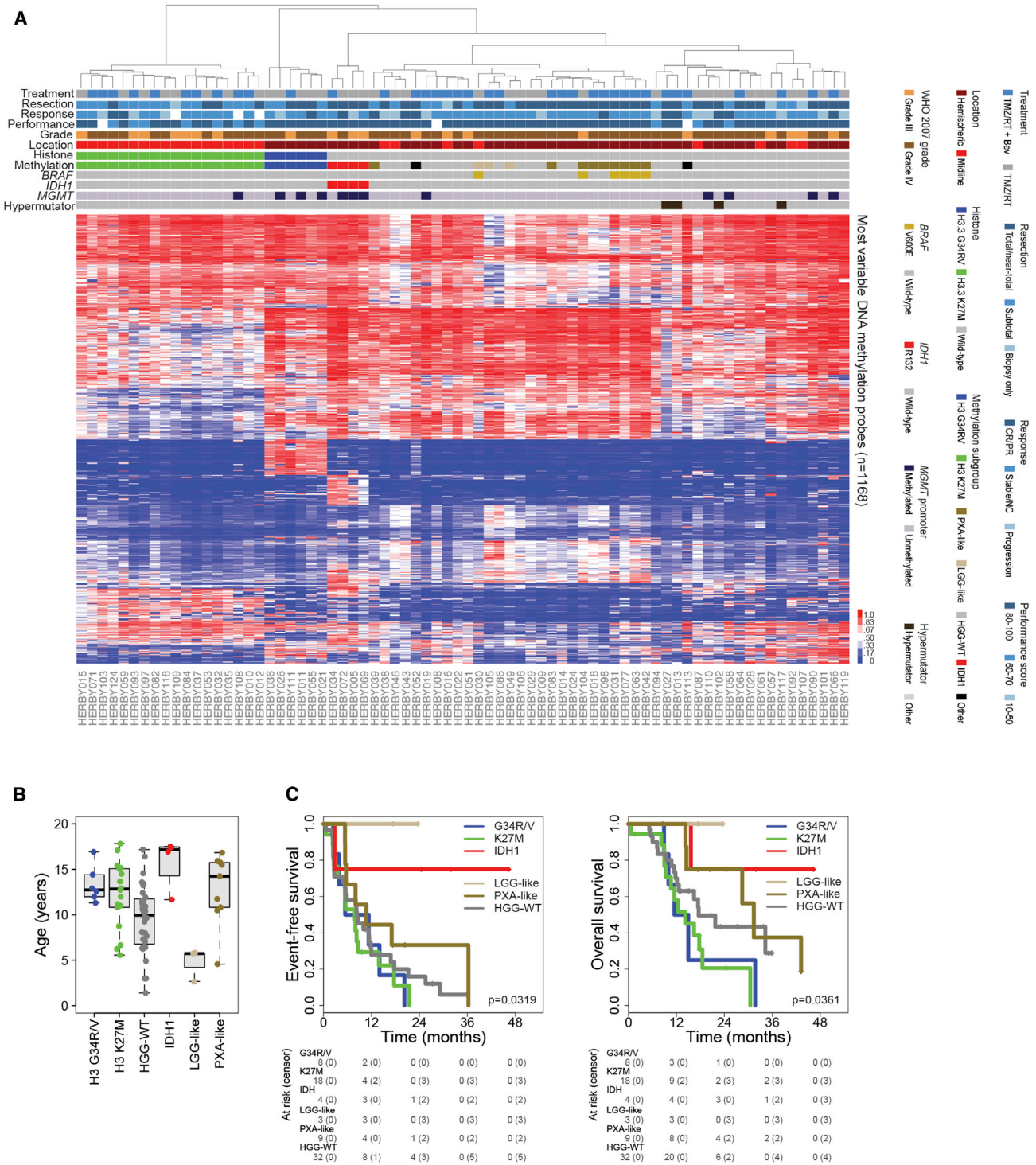


Figure 2. Methylation-Based Subclassification

(A) Heatmap representation of β values for 74 samples profiled on the Illumina 450k BeadChip platform (red, high; blue, low). Samples are arranged in columns clustered by probes with the largest median absolute deviation across the 10k predictor subset of probes. Clinicopathological and molecular annotations are provided as bars according to the included key. CR/PR, complete response or partial response; Stable/NC, stable disease or no change.

(B) Boxplot showing age at diagnosis of included cases, separated by methylation subclass. The thick line within the box is the median, the lower and upper limits of the boxes represent the first and third quartiles, and the whiskers 1.5× the interquartile range.

(C) Kaplan-Meier plot of event-free and overall survival of cases (y axis) separated by methylation subclass, time given in months (x axis) and overall p value calculated by the log rank test.

See also Figure S1 and Table S2.

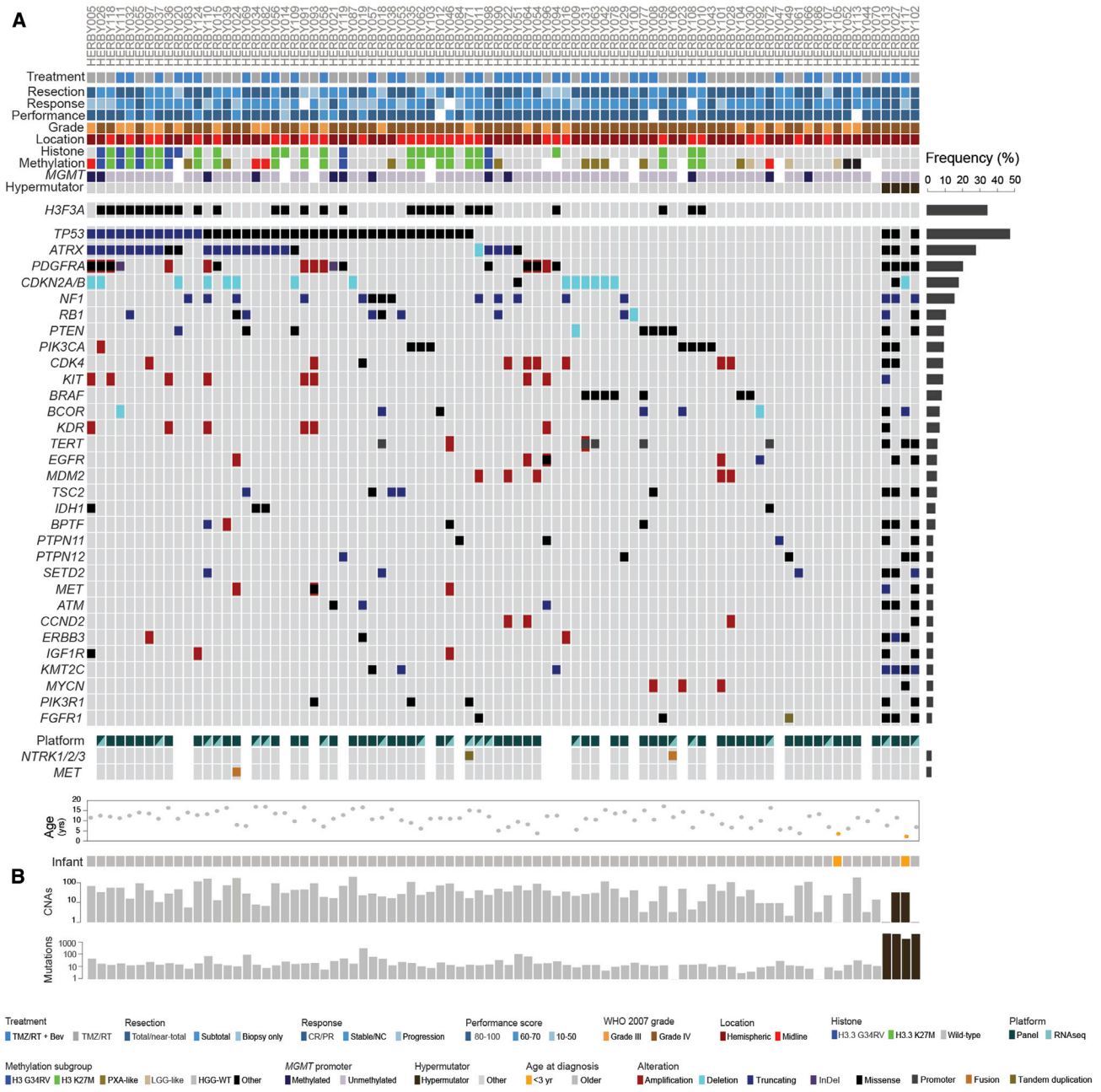


Figure 3. Somatic Mutations

(A) Oncoprint representation of an integrated annotation of somatic mutations and DNA copy-number changes for the 30 most frequently altered genes in 86 samples ($n \geq 3$, frequency barplot on the right, excluding hypermutator cases). Selected common fusion events are also shown where available. Samples are arranged in columns with genes labeled along rows.

(B) Barplots are provided on a $\log_{10}$ scale for numbers of copy-number aberrations and somatic mutations per case. Clinicopathological and molecular annotations are provided as bars according to the included key. CR/PR, complete response or partial response; Stable/NC, stable disease or no change.

See also Figures S2 and S3 and Table S3.

H3F3A Mutation Confers Poor Prognosis for Both K27M and G34R/V Substitutions

Histone mutations have been shown to be present in approximately half of all pHGGs (Mackay et al., 2017), with a clear negative impact on survival for K27M (Karremann et al., 2018; Khuong-Quang et al., 2012; Mackay et al., 2017), although the situation is less clear for G34R/V mutations (Bjerke et al., 2013; Korshunov et al., 2015; Mackay et al., 2017). *H3F3A*_G34R/V mutant tumors had a tendency to being diffusely infiltrative with predominant deep left temporo-parietal involvement (Figure 4A). There were seven cases with *H3F3A*_G34 substitutions (six G34R and one G34V), with six

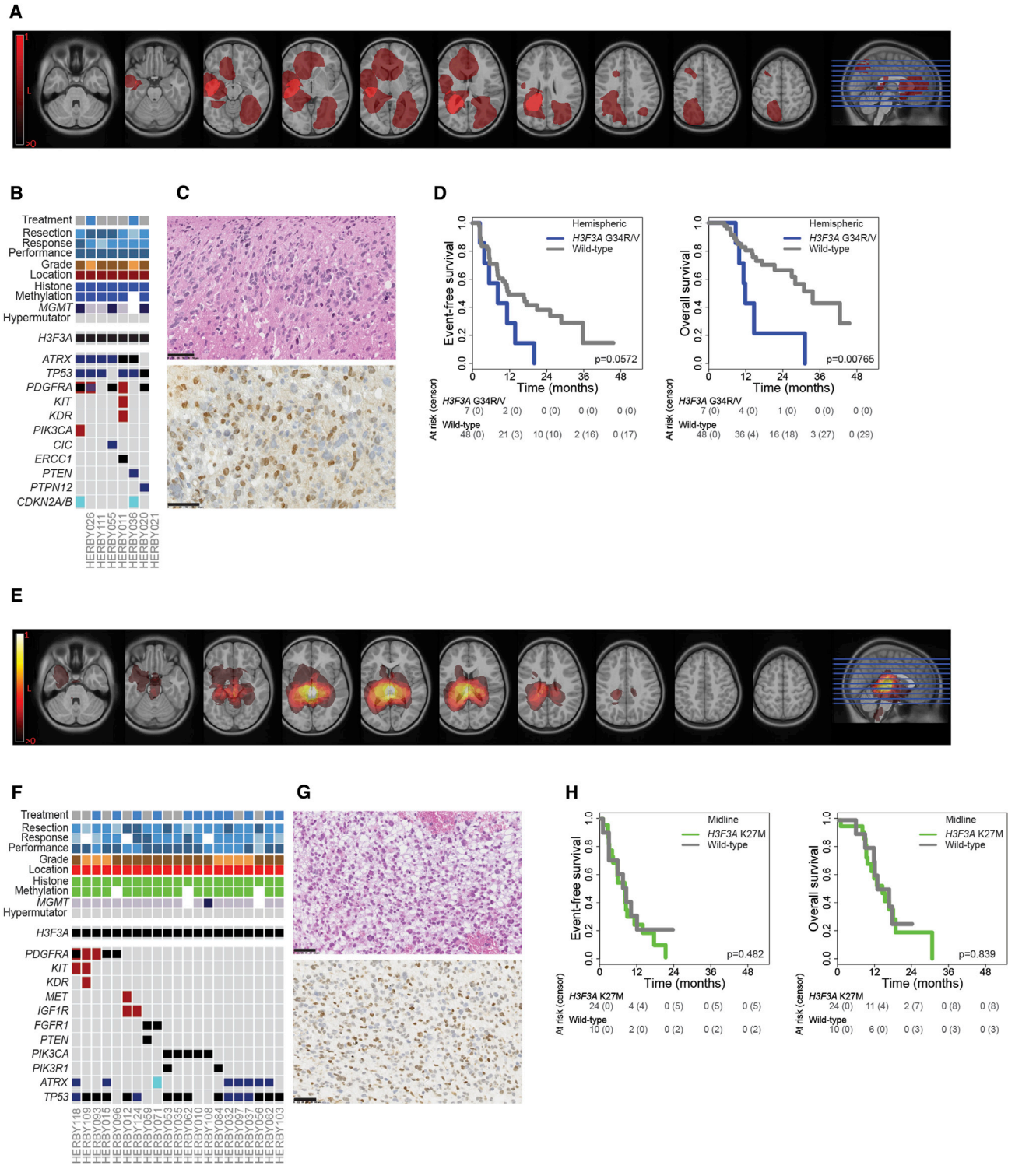


Figure 4. *H3F3A* Mutant Subgroups

(A–D) *H3F3A*_G34R/V. (A) Radiological tumor lesion map of 7 *H3F3A*_G34R/V cases. Brighter colored pixels indicate a higher probability of tumor incidence. (B) Integrated annotation of somatic mutations and DNA copy-number changes in *H3F3A*_G34R/V cases. Clinicopathological and molecular annotations are provided as bars according to the included key in Figure 3. (C) H&E (top) and immunohistochemistry (bottom) directed against H3.3G34R for HERBY020. Scale bars, 50 μm. (D) Kaplan-Meier plot of event-free and overall survival of 55 hemispheric cases (y axis) separated by *H3F3A* status, time given in months (x axis), p value calculated by the log rank test.

(legend continued on next page)

out of seven (86%) cases additionally harboring *TP53* and/or *ATRX* mutations (five out of seven, 71% both), while five out of seven (71%) also contained *PDGFRA* amplification and/or mutation (Figure 4B). There were no other recurrent mutations, although isolated instances of mutations in PI3K signaling (*PIK3CA* and *PTEN*) and DNA repair (*ERCC1*) were observed (Table S3). Histologically, there were four GBM, two AA with multinucleated cells, and one HGG, not otherwise specified (Figure 4C). Tumors were Olig2 negative (7/7) with strong nuclear accumulation of p53 (6/7). Across all tumors within this hemispheric subgroup, patients harboring these mutations trended toward a shorter EFS (median = 8.3 months; p = 0.0572, log rank test) and had a significantly shorter OS (median = 12.0 months; p = 0.00765, log rank test) (Figure 4D), although this association was lost when IDH1, PXA-like, and LGG-like tumors were excluded (p = 0.440 EFS and p = 0.139 OS, log rank test) (Figure S4A).

By contrast, K27M substitutions were restricted to midline regions (n = 24). Two patients had distinct, separate lesions in the thalamus and hypothalamus, while the remaining had central thalamic, midbrain, or cerebellar localization (Figure 4E). Fifteen out of 21 (71%) exome-sequenced cases carried additional amplifications/mutations in the receptor-tyrosine kinase (RTK)-PI3K pathway across a range of genes (*PDGFRA, MET, IGF1R, FGFR1, PTEN, PIK3CA,* and *PIK3R1*), with five out of six of the remaining tumors harboring *ATRX* mutation (Figure 4F; Table S3). There was strong immunoreactivity for H3K27M in 12/12 cases tested (Figure 4G). Although conferring a worse prognosis across the whole cohort (above), within midline locations there was no association with either EFS (median = 7.9 months; p = 0.482, log rank test) or OS (median = 14.2 months; p = 0.839, log rank test) (Figure 4H), nor any prognostic value for WHO-grade in K27M tumors (p = 0.646 EFS and p = 0.762 OS, log rank test) (Figure S4B).

Immune-Positive Subgroups Are Associated with MAPK and Benefit from the Addition of BEV

pHGGs with a high mutational burden had previously been described to have an elevated neoantigen load and a pronounced immune response (Bouffet et al., 2016). Four cases were classed as hypermutator, with a median somatic mutation count of 4,848 (range = 2,197–5,332; mutation rate 160–240 mutations/Mb) (Figure 5A, Table S1). Mutation signature analysis showed predominantly C > T transitions and hotspot somatic *POLE* mutations in three cases, and a somatic *POLD1* mutation in the fourth (Figure 5B). They were all categorized as GBM, and, in one case, the presumed mismatch repair deficiency was demonstrated by clear loss of MLH1 expression by immunohistochemistry (Figure 5C). This specimen had a heterogeneous immune phenotype, with a relatively high percentage of $CD8^+$ cells in the central area, and in three or four thin perivascular cuffs (Figure 5D; Table S4). These hypermutator cases had the highest percentage of $CD8^+$ cells ($p < 0.0001$ t test versus rest excluding PXA-like) (Figure 5E). Notably, PXA-like tumors were also significantly enriched for $CD8^+$ cells ($p < 0.0001$ t test versus rest excluding hypermutator). Of three HGG-WT cases with relatively high immune infiltrate, two had elevated somatic SNV counts (100–110 per case), while the third scored highly for the GBM_LYMPH_HI subgroup by methylation profiling. A formal histological assessment of tumor-infiltrating lymphocytes (TILs) confirmed the highest-scoring categories (lymphocytes scattered among tumor cells/in more than 50% of the tumor) to be almost exclusively present in hypermutator and PXA-like subgroups (Figure S5A; Table S4), as were those cases formally classified as an inflamed immune phenotype (Figure S5B; Table S4). Histone mutant tumors were notably immune cold as defined by a lack of CD8 immunoreactivity and an absence of TILs.

Nine cases classified by 450k methylation profiling as more similar to PXAs than HGGs. Five out of nine (55%) harbored *BRAF*_V600E mutations, with three out of five (60%) also containing *CDKN2A/B* deletions and/or *TERT* amplification or promoter mutation (C250T) (Figure 6A). Three of four of the remaining cases were instead found with somatic *NF1* mutation, often in concert with *TP53* (three out of four) and/or *ATRX* mutation (two out of four). Upon histological re-review according to WHO 2016 guidelines by the HERBY Pathology Committee, *BRAF*_V600E, cases were all found to comprise the epithelioid variant of grade IV GBM, while the *NF1* cases were all classified as the giant cell variant (Figure 6B). PXA-like tumors had a high degree of immune infiltrate, with cases exhibiting several perivascular cuffs of more than three layers of $CD8^+$ T cells, which were also scattered among tumor cells in more than 50% of the specimen (Figure 6C). They were found bilaterally restricted to temporo-parietal regions with a medial hemispheric predominance (Figure 6D).

Gene expression data from RNA-seq were available for a subset of samples, in which a CD8 T effector/T cell signature was found to correlate with CD8 positivity by immunohistochemistry ($r^2 = 0.49138$, p = 0.00523), with two hypermutator cases as outliers (Figure 7A). Although no PXA-like or *BRAF*_V600E mutant cases were included in this subset, these signatures were particularly evident in cases with predicted MAPK pathway-activating alterations in *NF1* (truncating frameshift/nonsense, disrupting translocation or predicted damaging missense), *NTRK2* (translocation or tandem duplication of kinase domains), and *FGFR1* (known activating hotspot mutation) (Figure 7B). Gene set enrichment analysis (GSEA) showed multiple enrichments for gene sets associated with T cell signaling and the immune response (e.g., PID_CD8_TCR_ DOWNSTREAM_PATHWAY, enrichment score = 0.594, nominal p = 0.0020; KEGG_T_CELL_RECEPTOR_SIGNALING_PATHWAY, enrichment score = 0.532, nominal p = 0.0297) and inflammatory-related MAPK signaling (e.g., ST_JNK_MAPK_PATHWAY, enrichment score = 0.466, nominal p = 0.0332; REACTOME_GRB2_SOS_PROVIDES_LINKAGE_TO_MAPK_SIGNALING_FOR_INTERGRINS,

(E–H) *H3F3A*_K27M. (E) Radiological tumor lesion map of 21 *H3F3A*_K27M cases. Brighter colored pixels indicate a higher probability of tumor incidence. (F) Integrated data from *H3F3A*_K27M cases. Clinicopathological and molecular annotations are provided as bars according to the included key in Figure 3. (G) H&E (top) and immunohistochemistry (bottom) directed against H3.3K27M for HERBY015. Scale bars 50 μm. (H) Kaplan-Meier plot of event-free and overall survival of 34 midline cases (y axis) separated by *H3F3A* status, time given in months (x axis), p value calculated by the log rank test.
See also Figure S4.

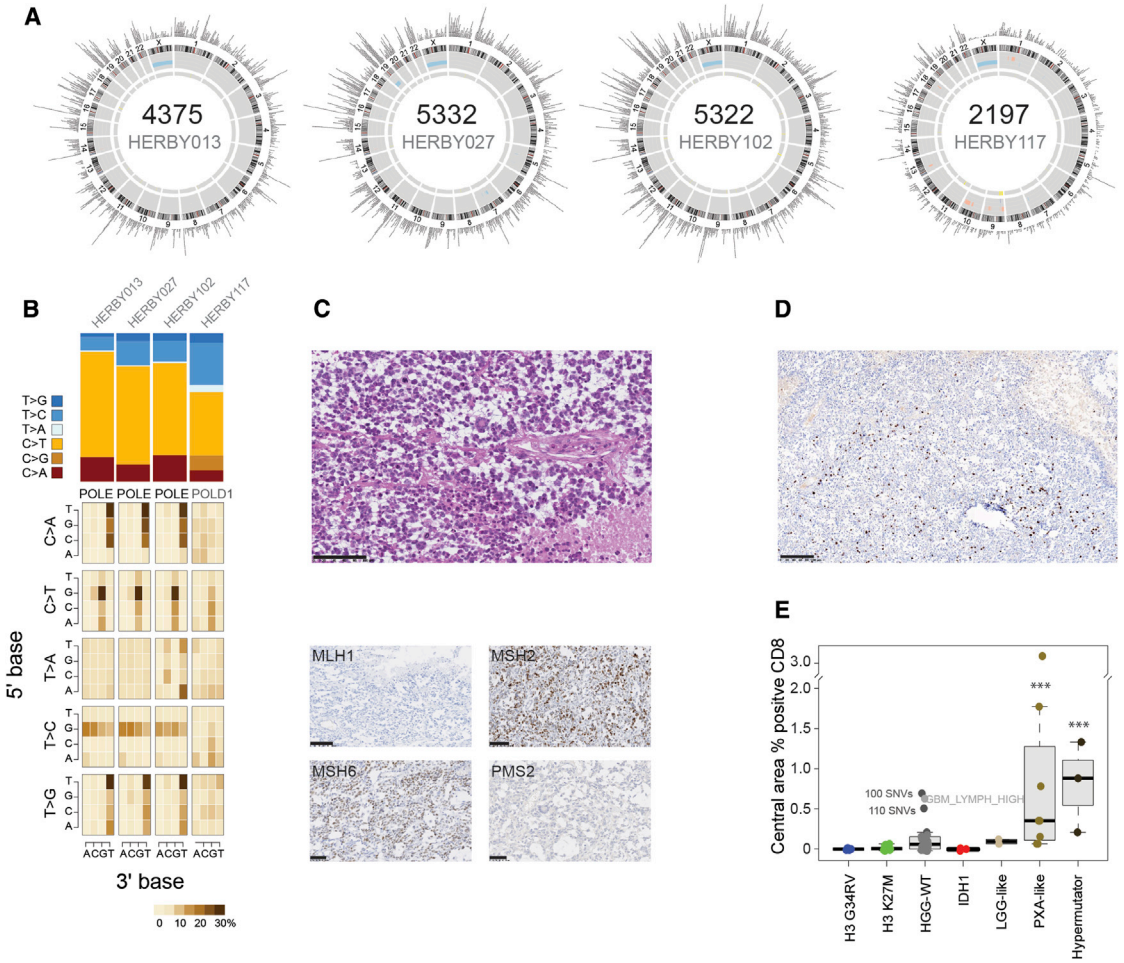


Figure 5. Hypermutator Cases

(A) Circos plots for four hypermutator cases. In each case, plots provide somatic SNVs and insertion/deletions on the outer ring, DNA copy-number changes (dark red, amplification; red, gain; dark blue, deletion; blue, loss), and loss of heterozygosity (yellow) on the inner rings, and intra- (orange) and inter- (blue) chromosomal translocations inside the circle.

(B) Mutation signatures. Top: simple stacked barplot representation of the proportion of mutation types observed in individual hypermutator cases and the remaining accumulated dataset. Base changes given in the key. Bottom: mutation context given for each of the 96 mutated trinucleotides, represented by heatmap. The base located 5′ to each mutated base is shown on the vertical axis, and the 3′ base is on the horizontal axis.

(C) H&E (top) and protein expression of mismatch repair proteins (bottom panel, clockwise from top left: MLH1, MSH2, MSH6, and PMS2) as assessed by immunohistochemistry in a glioblastoma with 5,322 somatic mutations (HERBY102). Scale bars, 100 μm (H&E, MLH1, MSH2) or 50 μm (MSH6, PMS2).

(D) CD8 expression in T cells by immunohistochemistry in HERBY102. Scale bar, 100 μm.

(E) Boxplot of percentage of $CD8^+$ cells in the central tumor region separated by subgroup. The thick line within the box is the median, the lower and upper limits of the boxes represent the first and third quartiles, and the whiskers 1.5× the interquartile range. ***$p < 0.0001$, t test.

See also Figure S5 and Table S4.

enrichment score = 0.581, nominal p = 0.0239) (Table S5), with mean T effector/T cell signature significantly higher than in cases without MAPK alterations (p = 0.0039, t test) (Figure 7C). This was validated in a restricted cohort of n = 59 patients from our retrospective analysis designed to approximate the HERBY cohort (i.e., non-brainstem pHGGs aged 3–18 years), in which we observed a significantly elevated T effector/T cell gene expression signature in MAPK-altered samples (p = 0.0018, t test) (Figures S6A and S6B), further demonstrated by GSEA (e.g., PID_CD8_TCR_ DOWNSTREAM_PATHWAY, enrichment score = 0.551, nominal p = 0.0099; BIOCARTA_TCR_PATHWAY enrichment score = 0.520, nominal p = 0.0305) (Table S5).

We also explored other immune-related gene expression signatures in the RNA-seq data, and noted a trend toward an elevated macrophage M2 response in MAPK-altered tumors (p = 0.0810, t test) (Figures S6C and S6D). We performed CD68 staining for a limited number of cases with a histologically defined immune response (n = 11), and found a heterogeneously distributed tumor-associated macrophages (TAM) component, comprising either TAM-free tumoral areas, or rich TAM areas especially around necrotic foci or associated with perivascular lymphocytes (Figure S6E). *BRAF*_V600E cases had $CD68^+$ TAM more diffusely intermingled with tumor cells (Figure S6E). Although the presence of macrophage infiltration has previously

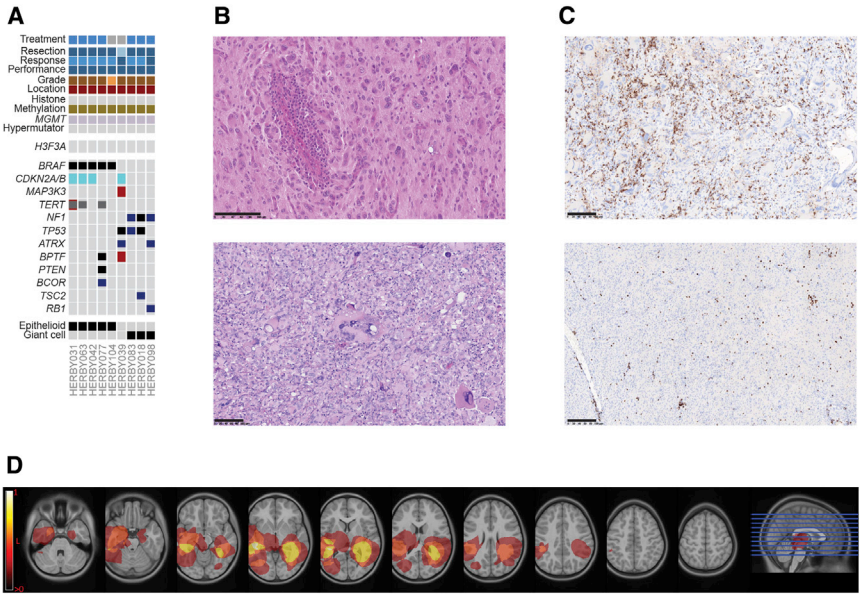


Figure 6. PXA-like Tumors

(A) Integrated annotation of somatic mutations and DNA copy-number changes in nine samples classifying as PXA-like. Clinicopathological and molecular annotations are provided as bars according to the included key in Figure 3.

(B) H&E staining of an epithelioid (top, HERBY104) and giant cell (bottom, HERBY098) variant of glioblastoma. Scale bars, 100 μm.

(C) CD8 immunohistochemistry for the same cases as (B), marking $CD8^+$ T cells. Scale bars, 100 μm.

(D) Radiological tumor lesion map of PXA-like cases. Brighter colored pixels indicate a higher probability of tumor incidence.

been reported in diffuse intrinsic pontine glioma (Caretti et al., 2014), we observed no association of the M2 gene expression signature with K27M midline tumors in our cohort ($p = 0.965$).

With the primary efficacy analysis failing to demonstrate survival differences between the two arms, we explored whether the molecular-profiling data could identify subsets of patients who may have benefited from the addition of BEV to the standard chemoradiotherapy. We performed a univariate Cox regression analysis on methylation-based molecular subgroups, as well as individual gene-level alterations, and found none which had significant difference in outcome between investigational arms (Figure S7A), although *NF1*, *PDGFRA*, and *TP53* were adverse prognostic markers across the whole cohort (Figures S7B–S7E). Notably, however, while not predictive in the TMZ/RT arm (Figure 7D), the presence of high levels of $CD8^+$ T cells within the central tumor area conferred a significantly better OS in children receiving the addition of BEV ($p = 0.0404$, log rank test) (Figure 7E). Fitting a Cox interaction model (with proportional hazards confirmed by calculating Schoenfeld residuals, $p = 0.268$), high CD8 levels trended toward predictivity of response to TMZ/RT + BEV, with a hazard ratio of 0.360 ($p = 0.066$). In keeping with the strong subgroup associations described, 17/18 cases with high levels of $CD8^+$ T cells are hemispheric (representing 38.6% cases in this location). Within hemispheric tumors only, there is a significant interaction between high CD8 and BEV (hazard ratio [HR] = 0.251, $p = 0.024$).

DISCUSSION

The HERBY trial opened in October 2011, prior to the discovery of histone gene mutations in 2012 (Schwartzentruber et al., 2012), and the more extensive genome sequencing published in 2014 (Wu et al., 2014), and there were no molecular markers incorporated into trial design. The present correlative biology study demonstrates the extent of the heterogeneous population of tumors, with widely differing biological drivers and clinicopathological features that were included in the cohort. An important example is the four adolescent patients with *IDH1* mutation, who had a significantly better clinical outcome, and may now be thought of as the lower age limit of an adult subgroup rather than pHGGs. Such patients will likely be included in future *IDH1*-focussed trials across adult and pediatric oncology services in order to maximize the possibility of positive trial outcomes. Similarly, within the HERBY cohort were seven patients with *BRAF*_V600E mutations, who would now be candidates for upfront trials of targeted inhibitors such as vemurafenib (Bautista et al., 2014; Robinson et al., 2014), and four hypermutator patients, with somatic *POLE*/*POLD1* mutations and likely harboring biallelic mismatch repair deficiency syndrome, who may benefit from immune checkpoint inhibitors such as nivolumab (Bouffet et al., 2016). In both instances, early clinical data show remarkably promising results and there would be little justification in their continued inclusion in catch-all HGG trials.

Notably, both hypermutator cases, and those biologically resembling PXAs, the latter of which harbored either *BRAF*- or *NF1*-driven MAP-kinase alterations, were found to be the most immunogenic in terms of $CD8^+$ T cells/TILs, including a significantly elevated CD8 effector T cell gene expression signature. This is important given reports in adult GBM of an immunosuppressive phenotype associated with an elevated $CD8^+$ regulatory T cell immune infiltrate (Kmiecik et al., 2013). The MAPK-altered hypermutator and PXA-like pediatric cases in the present cohort were found to benefit from the addition of BEV to TMZ/RT as regard to overall, although not EFS, the first and only such predictive biomarker identified in this patient population. The immunomodulatory effects of anti-angiogenic therapies have been previously demonstrated in other cancer types (Elamin et al., 2015), with VEGF-A shown to play an important role in the induction of an immunosuppressive environment (Gabrilovich et al., 1996), through increasing PD-L1 and other inhibitory checkpoints involved in $CD8^+$ T cell exhaustion.

An important implication of these observations is the possibility of overcoming resistance to BRAF/MEK-targeting therapies by combining them with BEV and/or other immune therapies in these subgroups, as has been suggested in melanoma (Hu-Lieskovan et al., 2015). Such a strategy is, however, complicated by the diverse roles on T cell function played by MAPK signaling. Although involved in the regulation of T cell proliferation and

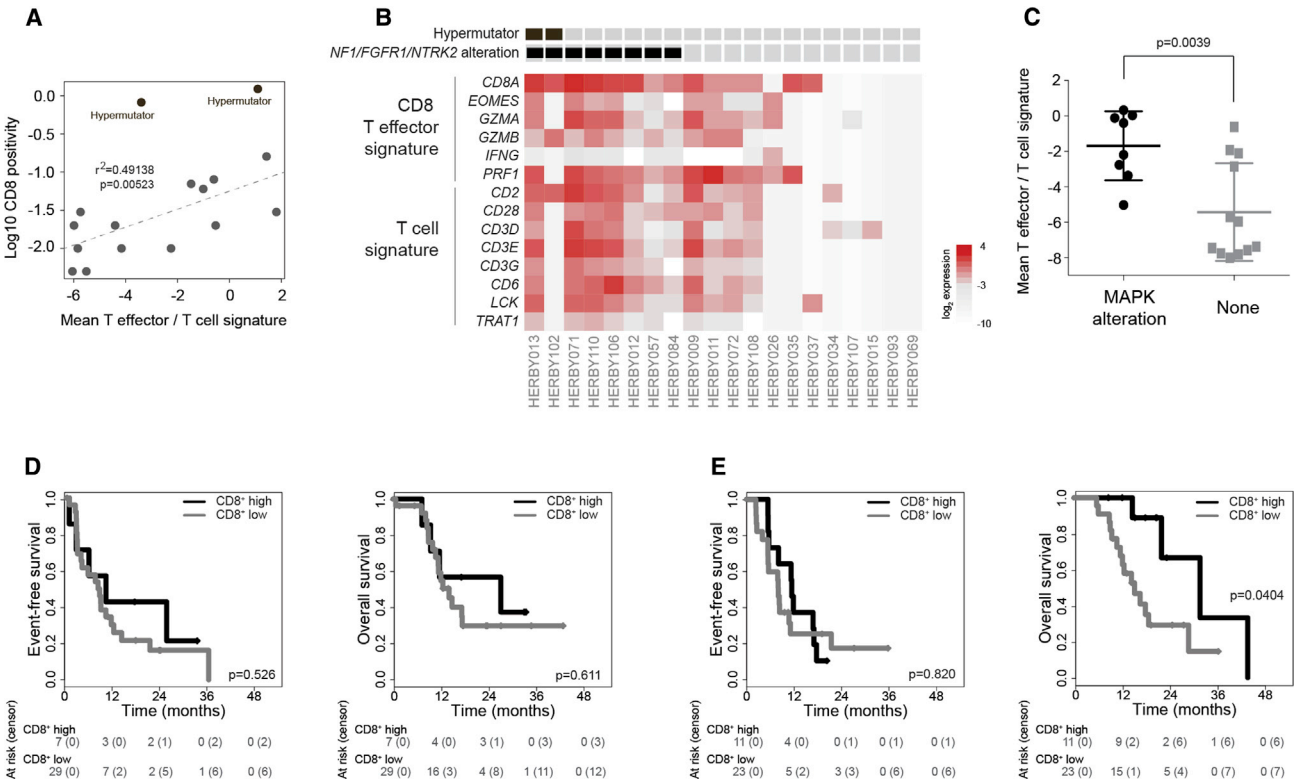


Figure 7. Immune Signatures and Response to BEV

(A) Mean T effector/T cell gene expression values (plotted as $\log_2$, x axis) were correlated with CD8 immunoreactivity (plotted as $\log_{10}$, y axis) for 18 cases. Two cases were scored as 0% by immunohistochemistry and were not plotted (expression values −7.37 and −7.58).

(B) Gene expression signatures for CD8 T effector and T cells plotted as a heatmap from 20 cases with RNA-seq data. Hypermutator cases and those with MAPK alterations including *NF1*, *FGFR1*, or *NTRK2* are annotated.

(C) Plot of T effector/T cell gene expression values in MAPK-altered samples compared with those without. Horizontal bar represents the mean, error bars the SD.

(D) Kaplan-Meier plot of event-free and overall survival of 36 cases (y axis) treated with TMZ/RT, separated by levels of $CD8^+$ T cells, time given in months (x axis), and p value calculated by the log rank test.

(E) Kaplan-Meier plot of event-free and overall survival of 34 cases (y axis) treated with TMZ/RT plus BEV, separated by levels of $CD8^+$ T cells, time given in months (x axis), and p value calculated by the log rank test.

See also Figures S6 and S7 and Table S5.

survival (D'Souza et al., 2008), selective BRAF inhibitors have been shown to increase $CD8^+$ lymphocytes in human metastatic melanoma models (Wilmott et al., 2012). Crucially, although MEK inhibition has been demonstrated to block naive $CD8^+$ T cell priming in a colon cancer model, the number of $CD8^+$ effector T cells within the tumor were increased, and could potentiate immune checkpoint therapy (Ebert et al., 2016). A limitation of the present study is the small numbers in this *post-hoc* analysis, and the benefit that patients with these biological subgroups may derive from an immunomodulatory mechanism of BEV would need to be tested in the prospective setting. This is a challenge given that these patients represent approximately 10%–15% of an already rare disease; however, international collaborative trials groups (such as those in HERBY represent), already recruit hypermutator and MAPK-altered HGGs in this population for appropriately targeted therapies (NCT02992964, NCT02684058). Equally importantly, the histone H3 mutant subgroups, which represent a substantial proportion of patients in this age group (Mackay et al., 2017) were found to be very poorly immunogenic, confirming a previous study in resectable malignant brainstem gliomas in children and adults with K27M mutations (Zhang et al., 2017), and further negating the likelihood of clinical response to such therapies.

A key observation is the high prevalence of tumors occurring outside the cerebral hemispheres harboring histone mutations, included in the most recent 2016 WHO classification system as a separate entity called diffuse midline glioma with H3K27M mutation. These represented 27% of the assessed population, and had a particularly poor outcome, as did histone WT midline cases. These tumors only rarely have *MGMT* promoter methylation, and have consistently proved refractory to TMZ and other chemotherapeutic agents (Jones et al., 2016). Critically, midline K27M tumors classified histologically as either grade 3 or 4 according to the WHO 2007 classification had no difference in clinical outcome within the HERBY cohort, and with the caveat of small numbers, support the current 2016 guidelines to assign all such tumors as grade 4 on the basis of their location and molecular findings.

More surprising was the poor outcome observed for patients with *H3F3A*_G34R/V mutations. A previous study reported G34R mutations to convey a better prognosis in respect of OS

(HR = 0.49, p = 0.01), although this was not significant in multivariate analysis (p = 0.84) (Korshunov et al., 2015), although the present study explores this in a consistently treated and well-annotated clinical trial setting. In the HERBY study, as well as in published work (Korshunov et al., 2015; Mackay et al., 2017), the *H3F3A*_G34R/V mutation is associated with a high frequency of *MGMT* methylation. Notably, *MGMT* promoter methylation itself was not predictive in this trial cohort of all patients receiving radiochemotherapy with TMZ, again demonstrating differences with the adult disease, and questioning the continued use of protocols extrapolated from the adult setting. It is clear that histone mutations represent clearly defined entities within an umbrella HGG classification in the pediatric setting, and given their profound impact on chromatin modifications, will require therapeutic development and clinical trials distinct from histone WT cases.

Within the remaining cases of pHGG were a small proportion whose methylation profiles were more similar to other lesions. The LGG-like cases were in young patients with longer EFS, while two additional cases had methylation classifier scores strongly suggestive of recently described entities coming from the study of tumors formally diagnosed as CNS primitive neuroectodermal tumors (Sturm et al., 2016). Integration with histological features and determination of the presence of marker gene fusions events for these entities (CNS HGNET-MN1, CNS NB-FOXR2) will be key in future studies.

Although there have been several early-phase and anecdotal studies of BEV in pHGG (Benesch et al., 2008; Friedman et al., 2013; Gururangan et al., 2010; Narayana et al., 2010; Salloum et al., 2015), none have included biological information on the patients treated. Aside from CD8 immunoreactivity, we did not identify any additional molecular markers of response. In similarly designed adult studies of BEV, gene expression-defined proneural (Sandmann et al., 2015) or mesenchymal (Sulman et al., 2013) subgroups of GBM have been reported to confer a significant OS advantage. Although we have not undertaken gene expression studies across our cohort, adult and pediatric cases expressing proneural genes appear to have distinct genetic and epigenetic drivers (Sturm et al., 2012), the mesenchymal subclass is rare in children (Sturm et al., 2012), and unlike the adult studies, we observed no EFS advantage of BEV in HERBY (Grill et al., 2018).

In conclusion, integrated molecular profiling of the HERBY sample cohort has demonstrated the biological and clinicopathological diversity of the term HGG in the pediatric setting, suspected but not confirmed at the onset of the trial. While there are several distinct subgroups for which there is strong rationale for bespoke future clinical studies, a large proportion of pHGG cases continue to defy improvements in survival and lack a clear path forward. Although BEV was not associated with better outcome in this trial, the extensive biological, pathological, and radiological ancillary research programs ongoing within HERBY aim to provide an integrated assessment of disease pathogenesis and treatment response.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and five tables and can be found with this article online at https://doi.org/10.1016/j.ccell.2018.04.004.

ACKNOWLEDGMENTS

The HERBY clinical trial was carried out as a collaboration between the European Innovative Therapies for Children with Cancer (ITCC) consortium, the SIOPE Brain Tumor Group, the Australian Children's Cancer Trials Group (ACCT), and the C-17 Council (Canada). The translational research study was funded by an unrestricted grant from F. Hoffmann-La Roche Ltd allied to the HERBY trial (study number BO25041; clinicaltrials.gov NCT01390948). The funder provided infrastructure for sample acquisition, shipping, storage, and record keeping. The funder had no role in study design, data collection, data analysis, data interpretation, or writing of the report. We acknowledge the excellent technical assistance of staff at the Tumor Profiling Unit at the Institute of Cancer Research, London, UK, Sainte-Anne Hospital, Paris, France, and HistoGeneX, Antwerp, Belgium. A.M., A.B., V.M., E.I.D., and C.J. acknowledge NHS funding to the NIHR Biomedical Research Center at The Royal Marsden and the ICR; T.S.J. acknowledges NHS funding to the NIHR Biomedical Research Center at Great Ormond Street Hospital. M.L.G. recognizes the support from the Associazione per la Ricerca sui Tumori Cerebrali del Bambino. We would like to thank the local investigators, pathologists and radiologists who took part in the HERBY study, and the patients and families who consented to inclusion in the translational research program.

AUTHOR CONTRIBUTIONS

C.J., P.V., J.Grill, T.W., S.M.P., R.T., M.D.T., and G.V. conceived the study. A.M., J.B.-V., and C.J. analyzed data. A.B., V.M., D.T., W.J., D.C., and E.I.D. performed molecular analysis. D.R., P.S.M., and T.J. carried out radiological analysis. F.G., C.H., T.P., T.S.J., D.F.-B., and P.V. carried out histopathological analysis. M.M., M.L.G., and J. Grill provided samples. P.R., A.J.W., and A.R. constructed analytical and visualization tools and databases. H.S. and J. Garcia provided logistical support. A.M. and C.J. wrote the manuscript. All authors approved the manuscript.

DECLARATION OF INTERESTS

C.J. received a research grant from F. Hoffmann-La Roche Ltd. C.J., P.V., M.M., J. Grill, G.V., D.R., P.M., and T.J. received consultancy fees from F.

Hoffmann-La Roche Ltd. F.G., C.H., T.P., T.S.J., D.F.-B., and P.V. received travel reimbursement from F. Hoffmann-La Roche. J.B.-V., H.S., and J. Garcia are employees of F. Hoffmann-La Roche Ltd. R.T. and M.D.T. are employees of Genentech.

Received: January 12, 2018
Revised: February 28, 2018
Accepted: April 10, 2018
Published: May 14, 2018

REFERENCES

Bautista, F., Paci, A., Minard-Colin, V., Dufour, C., Grill, J., Lacroix, L., Varlet, P., Valteau-Couanet, D., and Geoerger, B. (2014). Vemurafenib in pediatric patients with BRAFV600E mutated high-grade gliomas. Pediatr. Blood Cancer *61*, 1101–1103.

Benesch, M., Windelberg, M., Sauseng, W., Witt, V., Fleischhack, G., Lackner, H., Gadner, H., Bode, U., and Urban, C. (2008). Compassionate use of bevacizumab (Avastin) in children and young adults with refractory or recurrent solid tumors. Ann. Oncol. *19*, 807–813.

Bjerke, L., Mackay, A., Nandhabalan, M., Burford, A., Jury, A., Popov, S., Bax, D.A., Carvalho, D., Taylor, K.R., Vinci, M., et al. (2013). Histone H3.3. mutations drive pediatric glioblastoma through upregulation of MYCN. Cancer Discov. *3*, 512–519.

Bouffet, E., Larouche, V., Campbell, B.B., Merico, D., de Borja, R., Aronson, M., Durno, C., Krueger, J., Cabric, V., Ramaswamy, V., et al. (2016). Immune checkpoint inhibition for hypermutant glioblastoma multiforme resulting from germline biallelic mismatch repair deficiency. J. Clin. Oncol. *34*, 2206–2211.

Brouwer-Visser, J., Cheng, W.Y., Bauer-Mehren, A., Maisel, D., Lechner, K., Andersson, E., Dudley, J.T., and Milletti, F. (2018). Regulatory T-cell genes drive altered immune microenvironment in adult solid cancers and allow for immune contextual patient subtyping. Cancer Epidemiol. Biomarkers Prev. *27*, 103–112.

Capper, D., Jones, D.T.W., Sill, M., Hovestadt, V., Schrimpf, D., Sturm, D., Koelsche, C., Sahm, F., Chavez, L., Reuss, D.E., et al. (2018). DNA methylation-based classification of central nervous system tumours. Nature *555*, 469–474.

Caretti, V., Sewing, A.C., Lagerweij, T., Schellen, P., Bugiani, M., Jansen, M.H., van Vuurden, D.G., Navis, A.C., Horsman, I., Vandertop, W.P., et al. (2014). Human pontine glioma cells can induce murine tumors. Acta Neuropathol. *127*, 897–909.

Cohen, K.J., Pollack, I.F., Zhou, T., Buxton, A., Holmes, E.J., Burger, P.C., Brat, D.J., Rosenblum, M.K., Hamilton, R.L., Lavey, R.S., and Heideman, R.L. (2011). Temozolomide in the treatment of high-grade gliomas in children: a report from the Children's Oncology Group. Neuro Oncol. *13*, 317–323.

D'Souza, W.N., Chang, C.F., Fischer, A.M., Li, M., and Hedrick, S.M. (2008). The Erk2 MAPK regulates CD8 T cell proliferation and survival. J. Immunol. *181*, 7617–7629.

Ebert, P.J.R., Cheung, J., Yang, Y., McNamara, E., Hong, R., Moskalenko, M., Gould, S.E., Maecker, H., Irving, B.A., Kim, J.M., et al. (2016). MAP kinase inhibition promotes T cell and anti-tumor activity in combination with PD-L1 checkpoint blockade. Immunity *44*, 609–621.

Eisenstat, D.D., Pollack, I.F., Demers, A., Sapp, M.V., Lambert, P., Weisfeld-Adams, J.D., Burger, P.C., Gilles, F., Davis, R.L., Packer, R., et al. (2015). Impact of tumor location and pathological discordance on survival of children with midline high-grade gliomas treated on Children's Cancer Group high-grade glioma study CCG-945. J. Neurooncol. *121*, 573–581.

Elamin, Y.Y., Rafee, S., Toomey, S., and Hennessy, B.T. (2015). Immune effects of bevacizumab: killing two birds with one stone. Cancer Microenviron. *8*, 15–21.

Finlay, J.L., Boyett, J.M., Yates, A.J., Wisoff, J.H., Milstein, J.M., Geyer, J.R., Bertolone, S.J., McGuire, P., Cherlow, J.M., Tefft, M., et al. (1995). Randomized phase III trial in childhood high-grade astrocytoma comparing vincristine, lomustine, and prednisone with the eight-drugs-in-1-day regimen. Children's Cancer Group. J. Clin. Oncol. *13*, 112–123.

Fonov, V., Evans, A.C., Botteron, K., Almli, C.R., McKinstry, R.C., and Collins, D.L.; Brain Development Cooperative Group (2011). Unbiased average age-appropriate atlases for pediatric studies. Neuroimage *54*, 313–327.

Friedman, G.K., Spiller, S.E., Harrison, D.K., Fiveash, J.B., and Reddy, A.T. (2013). Treatment of children with glioblastoma with conformal radiation, temozolomide, and bevacizumab as adjuncts to surgical resection. J. Pediatr. Hematol. Oncol. *35*, e123–126.

Gabrilovich, D.I., Chen, H.L., Girgis, K.R., Cunningham, H.T., Meny, G.M., Nadaf, S., Kavanaugh, D., and Carbone, D.P. (1996). Production of vascular endothelial growth factor by human tumors inhibits the functional maturation of dendritic cells. Nat. Med. *2*, 1096–1103.

Gilles, F.H., Tavare, C.J., Becker, L.E., Burger, P.C., Yates, A.J., Pollack, I.F., and Finlay, J.L. (2008). Pathologist interobserver variability of histologic features in childhood brain tumors: results from the CCG-945 study. Pediatr. Dev. Pathol. *11*, 108–117.

Grill, J., Massimino, M., Bouffet, E., Azizi, A.A., McCowage, G., Canete, A., Saran, F., Le Deley, M.C., Varlet, P., Morgan, P.S., et al. (2018). Phase II, open-label, randomized, multicenter trial (HERBY) of bevacizumab in pediatric patients with newly diagnosed high-grade glioma. J. Clin. Oncol. *36*, 951–958.

Gururangan, S., Chi, S.N., Young Poussaint, T., Onar-Thomas, A., Gilbertson, R.J., Vajapeyam, S., Friedman, H.S., Packer, R.J., Rood, B.N., Boyett, J.M., and Kun, L.E. (2010). Lack of efficacy of bevacizumab plus irinotecan in children with recurrent malignant glioma and diffuse brainstem glioma: a Pediatric Brain Tumor Consortium study. J. Clin. Oncol. *28*, 3069–3075.

Hatae, R., Hata, N., Suzuki, S.O., Yoshimoto, K., Kuga, D., Murata, H., Akagi, Y., Sangatsuda, Y., Iwaki, T., Mizoguchi, M., and Iihara, K. (2016). A comprehensive analysis identifies BRAF hotspot mutations associated with gliomas with peculiar epithelial morphology. Neuropathology *37*, 191–199.

Hu-Lieskovan, S., Mok, S., Homet Moreno, B., Tsoi, J., Robert, L., Goedert, L., Pinheiro, E.M., Koya, R.C., Graeber, T.G., Comin-Anduix, B., and Ribas, A. (2015). Improved antitumor activity of immunotherapy with BRAF and MEK inhibitors in BRAF(V600E) melanoma. Sci. Transl. Med. *7*, 279ra241.

Jakacki, R.I., Cohen, K.J., Buxton, A., Krailo, M.D., Burger, P.C., Rosenblum, M.K., Brat, D.J., Hamilton, R.L., Eckel, S.P., Zhou, T., et al. (2016). Phase 2 study of concurrent radiotherapy and temozolomide followed by temozolomide and lomustine in the treatment of children with high-grade glioma: a report of the Children's Oncology Group ACNS0423 study. Neuro Oncol. *18*, 1442–1450.

Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage *17*, 825–841.

Jones, C., and Baker, S.J. (2014). Unique genetic and epigenetic mechanisms driving paediatric diffuse high-grade glioma. Nat. Rev. Cancer *14.*, 651–661.

Jones, C., Karajannis, M.A., Jones, D.T., Kieran, M.W., Monje, M., Baker, S.J., Becher, O.J., Cho, Y.J., Gupta, N., Hawkins, C., et al. (2016). Pediatric high-grade glioma: biologically and clinically in need of new thinking. Neuro Oncol. *19*, 153–161.

Karremann, M., Gielen, G.H., Hoffmann, M., Wiese, M., Colditz, N., Warmuth-Metz, M., Bison, B., Claviez, A., van Vuurden, D.G., von Bueren, A.O., et al. (2018). Diffuse high-grade gliomas with H3 K27M mutations carry a dismal prognosis independent of tumor location. Neuro Oncol. *20*, 123–131.

Khuong-Quang, D.A., Buczkowicz, P., Rakopoulos, P., Liu, X.Y., Fontebasso, A.M., Bouffet, E., Bartels, U., Albrecht, S., Schwartzentruber, J., Letourneau, L., et al. (2012). K27M mutation in histone H3.3 defines clinically and biologically distinct subgroups of pediatric diffuse intrinsic pontine gliomas. Acta Neuropathol. *124*, 439–447.

Kmiecik, J., Poli, A., Brons, N.H., Waha, A., Eide, G.E., Enger, P.O., Zimmer, J., and Chekenya, M. (2013). Elevated CD3+ and CD8+ tumor-infiltrating immune cells correlate with prolonged survival in glioblastoma patients despite integrated immunosuppressive mechanisms in the tumor microenvironment and at the systemic level. J. Neuroimmunol. *264*, 71–83.

Korshunov, A., Ryzhova, M., Hovestadt, V., Bender, S., Sturm, D., Capper, D., Meyer, J., Schrimpf, D., Kool, M., Northcott, P.A., et al. (2015). Integrated analysis of pediatric glioblastoma reveals a subset of biologically favorable

tumors with associated molecular prognostic markers. Acta Neuropathol. *129*, 669–678.

Mackay, A., Burford, A., Carvalho, D., Izquierdo, E., Fazal-Salom, J., Taylor, K.R., Bjerke, L., Clarke, M., Vinci, M., Nandhabalan, M., et al. (2017). Integrated molecular meta-analysis of 1,000 pediatric high-grade and diffuse intrinsic pontine glioma. Cancer Cell *32*, 520–537 e525.

Narayana, A., Kunnakkat, S., Chacko-Mathew, J., Gardner, S., Karajannis, M., Raza, S., Wisoff, J., Weiner, H., Harter, D., and Allen, J. (2010). Bevacizumab in recurrent high-grade pediatric gliomas. Neuro Oncol. *12*, 985–990.

Ostrom, Q.T., Gittleman, H., Fulop, J., Liu, M., Blanda, R., Kromer, C., Wolinsky, Y., Kruchko, C., and Barnholtz-Sloan, J.S. (2015). CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2008–2012. Neuro Oncol. *17* (*Suppl 4*), iv1–iv62.

Palma, L., Di Lorenzo, N., and Guidetti, B. (1978). Lymphocytic infiltrates in primary glioblastomas and recidivous gliomas. Incidence, fate, and relevance to prognosis in 228 operated cases. J. Neurosurg. *49*, 854–861.

Paugh, B.S., Qu, C., Jones, C., Liu, Z., Adamowicz-Brice, M., Zhang, J., Bax, D.A., Coyle, B., Barrow, J., Hargrave, D., et al. (2010). Integrated molecular genetic profiling of pediatric high-grade gliomas reveals key differences with the adult disease. J. Clin. Oncol. *28*, 3061–3068.

Pollack, I.F., Boyett, J.M., Yates, A.J., Burger, P.C., Gilles, F.H., Davis, R.L., and Finlay, J.L.; Children's Cancer Group (2003a). The influence of central review on outcome associations in childhood malignant gliomas: results from the CCG-945 experience. Neuro Oncol. *5*, 197–207.

Pollack, I.F., Finkelstein, S.D., Burnham, J., Hamilton, R.L., Yates, A.J., Holmes, E.J., Boyett, J.M., and Finlay, J.L. (2003b). Association between chromosome 1p and 19q loss and outcome in pediatric malignant gliomas: results from the CCG-945 cohort. Pediatr. Neurosurg. *39*, 114–121.

Pollack, I.F., Finkelstein, S.D., Woods, J., Burnham, J., Holmes, E.J., Hamilton, R.L., Yates, A.J., Boyett, J.M., Finlay, J.L., and Sposto, R.; Children's Cancer Group (2002). Expression of p53 and prognosis in children with malignant gliomas. N. Engl. J. Med. *346*, 420–427.

Pollack, I.F., Hamilton, R.L., Burger, P.C., Brat, D.J., Rosenblum, M.K., Murdoch, G.H., Nikiforova, M.N., Holmes, E.J., Zhou, T., Cohen, K.J., et al. (2010a). Akt activation is a common event in pediatric malignant gliomas and a potential adverse prognostic marker: a report from the Children's Oncology Group. J. Neurooncol. *99*, 155–163.

Pollack, I.F., Hamilton, R.L., Sobol, R.W., Burnham, J., Yates, A.J., Holmes, E.J., Zhou, T., and Finlay, J.L. (2006). O6-methylguanine-DNA methyltransferase expression strongly correlates with outcome in childhood malignant gliomas: results from the CCG-945 Cohort. J. Clin. Oncol. *24*, 3431–3437.

Pollack, I.F., Hamilton, R.L., Sobol, R.W., Nikiforova, M.N., Lyons-Weiler, M.A., LaFramboise, W.A., Burger, P.C., Brat, D.J., Rosenblum, M.K., Holmes, E.J., et al. (2011). IDH1 mutations are common in malignant gliomas arising in adolescents: a report from the Children's Oncology Group. Childs Nerv. Syst. *27*, 87–94.

Pollack, I.F., Hamilton, R.L., Sobol, R.W., Nikiforova, M.N., Nikiforov, Y.E., Lyons-Weiler, M.A., LaFramboise, W.A., Burger, P.C., Brat, D.J., Rosenblum, M.K., et al. (2010b). Mismatch repair deficiency is an uncommon mechanism of alkylator resistance in pediatric malignant gliomas: a report from the Children's Oncology Group. Pediatr. Blood Cancer *55*, 1066–1071.

Robinson, G.W., Orr, B.A., and Gajjar, A. (2014). Complete clinical regression of a BRAF V600E-mutant pediatric glioblastoma multiforme after BRAF inhibitor therapy. BMC Cancer *14*, 258.

Rorden, C., Karnath, H.O., and Bonilha, L. (2007). Improving lesion-symptom mapping. J. Cogn. Neurosci. *19*, 1081–1088.

Rutledge, W.C., Kong, J., Gao, J., Gutman, D.A., Cooper, L.A., Appin, C., Park, Y., Scarpace, L., Mikkelsen, T., Cohen, M.L., et al. (2013). Tumor-infiltrating lymphocytes in glioblastoma are associated with specific genomic alterations and related to transcriptional class. Clin. Cancer Res. *19*, 4951–4960.

Salloum, R., DeWire, M., Lane, A., Goldman, S., Hummel, T., Chow, L., Miles, L., Sutton, M., Stevenson, C., Fouladi, M., and Leach, J. (2015). Patterns of progression in pediatric patients with high-grade glioma or diffuse intrinsic pontine glioma treated with Bevacizumab-based therapy at diagnosis. J. Neurooncol. *121*, 591–598.

Sandmann, T., Bourgon, R., Garcia, J., Li, C., Cloughesy, T., Chinot, O.L., Wick, W., Nishikawa, R., Mason, W., Henriksson, R., et al. (2015). Patients with proneural glioblastoma may derive overall survival benefit from the addition of bevacizumab to first-line radiotherapy and temozolomide: retrospective analysis of the AVAglio trial. J. Clin. Oncol. *33*, 2735–2744.

Scholz, A., Harter, P.N., Cremer, S., Yalcin, B.H., Gurnik, S., Yamaji, M., Di Tacchio, M., Sommer, K., Baumgarten, P., Bahr, O., et al. (2016). Endothelial cell-derived angiopoietin-2 is a therapeutic target in treatment-naive and bevacizumab-resistant glioblastoma. EMBO Mol. Med. *8*, 39–57.

Schwartzentruber, J., Korshunov, A., Liu, X.Y., Jones, D.T., Pfaff, E., Jacob, K., Sturm, D., Fontebasso, A.M., Quang, D.A., Tonjes, M., et al. (2012). Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature *482*, 226–231.

Shlien, A., Campbell, B.B., de Borja, R., Alexandrov, L.B., Merico, D., Wedge, D., Van Loo, P., Tarpey, P.S., Coupland, P., Behjati, S., et al. (2015). Combined hereditary and somatic mutations of replication error repair genes result in rapid onset of ultra-hypermutated cancers. Nat. Genet. *47*, 257–262.

Sturm, D., Orr, B.A., Toprak, U.H., Hovestadt, V., Jones, D.T., Capper, D., Sill, M., Buchhalter, I., Northcott, P.A., Leis, I., et al. (2016). New brain tumor entities emerge from molecular classification of CNS-PNETs. Cell *164*, 1060–1072.

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

Sulman, E.P., Won, M., Blumenthal, D.T., Vogelbaum, M.A., Colman, H., Jenkins, R.B., Chakravarti, A., Jeraj, R., Brown, P.D., Jaeckle, K.A., et al. (2013). Molecular predictors of outcome and response to bevacizumab (BEV) based on analysis of RTOG 0825, a phase III trial comparing chemoradiation (CRT) with and without BEV in patients with newly diagnosed glioblastoma (GBM). J. Clin. Oncol. *31*, https://doi.org/10.1200/jco.2013.31.18_suppl.lba2010.

Wallin, J.J., Bendell, J.C., Funke, R., Sznol, M., Korski, K., Jones, S., Hernandez, G., Mier, J., He, X., Hodi, F.S., et al. (2016). Atezolizumab in combination with bevacizumab enhances antigen-specific T-cell migration in metastatic renal cell carcinoma. Nat. Commun. *7*, 12624.

Wilmott, J.S., Long, G.V., Howle, J.R., Haydu, L.E., Sharma, R.N., Thompson, J.F., Kefford, R.F., Hersey, P., and Scolyer, R.A. (2012). Selective BRAF inhibitors induce marked T-cell infiltration into human metastatic melanoma. Clin. Cancer Res. *18*, 1386–1394.

Wu, G., Broniscer, A., McEachron, T.A., Lu, C., Paugh, B.S., Becksfort, J., Qu, C., Ding, L., Huether, R., Parker, M., et al. (2012). Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. Nat. Genet. *44*, 251–253.

Wu, G., Diaz, A.K., Paugh, B.S., Rankin, S.L., Ju, B., Li, Y., Zhu, X., Qu, C., Chen, X., Zhang, J., et al. (2014). The genomic landscape of diffuse intrinsic pontine glioma and pediatric non-brainstem high-grade glioma. Nat. Genet. *46*, 444–450.

Zhang, Y., Pan, C., Wang, J., Cao, J., Liu, Y., Wang, Y., and Zhang, L. (2017). Genetic and immune features of resectable malignant brainstem gliomas. Oncotarget *8*, 82571–82582.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical Commercial Assays		
DNeasy blood & tissue kit	Qiagen	69504
QIAmp DNA FFPE tissue kit	Qiagen	56404
RNeasy mini kit	Qiagen	74104
QIAquick PCR purification kit	Qiagen	28104
BigDye terminator v3.1 mix	Thermo Fisher	4337455
SureSelect Human All Exon capture set V6	Agilent	5190-8863
SureSelect RNA Capture, 0.5-2.9Mb	Agilent	5190-4944
Deposited Data		
Exome and RNA sequencing of new samples	This paper	EGA: EGAS00001002328
Illumina methylation BeadChip profiling of new samples	This paper	ArrayExpress: E-MTAB-5552
Sequencing and methylation data	This paper	cavatica.org
Oligonucleotides		
Primer: *H3F3A*_forward FFPE TGGCTCGTACAAAGCAGACT	This paper	N/A
Primer: *H3F3A*_reverse FFPE ATATGGATACATACAAGAGAGACT	This paper	N/A
Primer: *H3F3A* _forward FROZEN GATTTTGGGTAGACGTAATCTTCA	This paper	N/A
Primer: *H3F3A* _ reverse FROZEN TTTCCTGTTATCCATCTTTTTGTT	This paper	N/A
Antibodies		
MLH1	BD Pharmingen	G168-728; RRID: AB_395227
PMS2	BD Pharmingen	A16-4; RRID: AB_396410
MSH2	DIAG-BIOSYSTEMS	25D12; RRID: AB_10978033
MSH6	DIAG-BIOSYSTEMS	44; RRID: AB_1958490
H3K27M	Merck	ABE419
H3G34R	University of Nottingham	richard.grundy@nottingham.ac.uk
CD8	Dako	C8/144B; RRID: AB_2075537
CD68	Glostrup	KP1; RRID: AB_2661840
Software and Algorithms		
Mutation Surveyor	SoftGenetics	softgenetics.com/mutationSurveyor.php
4Peaks	Nucleobytes	nucleobytes.com/4peaks/
minfi	BioConductor	bioconductor.org/packages/release/bioc/html/minfi.html
conumee	BioConductor	bioconductor.org/packages/release/bioc/html/conumee.html
BEDtools	University of Utah	github.com/arq5x/bedtools2
DNAcopy	BioConductor	bioconductor.org/packages/release/bioc/html/DNAcopy.html
CopyNumber450kData	BioConductor	bioconductor.org/packages/release/data/experiment/html/CopyNumber450kData.html
MNP	DKFZ Heidelberg	molecularneuropathology.org/mnp
Bowtie2	Johns Hopkins University	bowtie-bio.sourceforge.net/bowtie2/index.shtml
TopHat	Johns Hopkins University	ccb.jhu.edu/software/tophat/index.shtml

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
cufflinks	University of Washington	ole-trapnell-lab.github.io/cufflinks/cufflinks/
DESeq2	BioConductor	bioconductor.org/packages/release/bioc/html/DESeq2.html
Gene Set Enrichment Analysis	Broad Institute	http://software.broadinstitute.org/gsea
bwa	Sanger Institute	http://bio-bwa.sourceforge.net/
Genome Analysis Toolkit	Broad Institute	oftware.broadinstitute.org/gatk/
Variant Effect predictor	Ensembl tools	ensembl.org/info/docs/variation/vep
BCBio	Harvard TH Chan	bcb.io/
ANNOVAR	Children's Hospital of Philadelphia	annovar.openbioinformatics.org/en/latest/
ExAc	Broad Institute	exac.broadinstitute.org/
SIFT	J Craig Venter Institute	sift.jcvi.org
PolyPhen	Harvard	genetics.bwh.harvard.edu/pph2
Manta	Illumina	github.com/Illumina/manta
Oncoprinter	Memorial Sloan Kettering	cbioportal.org/oncoprinter.jsp
Circos	Michael Smith Genome Sciences Center	circos.ca
R	The Comprehensive R Archive Network	r-project.org
Other		
Integrated mutation, copy number, expression and methylation data	This paper and cited sources	pedcbioportal.org

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Chris Jones (chris.jones@icr.ac.uk).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Patient Samples

All patient samples were collected after signed consent to the HERBY translational research program, under full Research Ethics Committee approval at each participating center. Tumor material was available from 89 patients (out of a total of 113 providing consent) from 13 countries: France (n=27, Hospital Pour Enfants De La Timone, Marseille; Hôpital des Enfants, Toulouse; Centre Hospitalier Régional Universitaire, Tours; CHRU Rennes; CHRU Nancy; Institut Curie, Paris; CHRU Strasbourg; Cancer Research Center of Lyon; Institut Gustave Roussy, Villejuif; CHRU Clermont-Ferrand; Oscar Lambret Center, Lille; CHRU Angers), UK (n=17, Royal Victoria Infirmary, Newcastle; NHS Lothian, Edinburgh; Nottingham University Hospital, Nottingham; Great Ormond Street Hospital, London; Royal Hospital for Children, Bristol; Cambridge University Hospital, Cambridge; Leeds Teaching Hospitals, Leeds; Alder Hey Children's Hospital, Liverpool; Royal Marsden Hospital, Sutton), Italy (n=14, Istituto Giannina Gaslini, Genoa; Istituto Nazionale Tumori, Milan), Spain (n=6, Hospital San Joan de Deu, Barcelona; Hospital Universitari i Politècnic la Fe, Valencia), Canada (n=5, Hospital for Sick Children, Toronto; Alberta Children's Hospital, Calgary), Netherlands (n=4, Radboud University Medical Center, Nijmegen; Erasmus Medical Center, Rotterdam), Czech Republic (n=3, Faculty Hospital, Brno; Charles University Hospital, Prague), Denmark (n=3, Aarhus University Hospital, Aarhus; Rigshospitalet, Copenhagen), Hungary (n=3, Semmelweis University, Budapest), Sweden (n=3, University of Gothenburg, Gothenburg; Skåne University Hospital, Lund), Poland (n=2, Children's Memorial Health Institute, Warsaw), Austria (n=1, Kepler Universitätsklinikum, Linz), and Belgium (n=1, Universitaire Ziekenhuizen, Leuven).

Pathological Review

All patients had their initial local diagnosis of HGG confirmed according to the WHO 2007 classification by a central HERBY reference neuropathologist prior to enrollment. Subsequently, all specimens were further subjected to a consensus review by the HERBY panel of five independent expert paediatric neuropathologists, who also applied the diagnostic criteria of the WHO 2016 classification.

Radiological Anatomical Localization

Tumor localization was determined by the HERBY panel of paediatric neuroradiologists. Tumors were assigned to lobar, basal ganglia, thalamic, non-pontine brainstem or cerebellar locations. As many of the tumors spanned more than one of these locations, post hoc radiological analysis by one of the HERBY Neuroradiologists was undertaken to determine the epicenter of the tumor origin; this was used to classify the site of origin of the tumor in each case.

METHOD DETAILS

Nucleic Acid Extraction

DNA was extracted from frozen tissue by homogenisation prior to following the DNeasy Blood & Tissue kit protocol (Qiagen, Crawley, UK). DNA was extracted from formalin-fixed, paraffin-embedded (FFPE) pathology blocks after manual macrodissection using the QIAamp DNA FFPE tissue kit protocol (Qiagen). Matched normal DNA was extracted from blood samples using the DNeasy Blood & Tissue kit (Qiagen, Crawley, UK). Concentrations were measured using a Qubit fluorometer (Life Technologies, Paisley, UK). RNA was extracted by following the RNeasy Mini Kit protocol (Qiagen), and quantified on a 2100 Bioanalyzer (Agilent Technologies).

H3F3A Sanger Sequencing

PCR for *H3F3A* was carried out on 89 cases using primers obtained from Life Technologies (Paisley, UK) (FFPE: for-TGGCTCGTA CAAAGCAGACT; rev-ATATGGATACATACAAGAGAGACT; FROZEN: for-GATTTTGGGTAGACGTAATCTTCA; rev-TTTCCTGTTA TCCATCTTTTTGTT). Sequences were analysed using Mutation Surveyor (SoftGenetics, PN, USA) and manually with 4Peaks (Nucleobytes, Aalsmeer, Netherlands). Only three cases left insufficient DNA for exome sequencing, all of which were found to harbor K27M mutations, and thus additional Sanger sequencing for genes encoding H3.1 variants were not undertaken.

Methylation Profiling

50-500 ng DNA was bisulphite-modified and analyzed for genome-wide methylation patterns using the Illumina HumanMethylation450 BeadChip (450k) platform at either the DKFZ or the University College London Genomics Center, according the manufacturers instructions. All samples were checked for expected and unexpected genotype matches by pairwise correlation of the 65 genotyping probes on the 450k array.

MGMT Promoter Methylation

To evaluate the methylation status of the *MGMT* promoter region, we used either the MGMT_STP27 logistic regression model from Illumina 450k methylation array data, or methylation-specific (MS-) PCR. For MS-PCR, 300-1500 ng DNA was sodium bisulphite-treated and PCR products analyzed on an the ABI7900HT instrument (Applied Biosystems, Foster City, CA, USA) to quantify the copy number of MGMT/ACTB (MDxHealth, Irvine, CA, USA).

Next-Generation Sequencing

50-500 ng DNA from 86 cases was sent for exome sequencing at the Tumor Profiling Unit (ICR, London, UK) using the Agilent SureSelectXT Human All Exon V6 platform with additional customized coverage of all histone H3 genes and the *TERT* promoter (Agilent, Santa Clara, CA, USA), and paired-end-sequenced on an Illumina HiSeq2500 (Illumina, San Diego, CA, USA) with one single patient-matched tumor/normal pair per lane where possible. The average median coverage was 426x for the frozen tumors (range 351-598x), 321x for the FFPE tumors (range 115-519x) and 163x for normal samples (range 116-464x). A customized panel of biotinylated DNA probes (NimbleGen) was developed for the detection of structural variants (translocations and duplications) and potential amplifications. The panel capture a total of 22 genes recently implicated in brain tumors (*ALK, BCOR, BRAF, c11orf95, C19MC, CIC, ETV6, FGFR1-3, FOXR2, KIAA1549, MET, MN1, MYB, MYBL1, NTRK1-3, RAF, RELA, TPM3* and *YAP1*). Library preparation was performed using 50-200 ng of genomic DNA using the HyperPlus Kit (Kapa Biosystems, Wilmington MA, USA) and SeqCap EZ adapters (Roche). Following fragmentation, DNA was end-repaired, A-tailed and indexed adapters ligated. DNA was amplified, multiplexed and hybridized using 1 μg of total pre-capture library DNA to the design of DNA baits (NimbleGen SeqCap EZ Developer library, Roche). After hybridization, capture libraries were amplified and sequencing was performed on a NextSeq500 (Illumina) with 2 x 150 bp, paired-end reads following manufacturer's instructions. RNA from frozen tumors was sequenced on an Illumina HiSeq2500 as 100 bp paired end reads.

Immunohistochemistry

4 μM sections were stained by an automated Discovery XT (for H3G34R) or Benchmark XT (Ventana Medical Systems, Tucson, USA). A standard pre-treatment protocol included CC1 buffer (or CC2 for H3G34R) and then a primary antibody incubation for 32 minutes (92 minutes for MLH1, PMS2, MSH2 and MSH6) at room temperature (37°C for H3K27M and MLH1). Antibodies used were directed against MLH1 (BD Pharmingen, clone G168-728, 1/300), PMS2 (BD Pharmingen, clone A16-4, 1/150), MSH2 (DIAG-BIOSYSTEMS, clone 25D12, 1/10), MSH6 (DIAG-BIOSYSTEMS, clone 44, 1/50), H3K27M (Merck, polyclonal, 1/1000), CD8 (Dako, clone C8/144B, 1/100), CD68 (Glostrup, clone KP1, 1/400) and H3G34R (kind gift from Richard Grundy, Children's Brain Tumor Research Center, Nottingham, UK, polyclonal; 1/150). Antibody binding was visualized with an Optiview Kit (Roche-Ventana, Tucson, USA). Diamino-benzidine-tetra-hydrochloride (DAB, Ventana) was used as the chromogen.

Pathological Assessment of Immune Response

CD8 immunoreactivity was assessed as the percentage of the surface area of the tumor covered by $CD8^+$ cells present at a density belonging to one of four reference bins of increasing density (I0, I1, I2, I3). A further quantitative assessment of the percentage of the pathologist-defined central tumor area of $CD8^+$ cells was also performed. A qualitative categorization of the tumor as a whole as either 'inflamed', 'heterogeneous' or an immune 'desert' was also provided. Patients in the upper quartile of central tumor area

CD8 cell positivity with a heterogeneous or inflamed phenotype were classed as CD8-high. A histologically-defined assessment of tumor infiltrating lymphocytes (TILs) was carried out according to two distinct schema. Palma et al. (Palma et al., 1978) includes Category A (several perivascular cuffs of more than three layers of lymphocytes and often lymphocytes also scattered among tumor cells), Category B (three or four thin perivascular cuffs in the tumor) and Category C (no clear lymphocytes present); Rutledge et al. (Rutledge et al., 2013) includes Category 0 (absence of lymphocytes), Category 1+ (lymphocytes in less than 50% of the tumor) and Category 2+ (lymphocytes in more than 50% of the tumor).

QUANTIFICATION AND STATISTICAL ANALYSIS

Sequence Analysis

Exome capture reads were aligned to the hg19 build of the human genome using bwa v0.7.5a (bio-bwa.sourceforge.net), and PCR duplicates removed with PicardTools 1.5 (pcard.sourceforge.net). Somatic single nucleotide variants were called using the Genome Analysis Tool Kit v3.3-0 based upon current Best Practices using local re-alignment around InDels, downsampling and base recalibration with variants called by the Unified Genotyper (broadinstitute.org/gatk/). Somatic variants were covered by at least 20 reads in both tumor and normal sequences and carried at least 5 ALT reads in the tumor sequence; unmatched exomes (n=3) were annotated by ExAc and ANNOVAR. Variants were annotated using the Ensembl Variant Effect Predictor v74 (ensembl.org/info/docs/variation/vep) incorporating SIFT (sift.jcvi.org) and PolyPhen (genetics.bwh.harvard.edu/pph2) predictions, COSMIC v64 (sanger.ac.uk/genetics/CGP/cosmic/) and dbSNP build 137 (ncbi.nlm.nih.gov/sites/SNP) annotations. Copy number was obtained by calculating $\log_2$ ratios of tumor/normal coverage binned into exons of known Ensembl genes, smoothed using circular binary segmentation (DNAcopy, www.bioconductor.org) and processed using in-house scripts. Mutation signatures were ascertained by grouping somatic substitutions on the basis of their 3′ and 5′ bases into 96 possible trinucleotide categories (Shlien et al., 2015). NGS fusion panel alignment was performed against the human reference sequence GRCh37/Hg19. Quality control (QC), variant annotation, deduplication and metrics were generated for each sample. Manta (https://github.com/Illumina/manta) and Breakdancer (breakdancer.sourceforge.net) were used for the detection of structural variants.

Methylation Profiling

Methylation data from the Illumina Infinium HumanMethylation450 BeadChip was preprocessed using the minfi package in R. DNA copy number was recovered from combined intensities using the conumee package with reference to methylation profiles from normal individuals provided in the CopyNumber450kData package. We have used the Heidelberg brain tumor classifier (Capper et al., 2018) (molecularneuropathology.org) to assign subgroup scores for each tumor compared to 91 different brain tumor entities using a training set built from 2,801 tumors implemented in the MNP R package (v11b2). Simplified methylation subgroup assignments were then made to incorporate cases carrying G34R/V or K27M mutations in H3 histones, *IDH1* mutation at R132, low grade glioma-like profiles (predominantly diffuse infantile ganglioglioma and pilocytic astrocytoma) and those similar to pleomorphic xanthoastrocytoma (PXA). Low-scoring cases, or those with a high normal cell contamination were assigned to G34, K27 or IDH1 groups if the respective mutation was identified. Wild-type HGG encompassed many other methylation subgroups and were simply assigned by exclusion with the groups above. Clustering of beta values from methylation arrays was performed using the 10K probeset from the Heidelberg classifier based upon Euclidean distance with a ward algorithm. Methylation heatmaps show only the most variable probes of the classifier between simplified methylation subgroups.

RNAseq

RNA sequences were aligned to hg19 and organized into de-novo spliced alignments using bowtie2 and TopHat version 2.1.0 (ccb.jhu.edu/software/tophat). Fusion transcripts were detected using chimerascan version 0.4.5a filtered to remove common false positives. RNASeq raw count files were used to construct an expression matrix using Roche's internal pipeline. The expression matrix was normalized using edgeR and Voom in R (cran.rproject.org/), and a heatmap was created from the absolute gene expression data using Tibco Spotfire, as previously described (Brouwer-Visser et al., 2018). The mean expression of the signature genes was used to compare MAPK-altered to non-altered samples and statistical significance calculated using a two-tailed unpaired t-test. The mean expression was also used to correlate with CD8 positivity by IHC. Gene Set enrichment analysis was performed using the GSEA java application based upon pairwise comparisons of MAPK altered versus wild-type for curated canonical gene sets. All data are deposited in the European Genome-phenome Archive (ebi.ac.uk/ega/home) under accession number EGAS00001002328.

Tumor Lesion Maps

Pre-surgery tumor volume regions of interest (ROIs) were drawn on T2-weighted/Fluid-attenuated inversion recovery (FLAIR) magnetic resonance (MR) images by an experienced paediatric neuroradiologist (TJ). The images and corresponding ROIs were affinely registered, using FSL (Jenkinson et al., 2002), to a paediatric template (Left-Right Symmetric, 7.5–13.5 years old) from the Montreal Neurological Institute (http://www.bic.mni.mcgill.ca/ServicesAtlases/NIHPD-obj1) (Fonov et al., 2011). A further manual correction step was performed to limit tumor mass effects. Once registered to a common space, overall tumor lesion overlap maps were created using MRIcron (Rorden et al., 2007).

Statistical Analysis

Statistical analysis was carried out using R 3.3.0 (www.r-project.org) and GraphPad Prism 7. Categorical comparisons of counts were carried out using Fishers exact test, comparisons between groups of continuous variables employed Student's t-test or ANOVA. Univariate differences in survival were analysed by the Kaplan-Meier method and significance determined by the log-rank test. Exploratory Cox regression analyses were conducted to assess the impact of molecular prognostic and predictive factors. Confirmation of proportional hazards was assessed by calculating Schoenfeld residuals. All tests were two-sided and a p value of less than 0.05 was considered significant.

DATA AND SOFTWARE AVAILABILITY

All newly generated data have been deposited in the European Genome-phenome Archive (www.ebi.ac.uk/ega) with accession number EGAS00001002328 (sequencing) or ArrayExpress (www.ebi.ac.uk/arrayexpress/) with accession number E-MTAB-5552 (450k methylation).

ADDITIONAL RESOURCES

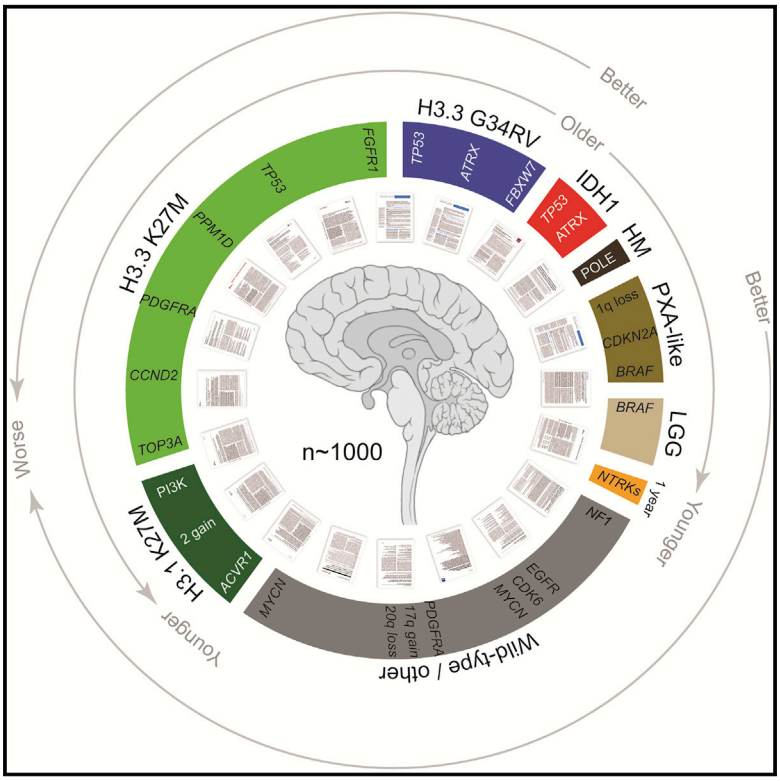
Curated gene-level copy number, mutation data and RNAseq data are provided as part of the paediatric-specific implementation of the cBioPortal genomic data visualization portal (pedcbioportal.org). Raw data files are also made available through the Cavatica NIH-integrated cloud platform (cavatica.org).

Article

Cancer Cell

Integrated Molecular Meta-Analysis of 1,000 Pediatric High-Grade and Diffuse Intrinsic Pontine Glioma

Graphical Abstract



Highlights

- Pediatric HGG and DIPG comprise a diverse set of clinical and biological subgroups
- Somatic coding mutations per tumor range from none to among the highest seen in human cancer
- Histone mutations co-segregate with distinct alterations and downstream pathways
- H3/IDH1 WT tumors may resemble low-grade lesions and have targetable alterations

Authors

Alan Mackay, Anna Burford, Diana Carvalho, ..., Michael Baudis, Adam Resnick, Chris Jones

Correspondence

chris.jones@icr.ac.uk

In Brief

Mackay et al. perform an integrated analysis of >1,000 cases of pediatric high-grade glioma and diffuse intrinsic pontine glioma. They identify co-segregating mutations in histone-mutant subgroups and show that histone wild-type subgroups are molecularly more similar to lower-grade tumors.

Mackay et al., 2017, Cancer Cell *32*, 520–537
October 9, 2017
http://dx.doi.org/10.1016/j.ccell.2017.08.017

CellPress

Cancer Cell
Article

Integrated Molecular Meta-Analysis of 1,000 Pediatric High-Grade and Diffuse Intrinsic Pontine Glioma

Alan Mackay,[1,2] Anna Burford,[1,2] Diana Carvalho,[1,2] Elisa Izquierdo,[1,2] Janat Fazal-Salom,[1,2] Kathryn R. Taylor,[1,2,3] Lynn Bjerke,[1,2] Matthew Clarke,[1,2] Mara Vinci,[1,2] Meera Nandhabalan,[1,2] Sara Temelso,[1,2] Sergey Popov,[1,2,4] Valeria Molinari,[1,2] Pichai Raman,[5,6] Angela J. Waanders,[5,7] Harry J. Han,[5,7] Saumya Gupta,[8] Lynley Marshall,[9] Stergios Zacharoulis,[9] Sucheta Vaidya,[9] Henry C. Mandeville,[10] Leslie R. Bridges,[11] Andrew J. Martin,[12]

(Author list continued on next page)

[1]Division of Molecular Pathology
[2]Division of Cancer Therapeutics
The Institute of Cancer Research, London, UK
[3]Department of Neurology, Stanford University School of Medicine, Stanford, CA, USA
[4]Department of Cellular Pathology, University Hospital of Wales, Cardiff, UK
[5]The Center for Data Driven Discovery in Biomedicine (D^3b)
[6]Division of Neurosurgery
[7]Division of Oncology
Children's Hospital of Philadelphia, Philadelphia, PA, USA
[8]Institute of Molecular Life Sciences, Swiss Institute of Bioinformatics, University of Zürich, Zürich, Switzerland
[9]Pediatric Oncology Drug Development Team, Children and Young People's Unit, Royal Marsden Hospital, Sutton, UK
[10]Department of Radiotherapy, Royal Marsden Hospital, Sutton, UK
[11]Department of Cellular Pathology
[12]Department of Neurosurgery
St George's Hospital NHS Trust, London, UK

(Affiliations continued on next page)

SUMMARY

We collated data from 157 unpublished cases of pediatric high-grade glioma and diffuse intrinsic pontine glioma and 20 publicly available datasets in an integrated analysis of >1,000 cases. We identified co-segregating mutations in histone-mutant subgroups including loss of *FBXW7* in H3.3G34R/V, *TOP3A* rearrangements in H3.3K27M, and *BCOR* mutations in H3.1K27M. Histone wild-type subgroups are refined by the presence of key oncogenic events or methylation profiles more closely resembling lower-grade tumors. Genomic aberrations increase with age, highlighting the infant population as biologically and clinically distinct. Uncommon pathway dysregulation is seen in small subsets of tumors, further defining the molecular diversity of the disease, opening up avenues for biological study and providing a basis for functionally defined future treatment stratification.

INTRODUCTION

Pediatric glioblastoma (pGBM) and diffuse intrinsic pontine glioma (DIPG) are high-grade glial tumors of children with a median overall survival of 9–15 months, a figure that has remained unmoved for decades (Jones et al., 2012). Although relatively rare in this age group (1.78 per 100,000 population), taken together, gliomas are nonetheless the most common malignant

Significance

High-grade and diffuse intrinsic pontine glioma in children are rare, incurable brain tumors with differing biology to adult cancers. An integrated genomic, epigenomic and transcriptomic analysis of >1,000 cases across all anatomical compartments of the CNS defines robust clinicopathological and molecular subgroups with distinct biological drivers. As modern classification schemes begin to recognize the diversity of this disease in the pediatric population, we provide a framework for meaningful further subcategorization and identify subgroup-restricted therapeutic targets.

Safa Al-Sarraj,[13] Christopher Chandler,[14] Ho-Keung Ng,[15] Xingang Li,[16] Kun Mu,[17] Saoussen Trabelsi,[18] Dorra H'mida-Ben Brahim,[18] Alexei N. Kisljakov,[19] Dmitry M. Konovalov,[20] Andrew S. Moore,[21,22,23] Angel Montero Carcaboso,[24] Mariona Sunol,[24] Carmen de Torres,[24] Ofelia Cruz,[24] Jaume Mora,[24] Ludmila I. Shats,[25] João N. Stavale,[26] Lucas T. Bidinotto,[27] Rui M. Reis,[27,28] Natacha Entz-Werle,[29] Michael Farrell,[30] Jane Cryan,[30] Darach Crimmins,[31] John Caird,[31] Jane Pears,[32] Michelle Monje,[3] Marie-Anne Debily,[33] David Castel,[33] Jacques Grill,[33] Cynthia Hawkins,[34] Hamid Nikbakht,[35] Nada Jabado,[5] Suzanne J. Baker,[36] Stefan M. Pfister,[37,38,43] David T.W. Jones,[37,43] Maryam Fouladi,[39] André O. von Bueren,[40,41,42] Michael Baudis,[8] Adam Resnick,[5,6,7] and Chris Jones[1,2,44,*]

[13]Department of Neuropathology
[14]Department of Neurosurgery
Kings College Hospital, London, UK
[15]Department of Anatomical and Cellular Pathology, The Chinese University of Hong Kong, Hong Kong, China
[16]Department of Neurosurgery, Qilu Hospital of Shandong University and Brain Science Research Institute, Shandong University, Jinan, China
[17]Department of Pathology, Shandong University School of Medicine, Jinan, China
[18]Department of Cytogenetics and Reproductive Biology, Farhat Hached Hospital, Sousse, Tunisia
[19]Department of Pathology, Morozov Children's Hospital, Moscow, Russian Federation
[20]Department of Pathology, Dmitrii Rogachev Research and Clinical Centre of Pediatric Hematology, Oncology and Immunology, Moscow, Russian Federation
[21]UQ Child Health Research Centre, The University of Queensland, Brisbane, Australia
[22]Oncology Services Group, Children's Health Queensland Hospital and Health Service, Brisbane, Australia
[23]The University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, Australia
[24]Institut de Recerca Sant Joan de Deu, Barcelona, Spain
[25]Division of Oncology, Pediatric Oncology and Radiotherapy, St Petersburg State Pediatric Medical University, St Petersburg, Russian Federation
[26]Department of Pathology, Federal University of São Paulo, São Paulo, São Paulo, Brazil
[27]Molecular Oncology Research Centre, Barretos Cancer Hospital, Barretos, São Paulo, Brazil
[28]Life and Health Sciences Research Institute (ICVS), Medical School, University of Minho, Braga, Portugal and ICVS/3B's-PT Government Associate Laboratory, Braga/Guimarães, Portugal
[29]Pédiatrie Onco-Hématologie - Pédiatrie III, Centre Hospitalier Régional et Universitaire Hautepierre, Strasbourg, France
[30]Histopathology Department, Beaumont Hospital, Dublin, Ireland
[31]Department of Neurosurgery, Temple Street Children's University Hospital, Dublin, Ireland
[32]Department of Paediatric Oncology, Our Lady's Children's Hospital, Dublin, Ireland
[33]Département de Cancerologie de l'Enfant et de l'Adolescent, Institut Gustav Roussy, Villejuif, France
[34]Pediatric Laboratory Medicine, Hospital for Sick Children, Toronto, Canada
[35]Department of Pediatrics, McGill University, Montreal, Canada
[36]Department of Developmental Neurobiology, St Jude Children's Research Hospital, Memphis, TN, USA
[37]Division of Pediatric Neuro-oncology, German Cancer Research Center (DKFZ), Heidelberg, Germany
[38]Department of Pediatric Hematology and Oncology, Heidelberg University Hospital, Heidelberg, Germany
[39]Department of Pediatrics, Cancer and Blood Diseases Institute, Cincinnati Children's Hospital, Cincinnati, OH, USA
[40]Department of Pediatrics, Division of Pediatric Hematology and Oncology, University Medical Center Goettingen, Goettingen, Germany
[41]Department of Pediatrics and Adolescent Medicine, Division of Pediatric Hematology and Oncology, University Hospital of Geneva, Geneva, Switzerland
[42]Department of Pediatrics, CANSEARCH Research Laboratory, Faculty of Medicine, University of Geneva, Geneva, Switzerland
[43]Hopp-Children's Cancer Center at the NCT Heidelberg (KiTZ), Heidelberg, Germany
[44]Lead Contact
*Correspondence: chris.jones@icr.ac.uk
http://dx.doi.org/10.1016/j.ccell.2017.08.017

brain tumors in children, and represent the greatest cause of cancer-related deaths under the age of 19 years (Ostrom et al., 2015). Unlike histologically similar lesions in adults, which tend to be restricted to the cerebral hemispheres, diffuse high-grade gliomas in childhood (pHGG) can occur throughout the CNS, with around half occurring in midline locations, in particular the thalamus and the pons (Jones and Baker, 2014), where the lack of available surgical options confers an especially poor prognosis (Kramm et al., 2011). Numerous clinical trials of chemotherapeutics and targeted agents extrapolated from adult GBM studies have failed to show a survival benefit, and more rationally derived approaches based upon an understanding of the childhood diseases are urgently needed (Jones et al., 2016).

It has become increasingly apparent that pHGG differ from their adult counterparts, with molecular profiling studies carried out over the last 6–7 years having incrementally identified key genetic and epigenetic differences in pHGG associated with distinct ages of onset, anatomical distribution, clinical outcome, and histopathological and radiological features (Jones and Baker, 2014; Sturm et al., 2014). In particular, the identification of unique recurrent mutations in genes encoding histones H3.3 and H3.1 (Schwartzentruber et al., 2012; Wu et al., 2012) have demonstrated the distinctiveness of the pediatric disease, with the G34R/V and K27M variants appearing to represent different clinicopathological and biological subgroups. This has been recognized by the World Health Organization (WHO) classification of CNS tumors, with the latest version including the novel entity, *diffuse midline glioma with H3K27 mutation* (Louis et al., 2016). Further refinements incorporating other clearly delineated subsets of the disease in future iterations appear likely and might prove clinically useful.

In addition to these uniquely defining histone mutations, detailed molecular profiling has served to identify numerous targets for therapeutic interventions. These include known oncogenes in adult glioma and other tumors with an elevated frequency in the childhood setting (e.g., *PDGFRA*) (Paugh et al., 2013; Puget et al., 2012) or certain rare histological variants (e.g., *BRAF* V600E) (Nicolaides et al., 2011; Schiffman et al., 2010), as well as others seemingly unique to DIPG (e.g., *ACVR1*) (Buczkowicz et al., 2014; Fontebasso et al., 2014; Taylor et al., 2014a; Wu et al., 2014). Future trials will need to exploit these targets, but also incorporate innovative designs that allow for selection of the patient populations within the wide spectrum of disease who are most likely to benefit from any novel agent (Jones et al., 2016).

Despite these advances, driven by the efforts of several international collaborative groups to collect and profile these rare tumors, individual publications remain necessarily modestly sized, involving a range of different platforms and analytical techniques. This leaves certain subgroups poorly represented across studies, widely differing individual gene frequencies in different cohorts, and an inability to draw robust conclusions across the whole spectrum of the disease. We have gathered together publicly available data, supplemented with 157 new cases, in order to provide a statistically robust, manually annotated resource cohort of >1,000 such tumors for interrogation.

RESULTS

Sample Cohort

In total, we obtained data from clinically annotated high-grade glioma (WHO, 2007, grade III or IV) or DIPG (radiologically diagnosed, grades II–IV) in 1,067 unique cases (Figure S1A). These were predominantly from children but also included young adults, in order to capture more *H3F3A* G34R/V mutations, as well as to explore an otherwise under-studied population. There was a median age at diagnosis of 9.8 years, and 982 cases aged 21 years or younger (Table S1). These included 910 taken from 20 published series (Barrow et al., 2011; Bax et al., 2010; Buczkowicz et al., 2014; Carvalho et al., 2014; Castel et al., 2015; Fontebasso et al., 2013, 2014; Grasso et al., 2015; International Cancer Genome Consortium PedBrain Tumor Project, 2016; Khuong-Quang et al., 2012; Korshunov et al., 2015; Paugh et al., 2010, 2011; Puget et al., 2012; Schwartzentruber et al., 2012; Sturm et al., 2012; Taylor et al., 2014a; Wu et al., 2012; Wu et al., 2014; Zarghooni et al., 2010) and 157 unpublished cases. The vast majority of samples were obtained pre-treatment (biopsy or resection, n = 913), as opposed to post-therapy (relapse or autopsy, n = 146). Samples were classified as occurring within the cerebral hemispheres (n = 482), brainstem (n = 323 in pons, of which 322 were DIPG; three additional cases were in the midbrain and one in the medulla) or other non-brainstem midline locations (n = 224, predominantly thalamus, but also cerebellum, spinal cord, ventricles, and others; referred to as "midline" for simplicity throughout) (Figure 1A). There was a significant association of anatomical location with age of diagnosis, with medians of 13.0 years for hemispheric, 10.0 years for midline, and 6.5 years for DIPG, respectively ($p < 0.0001$, ANOVA; all pairwise comparisons adjusted $p < 0.0001$, t test) (Figure 1B), in addition to clinical outcome, with a median overall survival of 18.0 months for hemispheric tumors (2 year overall survival 32%), 13.5 months for midline (2 year overall survival 21.4%), and 10.8 months for DIPG (2 year overall survival 5.2%; $p < 0.0001$ for all pairwise comparisons, log rank test) (Figure 1C). Children 3 years of age and younger had a markedly improved clinical outcome ($p = 0.0028$, log rank test), although this benefit was largely restricted to children 1 year and under (n = 40, 2 year survival 61%, $p < 0.0001$, log rank test), with this association significant in all anatomical locations ($p = 0.0402$, hemispheric; $p < 0.0001$, midline; $p = 0.00286$, pons, log rank test). There were, however, proportionally fewer midline and pontine tumors in <1-year-olds compared with 1- to 3-year-olds (12/40, 30.0% versus 46/85, 45.9%, $p = 0.0131$ Fisher's exact test) (Figures S1B and S1C).

Molecular Subgrouping

Hotspot mutation data for the genes encoding histone H3 were available or newly generated for 903 cases. At minimum, this included Sanger sequencing for *H3F3A* (H3.3) and *HIST1H3B* (H3.1); however, the absence of wider screening or next-generation sequencing data for 310 cases annotated as H3 wild-type (WT) means we cannot rule out rare variants in other H3.1 or H3.2 genes in those cases. In total, the cohort comprised 67 H3.3G34R/V (n = 63 G34R, n = 4 G34V), 316 H3.3K27M, 66 H3.1/3.2K27M (n = 62 *HIST1H3B*, n = 2 *HIST1H3C*, n = 2 *HIST2H3C*), and 454 WT. There were profound distinctions in anatomical location ($p < 0.0001$, Fisher's exact test) (Figure 1D), age at diagnosis ($p < 0.0001$, ANOVA; all pairwise comparisons adjusted $p < 0.0001$, t test) (Figure 1E), and overall survival ($p < 0.0001$, log rank test) (Figure 1F). H3.3G34R/V tumors were almost entirely restricted to the cerebral hemispheres (accounting for 16.2% total in this location, particularly parietal and temporal lobes), were found predominantly in adolescents and young adults (median 15.0 years), and had a longer overall survival compared with other H3 mutant groups (median 18.0 months, 2 year overall survival 27.3%, $p < 0.0001$ versus H3.3K27M, $p = 0.00209$ versus H3.1H27M, log rank test). H3.3K27M were spread throughout the midline and pons, where they account for 63.0% DIPG and 59.7% non-brainstem midline tumors. In all locations (including ten cases reported to present in the cortex), these mutations conferred a significantly shorter time to death from disease (overall median 11 months, 2 year overall survival 4.7%) (Figures S1D–S1F). H3.1/3.2K27M were highly specific to the pons (21.4% total) where they represent a younger age group (median 5.0 years) with a significantly longer overall survival (median 15.0 months) than H3.3K27M ($p = 0.00017$, log rank test) (Figure S1F). In multivariate analysis incorporating the histone mutations alongside age, WHO grade, and gender, K27M mutations in both H3.3 and H3.1 are independently associated with shorter survival ($p < 0.0001$, Cox proportional hazards model).

BRAF V600E status was available for 535 cases, with mutant cases (n = 32, 6.0%) present only in midline and hemispheric locations, and conferring a significantly improved prognosis (2 year survival 67%, $p < 0.0001$, log rank test) (Figures S1G–S1I). There was additional annotation for *IDH1* R132 mutation status in 640 cases (n = 40, 6.25%), representing a forebrain-restricted, significantly older group of patients (median 17.0

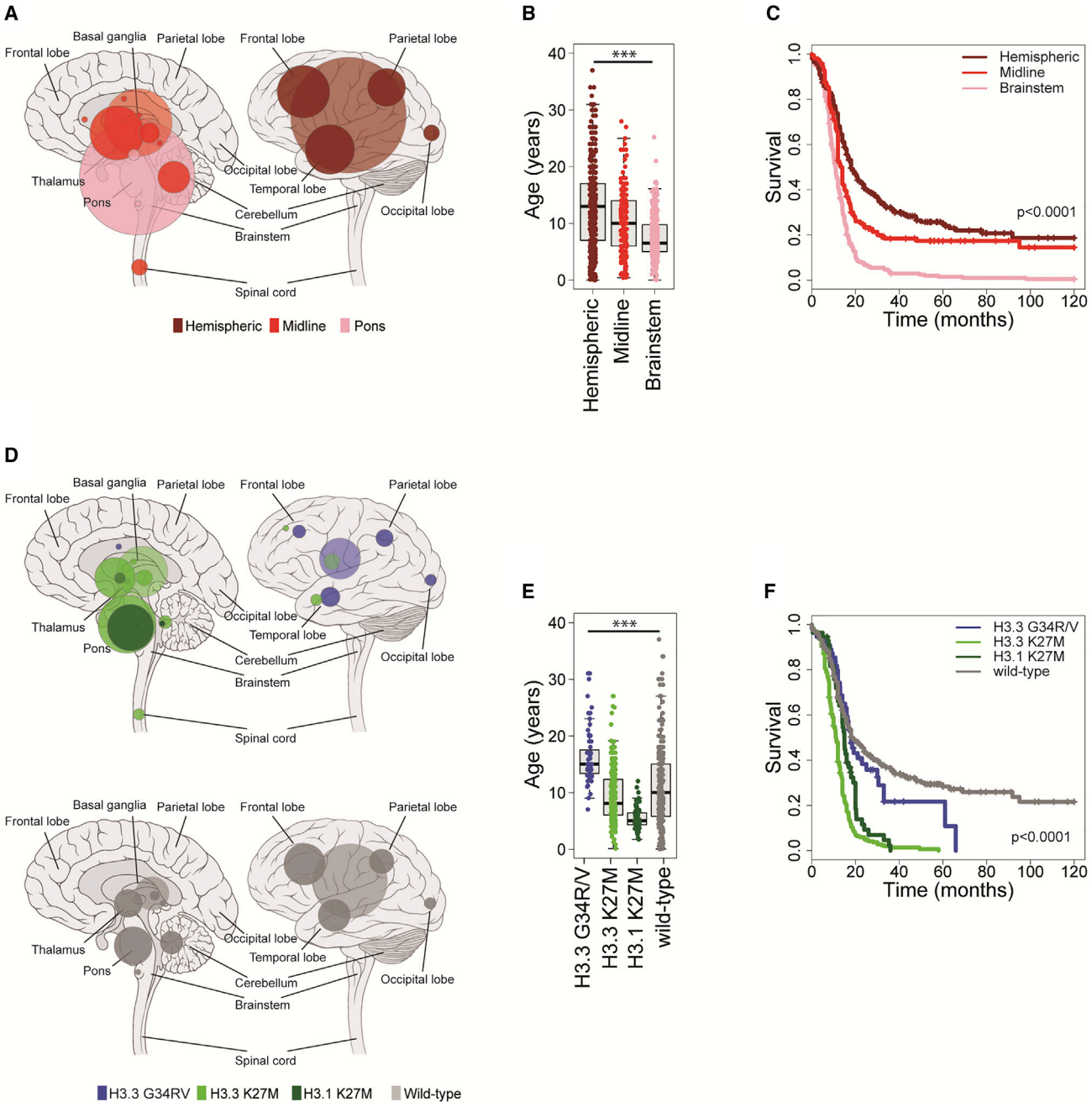


Figure 1. Clinicopathological and Molecular Subgroups of pHGG/DIPG

(A) Anatomical location of all high-grade glioma cases included in this study, taken from original publications (n = 1,033). Left, sagittal section showing internal structures; right, external view highlighting cerebral lobes. Hemispheric, dark red; non-brainstem midline structures, red; pons, pink. Radius of circle is proportional to the number of cases. Lighter shaded circles represent a non-specific designation of hemispheric, midline, or brainstem.

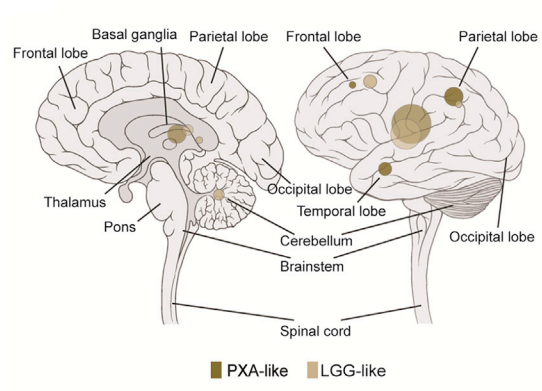
(B) Boxplot showing age at diagnosis of included cases, separated by anatomical location (n = 1,011). The thick line within the box is the median, the lower and upper limits of the boxes represent the first and third quartiles, and the whiskers 1.5× the interquartile range. ***Adjusted $p < 0.0001$ for all pairwise comparisons, t test.

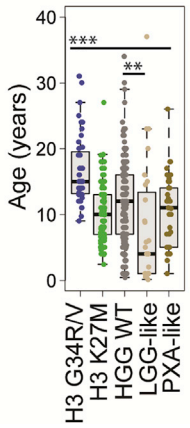
(C) Kaplan-Meier plot of overall survival of cases separated by anatomical location, p value calculated by the log rank test (n = 811).

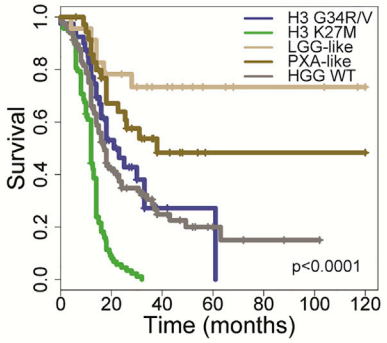
(D) Anatomical location of all cases separated by histone mutation (top, n = 441) and histone WT (bottom, n = 314). Left, sagittal section showing internal structures; right, external view highlighting cerebral lobes. Blue, H3.3G34R/V; green, H3.3K27M; dark green, H3.1K27M. Radius of circle is proportional to the number of cases. Lighter shaded circles represent a non-specific designation of hemispheric, midline, or brainstem.

(E) Boxplot showing age at diagnosis of included cases, separated by histone mutation (n = 753). The thick line within the box is the median, the lower and upper limits of the boxes represent the first and third quartiles, and the whiskers 1.5× the interquartile range. ***Adjusted $p < 0.0001$ for all pairwise comparisons, t test.

(F) Kaplan-Meier plot of overall survival of cases separated by histone mutation, p value calculated by the log rank test (n = 693). See also Figure S1 and Table S1.







(legend on next page)

years, $p < 0.0001$, t test) with longer overall survival (2 year survival 59%, $p < 0.0001$, log rank test) (Figures S1J–S1L).

For 441 cases, Illumina 450k methylation BeadArray data was available, which provides robust classification into clinically meaningful epigenetic subgroups marked by recurrent genetic alterations (Korshunov et al., 2015, 2017). We used the Heidelberg brain tumor classifier to assign tumors into following subgroups: H3G34R/V (n = 51), H3K27M (n = 119), HGG WT (n = 156), IDH1 (n = 36), low-grade glioma (LGG)-like (n = 27), pleomorphic xanthoastrocytoma (PXA)-like (n = 43), and "other" (n = 9) (Figure S2A), visualized by hierarchical clustering (Figure 2A) (Table S2). As reported previously, these subgroups have profound differences in anatomical location (Figure 2B), age at diagnosis ($p < 0.00001$ ANOVA) (Figure 2C), and overall survival ($p < 0.0001$, log rank test) (Figure 2D), with LGG-like group representing a younger cohort (median 4.0 years, 10/16 infant cases under 1 year, $p < 0.0001$ Fisher's exact test) with excellent prognosis (2 year survival 74%, $p < 0.0001$ versus WT, log rank test), while the PXA-like group are enriched for *BRAF* V600E mutations (19/34, 56%) and carry an intermediate risk (median 38 months, 2 year survival 56%, $p = 0.00423$ versus WT, log rank test). After removing the PXA- and LGG-like groups, the remaining histone H3/IDH1 WT tumors had a 2 year survival of 23.5% (median overall survival 17.2 months). *MGMT* promoter methylation was significantly enriched in the H3G34R/V (65.1%, globally hypomethylated) and IDH1 (78.1%, globally hypermethylated) groups, and largely absent from H3K27M tumors (4.5%, all tests versus rest, $p < 0.0001$ Fisher's exact test) (Figure S2B). Total methylation was lowest in H3G34R/V (median beta value 0.452), and highest in the IDH subgroup (median beta value 0.520), as reported previously (Sturm et al., 2012); however it was also found to be significantly elevated in PXA-like tumors (median beta value 0.501, $p < 0.0001$ t test) (Figure S2C).

DNA Copy Number

High-quality DNA copy-number profiles were obtained from 834 unique cases of pHGG/DIPG, taken from BAC and oligonucleotide arrays (n = 112), SNP arrays (n = 128), 450k methylation arrays (n = 428), and whole-genome or exome sequencing (n = 325) (Table S3). Clustering on the basis of segmented $\log_2$ ratios highlighted some of the defining chromosomal features of the pediatric disease, including recurrent gains of chromosome 1q, and losses of chromosomes 13q and 14q (Figure 3A). There are also a significant proportion of tumors (n = 147, 17.6%) with few if any DNA copy-number changes, with no bias toward lower-resolution platforms ($p = 0.134$, Fisher's exact test), and the presence of other molecular markers obviating concerns of a substantial normal tissue contamination. These cases were found throughout the CNS, were younger at diagnosis (7.0 versus 10.3 years, $p < 0.0001$, t test) and had a longer overall survival (median 18.0 versus 14.0 months, $p = 0.0107$ log rank test) (Figure S3A). Common large-scale chromosomal alterations with prognostic significance included loss of 17p (n = 156), which targets *TP53* at 17p13.1 and confers a shorter overall survival in tumors of all locations and all subgroups (Figure S3B), and gains of 9q (n = 108), more broadly encompassing a region of structural rearrangement on 9q34 in medulloblastoma (Northcott et al., 2014), and correlating with shorter overall survival in multiple pHGG/DIPG subgroups (Figure S3C).

We used GISTIC (genomic identification of significant targets in cancer) in order to determine subgroup-specific copy-number drivers based on focality, amplitude, and recurrence of alterations. The most common focal events were the previously described high-level amplifications (Figure 3B) at 4q12 (*PDGFRA/KIT/KDR*, n = 77), 2p24.3 (*MYCN/ID2,* n = 42), chromosome 7 (7p11.2 (*EGFR,* n = 32), 7q21.2 (*CDK6,* n = 14), and 7q31.2 (*MET,* n = 19)) (Figures S3D–S3F), as well as focal deletions (Figure 3C) at 9p21.3 (*CDKN2A/CDKN2B,* n = 102) (Figure S3G). Amplifications conferred a shorter overall survival, and *CDKN2A/CDKN2B* deletion a better prognosis, either across the whole cohort or selected subgroups (Figures S3B–S3G). In addition, the aggregated data identified less-frequent alterations, recurrent across multiple studies, identifying pHGG/DIPG candidates including *NFIB* (nuclear factor I B, 9p23-p22.3, n = 4), *GAB2* (GRB2-associated binding protein 2, 11q14.1, n = 4), *SMARCE1* (SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily e, member 1, 17q21.2, n = 4), and others (Figures 3B and 3C).

Subgroup-Specific Alterations

When *IDH1*-mutant tumors were removed and the cohort restricted to those cases for which histone H3 status was available, we were able to investigate subgroup-specific DNA copy-number changes in 705 pHGG/DIPG (Figure 4A). Applying GISTIC within these case sets revealed specific focal events enriched within individual subgroups, including *AKT1* amplifications in H3.3G34R/V, *MYC* and *CCND2* amplification in H3.3K27M, and *MYCN/ID2*, *MDM4/PIK3C2B*, and *KRAS* amplification in H3 WT (Figure 4B) (Table S4). These latter events were generally restricted to hemispheric tumors, while *MYCN/ID2* were enriched in H3 WT DIPG (Figure S4A). H3.1K27M tumors generally lacked amplifications/deletions, but were instead characterized by frequent gains of 1q and the whole of chromosome

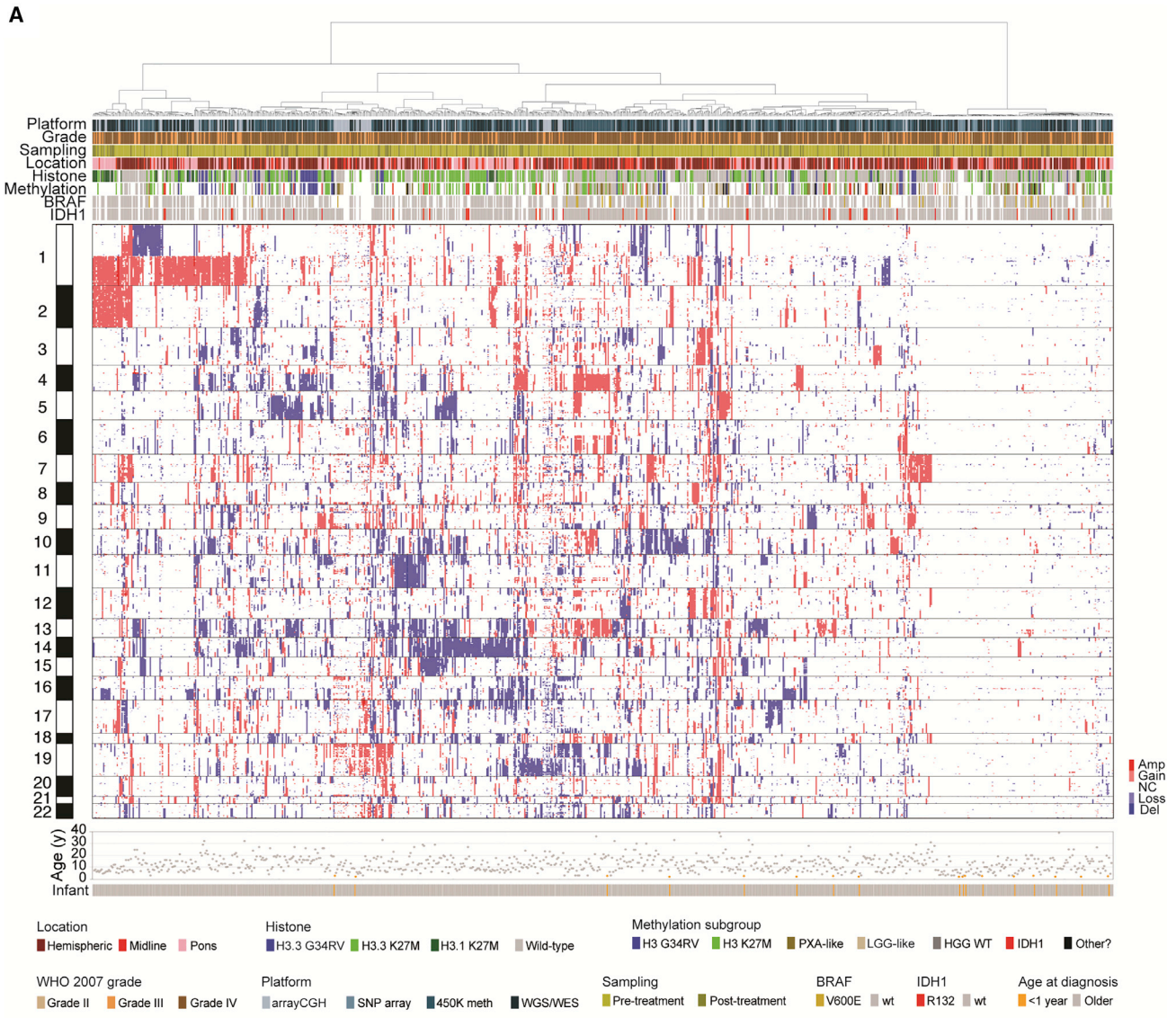
Figure 2. Methylation-based Subclassification of pHGG/DIPG

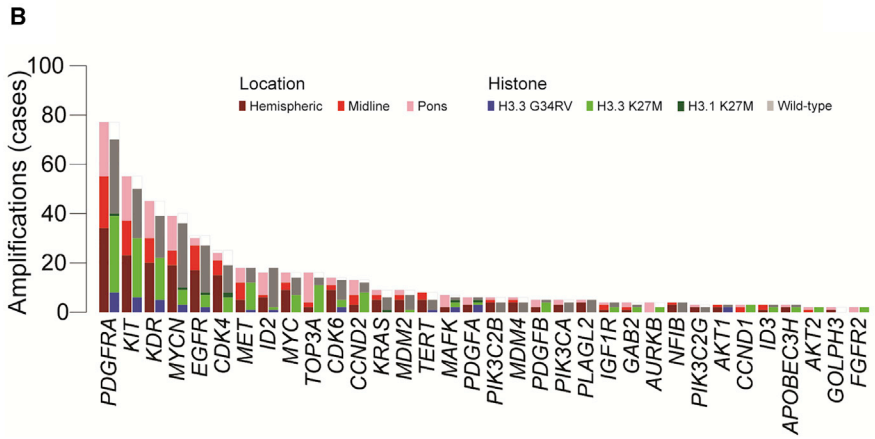
(A) Unsupervised hierarchical clustering and heatmap representation of β values for 441 samples profiled on the Illumina 450k BeadArray platform (red, high; blue, low). Samples are arranged in columns clustered by most variable 1,381 classifier probes. Age at diagnosis is provided below. Clinicopathological and molecular annotations are provided as bars according to the included key.

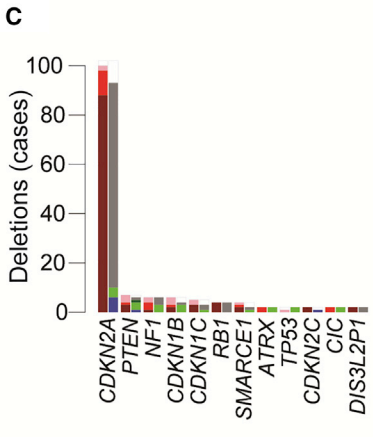
(B) Anatomical location of methylation-defined PXA-like (n = 43) and LGG-like (n = 27) cases. Left, sagittal section showing internal structures; right, external view highlighting cerebral lobes. Dark gold, PXA-like; tan, LGG-like. Radius of circle is proportional to the number of cases. Lighter shaded circles represent a non-specific designation of hemispheric, midline, or brainstem.

(C) Boxplot showing age at diagnosis of included cases, separated by simplified methylation subclass (n = 440). The thick line within the box is the median, the lower and upper limits of the boxes represent the first and third quartiles, and the whiskers 1.5× the interquartile range. ***Adjusted $p < 0.0001$ for all H3 G34R/V pairwise comparisons, t test; **adjusted $p < 0.01$ for LGG-like versus WT, t test.

(D) Kaplan-Meier plot of overall survival of cases separated by simplified methylation subclass, p value calculated by the log rank test (n = 307). See also Figure S2 and Table S2.







(legend on next page)

2, and the loss of 16q (Figure 4C). PXA-like tumors had frequent *CDKN2A/B* deletions and a unique loss at 1q, associated with shorter overall survival within this group (Figure S4B).

Whole-arm losses were also enriched in H3.3G34R/V tumors, specifically 3q, 4q, 5q, and 18q, where smallest regions of overlap were in some instances able to narrow the common region to a handful of candidate genes (Figure S4C). On chromosome 4q this appeared to target *FBXW7* at 4q31.3, also aligning with the GISTIC data (Figure 5A). Across three independent platforms, gene expression over the whole arm was significantly lower when 4q was lost (Agilent, $p = 0.00231$; Affymetrix, $p = 0.000102$; RNA sequencing (RNA-seq), $p = 0.0398$; Wilcoxon signed-rank test) (Figure S5A). (Table S5). There were also four patients with three different somatic coding mutations identified (below), two truncating and one missense, three of which were in hemispheric tumors, and two with *H3F3A* G34R (Figure 5B). In cases with 4q loss, median *FBXW7* gene expression was reduced compared with those with normal copy number (Agilent, $p = 0.029$; Affymetrix, $p = 0.015$; RNA-seq, $p = 0.4$; Mann-Whitney U test) (Figure 5C).

Within H3.3K27M tumors, we identified a recurrent amplification at 17p11.2 (n = 17; 170 kb to 11.96 Mb), across multiple platforms and significantly enriched in DIPGs, which appears to target *TOP3A* within these tumors (Figure 5D). Where available (n = 6) (Figure S5B), whole-genome sequencing data reveals this occurs via complex intra- and inter-chromosomal rearrangements (Figure 5E) leading to increased mRNA expression of *TOP3A* in amplified versus non-amplified cases in Agilent (n = 1), and Affymetrix and RNA-seq ($p = 0.011$ and $p = 0.016$, respectively, Mann-Whitney U test) data (Figure 5F) (Table S5). In an integrated dataset, *TOP3A* was the most differentially expressed gene in the region in amplified cases (adjusted $p = 0.00856$ Mann-Whitney U test). We further identified a single somatic missense mutation (C658Y) in an additional case of DIPG, and, taken together, *TOP3A* alterations were mutually exclusive with *ATRX* deletion/mutations found in H3.3K27M DIPG (0/13).

Whole-Genome and Exome Sequencing

Out of 372 sequenced cases (n = 118 whole genome, n = 247 exome, 7 both), we were able to retrieve raw data from 351 for integration of somatic variant calling (Table S6) and DNA copy-number changes. Of these, RNA-seq data was available for 47, allowing for candidate fusion gene nomination in 150 cases (RNA-seq or whole-genome sequencing restricted to high-confidence nominations in relevant pathway-associated genes, Table S6). Taking a conservative approach to variant calling given the disparate sequencing coverage (median 88×, range 16–295×), capture platforms, and availability of germline data, we report a median number of somatic single nucleotide variants (SNVs) and insertion/deletions (InDels) of 12, with 97% cases in the range 0–305. DNA copy neutral cases had significantly fewer somatic mutations (median 8.37 versus 21.32 SNVs/InDels per case), with those copy neutral cases also having no detectable mutations falling into the youngest age group (median = 3.9 years). There was only a modestly elevated mutation rate between samples taken post- compared with pre-treatment (1.5-fold, $p = 0.056$, t test) (Figure S6A). However, 11 cases had a vastly increased mutational burden, described as a hypermutator phenotype (median 13,735 SNVs/InDels, range 852–38,250), with distinctive mutational spectra from non-hypermutated pHGG/DIPG (Shlien et al., 2015), including three cases with plausible somatic activating *POLE* mutations (Figure S6B). *IDH1*-mutant cases were again excluded (n = 14), with genetic profiles identical to that described in adults for astrocytic tumors (13/14 *TP53*, 7/14 *ATRX* mutations), and oligodendroglial tumors (1p19q co-deletion, *TERT* promoter, *CIC*, *FUBP1* mutations) entirely absent (Figure S6C). We were thus left with an integrated dataset of 326 pHGG/DIPG, providing robust annotation of the most frequently altered genes across histone H3 subgroups and anatomical locations (Figure 6A). As well as known associations such as hemispheric H3.3G34R/V and *TP53/ATRX* (18/20, 90%; $p = 0.0001$), midline H3.3K27M and *FGFR1* (8/39, 20.5%; $p = 0.212$, not significant), pontine H3.1K27M and *ACVR1* (28/33, 84.8%; $p < 0.0001$), and PXA-like GBM with *BRAF* V600E (17/28, 60.7%; $p < 0.0001$), we also identified previously unrecognized co-segregating mutations including H3.3G34R/V and *ARID1B* (2/20, 10%; $p = 0.0720$), H3.3K27M DIPG and *ATM* and *ASXL1* (5/93, 10.7%; $p = 0.0473$), and H3.1K27M and *BCOR* (6/37, 16.2%; $p = 0.0022$, all Fisher's exact test) (Figure S6D). We also identified recurrent events in genes such as *PTPN23* (protein tyrosine phosphatase, non-receptor type 23, n = 5), *SOX10* (SRY-box 10, n = 5), *SRCAP* (Snf2-related CREBBP activator protein, n = 5), *DEPDC5* (DEP domain-contain 5, member of GATOR complex, n = 4), *SGK223* (PEAK1-related kinase activating pseudokinase, n = 4), and others (Figure 6B).

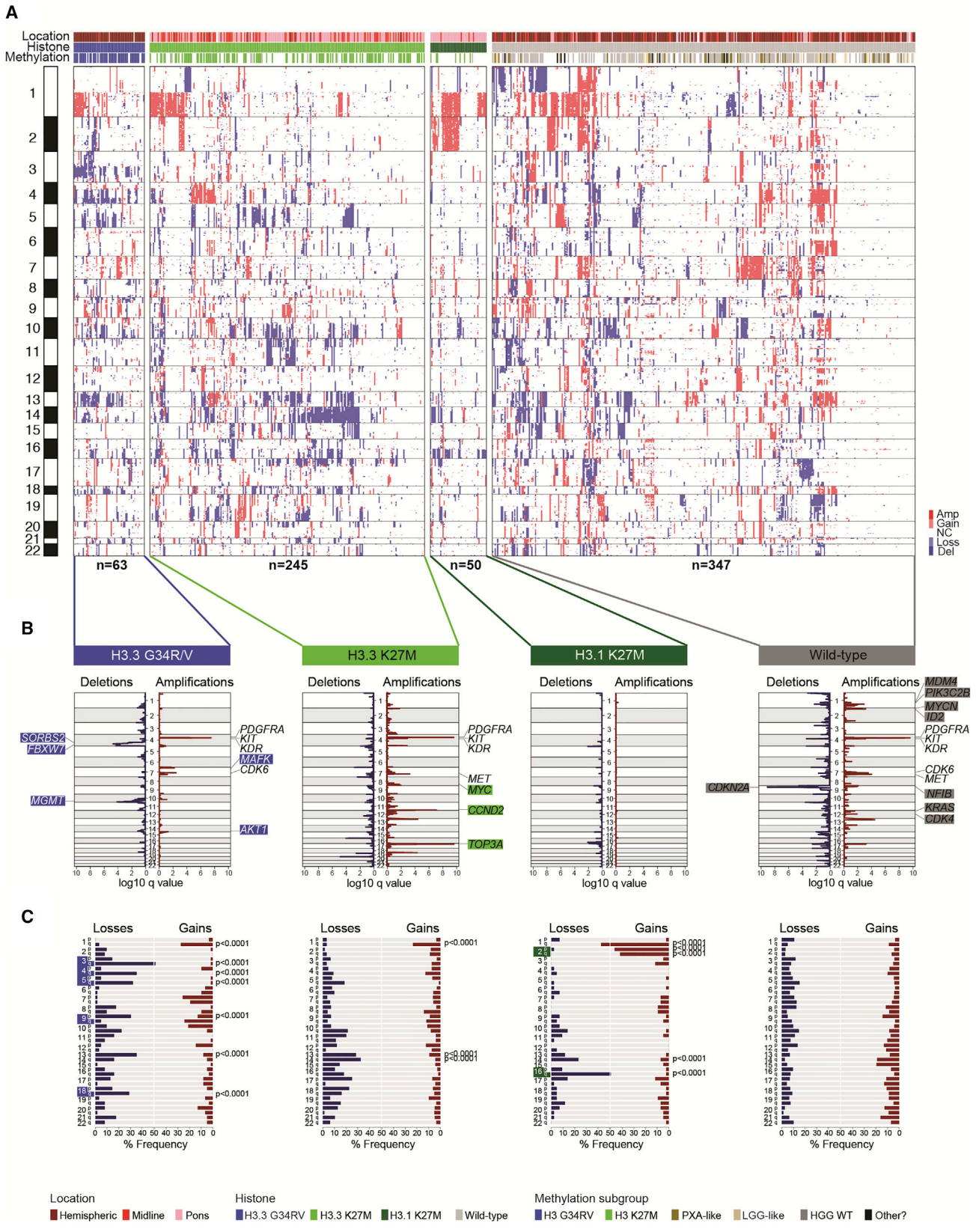
Integrated Pathway Analysis

Many of the rare variants we identified (Figure S6E) were found in genes associated with intracellular signaling pathways and processes more commonly targeted by high-frequency events, often in a mutually exclusive manner. In total, 297/326 (91.1%) of cases were found to harbor genetic alterations in one or more of nine key biological processes (Figure 7A). These included well-recognized pathways such as DNA repair (198/326, 60.7%), largely driven by *TP53* mutations (n = 160), but also by common mutually exclusive ($p < 0.0001$, Fisher's exact test) activating truncating alterations in *PPM1D* (n = 18), as well as heterozygous mutations in a diverse set of genes including those involved in homologous recombination (*ATM, BRCA2, BLM, ATR, PALB2, RAD50,* and *RAD51C*) and numerous Fancomi anemia genes (*BRIP1, FANCM, FANCA,* and *FANCG*), among others (Figure S7A). Although *TP53* is almost always found in concert with H3.3G34R/V in the cerebral hemispheres, these additional DNA repair pathway mutations were enriched in

Figure 3. DNA Copy-Number Aberrations in pHGG/DIPG

(A) Heatmap representation of segmented DNA copy number for 834 pHGG/DIPG profiled across one or more of seven different platforms (dark red, amplification; red, gain; dark blue, deletion; blue, loss). Samples are arranged in columns clustered by gene-level data across the whole genome. Age at diagnosis is provided below. Clinicopathological and molecular annotations are provided as bars according to the included key.

(B and C) Barplot of all recurrent focal amplifications (B) and deletions (C) across all 834 cases, in order of frequency, and colored independently by both anatomical location and histone mutation. See also Figure S3 and Table S3.



(legend on next page)

H3.3K27M DIPG (36/68, 52.9%). Also co-segregating with H3.3G34R/V and *TP53* is *ATRX*, although mutations/deletions of the latter gene are also frequently found in conjunction with H3.3K27M (28/54, 51.8%). *ATRX* accounts for a large proportion of the cases harboring mutations in genes coding for chromatin modifiers (54/118, 45.8%); however, there is a diverse set of readers, writers, and erasers also targeted at lower frequency, especially in DIPG, including the previously mentioned *BCOR* (n = 14) and *ASXL1* (n = 6) in addition to *SETD2* (n = 8), *KDM6B* (n = 6), *SETD1B* (n = 5), and *ARID1B* (n = 5) among many others (Figure S7B).

While *CDKN2A/CDKN2B* deletions were almost entirely absent from DIPG (1/154, 0.65%), dysregulation of the G_1/S cell-cycle checkpoint was common throughout anatomical locations and subgroups (82/326, 25.2%), with amplifications of *CCND2* and deletions of *CDKN2C* predominating in the pons (n = 5/7 and 5/5 DIPG, respectively), in contrast to recurrent homozygous *RB1* deletions and *CDK6* amplifications (n = 6/7 and 4/6 hemispheric) (Figure S4A) (Figure S7C).

Subgroup-specific dysregulation was also observed when considering discrete components of the RTK-PI3K-MAPK pathway. In total, 201/326 (61.7%) cases harbored alterations in any given node; however, for H3.3G34R/V this was predominantly at the RTK level (11/20, n = 9 *PDGFRA*) (Figure S7D), whereas H3.1K27M cases were enriched for PI3K/mTOR alterations (17/37, n = 9 *PIK3CA,* n = 5 *PIK3R1*) (Figure S7E), and H3 WT cases harbored the highest frequency of MAPK alterations (mainly *BRAF* V600E in PXA-like, n = 5/10 plus one *NF1*) (Figure S7F). *NRTK1-NRTK3* fusions were enriched in the infant group (4/6 fusions under 1 year old, median age 3.25 versus 8.5 years, p = 0.00033, t test) (Figure S7D). We further identified mutations in genes regulating mTOR signaling, including *TSC2* (n = 3), *RPTOR,* and *MTOR* itself (both n = 2), as well as a diverse series of SNVs and fusion candidates in MAPKs across all subgroups and locations (*MAP2K7*, *MAP3K15*, *MAP3K4* and others) (Figure S7F).

BMP signaling was significantly enriched in H3.1K27M DIPG due to the strong correlation with *ACVR1* mutations; however, alterations in other pathway members such as amplification of *ID2* (n = 10) or *ID3* (n = 3) and mutations in *BMP3* (n = 5), *BMP2K* (n = 3), and others across locations and subgroups, extends the proportion of tumors for which this pathway may be relevant (62/326, 18.7%) (Figure S7G). There was also a subset of cases harboring alterations in members of the WNT signaling pathway (16/326, 4.9%), including *AMER1, APC* (both n = 3), and *WNT8A*, *WNT9A, PLAGL2*, and *TCF7L2* (all n = 2) (Figure S7H).

Uniquely, the accumulated data uncovered a series of additional processes involved in maintenance of DNA replication, genome integrity, or transcriptional fidelity, targeted by infrequent but mutually exclusive alterations in pHGG and DIPG. These included mutations in splicing factors (*SF3A1, SF3A2, SF3A3, SF3B1, SF3B2,* and *SF3B3,* total n = 10), sister chromatid segregation (*STAG2, STAG3*, and *ESPL*1, total n = 9), pre-miRNA processing (*DICER* and *DROSHA*, total n = 4), DNA polymerases (*POLK, POLQ,* and *POLR1B,* total n = 4), as well as genes involved in centromere (*CENPB*, n = 3) and telomere maintenance (*PML*, n = 2; *TERT*, n = 7) (Figure S7I). *TERT* promoter mutations were found in 5/326 (1.5%) cases; however, alternative lengthening of telomeres (ALT) status was only available for 26 cases, although the 5 ALT-positive samples (19.2%) were mutually exclusive with *TERT* alterations.

We incorporated the integrated dataset into a pathway enrichment analysis (significant gene sets, false discovery rate [FDR] < 0.05, visualized as interaction networks by Cytoscape Enrichment Map) in order to gain additional insight into dysregulated biological processes. In addition to the subgroup-specific differential targeting of distinct nodes within common signaling pathways already described (e.g., RTK, PI3K/mTOR, and MAPK), additional dysregulated processes across the diversity of the disease were identified (Figure 7B). This revealed the perhaps not unexpected dysregulation of numerous developmental and CNS-associated gene sets (various immature organ systems, neuronal communication), but also previously unrecognized areas such as nuclear transport, cell migration, and the immune response (Table S7), which may provide further insight into disease biology as well as represent potential therapeutic strategies targeting key regulators of tumor phenotype. Indeed, neuronal communication with pGBM and DIPG cells is a recently demonstrated microenvironmental driver of pediatric glioma growth (Qin et al., 2017; Venkatesh et al., 2015).

Histone H3/IDH1 WT Subgroups

Finally, we wanted to explore those cases absent of any histone H3 or *IDH1* mutations in more depth. Using a t statistic-based stochastic neighbor embedding projection of the 450k methylation data, we identified three distinct clusters of tumors separate from the G34, K27, and IDH1 groups (Figure 8A). Consensus clustering of the H3/IDH1 WT cases alone confirmed the presence of three robust subgroups (Figure 8B), which were also recapitulated by unsupervised hierarchical clustering of the 10,000 probe classifier subset (Figure 8C). These groups included a largely hemispheric set of tumors containing, but not restricted to, the PXA- and LGG-like subgroups (WT-A). These tumors were driven by *BRAF* V600E, *NF1* mutations, or fusions in RTKs including *MET*, *FGFR2*, and *NTRK2,3* (Figure 8D). Although including many younger patients, the ages varied widely (Figure 8E). Regardless, this group had the best overall survival (median = 63 months, p < 0.0001 versus rest, log rank test) (Figure 8F), with the non-PXA/LGG-like tumors within this group themselves having an extended median survival time of

Figure 4. Subgroup-specific Copy-Number Changes in pHGG/DIPG

(A) Heatmap representation of segmented DNA copy number for 705 pHGG/DIPG separated for known histone mutation subgroup (dark red, amplification; red, gain; dark blue, deletion; blue, loss). Samples are arranged in columns clustered by gene-level data across the whole genome. Clinicopathological and molecular annotations are provided as bars according to the included key.

(B) GISTIC analysis of focal amplifications and deletions for histone mutation subgroups. Log_{10} values are plotted across the genome for both amplifications (dark red) and deletions (dark blue), with significantly enriched events labeled by likely driver genes. Subgroup-specific genes are highlighted by the appropriate color.

(C) Barplot of frequency of whole chromosomal arm gains (red) and losses (blue) for each subgroup. Significantly enriched alterations (p < 0.0001, Fisher's exact test) are labeled, with subgroup-specific arm changes highlighted by the appropriate color. See also Figure S4 and Table S4.

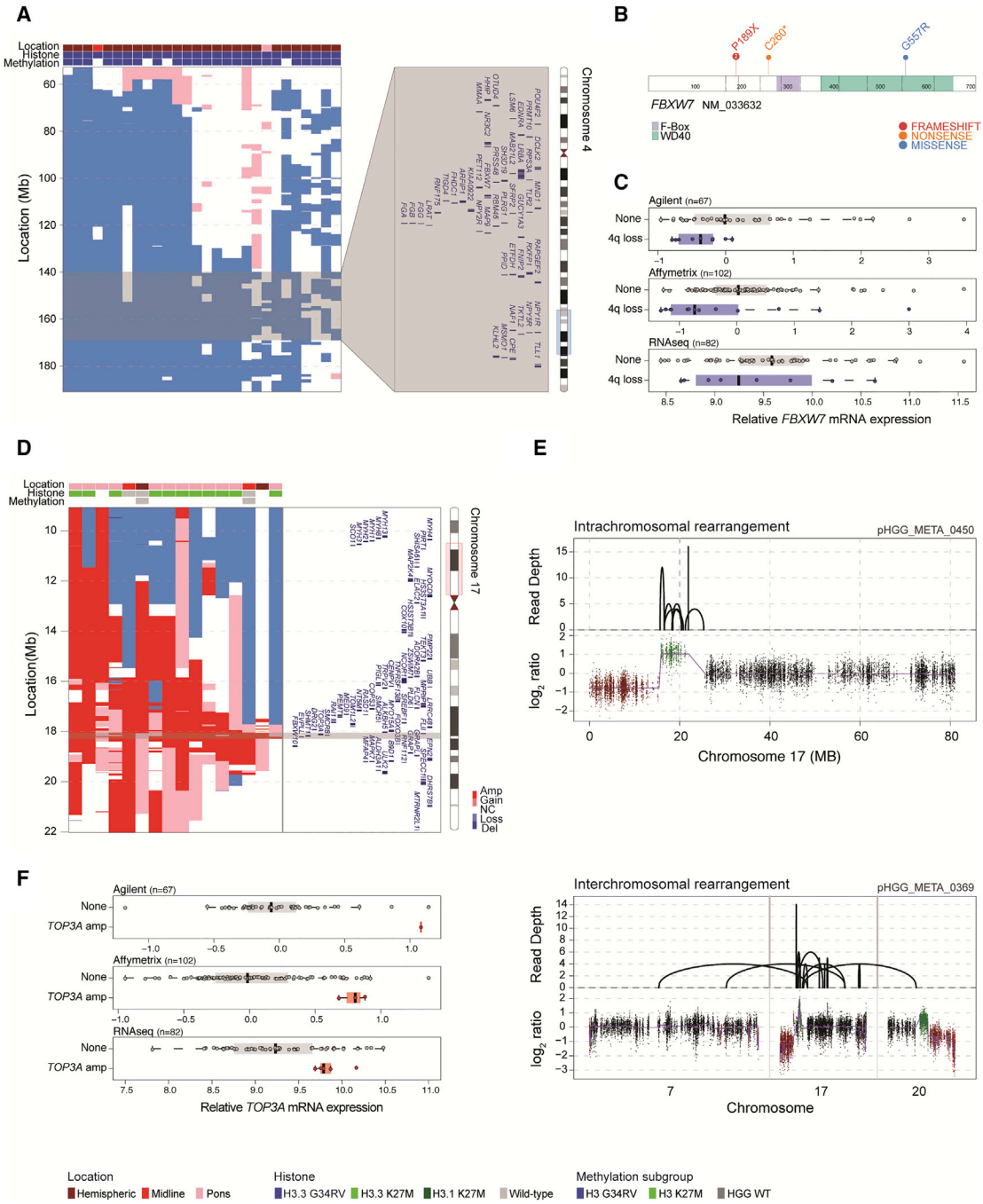


Figure 5. Alterations Targeting *FBXW7* in H3.3G34R/V pHGG and *TOP3A* in H3.3K27M DIPG

(A) Segmented exon-level DNA copy-number heatmaps for 4q loss in H3.3G34R/V tumors (dark red, amplification; red, gain; dark blue, deletion; blue, loss; n = 28). An ideogram of chromosome 4 is provided indicating enlarged genome browser view and genes within common regions targeted across samples (gray). Clinicopathological and molecular annotations are provided as bars according to the included key.

(B) Cartoon representation of amino acid position for four somatic mutations found in *FBXW7*, colored by annotated functional domains and numbers provided for recurrent variants.

(C) Boxplots representing gene expression differences between *FBXW7* lost/mutated cases (blue) and those with normal copy/WT (gray) in three independent gene expression platform datasets. The thick line within the box is the median, the lower and upper limits of the boxes represent the first and third quartiles, and the whiskers 1.5× the interquartile range.

(D) Segmented exon-level DNA copy-number heatmaps for 17p11.2 amplification in predominantly H3.3K27M DIPG (dark red, amplification; red, gain; dark blue, deletion; blue, loss; n = 17). Chromosome 17 ideogram is provided indicating enlarged genome browser view and genes within common regions targeted across samples (gray). Clinicopathological and molecular annotations are provided as bars according to the included key.

(legend continued on next page)

38 months (p = 0.00928 versus other H3/IDH1 WTs, log rank test). Taking an integrated gene expression profiling dataset (Figures S8A–S8E), these tumors were found to have upregulation of gene signatures associated with cytokine signaling and cell junction organization (Figures S8F and S8G). A second group of tumors (WT-B) were found in all anatomical compartments, and were distinguished by chromosome 2 gains (Figure 8C) and, most notably, by high-level amplifications in *EGFR*, *CDK6*, and *MYCN* (p = 0.00033, p = 0.0299, p = 0.00037, respectively, Fisher's exact test), with an imperfect overlay to the classifier "GBM_pedRTK" and "GBM_MYCN" groups (Figure 8D). This group had strong upregulation of MYC target genes, and had the poorest overall survival (median = 14 months) (Figure 8F). The remaining cases encompassed a methylation classifier group described as "HGG_MID," although in fact were split 80:20 hemispheric:midline (WT-C) (Figure 8C). This group was enriched for chromosome 1p and 20q loss, 17q gain (p = 0.00595, p = 0.0286, p = 0.0478, respectively, Fisher's exact test) (Figure 8C), harbored *PDGFRA* and *MET* amplifications (p = 0.0159, Fisher's exact test) (Figure 8D), and was strongly associated with the adult GBM-defined "Proneural" gene signature. These patients had a median survival of 18 months.

Although there remain tumors without detectable genetic alterations, we are nonetheless able to assign clinically meaningful subgroups with plausible driver alterations to the vast majority of pediatric HGG/DIPG.

DISCUSSION

Integrated molecular profiling has revolutionized the study of diffusely infiltrating high-grade glial tumors in children, providing evidence for unique mechanisms of molecular pathogenesis reflecting their distinct developmental origins (Baker et al., 2015; Jones and Baker, 2014). Although they are relatively rare, the present study accumulates 1,067 unique cases, a number similar to the aggregated analysis of the The Cancer Genome Atlas adult LGG/GBM cohorts (n = 1122, with grade III included in the "lower-grade" series) (Ceccarelli et al., 2016). Although there are clearly the usual caveats with such retrospective analyses of inconsistently annotated and treated cases, the cohort appears to represent a clinically useful approximation of the diversity of the pHGG/DIPG population.

In adults, the key distinction is between *IDH1* mutant (G-CIMP/*ATRX*/*TP53* or 1p19q co-deleted/*TERT* promoter mutated) and WT (classical, mesenchymal, PA-like) (Ceccarelli et al., 2016), whereas in the childhood setting *IDH1* mutations were restricted to a small proportion (6.25%) of tumors mostly in adolescents (representing the tail end of an overwhelmingly adult disease), and harbored only rare examples of the common alterations seen in WT adult GBM (e.g., 4.9% *EGFR* mutation/amplification). Instead, most prominent among the differences between pediatric and adult studies is the frequency of hotspot mutations in genes encoding histone H3 variants: 2/820 (0.2%) in adults (Ceccarelli et al., 2016) versus 449/893 (50.3%) in the present pHGG/DIPG series.

The importance of recurrent H3 mutations in the childhood setting has become increasingly clear since their unexpected discovery in 2012 (Schwartzentruber et al., 2012; Wu et al., 2012), with clear clinicopathological differences associated with distinct variants (Jones and Baker, 2014; Jones et al., 2016; Sturm et al., 2014), and fundamental insights into mechanisms of epigenetically linked tumorigenesis (Bender et al., 2013; Bjerke et al., 2013; Chan et al., 2013; Funato et al., 2014). Despite this, precisely how we can target these mutations clinically remains elusive (Grasso et al., 2015; Hennika et al., 2017). Data from such a large series of tumors demonstrates the robustness of the histone-defined subgroups in terms of anatomical location, age of incidence, clinical outcome, methylation and gene expression profiles, copy-number changes, co-segregating somatic mutations, and pathway dysregulation. As most of the non-histone molecular alterations previously reported in pHGG/DIPG have been relatively infrequent, it is only through this accumulated dataset that we have been able to uncover subgroup-specific genes/processes that may play a role as diagnostic, prognostic, or predictive markers or drug targets in these diseases.

H3.3G34R/V-mutant tumors are restricted to the cerebral hemispheres and co-segregate with *ATRX* and *TP53* mutations; they are also the only pediatric subgroup to harbor frequent *MGMT* promoter methylation (Korshunov et al., 2015). Copy-number profiling of 63 cases highlighted a significant enrichment of chromosomal arm losses at 3q, 4q, 5q, and 18q, further refined by smallest region of overlap and GISTIC analyses. At 4q31.3, this identified *FBXW7* as a candidate gene target of the loss. *FBXW7* encodes a member of the F box protein family and is frequently deleted/mutated in cancer, supporting its tumor-suppressive function (Davis et al., 2014); notably in relation to H3.3G34R/V it has been reported to play a role in MYC/MYCN stabilization through its action as a component of the SCF-like ubiquitin ligase complex that targets MYC/MYCN for proteasomal degradation (Welcker et al., 2004; Yada et al., 2004). With MYCN upregulated in H3.3G34R/V tumors through differential H3K36me3 binding (Bjerke et al., 2013), this observation adds to the mechanisms by which Myc proteins exert their influence in this subgroup, and provide further rationale for the observed effects of disrupting these interactions, such as with Aurora kinase A inhibitors which target the direct interaction between the catalytic domain of Aurora A and a site flanking Myc Box I that also binds SCF/FbxW7 (Richards et al., 2016).

H3.3K27M tumors are found in two-thirds of DIPG and non-brainstem midline pHGG alike, where they are associated with a shorter overall survival in both locations, as well as in the small number of cases reported in the cortex. Although presumably reflecting a common or overlapping origin, the pattern of co-segregating mutations differ, e.g., *PDGFRA* alterations predominating

(E) Sequencing coverage (top) and $\log_2$ ratio plot (bottom) for chromosomes 7, 17, and 20 for two cases, showing complex intra- or inter-chromosomal rearrangements leading to specific copy-number amplification of *TOP3A*.

(F) Boxplots representing gene expression differences between *TOP3A* amplified cases (red) and those with normal copy (gray) in three independent gene expression platform datasets. The thick line within the box is the median, the lower and upper limits of the boxes represent the first and third quartiles, and the whiskers 1.5× the interquartile range. See also Figure S5 and Table S5.

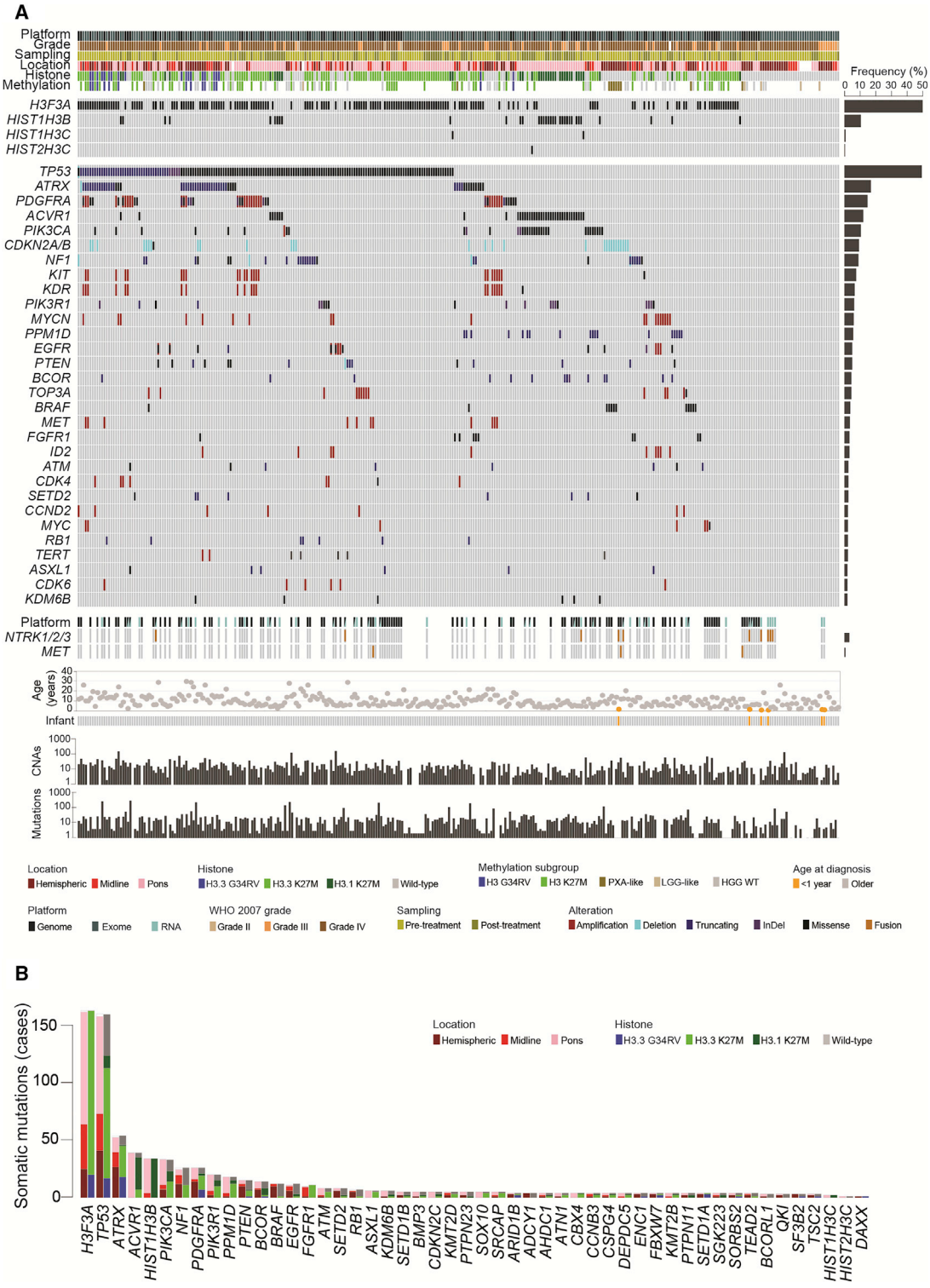


Figure 6. Somatic Mutations in pHGG/DIPG

(A) Oncoprint representation of an integrated annotation of somatic mutations and DNA copy-number changes for the 30 most frequently altered genes in 326 pHGG/DIPG ($n \geq 6$, frequency barplot on the right). Selected common fusion events are also shown where available. Samples are arranged in columns with genes labeled along rows. Age at diagnosis is provided below. Underneath, barplots are provided on a $\log_{10}$ scale for numbers of copy-number aberrations and somatic mutations per case. Clinicopathological and molecular annotations are provided as bars according to the included key.

(B) Barplot of all recurrent somatic mutations across all 326 cases, in order of frequency, and colored independently by both anatomical location and histone mutation. See also Figure S6 and Table S6.

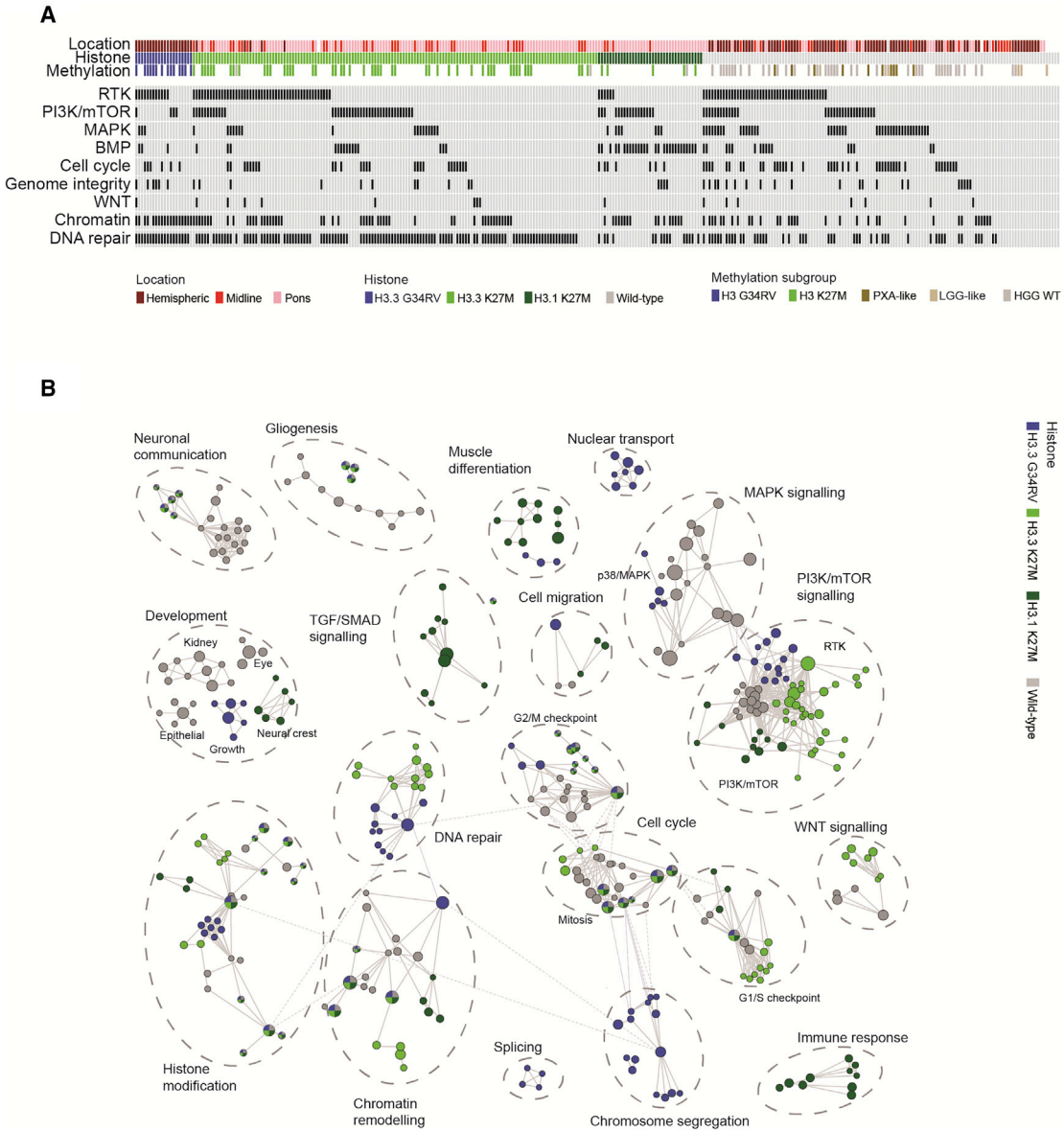


Figure 7. Integrated Pathway Analysis of pHGG/DIPG

(A) Oncoprint-style representation of an integrated annotation of somatic mutations and DNA copy-number changes in one or more of nine commonly targeted pathways in 326 pHGG/DIPG ($n \geq 6$, frequency barplot on the right). Samples are arranged in columns with pathways labeled along rows. Clinicopathological and molecular annotations are provided as bars according to the included key.

(B) Pathway enrichment analysis of pHGG/DIPG subgroups. Distinct pathways and biological processes between the subgroups are colored appropriately (FDR $q < 0.01$). Nodes represent enriched gene sets, which are grouped and annotated by their similarity according to related gene sets. Node size is proportional to the total number of genes within each gene set. The illustrated network map was simplified by manual curation to remove general and uninformative sub-networks. See also Figure S7 and Table S7.

in the pons, and *FGFR1* variants being largely restricted to the thalamus (Fontebasso et al., 2014). Our analysis of more than 300 cases further identifies differential amplification of *CCND2* (DIPG) and *CDK4* (non-brainstem midline), and, most strikingly, an amplification at 17p11.2 involving *TOP3A* in H3.3K27M DIPG. This complex rearrangement often involves loss of the more distal part of 17p involving *TP53*, along with intra- or inter-chromosomal translocations to deliver an increase in *TOP3A* copy number and gene expression. *TOP3A* encodes DNA topoisomerase III alpha, which forms a complex with BLM (Wu et al., 2000), has an important role in homologous recombination (Yang et al., 2010), and has been implicated in maintenance of the ALT phenotype (Temime-Smaali et al., 2009). Notably, *TOP3A* amplification/mutation was found to be mutually exclusive with *ATRX* mutation in H3.3K27M DIPG, with depletion by small interfering RNA reducing ALT cell survival (Temime-Smaali et al., 2008), and therefore represents a potential therapeutic target in this subgroup.

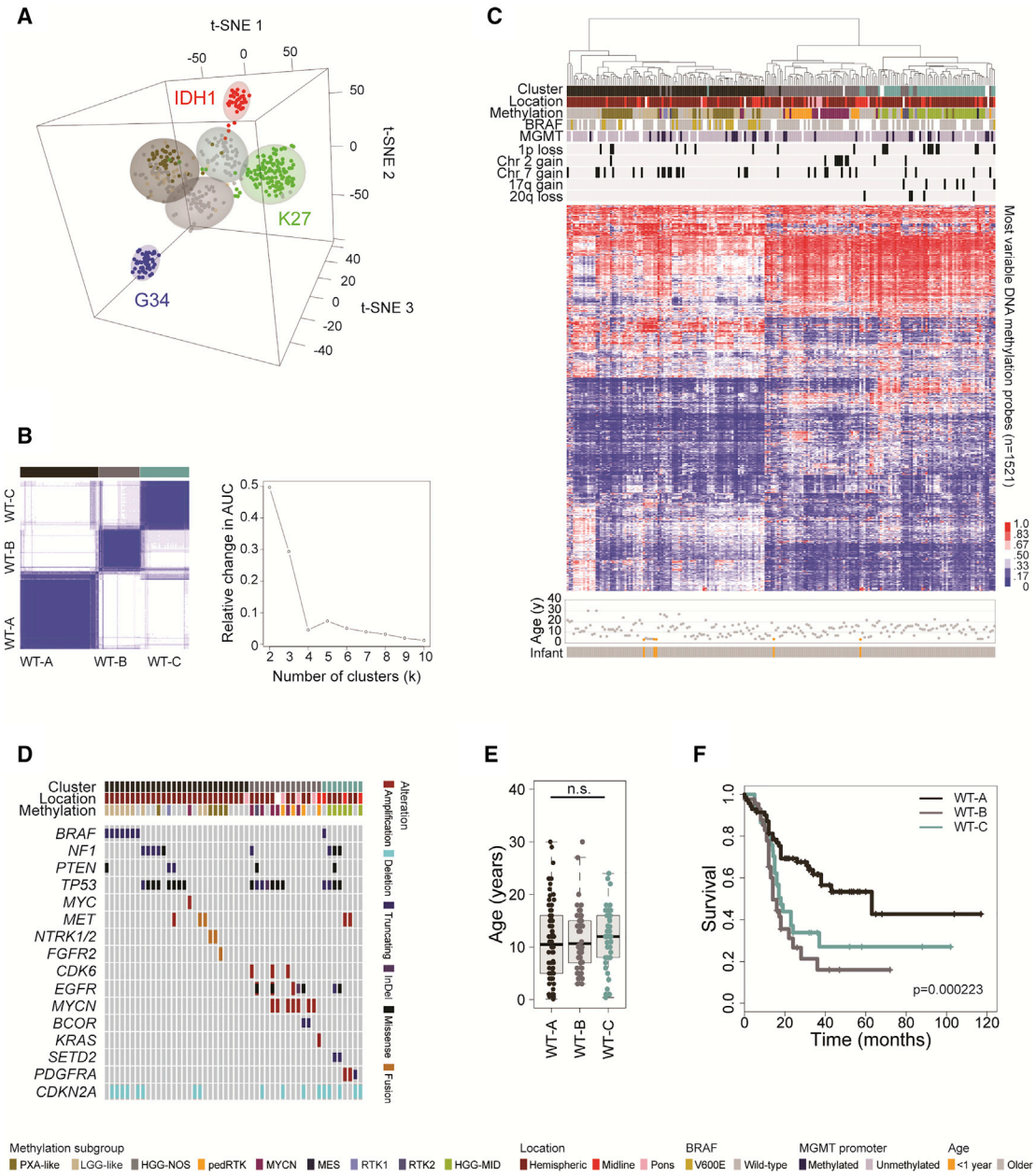


Figure 8. Integrated Analysis of H3/IDH1 WT pHGG/DIPG

(A) t Statistic-based stochastic neighbor embedding (t-SNE) projection of the combined 450k methylation dataset (n = 441). The first three projections are plotted in the x, y, and z axes, with samples represented by dots colored by histone H3G34 (blue), H3K27 (green), IDH1 (red), PXA-like (dark gold), LGG-like (tan), and "others" (gray).

(B) K means consensus clustering on the H3/IDH1 WT cases highlights three stable clusters (left, black/brown [WT-A], gray/pink [WT-B], and dark cyan [WT-C]) as the most robust subdivision of the data (right, area under the curve analysis for different cluster numbers).

(C) Unsupervised hierarchical clustering and attendant heatmap of the H3/IDH1 WT cases (n = 219). Samples are arranged in columns clustered by the most variable 1,521 classifier probes. Age at diagnosis is provided below. Clinicopathological and molecular annotations are provided as bars according to the included key.

(D) Oncoprint representation of an integrated annotation of somatic mutations and DNA copy-number changes for the H3/IDH1 WT cases (n = 50). Samples are arranged in columns with genes labeled along rows. Age at diagnosis is provided below. Clinicopathological and molecular annotations are provided as bars according to the included key.

(E) Boxplot showing age at diagnosis of H3/IDH1 WT subgroups, separated by anatomical location (n = 190). The thick line within the box is the median, the lower and upper limits of the boxes represent the first and third quartiles, and the whiskers 1.5× the interquartile range.

(F) Kaplan-Meier plot of overall survival of H3/IDH1 WT subgroups separated by anatomical location, p value calculated by the log rank test (n = 150). See also Figure S8 and Table S8.

H3.1K27M tumors by contrast are restricted to the pons, patients are younger and with a slightly longer survival (Castel et al., 2015), and are largely defined at the copy-number level by whole chromosomal arm gains and losses (Taylor et al., 2014a). They have the well-recognized association with *ACVR1* mutation (Taylor et al., 2014b); however, we also identify an enrichment of downstream PI3K pathway mutations (*PIK3CA* and *PIK3R1*) in comparison with the largely upstream RTK alterations present in H3.3K27M DIPGs, important in designing stratified trials and combinatorial therapies. Further association with mutations of the BCL6 repressor gene *BCOR*, commonly altered in medulloblastomas, neuroepithelial tumors, and sarcomas, highlights a further avenue for interventional study through its regulation of the SHH pathway (Tiberi et al., 2014).

In H3/IDH1 WT cases, methylation profiling refines the heterogeneous collection of tumors, particularly identifying two predominantly hemispheric intermediate risk subgroups that classify alongside other entities (PXA- and LGG-like) in a larger series of better outcome tumors (WT-A). These had already been strongly linked with dysregulation of the MAPK pathway (*BRAF* V600E) (Korshunov et al., 2015) along with *CDKN2A/CDKN2B* deletion (Nicolaides et al., 2011). However, with molecular markers such as losses at 1q and 17p appearing to confer a worse outcome there may be more than one subgroup within this entity, and a co-clustering group of H3/IDH1 WT tumors appeared distinctly driven by somatic *NF1* mutation. The LGG-like tumors generally occur in very young patients, where the appearance of few genetic alterations and a significantly better prognosis is shared by the majority of infant HGG. Gene fusion events, including those targeting *NTRKs1-NTRK3*, are common in this age range. Notably this enhanced survival is restricted to patients diagnosed under 12 months of age, and is not recapitulated in the 1–3 year age group, although this is the common clinical definition of ''infants'' in many centers.

Excluding these morphologically high-grade but biologically and clinically low-grade tumors, the remaining H3/IDH1 WT cases can be further split into two poor-outcome groups driven by *EGFR/MYCN/CDK6* (WT-B) or *PDGFRA/MET* (WT-C) or amplifications. These groups overlap with other methylation-based classification groups (PDGFRA versus EGFR versus MYCN (Korshunov et al., 2017); ''GBM_pedRTK'' versus ''GBM_MYCN'' versus ''HGG_MID'' (molecularneuropathology.org/mnp), however, are uniquely defined here spanning anatomical locations and integrated with sequencing data. Further exploration of these heterogeneous subgroups in order to refine integrated molecular diagnostics to prioritize patient subpopulations for stratified treatment remains a priority.

The remarkable biological diversity spanning pediatric malignant glioma is finally demonstrated by the <5% tumors with a hypermutator phenotype, some of the greatest mutational burdens in all human cancer, and candidates for immune checkpoint inhibitors (Bouffet et al., 2016). Previously unrecognized processes altered in small subsets of tumors identified through this meta-analysis, such as the splicing machinery, miRNA regulation, and the WNT pathway offer further areas for exploration. The thorough cataloging of dysregulated molecular pathways across the whole spectrum of pediatric diffusely infiltrating gliomas in the present study provides the basis for novel therapeutic development.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Patient Samples
- METHOD DETAILS
  - Nucleic Acid Extraction
  - Sanger Sequencing of *H3F3A* / *HIST1H3B*
  - Methylation Profiling
  - Exome and RNA Sequencing
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Published Data Sources
  - DNA Copy Number
  - DNA Methylation
  - mRNA Expression
  - Sequence Analysis
  - Candidate Fusion Gene Nomination
  - Inferred Tumor Purity
  - Integrated Analysis of Driver Events
  - Pathway Analysis
  - Statistical Analysis
- DATA AND SOFTWARE AVAILABILITY
- ADDITIONAL RESOURCES

SUPPLEMENTAL INFORMATION

Supplemental Information includes eight figures and eight tables and can be found with this article online at http://dx.doi.org/10.1016/j.ccell.2017.08.017.

AUTHOR CONTRIBUTIONS

A.M., M.F., A.O.v.B., M.B., and C.J. conceived the study. A.M. and C.J. analyzed data and wrote the manuscript. A.B., D.C., E.I.D., J.F.S., K.T., L.B., M.V., M.N., S.T., and V.M. performed molecular analysis of unpublished samples. M.C., S.P., L.R.B., S.A.-S., A.N.K., D.M.K., K.M., K.-K.N., M.S., and C.K. carried out histopathological assessment of cases. M.M., J.G., C.H., N.J., S.J.B., S.M.P., and D.T.W.J. provided data. L.M., S.Z., S.V., H.C.M., A.J.M., C.C., N.E.-W., J.P., J.S., R.M.R., A.S.M., L.S., S.T., D.H.-B.B., A.M.C., C.d.T., O.C., J.M., and M.M. provided samples and clinical annotation. M.B., P.R., A.J.W., H.J.H., S.G., and A.R. constructed analytical and visualization tools and databases. All authors approved the manuscript.

ACKNOWLEDGMENTS

This work was supported by Cancer Research UK (grants C13468/A13982 and C13468/A23536), CRIS Cancer Foundation, Abbie's Army and the DIPG Collaborative, the Cure Starts Now Foundation, Christopher's Smile, McKenna Claire Foundation, Lyla Nsouli Foundation, National Institutes of Health (grants R01NS085336 and R01NS091620), The Dragon Master Foundation, The Kortney Rose Foundation, The Musella Foundation For Brain Tumor Research & Information, Gray Matters Foundation, Pediatric Brain Tumor Foundation, and the INSTINCT network funded by The Brain Tumour Charity, Great Ormond Street Children's Charity, and Children with Cancer UK. The authors acknowledge NHS funding to the NIHR Biomedical Research Centre at The Royal Marsden and the ICR. LTB was recipient of São Paulo Research Foundation fellowships (2011/08523-7 and 2012/08287-4). This study makes use of data generated by the St. Jude Children's Research Hospital – Washington University Pediatric Cancer Genome Project, the Hospital for Sick Children, the McGill University-DKFZ Pediatric Brain Tumor Consortium, the International Cancer Genomics Consortium PedBrain Project and the Cancer Research

UK Genomics Initiative (C13468/A14078) and supports the Children's Brain Tumor Tissue Consortium and The Pacific Pediatric Neuro-Oncology Consortium (PNOC) CAVATICA: Project OPEN DIPG initiative.

Received: May 5, 2017
Revised: July 14, 2017
Accepted: August 29, 2017
Published: September 28, 2017

REFERENCES

Baker, S.J., Ellison, D.W., and Gutmann, D.H. (2015). Pediatric gliomas as neurodevelopmental disorders. Glia *64*, 879–895.

Barrow, J., Adamowicz-Brice, M., Cartmill, M., MacArthur, D., Lowe, J., Robson, K., Brundler, M.A., Walker, D.A., Coyle, B., and Grundy, R. (2011). Homozygous loss of ADAM3A revealed by genome-wide analysis of pediatric high-grade glioma and diffuse intrinsic pontine gliomas. Neuro Oncol. *13*, 212–222.

Bax, D.A., Mackay, A., Little, S.E., Carvalho, D., Viana-Pereira, M., Tamber, N., Grigoriadis, A.E., Ashworth, A., Reis, R.M., Ellison, D.W., et al. (2010). A distinct spectrum of copy number aberrations in pediatric high-grade gliomas. Clin. Cancer Res. *16*, 3368–3377.

Bender, S., Tang, Y., Lindroth, A.M., Hovestadt, V., Jones, D.T., Kool, M., Zapatka, M., Northcott, P.A., Sturm, D., Wang, W., et al. (2013). Reduced H3K27me3 and DNA hypomethylation are major drivers of gene expression in K27M mutant pediatric high-grade gliomas. Cancer Cell *24*, 660–672.

Bjerke, L., Mackay, A., Nandhabalan, M., Burford, A., Jury, A., Popov, S., Bax, D.A., Carvalho, D., Taylor, K.R., Vinci, M., et al. (2013). Histone H3.3. mutations drive pediatric glioblastoma through upregulation of MYCN. Cancer Discov. *3*, 512–519.

Bouffet, E., Larouche, V., Campbell, B.B., Merico, D., de Borja, R., Aronson, M., Durno, C., Krueger, J., Cabric, V., Ramaswamy, V., et al. (2016). Immune checkpoint inhibition for hypermutant glioblastoma multiforme resulting from germline biallelic mismatch repair deficiency. J. Clin. Oncol. *34*, 2206.

Buczkowicz, P., Hoeman, C., Rakopoulos, P., Pajovic, S., Letourneau, L., Dzamba, M., Morrison, A., Lewis, P., Bouffet, E., Bartels, U., et al. (2014). Genomic analysis of diffuse intrinsic pontine gliomas identifies three molecular subgroups and recurrent activating ACVR1 mutations. Nat. Genet. *46*, 451–456.

Cai, H., Kumar, N., and Baudis, M. (2012). arrayMap: a reference resource for genomic copy number imbalances in human malignancies. PLoS One *7*, e36944.

Cai, H., Kumar, N., Ai, N., Gupta, S., Rath, P., and Baudis, M. (2014). Progenetix: 12 years of oncogenomic data curation. Nucleic Acids Res. *42*, D1055–D1062.

Carvalho, D., Mackay, A., Bjerke, L., Grundy, R.G., Lopes, C., Reis, R.M., and Jones, C. (2014). The prognostic role of intragenic copy number breakpoints and identification of novel fusion genes in paediatric high grade glioma. Acta Neuropathol. Commun. *2*, 23.

Castel, D., Philippe, C., Calmon, R., Le Dret, L., Truffaux, N., Boddaert, N., Pages, M., Taylor, K.R., Saulnier, P., Lacroix, L., et al. (2015). Histone H3F3A and HIST1H3B K27M mutations define two subgroups of diffuse intrinsic pontine gliomas with different prognosis and phenotypes. Acta Neuropathol. *130*, 815–827.

Ceccarelli, M., Barthel, F.P., Malta, T.M., Sabedot, T.S., Salama, S.R., Murray, B.A., Morozova, O., Newton, Y., Radenbaugh, A., Pagnotta, S.M., et al. (2016). Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. Cell *164*, 550–563.

Chan, K.M., Fang, D., Gan, H., Hashizume, R., Yu, C., Schroeder, M., Gupta, N., Mueller, S., James, C.D., Jenkins, R., et al. (2013). The histone H3.3K27M mutation in pediatric glioma reprograms H3K27 methylation and gene expression. Genes Dev. *27*, 985–990.

Davis, R.J., Welcker, M., and Clurman, B.E. (2014). Tumor suppression by the Fbw7 ubiquitin ligase: mechanisms and opportunities. Cancer Cell *26*, 455–464.

Fontebasso, A.M., Schwartzentruber, J., Khuong-Quang, D.A., Liu, X.Y., Sturm, D., Korshunov, A., Jones, D.T., Witt, H., Kool, M., Albrecht, S., et al. (2013). Mutations in SETD2 and genes affecting histone H3K36 methylation target hemispheric high-grade gliomas. Acta Neuropathol. *125*, 659–669.

Fontebasso, A.M., Papillon-Cavanagh, S., Schwartzentruber, J., Nikbakht, H., Gerges, N., Fiset, P.O., Bechet, D., Faury, D., De Jay, N., Ramkissoon, L.A., et al. (2014). Recurrent somatic mutations in ACVR1 in pediatric midline high-grade astrocytoma. Nat. Genet. *46*, 462–466.

Funato, K., Major, T., Lewis, P.W., Allis, C.D., and Tabar, V. (2014). Use of human embryonic stem cells to model pediatric gliomas with H3.3K27M histone mutation. Science *346*, 1529–1533.

Grasso, C.S., Tang, Y., Truffaux, N., Berlow, N.E., Liu, L., Debily, M.A., Quist, M.J., Davis, L.E., Huang, E.C., Woo, P.J., et al. (2015). Functionally defined therapeutic targets in diffuse intrinsic pontine glioma. Nat. Med. *21*, 555–559.

Hennika, T., Hu, G., Olaciregui, N.G., Barton, K.L., Ehteda, A., Chitranjan, A., Chang, C., Gifford, A.J., Tsoli, M., Ziegler, D.S., et al. (2017). Pre-clinical study of panobinostat in xenograft and genetically engineered murine diffuse intrinsic pontine glioma models. PLoS One *12*, e0169485.

International Cancer Genome Consortium PedBrain Tumor Project (2016). Recurrent MET fusion genes represent a drug target in pediatric glioblastoma. Nat. Med. *22*, 1314–1320.

Jones, C., and Baker, S.J. (2014). Unique genetic and epigenetic mechanisms driving paediatric diffuse high-grade glioma. Nat. Rev. Cancer *14*, 651–661.

Jones, C., Perryman, L., and Hargrave, D. (2012). Paediatric and adult malignant glioma: close relatives or distant cousins? Nat. Rev. Clin. Oncol. *9*, 400–413.

Jones, C., Karajannis, M.A., Jones, D.T., Kieran, M.W., Monje, M., Baker, S.J., Becher, O.J., Cho, Y.J., Gupta, N., Hawkins, C., et al. (2016). Pediatric high-grade glioma: biologically and clinically in need of new thinking. Neuro Oncol. *19*, 153.

Khuong-Quang, D.A., Buczkowicz, P., Rakopoulos, P., Liu, X.Y., Fontebasso, A.M., Bouffet, E., Bartels, U., Albrecht, S., Schwartzentruber, J., Letourneau, L., et al. (2012). K27M mutation in histone H3.3 defines clinically and biologically distinct subgroups of pediatric diffuse intrinsic pontine gliomas. Acta Neuropathol. *124*, 439–447.

Korshunov, A., Ryzhova, M., Hovestadt, V., Bender, S., Sturm, D., Capper, D., Meyer, J., Schrimpf, D., Kool, M., Northcott, P.A., et al. (2015). Integrated analysis of pediatric glioblastoma reveals a subset of biologically favorable tumors with associated molecular prognostic markers. Acta Neuropathol. *129*, 669–678.

Korshunov, A., Schrimpf, D., Ryzhova, M., Sturm, D., Chavez, L., Hovestadt, V., Sharma, T., Habel, A., Burford, A., Jones, C., et al. (2017). H3-/IDH-wild type pediatric glioblastoma is comprised of molecularly and prognostically distinct subtypes with associated oncogenic drivers. Acta Neuropathol. http://dx.doi.org/10.1007/s00401-017-1710-1.

Kramm, C.M., Butenhoff, S., Rausche, U., Warmuth-Metz, M., Kortmann, R.D., Pietsch, T., Gnekow, A., Jorch, N., Janssen, G., Berthold, F., et al. (2011). Thalamic high-grade gliomas in children: a distinct clinical subset? Neuro Oncol. *13*, 680–689.

Louis, D.N., Perry, A., Reifenberger, G., von Deimling, A., Figarella-Branger, D., Cavenee, W.K., Ohgaki, H., Wiestler, O.D., Kleihues, P., and Ellison, D.W. (2016). The 2016 World Health Organization classification of tumors of the central nervous system: a summary. Acta Neuropathol. *131*, 803–820.

Nicolaides, T.P., Li, H., Solomon, D.A., Hariono, S., Hashizume, R., Barkovich, K., Baker, S.J., Paugh, B.S., Jones, C., Forshew, T., et al. (2011). Targeted therapy for BRAFV600E malignant astrocytoma. Clin. Cancer Res. *17*, 7595–7604.

Northcott, P.A., Lee, C., Zichner, T., Stutz, A.M., Erkek, S., Kawauchi, D., Shih, D.J., Hovestadt, V., Zapatka, M., Sturm, D., et al. (2014). Enhancer hijacking activates GFI1 family oncogenes in medulloblastoma. Nature *511*, 428–434.

Ostrom, Q.T., Gittleman, H., Fulop, J., Liu, M., Blanda, R., Kromer, C., Wolinsky, Y., Kruchko, C., and Barnholtz-Sloan, J.S. (2015). CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2008–2012. Neuro Oncol. *17* (*Suppl 4*), iv1–iv62.

Paugh, B.S., Qu, C., Jones, C., Liu, Z., Adamowicz-Brice, M., Zhang, J., Bax, D.A., Coyle, B., Barrow, J., Hargrave, D., et al. (2010). Integrated molecular genetic profiling of pediatric high-grade gliomas reveals key differences with the adult disease. J. Clin. Oncol. *28*, 3061–3068.

Paugh, B.S., Broniscer, A., Qu, C., Miller, C.P., Zhang, J., Tatevossian, R.G., Olson, J.M., Geyer, J.R., Chi, S.N., da Silva, N.S., et al. (2011). Genome-wide analyses identify recurrent amplifications of receptor tyrosine kinases and cell-cycle regulatory genes in diffuse intrinsic pontine glioma. J. Clin. Oncol. *29*, 3999–4006.

Paugh, B.S., Zhu, X., Qu, C., Endersby, R., Diaz, A.K., Zhang, J., Bax, D.A., Carvalho, D., Reis, R.M., Onar-Thomas, A., et al. (2013). Novel oncogenic PDGFRA mutations in pediatric high-grade gliomas. Cancer Res. *73*, 6219–6229.

Puget, S., Philippe, C., Bax, D.A., Job, B., Varlet, P., Junier, M.P., Andreiuolo, F., Carvalho, D., Reis, R., Guerrini-Rousseau, L., et al. (2012). Mesenchymal transition and PDGFRA amplification/mutation are key distinct oncogenic events in pediatric diffuse intrinsic pontine gliomas. PLoS One *7*, e30313.

Qin, E.Y., Cooper, D.D., Abbott, K.L., Lennon, J., Nagaraja, S., Mackay, A., Jones, C., Vogel, H., Jackson, P.K., and Monje, M. (2017). Neural Precursor-Derived Pleiotrophin Mediates Subventricular Zone Invasion by Glioma. Cell *170* (5), 845–859.

Richards, M.W., Burgess, S.G., Poon, E., Carstensen, A., Eilers, M., Chesler, L., and Bayliss, R. (2016). Structural basis of N-Myc binding by Aurora-A and its destabilization by kinase inhibitors. Proc. Natl. Acad. Sci. USA *113*, 13726–13731.

Schiffman, J.D., Hodgson, J.G., VandenBerg, S.R., Flaherty, P., Polley, M.Y., Yu, M., Fisher, P.G., Rowitch, D.H., Ford, J.M., Berger, M.S., et al. (2010). Oncogenic BRAF mutation with CDKN2A inactivation is characteristic of a subset of pediatric malignant astrocytomas. Cancer Res. *70*, 512–519.

Schwartzentruber, J., Korshunov, A., Liu, X.Y., Jones, D.T., Pfaff, E., Jacob, K., Sturm, D., Fontebasso, A.M., Quang, D.A., Tonjes, M., et al. (2012). Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature *482*, 226–231.

Shlien, A., Campbell, B.B., de Borja, R., Alexandrov, L.B., Merico, D., Wedge, D., Van Loo, P., Tarpey, P.S., Coupland, P., Behjati, S., et al. (2015). Combined hereditary and somatic mutations of replication error repair genes result in rapid onset of ultra-hypermutated cancers. Nat. Genet. *47*, 257–262.

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

Sturm, D., Bender, S., Jones, D.T., Lichter, P., Grill, J., Becher, O., Hawkins, C., Majewski, J., Jones, C., Costello, J.F., et al. (2014). Paediatric and adult glioblastoma: multiform (epi)genomic culprits emerge. Nat. Rev. Cancer *14*, 92–107.

Taylor, K.R., Mackay, A., Truffaux, N., Butterfield, Y.S., Morozova, O., Philippe, C., Castel, D., Grasso, C.S., Vinci, M., Carvalho, D., et al. (2014a). Recurrent activating ACVR1 mutations in diffuse intrinsic pontine glioma. Nat. Genet. *46*, 457–461.

Taylor, K.R., Vinci, M., Bullock, A.N., and Jones, C. (2014b). ACVR1 mutations in DIPG: lessons learned from FOP. Cancer Res. *74*, 4565–4570.

Temime-Smaali, N., Guittat, L., Wenner, T., Bayart, E., Douarre, C., Gomez, D., Giraud-Panis, M.J., Londono-Vallejo, A., Gilson, E., Amor-Gueret, M., et al. (2008). Topoisomerase IIIalpha is required for normal proliferation and telomere stability in alternative lengthening of telomeres. EMBO J. *27*, 1513–1524.

Temime-Smaali, N., Guittat, L., Sidibe, A., Shin-ya, K., Trentesaux, C., and Riou, J.F. (2009). The G-quadruplex ligand telomestatin impairs binding of topoisomerase IIIalpha to G-quadruplex-forming oligonucleotides and uncaps telomeres in ALT cells. PLoS One *4*, e6919.

Tiberi, L., Bonnefont, J., van den Ameele, J., Le Bon, S.D., Herpoel, A., Bilheu, A., Baron, B.W., and Vanderhaeghen, P. (2014). A BCL6/BCOR/SIRT1 complex triggers neurogenesis and suppresses medulloblastoma by repressing Sonic Hedgehog signaling. Cancer Cell *26*, 797–812.

Venkatesh, H.S., Johung, T.B., Caretti, V., Noll, A., Tang, Y., Nagaraja, S., Gibson, E.M., Mount, C.W., Polepalli, J., Mitra, S.S., et al. (2015). Neuronal activity promotes glioma growth through Neuroligin-3 secretion. Cell *161*, 803–816.

Welcker, M., Orian, A., Jin, J., Grim, J.E., Harper, J.W., Eisenman, R.N., and Clurman, B.E. (2004). The Fbw7 tumor suppressor regulates glycogen synthase kinase 3 phosphorylation-dependent c-Myc protein degradation. Proc. Natl. Acad. Sci. USA *101*, 9085–9090.

Wu, L., Davies, S.L., North, P.S., Goulaouic, H., Riou, J.F., Turley, H., Gatter, K.C., and Hickson, I.D. (2000). The Bloom's syndrome gene product interacts with topoisomerase III. J. Biol. Chem. *275*, 9636–9644.

Wu, G., Broniscer, A., McEachron, T.A., Lu, C., Paugh, B.S., Becksfort, J., Qu, C., Ding, L., Huether, R., Parker, M., et al. (2012). Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. Nat. Genet. *44*, 251–253.

Wu, G., Diaz, A.K., Paugh, B.S., Rankin, S.L., Ju, B., Li, Y., Zhu, X., Qu, C., Chen, X., Zhang, J., et al. (2014). The genomic landscape of diffuse intrinsic pontine glioma and pediatric non-brainstem high-grade glioma. Nat. Genet. *46*, 444–450.

Yada, M., Hatakeyama, S., Kamura, T., Nishiyama, M., Tsunematsu, R., Imaki, H., Ishida, N., Okumura, F., Nakayama, K., and Nakayama, K.I. (2004). Phosphorylation-dependent degradation of c-Myc is mediated by the F-box protein Fbw7. EMBO J. *23*, 2116–2125.

Yang, J., Bachrati, C.Z., Ou, J., Hickson, I.D., and Brown, G.W. (2010). Human topoisomerase IIIalpha is a single-stranded DNA decatenase that is stimulated by BLM and RMI1. J. Biol. Chem. *285*, 21426–21436.

Zarghooni, M., Bartels, U., Lee, E., Buczkowicz, P., Morrison, A., Huang, A., Bouffet, E., and Hawkins, C. (2010). Whole-genome profiling of pediatric diffuse intrinsic pontine gliomas highlights platelet-derived growth factor receptor alpha and poly (ADP-ribose) polymerase as potential therapeutic targets. J. Clin. Oncol. *28*, 1337–1344.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical Commercial Assays		
DNeasy blood & tissue kit	Qiagen	69504
QIAmp DNA FFPE tissue kit	Qiagen	56404
RNeasy mini kit	Qiagen	74104
QIAquick PCR purification kit	Qiagen	28104
BigDye terminator v3.1 mix	Thermo Fisher	4337455
SureSelect Human All Exon capture set V4 SureSelect Human All Exon capture set V5	Agilent	5190-4666 5190-6208
SureSelect RNA Capture, 0.5-2.9Mb	Agilent	5190-4944
Deposited Data		
Exome and RNA sequencing of new samples	This paper	EGA: EGAS00001002314
Illumina methylation BeadChip profiling of new samples	This paper	ArrayExpress: E-MTAB-5528
Sequencing and methylation data	This paper	cavatica.org
Oligonucleotides		
Primer: H3F3A_forward TGGCTCGTACAAAGCAGACT	This paper	N/A
Primer: H3F3A_reverse ATATGGATACATACAAGAGAGACT	This paper	N/A
Primer: HIST1H3B_forward GGGCAGGAGCCTCTCTTAAT	This paper	N/A
Primer: HIST1H3B _ reverse ACCAAGTAGGCCTCACAAGC	This paper	N/A
Software and Algorithms		
Mutation Surveyor	SoftGenetics	softgenetics.com/mutationSurveyor.php
4Peaks	Nucleobytes	http://nucleobytes.com/4peaks/
limma	BioConductor	bioconductor.org/packages/release/bioc/html/limma.html
marray	BioConductor	bioconductor.org/packages/release/bioc/html/marray.html
aroma.affymetrix	The Comprehensive R Archive Network	cran.rstudio.com/web/packages/aroma.affymetrix/index.html
aroma.cn	The Comprehensive R Archive Network	cran.r-project.org/web/packages/aroma.cn/index.html
minfi	BioConductor	bioconductor.org/packages/release/bioc/html/minfi.html
conumee	BioConductor	bioconductor.org/packages/release/bioc/html/conumee.html
BEDtools	University of Utah	github.com/arq5x/bedtools2
DNAcopy	BioConductor	bioconductor.org/packages/release/bioc/html/DNAcopy.html
gviz	BioConductor	bioconductor.org/packages/release/bioc/html/Gviz.html
GISTIC	Broad Institute	oftware.broadinstitute.org/software/cprg/?q=node/31
CopyNumber450kData	BioConductor	bioconductor.org/packages/release/data/experiment/html/CopyNumber450kData.html
MNP	DKFZ Heidelberg	molecularneuropathology.org/mnp

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
tSNE	The Comprehensive R Archive Network	cran.r-project.org/web/packages/Rtsne/index.html
rgl	The Comprehensive R Archive Network	cran.r-project.org/web/packages/rgl/index.html
affy	BioConductor	bioconductor.org/packages/release/bioc/html/affy.html
Bowtie2	Johns Hopkins University	bowtie-bio.sourceforge.net/bowtie2/index.shtml
TopHat	Johns Hopkins University	ccb.jhu.edu/software/tophat/index.shtml
cufflinks	University of Washington	ole-trapnell-lab.github.io/cufflinks/cufflinks/
DESeq2	BioConductor	bioconductor.org/packages/release/bioc/html/DESeq2.html
Gene Set Enrichment Analysis	Broad Institute	http://software.broadinstitute.org/gsea
bwa	Sanger Institute	http://bio-bwa.sourceforge.net/
Genome Analysis Toolkit	Broad Institute	oftware.broadinstitute.org/gatk/
Variant Effect predictor	Ensembl tools	ensembl.org/info/docs/variation/vep
ANNOVAR	Children's Hospital of Philadelphia	annovar.openbioinformatics.org/en/latest/
ExAc	Broad Institute	exac.broadinstitute.org/
BCBio	Harvard TH Chan	bcb.io/
SIFT	J Craig Venter Institute	sift.jcvi.org
PolyPhen	Harvard	genetics.bwh.harvard.edu/pph2
ChimeraScan	University of Michigan	omictools.com/chimerascan-tool
Breakdancer	Washington University of St Louis	breakdancer.sourceforge.net
ASCAT	Francis Crick Institute	rick.ac.uk/peter-van-loo/software/ASCAT
Oncoprinter	Memorial Sloan Kettering	cbioportal.org/oncoprinter.jsp
ProteinPaint	St Jude	pecan.stjude.org/#/proteinpaint
Circos	Michael Smith Genome Sciences Center	circos.ca
MSigDB	Broad Institute	http://software.broadinstitute.org/gsea/msigdb
CytoScape	National Institute of General Medical Sciences	cytoscape.org
R	The Comprehensive R Archive Network	r-project.org
Other		
Processed DNA copy number profiles	This paper and cited sources	dipg.progenetix.org arraymap.org
Integrated mutation, copy number, expression and methylation data	This paper and cited sources	pedcbioportal.org

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Chris Jones (chris.jones@icr.ac.uk).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Patient Samples

All new patient material was collected after informed consent and subject to local research ethics committee approval. We collated and profiled 157 unpublished cases of HGG in children and young adults up to the age of 30 years at diagnosis obtained from the Royal Marsden, St Georges and Kings College Hospitals, (n=39, all London, UK), Chinese University of Hong Kong (n=24,

Hong Kong, China), Qilu University Hospital (n=23, Jinan, China), Farhad Hatched Hospital (n=14, Sousse, Tunisia), Federal University of São Paolo (n=14, São Paulo, Brazil), Morozov Children's and Dmitri Rogachev Hospitals (n=12, Moscow, Russia), Queensland Children's Tumor Bank (n=8, Brisbane, Australia), Hospital San Joan de Déu (n=8, Barcelona, Spain), City Hospital #31 (n=6, St Petersburg, Russia), Barretos Cancer Hospital (n=4, Barretos, Brazil), Centre Hospitalier Régional et Universitaire Hautepierre (n=3, Strasbourg, France), and Our Lady Children's Hospital Crumlin (n=2, Dublin, Ireland). A full description of the samples included are provided in Table S1.

METHOD DETAILS

Nucleic Acid Extraction

DNA was extracted from frozen tissue by homogenisation prior to following the DNeasy Blood & Tissue kit protocol (Qiagen, Crawley, UK). DNA was extracted from formalin-fixed, paraffin-embedded (FFPE) pathology blocks after manual macrodissection using the QIAamp DNA FFPE tissue kit protocol (Qiagen). Matched normal DNA was extracted from blood samples using the DNeasy Blood & Tissue kit (Qiagen, Crawley, UK). Concentrations were measured using a Qubit fluorometer (Life Technologies, Paisley, UK). RNA was extracted by following the RNeasy Mini Kit protocol (Qiagen), and quantified on a 2100 Bioanalyzer (Agilent Technologies).

Sanger Sequencing of *H3F3A / HIST1H3B*

PCR for *H3F3A* and *HIST1H3B* was carried out using primers obtained from Life Technologies (Paisley, UK). Products were purified using the QIAquick PCR purification kit (Qiagen), subjected to bidirectional sequencing using BigDye Terminator mix 3.1 (Applied Biosystems, Foster City, CA, USA), with capillary sequencing was done on an ABI 3100 genetic analyzer (Applied Biosystems, Foster City, CA, USA). Sequences were analysed using Mutation Surveyor (SoftGenetics, PN, USA) and manually with 4Peaks (Nucleobytes, Aalsmeer, Netherlands).

Methylation Profiling

50-500 ng DNA was bisulphite-modified and analyzed for genome-wide methylation patterns using the Illumina HumanMethylation450 BeadArray (450k) platform at either the DKFZ or the University College London Genomics Centre, according the manufacturer's instructions. All samples were checked for expected and unexpected genotype matches by pairwise correlation of the 65 genotyping probes on the 450k array.

Exome and RNA Sequencing

50-500 ng DNA was sequenced at the Tumor Profiling Unit, ICR, London, UK using the SureSelect Human All Exon capture sets V4 or V5 (Agilent, Santa Clara, CA, USA), and paired-end-sequenced on an Illumina HiSeq2000 (Illumina, San Diego, CA, USA) with a 100 bp read length. Coverage ranged from 29-295x (median=105x). RNA was sequenced at the ICR Tumor Profiling Unit after SureSelect RNA capture on an Illumina HiSeq2500 with a 125 bp read length.

QUANTIFICATION AND STATISTICAL ANALYSIS

Published Data Sources

These data were combined with those obtained directly from the authors or from public data repositories representing 20 published studies (Barrow et al., 2011; Bax et al., 2010; Buczkowicz et al., 2014; Carvalho et al., 2014; Castel et al., 2015; Fontebasso et al., 2013, 2014; Grasso et al., 2015; International Cancer Genome Consortium PedBrain Tumor Project, 2016; Khuong-Quang et al., 2012; Korshunov et al., 2015; Paugh et al., 2010, 2011; Puget et al., 2012; Schwartzentruber et al., 2012; Sturm et al., 2012; Taylor et al., 2014a; Wu et al., 2012; Wu et al., 2014; Zarghooni et al., 2010) with the following accession numbers: EGA - EGAS00001000226, EGAS0000100192, EGAS00001000575, EGAS00001000720, EGAS00001001139; the Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo/) - GSE19578, GSE26576, GSE21420, GSE34824, GSE36245, GSE36278 GSE50022, GSE50021, GSE50024, GSE55712; ArrayExpress - E-TABM-857 , E-TABM-1107 . The full cohort included a total of 1254 molecular profiles from 955 samples across 12 platforms, which after quality control and manual annotation to remove duplicates, and supplemented with targeted sequencing of an additional 158 cases, resulted in a total dataset comprised of 1067 individual patients. The full dataset comprises genomic profiles from DNA copy number arrays (Agilent 44K, n=127; Affymetrix 500K, n=100; Affymetrix SNP6.0, n=78; 32k BAC, n=61), Illumina 450k methylation arrays (n=441), whole exome (n=254), genome (n=125), targeted (n=212) and RNA sequencing (n=82), as well as gene expression from Affymetrix U133Plus2 (n=102) and Agilent WG2.5 (n=67) platforms.

DNA Copy Number

DNA copy number data was obtained as array CGH (Agilent 44k and 32K BAC), SNP arrays (Affymetrix 500k and SNP6.0), 450k methylation arrays (Illumina) and/or sequencing data (whole genome and exome). Two color aCGH data was read and normalized using the R packages limma and marray. Log intensity data from Affymetrix SNP arrays was derived using the aroma.affymetrix and aroma.cn package. Combined $\log_2$ intensity data from Illumina 450K methylation arrays was processed using the R packages minfi and conumee. For sequenced samples, coverage of aligned reads was binned into known genes and exons with BEDTools and

log_2 ratios of median coverage in tumor and normal sequences were processed with in-house scripts. To combine copy number platforms, median log_2 ratios were recovered within all known genes and exons and normalized such that the median displacement of X in male:female comparisons was rescaled to an average of -1. Exon-level median log ratios and smoothed values were then combined across platforms and thresholded to call gains and losses above and below log_2 ratios of ±0.3 with a contig of ~1MB and amplifications and deletions above and below a threshold of ±1.5 with a minimum of 3 contiguous exons.

CBS binary segmentation from the DNAcopy package was applied to each dataset to provide smoothed log_2 ratios. Genes within common CNVs in normal individuals were excluded from further analysis with reference to the CNV map of the human genome. DNA copy number data was clustered based upon categorical states (deep deletion, loss, no change, gain and amplification) based upon the Euclidean distance method with a Ward algorithm. Gains and losses in chromosomal arms were called based upon contiguous regions covering more than one third of the exonic regions within each arm. For regions of focal copy number change cases carrying copy number alterations were ranked according to the length of the largest CNA in each case and are plotted as heatmaps aligned to precise genomic coordinates alongside genomic tracks based upon hg19 made with the R package gviz. Minimal regions of copy number alteration were assigned based on the frequency of categorical states within each region. Focal amplifications and deletions were identified in CBS segmented data using the GISTIC algorithm in MATLAB on the exon-level data, with thresholds for gain and loss of 0.3 and gene-level filters to remove regions of common copy number variation in normal individuals based on the CNV map of the human genome.

DNA Methylation

Methylation data from the Illumina Infinium HumanMethylation450 BeadChip was preprocessed using the minfi package in R. DNA copy number was recovered from combined intensities using the conumee package with reference to methylation profiles from normal individuals provided in the CopyNumber450kData package. We have used the Heidelberg brain tumor classifier (molecularneuropathology.org) to assign subtype scores for each tumor compared to 91 different brain tumor entities using a training set built from more than 2000 tumors implemented in the MNP R package. Simplified methylation subgroup assignments were then made to incorporate cases carrying G34R/V or K27M mutations in H3 histones, IDH1 mutation at R132, low grade glioma-like profiles (predominantly diffuse infantile ganglioglioma and pilocytic astrocytoma) and those similar to pleomorphic xanthoastrocytoma (PXA). Wild-type HGG encompassed many other methylation subgroups and were simply assigned by exclusion with the groups above. Clustering of beta values from methylation arrays was performed using the 10K probeset from the Heidelberg classifier based upon Euclidean distance with a ward algorithm. Methylation heatmaps show only the most variable probes of the classifier between simplified methylation subgroups. Overall methylation was calculated as the mean of the 10K classifier probeset for each subgroup and MGMT promoter methylation was calculated based upon the MGMT-SPT27 model implemented in the MNP package. t- stochastic neighbor embedding (tSNE) was used to project the methylation clustering in three dimensions using the Rtsne package. A Pearson correlation matrix of the 10K probeset was subjected to tSNE using a theta value of zero over 10,000 iterations as previously described and plotted using the rgl package.

mRNA Expression

Gene expression data was obtained from Agilent WG2.5, Affymetrix U133Plus2.0 or RNA sequencing platforms. Gene expression was processed from two color Agilent microarrays using the R packages marray and limma and from single channel Affymetrix arrays using the affy package. Differential expression was assigned for microarray data using the limma package based upon a false discovery rate of 5%. RNASeq was aligned with Bowtie2 and TopHat and summarized as gene level fragments per kilobase per million reads sequenced using BEDTools and cufflinks/cuffnorm. Following rlog transformation and normalization, differential expression was assigned with DESeq.2. Known Ensembl genes were further filtered to remove low abundance genes in all three datasets whose maximal expression was within the lowest 20% of all expression values based upon probe intensities or read depth. Replicate probes/features for each gene were removed by selecting those with the greatest median absolute deviation (MAD) in each dataset. Following centering within each dataset, log-transformed expression measures were combined and further normalized using pairwise loess normalization. Gene Set Enrichment Analysis was performed using the GSEA java application based upon pairwise comparisons of the major subgroups in the merged dataset. Heatmaps of gene expression across chromosomal arms were made using centered expression values rescaled across each chromosomal arm based upon the median absolute deviation of each probe. Differential expression analysis of *TOP3A* and *FBXW7* was based on a Mann-Whitney U test of centred expression values between cases with and without losses and amplifications respectively in each case.

Sequence Analysis

Sequencing data was available as whole genome and/or whole exome (predominantly using Agilent's SureSelect whole exome capture sets v4 and v5) Short read sequences from whole exome or whole genome sequencing were aligned to the hg19 assembly of the human genome using bwa. Following duplicate removal with Picard tools variants were called using the Genome Analysis toolkit according to standard Best Practices (Broad) including local re-alignment around Indels, downsampling and variant calling with the Unified Genotyper. Variants were annotated with the variant Effect predictor v74 from Ensembl tools and ANNOVAR to include annotations for variant allele frequency in 1000 genomes dbSNP v132 and the ExAc database as well as functional annotation tools SIFT and Polyphen). Depth of coverage varied from 16-295x (median 88x), with the greatest variation unsurprisingly in the exome data (whole genome range 50-150x, median=85x). Somatic variants were identified in regions covered by at least 10 reads in normal and

tumor sequences carrying at least 3 variant reads in the tumor and less than 2 in normal sequences. Hotspot *TERT* promoter mutations C228T and C250T were incidentally captured by the various exome platforms as they are located only 114 and 146 bp upstream of the translation start site, and were called even if only covered by a few reads. Mutation signatures were ascertained by grouping somatic substitutions on the basis of their 3′ and 5′ bases into 96 possible trinucleotide categories.

Candidate Fusion Gene Nomination

Structural variants were called from whole genome data using Breakdancer (breakdancer.sourceforge.net) filtered to remove commonly multi-mapped regions to identify somatic breakpoints separated by a minimum of 10 kbp involving at least one Ensembl gene. Fusion transcripts were detected from RNAseq data using chimerascan version 0.4.5a filtered to remove common false positives. To minimize unverified false positives, reporting of nominated fusions was restricted to genes within the core functional pathways and processes identified through integrated DNA copy number and somatic variant calling.

Inferred Tumor Purity

We used determined the somatic allele-specific copy number profiles using read depth from whole genome / exome sequencing, and used ASCAT (rick.ac.uk/peter-van-loo/software/ASCAT) to provide for an estimate of the non-neoplastic cell contamination of the sample as well as the overall ploidy of the tumor. Values ranged from 36-100%, with a median of 83%.

Integrated Analysis of Driver Events

Somatic non-synonymous coding mutations were filtered to remove common passenger mutations, polymorphisms and false positives in exome sequencing. Data were integrated with focal DNA copy number calls by GISTIC to provide gene-level binary alteration calls which were further selected for putative drive status on the basis of functional annotation. Oncoprint representations of integrated mutations, gene-level copy number alterations and fusion events were made using the online tool available at cBioportal (cbioportal.org). For the most commonly mutated genes mutations were mapped to the canonical transcript and plotted according to their predicted protein position using the Protein Painter (pecan.stjude.org). Integrated views of copy number alterations, structural variants and somatic mutations were made using CIRCOS (circos.ca) and rearrangements within *TOP3*A amplified regions in whole genome sequenced cases were identified using Breakdancer and aligned with copy number breakpoints in R.

Pathway Analysis

Pathway assignments were made for all genes carrying copy number alterations, structural variations or somatic mutations based on pathways in the MSigDB molecular signatures databases (Broad) as well as Gene Ontologies for Biological Processes and Molecular Functions (Gene Ontology consortium) and canonical pathways from KEGG, NetPath and Reactome. Genes within known CNVs and common false positives in exomic sequencing were excluded with reference to large scale genome profiling studies (CNVmap, ExAc, BCBio) Pathway analysis of genes carrying mutations, gene fusions and copy number aberrations was based on the pathways defined by these combined databases and subjected to enrichment analysis using the EnrichmentMap module within CytoScape.

Statistical Analysis

Statistical analysis was carried out using R 3.3.1 (www.r-project.org). Categorical comparisons of counts were carried out using Fishers exact test, comparisons between groups of continuous variables employed Student's t-test, Wilcoxon signed –rank test, ANOVA or Mann-Whitney U test. Differences in survival were analysed by the Kaplan-Meier method and significance determined by the log-rank test. All tests were two-sided and a p value of less than 0.05 was considered significant. Multiple testing was accounted for using false discovery rate q values or the Bonferroni adjustment.

DATA AND SOFTWARE AVAILABILITY

All newly generated data have been deposited in the European Genome-phenome Archive (www.ebi.ac.uk/ega) with accession number EGAS00001002314 (sequencing) or ArrayExpress (www.ebi.ac.uk/arrayexpress/) with accession number E-MTAB-5528 (450k methylation).

ADDITIONAL RESOURCES

Processed copy number profiles are hosted as a disease-specific project within the Progenetix framework for annotated genomic analyses (dipg.progenetix.org) (Cai et al., 2014), and represented in the arrayMap resource (arraymap.org) (Cai et al., 2012). Curated gene-level copy number and mutation data are provided as part of the pediatric-specific implementation of the cBioPortal genomic data visualisation portal (pedcbioportal.org). Newly-generated raw data files are housed alongside published datasets made available to the Cavatica NIH-integrated cloud platform (www.cavatica.org).

AACR American Association for Cancer Research

CANCER DISCOVERY

Infant high grade gliomas comprise multiple subgroups characterized by novel targetable gene fusions and favorable outcomes

Matthew Clarke, Alan Mackay, Britta Ismer, et al.

Cancer Discov Published OnlineFirst April 1, 2020.

Infant high grade gliomas comprise multiple subgroups characterized by novel targetable gene fusions and favorable outcomes

Matthew Clarke[1], Alan Mackay[1*], Britta Ismer[2,3,4*], Jessica C Pickles[5], Ruth G Tatevossian[6], Scott Newman[7], Tejus A Bale[8], Iris Stoler[9], Elisa Izquierdo[1], Sara Temelso[1], Diana M Carvalho[1], Valeria Molinari[1], Anna Burford[1], Louise Howell[1], Alex Virasami[5], Amy R Fairchild[5], Aimee Avery[5], Jane Chalker[5], Mark Kristiansen[5], Kelly Haupfear[6], James D Dalton[6], Wilda Orisme[6], Ji Wen[6], Michael Hubank[10], Kathreena M Kurian[11], Catherine Rowe[11], Mellissa Maybury[12,13,14], Stephen Crosier[15], Jeffrey Knipstein[16], Ulrich Schüller[17,18], Uwe Kordes[18], David E Kram[19], Matija Snuderl[20], Leslie Bridges[21], Andrew J Martin[22], Lawrence J Doey[23], Safa Al-Sarraj[23], Christopher Chandler[24], Bassel Zebian[24], Claire Cairns[24], Rachael Natrajan[25], Jessica KR Boult[26], Simon P Robinson[26], Martin Sill[2], Ira J Dunkel[27], Stephen W Gilheeney[27], Marc K Rosenblum[8], Debbie Hughes[10], Paula Z Proszek[10], Tobey J Macdonald[28], Matthias Preusser[29], Christine Haberler[29,30], Irene Slavc[31], Roger Packer[32], Ho-keung Ng[33], Shani Caspi[34], Mara Popović[35], Barbara Faganel Kotnik[36], Matthew D Wood[37], Lissa Baird[38], Monika Ashok Davare[39], David A Solomon[40,41], Thale Kristin Olsen[42], Petter Brandal[43], Michael Farrell[44], Jane B Cryan[44], Michael Capra[45], Michael Karremann[46], Jens Schittenhelm[47], Martin U Schuhmann[48], Martin Ebinger[49], Winand NM Dinjens[50], Kornelius Kerl[51], Simone Hettmer[52], Torsten Pietsch[53], Felipe Andreiuolo[53], Pablo Hernáiz Driever[54], Andrey Korshunov[55], Lotte Hiddingh[2], Barbara C Worst[2,4,56], Dominik Sturm[2,4,56], Marc Zuckermann[2,4], Olaf Witt[2,4,56], Tabitha Bloom[57], Clare Mitchell[57], Evelina Miele[58], Giovanna Stefania Colafati[59], Francesca Diomedi-Camassei[60], Simon Bailey[15], Andrew S Moore[12,13,14], Timothy EG Hassall[13,61,62], Stephen P Lowis[11], Maria Tsoli[63,64], Mark J Cowley[63.64], David S Ziegler[63,64], Matthias A Karajannis[27], Kristian Aquilina[65], Darren R Hargrave[66], Fernando Carceller[67,68], Lynley V Marshall[67,68], Andreas von Deimling[55,69], Christof M Kramm[70], Stefan M Pfister[2,4,56], Felix Sahm[54,69], Suzanne J Baker[71], Angela Mastronuzzi[72], Andrea Carai[73], Maria Vinci[58], David Capper[9,74], Sergey*

Popov[1,75], David W Ellison[6#], Thomas S Jacques[5#], David TW Jones,[2,4#], and Chris Jones[1#]

[1]Division of Molecular Pathology, Institute of Cancer Research, London, UK; [2]German Cancer Research Center (DKFZ); [3]Faculty of Biosciences, Heidelberg University, Heidelberg, Germany; [4]Hopp Children's Cancer Center Heidelberg (KiTZ), Heidelberg, Germany; [5]UCL Great Ormond Street Institute of Child Health, London, UK; [6]Department of Neuropathology, St Jude Children's Research Hospital, Memphis TN, USA; [7]Department of Computational Biology, St Jude Children's Research Hospital, Memphis TN, USA; [8]Department of Neuropathology, Memorial Sloan-Kettering Cancer Center, New York NY, USA; [9]Charité Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Department of Neuropathology, Berlin, Germany; [10]Molecular Diagnostics, Royal Marsden Hospital NHS Trust, Sutton, UK; [11]Brain Tumour Research Centre, University of Bristol, Bristol, UK; [12]The University of Queensland Diamantina Institute, The University of Queensland, Woolloongabba, Australia; [13]Oncology Service, Queensland Children's Hospital, Brisbane, Australia; [14]Child Health Research Centre, The University of Queensland, South Brisbane, Australia; [15]Newcastle Hospitals NHS Foundation Trust, Newcastle, UK; [16]Division of Pediatric Hematology/Oncology/BMT, Medical College of Wisconsin, Milwaukee WI, USA; [17]Department of Neuropathology, University Hospital Hamburg-Eppendorf, and Research Institute Children's Cancer Center, Hamburg, Germany; [18]Pediatric Hematology and Oncology, University Hospital Hamburg-Eppendorf, Hamburg, Germany; [19]Section of Pediatric Hematology-Oncology, Wake Forest School of Medicine, Winston-Salem NC, USA; [20]Department of Neuropathology, NYU Langone Health, New York NY, USA; [21]Department of Neuropathology, St George's Hospital NHS Trust, London, UK; [22]Department of Neurosurgery, St George's Hospital NHS Trust, London, UK; [23]Department of Clinical Neuropathology, Kings College Hospital NHS Trust, London, UK; [24]Department of Neurosurgery, Kings College Hospital NHS Trust, London, UK; [25]The Breast Cancer Now Toby Robins Research Centre, The Institute of Cancer Research, London, UK; [26]Division of Radiotherapy and Imaging, The Institute of Cancer Research, London, UK; [27]Department of Pediatrics, Memorial Sloan-Kettering Cancer Center, New York NY, USA; [28]Aflac Cancer and Blood Disorders Center, Department of Pediatrics, Emory University School of Medicine, Atlanta GA, USA; [29]Comprehensive Cancer Center, Medical University of Vienna, Vienna, Austria; [30]Institute of Neurology, Medical University of Vienna, Vienna, Austria; [31]Department of Pediatrics and Adolescent Medicine, Medical University of Vienna, Vienna, Austria; [32]Center for Neuroscience and Behavioural Medicine, Children's National Medical Center, Washington DC, USA; [33]Department of Anatomical and Cellular Pathology, The Chinese University of Hong Kong,

China; [34]Cancer Research Center, Sheba Medical Center, Tel Aviv, Israel; [35]Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia; [36]Department of Hematology and Oncology, University Children's Hospital, Ljubljana, Slovenia; [37]Department of Pathology, Oregon Health & Science University, Portland OR, USA; [38]Department of Neurosurgery, Oregon Health & Science University, Portland OR, USA; [39]Department of Pediatrics, Oregon Health & Science University, Portland OR, USA; [40]Department of Pathology, University of California, San Francisco, CA, United States; [41]Clinical Cancer Genomics Laboratory, University of California, San Francisco, CA, United States; [42]Department of Women's and Children's Health, Karolinska Institute, Stockholm, Sweden; [43]Department of Oncology, Oslo University Hospital, Oslo, Norway; [44]Department of Histopathology, Beaumont Hospital, Dublin, Ireland; [45]Paediatric Oncology, Our Lady's Children's Hospital, Dublin, Ireland; [46]Department of Pediatrics, University Medical Center Mannheim, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany; [47]Institute of Pathology and Neuropathology, University Hospital Tübingen, Germany; [48]Department of Neurosurgery, University Hospital Tübingen, Germany; [49]Department of Pediatric Hematology and Oncology, University Hospital Tübingen, Germany; [50]Department of Pathology, Erasmus Medical Center, Rotterdam, Netherlands; [51]Department of Pediatric Hematology and Oncology, University Hospital Muenster, Germany; [52]Department of Pediatric Hematology and Oncology, University Hospital Freiburg, Germany; [53]Institute of Neuropathology, DGNN Brain Tumor Reference Center, University of Bonn Medical Center, Bonn, Germany; [54]Department of Paediatric Haematology/Oncology Charité Universitätsmedizin, Berlin, Germany; [55]Department of Neuropathology, University Hospital Heidelberg, Germany; [56]Department of Pediatric Oncology, Hematology, Immunology and Pulmonology, Heidelberg University Hospital, 69120 Heidelberg, Germany; [57]BRAIN UK, University of Southampton, Southampton, UK; [58]Department of Onco-haematology, Cell and Gene Therapy, Bambino Gesù Children's Hospital-IRCCS, Rome, Italy; [59]Oncological Neuroradiology Unit, Department of Diagnostic Imaging, Bambino Gesù Children's Hospital-IRCCS, Rome, Italy; [60]Department of Laboratories - Pathology Unit, Bambino Gesù Children's Hospital-IRCCS, Rome, Italy; [61]Institute for Molecular Bioscience, The University of Queensland, St Lucia, Australia; [62]School of Biomedical Sciences, Faculty of Health, Queensland University of Technology, Brisbane, Australia; [63]Children's Cancer Institute, University of New South Wales, Sydney, Australia; [64]Kids Cancer Centre, Sydney Children's Hospital, Randwick, Australia; [65]Department of Neurosurgery, Great Ormond Street Hospital NHS Foundation Trust, London, UK; [66]Department of Paediatric Oncology, Great Ormond Street Hospital NHS Foundation Trust, London, UK; [67]Division of Clinical Studies, The Institute of Cancer Research, London, UK; [68]Children & Young People's Unit, Royal Marsden Hospital NHS Trust, Sutton, UK; [69]Clinical Cooperation Unit Neuropathology, German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany; [70]Division of Pediatric Hematology and

Oncology, University Medical Centre Göttingen, Germany; [71]Department of Neurobiology, St Jude Children's Research Hospital, Memphis TN, USA; [72]Neuro-oncology Unit, Department of Onco-haematology, Cell and Gene Therapy, Bambino Gesù Children's Hospital-IRCCS, Rome, Italy; [73]Oncological Neurosurgery Unit, Department of Neuroscience and Neurorehabilitation, Bambino Gesù Children's Hospital-IRCCS, Rome, Italy; [74]German Cancer Consortium (DKTK), Partner Site Berlin, German Cancer Research Center (DKFZ), Heidelberg, Germany; [75]Department of Pathology, University of Wales Hospital NHS Trust, Cardiff, UK

* *These authors contributed equally*

[#]*Correspondence to:*

Chris Jones,

Division of Molecular Pathology, The Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey, SM2 5NG, UK, chris.jones@icr.ac.uk, +44 (0)20 8722 4416

David TW Jones,

Hopp Children's Cancer Center Heidelberg (KiTZ) and German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany, david.jones@kitz-heidelberg.de, +49 6221 42-4675

Thomas S Jacques,

UCL GOS Institute of Child Health, 30 Guilford Street, London WC1N 1EH, UK, t.jacques@ucl.ac.uk, +44 (0)20 7762 6108

David W Ellison

Department of Pathology, St Jude Children's Research Hospital, 262 Danny Thomas Place, Memphis, TN 38105-3100, USA, david.ellison@stjude.org, +1 901 595-3533

Conflict of interest statement

The authors declare no conflict of interest

Keywords: infant, glioma, methylation, sequencing, NTRK, ALK, ROS1

Running title: Classification of infant glioma

Abstract

Infant high grade gliomas appear clinically distinct from their counterparts in older children, indicating that histopathologic grading may not accurately reflect the biology of these tumors. We have collected 241 cases under 4 years of age, and carried out histological review, methylation profiling, custom panel and genome/exome sequencing. After excluding tumors representing other established entities or subgroups, we identified 130 cases to be part of an 'intrinsic' spectrum of disease specific to the infant population. These included those with targetable MAP-kinase alterations, and a large proportion of remaining cases harboring gene fusions targeting *ALK* (n=31), *NTRK1/2/3* (n=21), *ROS1* (n=9) and *MET* (n=4) as their driving alterations, with evidence of efficacy of targeted agents in the clinic. These data strongly supports the concept that infant gliomas require a change in diagnostic practice and management.

Statement of Significance

Infant high grade gliomas in the cerebral hemispheres comprise novel subgroups, with a prevalence of *ALK*, *NTRK1-3*, *ROS1* and *MET* gene fusions. Kinase fusion-positive tumors have better outcome and respond to targeted therapy clinically. Other subgroups have poor outcome, with fusion-negative cases possibly representing an epigenetically-driven pluripotent stem cell phenotype.

Introduction

The prognosis of paediatric high grade gliomas (HGG) remains dismal, with a 5-year survival rate of only ~20% for children aged 0-14 years (1). They are strongly associated with unique location-dependent mutations in histone H3 variants H3.3 (*H3F3A*) and H3.1 (*HIST1H3B/C*) including two recurrent amino acid substitutions (K27M and G34R/V) (2,3) which together account for nearly half of all paediatric HGG and identify robust biological subgroups (4,5). Histone wild-type cases are comprised of a highly diverse set of tumors, ranging from those with some of the highest somatic mutational burdens in human cancer (patients with biallelic mismatch repair deficiency syndrome) (6,7) to others seemingly driven by single genetic events, often gene fusions (8). The latter are particularly found in cases originally diagnosed as high grade glioma at an infant age (9).

The definition of an infant used in paediatric neuro-oncology varies, but typically refers to children under 3-5 years (10); congenital cases are generally defined as being present at birth (11). The most frequent types of infant brain tumor are medulloblastoma, ependymomas and low grade gliomas (LGG) (12). The latter include the relatively common pilocytic astrocytomas, but also other rarer entities such as desmoplastic infantile ganglioglioma/astrocytoma (DIGG/DIA) (13). Tumors reported as HGG appear to be associated with significant differences in clinical outcome, with infant HGG (even with incomplete resection and without irradiation) showing a significantly improved survival compared to those in older children (8,14-17), which may indicate the presence of a distinct, overlapping group of tumors where histopathologic grading may not be representative of clinical behavior.

Treatment outcomes also reflect these differences; the Baby POG I study found four children under 3 years of age who were diagnosed with a malignant glioma and underwent 24 months of chemotherapy without radiation treatment, and did not

develop recurrent disease (10). A 5-year overall survival rate of 59% was reported in infants with HGG after prolonged chemotherapy treatment alone, and in another study, 16 patients diagnosed with HGG and treated with focal radiation therapy showed a 5-year overall survival rate of 66% (11). Five reported cases of congenital glioblastomas who survived surgery (with only one patient receiving a gross total resection) all showed a better outcome than expected (18), whilst two infant cases who both underwent subtotal resection of their tumors and did not receive any adjuvant therapy post-operatively saw regression of the residual tumors (19). The improved outcome both with chemotherapy and with surgery alone is particularly significant in this age group when considering the risk of declining cognition (13) and the development of leukoencephalopathy post radiation treatment (11).

Previous studies have hinted at different histological features within infant high grade gliomas. High densities of ‘minigemistocytic shaped’ cells with abundant mitoses and absent necrosis were described (20), with others showing moderately hypercellular, mitotic and necrotic tumors with cellular monotony and a lack of significant pleomorphism, and some showing a more spindled appearance (18,19).

Current molecular data is limited, but epidermal growth factor receptor (*EGFR*) and platelet-derived growth factor receptor A (*PDGFRA*) expression is reported as uniformly low in congenital glioblastomas (GBM), with a low level or absence of copy number alterations in these genes (18,21). *TP53* and *PTEN* mutations, *CDKN2A/B* deletions, and other copy number alterations often seen in older children are also not typically found in infant HGG (22). Occasional *BRAF* V600E mutations are found, particularly in DIGG/DIA (23), while histone and *IDH1* mutations are rare. Methylation profiling indicates that the infant group may display a more LGG-like methylation pattern, with a 2-year survival of 74% (8). The most common somatic alterations seen in infants appear to be gene fusion events, particularly *NTRK1/2/3.* Although

not specific to brain tumors (24), these were found to span both LGG and HGG in large-scale studies in children, with novel *QKI-NTRK2* and *NACC2-NTRK2* fusions found in pilocytic astrocytomas (25,26), and *AGBL4:NTRK2*, *TPM3:NTRK1*, and *ETV6:NTRK3* fusions found in HGG patients aged less than 3 years (9). More recently, several case reports have identified additional receptor tyrosine kinase (RTK) gene fusions in infant glioma of differing histologies (17,27-35).

In the present study, we collected the largest series of infant gliomas (exclusive of pilocytic astrocytomas) assembled to date and present a classification system based on integrated methylation profiling, fusion gene analysis, mutation detection, and histological review, with preclinical and clinical evidence of effective targeting of the driving alterations in these unique entities.

Results

Refinement of an intrinsic set of infant hemispheric gliomas

We collected a unique series of 241 gliomas, from patients under the age of four years at diagnosis from multiple centres around the world, with a view to exclude *a priori* pilocytic astrocytomas and other well characterised, low grade lesions with clear molecular markers (Figure 1A). To ensure this, we searched for pathognomonic structural variants using a variety of sequencing platforms including whole genome, exome, RNAseq and a novel custom capture panel (Figure 1B). We identified 28 cases to be excluded, mostly due to presence of *KIAA1549:BRAF* fusions (n=22), the vast majority of which were collected as an otherwise unspecified cerebellar astrocytoma (Supplementary Table S1). We also identified three cases of *FGFR1* tandem duplication (including glioneuronal tumors), two *MYB/MYBL1* fusions, and a case with *MN1:BEND2* (representing the novel entity of HGNET-MN1 (36)). Of the remaining 213 cases, a further 13 were excluded based on clear Heidelberg classifier matches to other non-glioma CNS tumors from methylation array profiling data (Figure 1C). These included two ependymomas, two HGNET-BCORs, an ETMR and others (Supplementary Figure S1). A further 9 cases failed array QC and were excluded from further analysis. Finally, our series of 191 cases were projected onto a reference set of gliomas comprising multiple entities. Sixty-one of these infant samples most readily clustered with a known high or low grade subtype, leaving us 130 infant gliomas for further analysis that we define as our ‘intrinsic set’ (Figure 1D), as they comprise a novel grouping of tumours with key clinical and molecular features in common, as we describe below.

The infant glioma cases excluded on the basis of methylation profiling (n=61) were found to have arisen in anatomical areas of the CNS appropriate for the relative subgroup assignment, such as diffuse midline glioma K27M mutant cases in the pons, pilocytic astrocytoma-like cases in the cerebellum, and PXA-like cases in the

cerebral hemispheres (Figure 1E), and were often accompanied by the expected genetic alteration. Interestingly, the remaining intrinsic set included the vast majority of those patients diagnosed under the age of 1 year (49/63, 78%; overall median of intrinsic set = 7.2 months). These cases scored most highly as two named subgroups in the current version (v11b4) of the methylation classifier – desmoplastic infantile ganglioglioma / astrocytoma (DIGG/DIA) and the poorly defined infant hemispheric glioma (IHG) (Supplementary Table S2). The vast majority of these cases were found in the cortex, DIGG/DIAs particularly in the frontal lobe. These cases were found to have a significantly improved outcome compared to cases classified as high grade gliomas (HGG), with a median overall survival similar to those considered as low grade gliomas (LGG) (Figure 1F), with the important caveat that detailed treatment information was not available across the cohort. The HGG subtype exclusions were predominantly >1 year old and showed a tendency towards a worse outcome than the other infant tumors (p=0.0567, log-rank test). This remaining intrinsic group of tumors formed a continuum which clustered clearly apart from other glioma subgroups in a tSNE projection based upon methylation array data from the glioma reference set (n=1652) (Figure 1G). Many of these cases did not unequivocally classify as either IHG or DIGG/DIA despite their tight clustering, suggesting that the reference classes for these tumors likely needs expanding and updating.

Infant hemispheric gliomas are defined by presence or absence of receptor tyrosine kinase fusions

Additional gene sequencing (panel, exome or genome) was available for 65 cases, including 41 of the intrinsic set, all of whom had fusion analysis by panel or RNAseq. Samples excluded as representing other glioma subtypes were found to harbor mutations consistent with such tumors, including *IDH1* R132H, *H3F3A* and *HIST1H3B* K27M, as well as common co-segregating variants in *TP53*, *NF1*, *PTEN*,

PIK3CA and *ACVR1*, deletions of *CDKN2A/B* and amplification of *PDGFRA* (Figure 2A). These were almost entirely absent from the intrinsic set. Instead, 25/41 cases (61%) harbored fusions in either *ALK* (n=10), *NTRK1/2/3 (n*=2, 2 and 8, respectively) *ROS1* (n=2) or *MET* (n=1), usually in the absence of other alterations (Figure 2B). The fusion-positive cases were mostly classified as IHGs (n=21) or low-scoring DIGG/DIAs (n=4). Although *ALK* fusions were restricted to the intrinsic set, we observed *NTRK* fusions in other glioma subtypes (especially *NTRK2*, n=3). We additionally observed an *FGFR1:TACC1* fusion in the *IDH1* / *TP53* case (Figure 2A). High-scoring DIGG/DIAs and 'DIGG/DIA-like' tumors were found with *BRAF* V600E (n=3) or *PIK3R1* mutations and isolated mutations in bromodomain-containing genes (*BRD8*, *BRD4*, *BRD2*) and others (Figure 2B). A single case harbored amplifications in both *MYC* and *MYCN*, in addition to *TP53* and *PIK3CA* mutations. Although a proportion (<25%) of tumors were found with whole-arm DNA copy number changes, the majority of intrinsic cases harbored few if any large-scale copy number alterations (Supplementary Figure S2A) (Supplementary Table S3).

There were no differences in the number of copy number changes between fusion-positive and fusion-negative cases (p=0.567, t-test) (Supplementary Figure S2B). Notably, the only significant focal differences were those marking common gene fusions at the *ALK* and *NTRK3* loci (Supplementary Figure S2C). A novel and refined copy number analysis from the methylation array data identified such breakpoints in either intra-chromosomal (short gains or losses) or inter-chromosomal (imbalances) RTK fusion events in 53/71 (75%) cases across the whole cohort (Figure 3A-C). Across the intrinsic set as a whole, 65/130 (50%) cases were found to harbor structural variants targeting *ALK*, *NTRK1/2/3*, *ROS1* or *MET* (46/80, 57.5% IHGs), compared with 18 of the other 111 cases in the original series (16%; p<0.0001, Fisher's exact test) (Supplementary Figure S3A-C) (Supplementary Table S4). Where possible, these were validated through a combination of genome, RNA and/or Sanger

sequencing, and were frequently accompanied by detectable focal DNA copy number breakpoints within the fusion partners, as exemplified for *ETV6:NTRK3* (Figure 3D) and the novel *ZC3H7A:ALK* fusions (Figure 3E). The most commonly targeted genes in the intrinsic set included *NTRK1/2/3*, predominantly *ETV6:NTRK3*, but also recurrent *EML4:NTRK3* and *TPM3:NTRK1* fusions (Figure 3F). *NTRK2* was found with numerous novel partners (e.g. *KCTD16:NTRK2* and *AGBL4:NTRK2*) but were largely seen in other glioma subtypes occuring in the appropriate anatomical locations (e.g. H3K27M in midline regions) (Supplementary Table S4), suggesting an important difference in *NTRK2* compared to *NTRK1/3* fusion-positive cases. *ALK* fusions were the most common (n=39), were largely restricted to the intrinsic set, and included both intra- and inter-chromosomal rearrangements (Figure 3G), including both previously reported (*PPP1CB:ALK*, *EML4:ALK*, *HIP1:ALK*, *PRKAR2A:ALK, SPTBN1:ALK*) and novel fusions (*MAD1L1:ALK*, *MAP2:ALK*, *MSI2:ALK*, *SPECC1L1:ALK*, *SYNDIG1L:ALK*, *ZC3H7A:ALK*, *CLIP2A:ALK*) (Supplementary Table S4). Within the intrinsic set, there was a trend towards the presence of any fusion conferring a longer overall survival compared to those without (p=0.0687, log-rank test) (Figure 3H).

With whole genome sequencing of fusion-negative cases failing to identify consistent genetic drivers of this subtype of the disease (Supplementary Figure S4), we turned to the methylation data in order to further explore the heterogeneity within infant HGG. Hierarchical clustering on the basis of differential probes associated with the most common genetic alterations found, resulting in the separation of distinct sets of IHG subgroups in addition to clear DIGG/DIA and ‘DIGG/DIA-like’ tumors (Supplementary Figure S5A). Despite the presence of recurrent *NTRK* fusions, these infant gliomas clustered apart from mesenchymal tumors harbouring *ETV6:NTRK3*, including infantile fibrosarcoma and congenital mesoblastic nephroma (37) (Supplementary Figure S5B). Running methylation-based gene ontology analysis on

the differentially methylated regions (Supplementary Table S5) highlighted little overlap between *ALK* fusion, *NTRK* fusion and fusion-negative cases (total 9.5%) (Supplementary Figure S5C). *ALK* fusion cases were significantly associated with dysregulation of genes associated with glutamate receptors, synapses, signal transduction and morphogenic stages of development (Figure 4A), whilst *NTRK* fusion cases were linked with genes controlling neuronal differentiation and the earliest stages of embryogenesis, as well as signalling via the JNK cascade (Figure 4B). By contrast, fusion-negative cases were predominantly associated with the response to multiple endogenous stimuli, particularly the TGFβ pathway, and the regulation of stem cell pluripotency and cell fate (Figure 4C). Although only exploratory due to the small sample sizes, and needing independent validation in an independent cohort, as exemplars of the differential epigenetic regulation of key genes controlling these processes in the distinct subgroups, we observed consistent reduction in methylation at CpG sites governing expression of *WNT5A* in *ALK* fusion cases (Figure 4D), *STAT1* in *NTRK* fusion cases (Figure 4E) and *TP63* in fusion-negative samples (Figure 4F) (Supplementary Table S5). This resulted in differential protein expression as assessed by multi-labeled immunofluorescence with antibodies directed against these targets, with representative examples shown for WNT5A and STAT1 in *ALK*-fusion (Figure 4G) and *NTRK*-fusion (Figure 4H) cases respectively. Using a NanoString assay for the 30 most differentially methylated genes between subgroups, we were able to distinguish ALK- / NTRK- fusion positive and -negative subgroups in a series of 21 infant HGG for which we had sufficient material (Figure 4I). Notably, we did not observe TP63 protein expression in any of our samples, although differential overexpression of the transcript was observed for fusion-negative cases.

Histological examination of those tumors classified as IHG revealed highly cellular astrocytic tumors with cells arranged in uniform sheets throughout the section (Supplementary Figure S6A,B,C). Cytologically, spindled nuclei (Supplementary

Figure S6D), an occasional ganglion cell component (Supplementary Figure S6E), or gemistocyte-like cells (Supplementary Figure S6F) could be seen either focally or throughout the tumor. Tumors frequently showed a superficial hemispheric location often involving the meninges, and had a well-defined border with adjacent normal brain. Palisading necrosis (Supplementary Figure S6G), microvascular proliferation and mild-moderate nuclear pleomorphism were almost universally seen. In some cases, a more nodular architecture was observed (Supplementary Figure S6H,I). Rarely, some showed less cellularity (Supplementary Figure S6J), and mineralisation, calcification or xanthomatous change could be observed (Supplementary Figure S6K). Consistent with these features, 67/80 (84%) of IHG cases were originally diagnosed as a high-grade glioma, although a variety of other diagnoses were included in the original pathology reports (Supplementary Figure S6L). A summary of the histological findings is given in Supplementary Table S6, with no statistically significant difference of features assessed between fusion-positive and-negative subgroups. The number of mitoses observed was highly variable, and proliferation as assessed by Ki67 staining highlighted cases presenting with both frequent (Supplementary Figure S6M) and sparsely positive nuclei (Supplementary Figure S6N). There was a significantly elevated Ki67 index in *NTRK* fusion-positive compared to fusion-negative IHG cases (p=0.0479, t-test), though not for *ALK* (p=0.3622, t-test) (Supplementary Figure S6O). Notably, the *NTRK* (median=22.5) and *ALK* (median=15.6) fusion-positive indices are at the upper end of values reported (38) for older patients with grade IV (median=15.8) and grade III (median=11.8) glioblastomas and anaplastic astrocytomas, with fusion-negative cases (median=5.6) closer to grade II astrocytomas (median=3.0).

Generation and pre-clinical testing of an ALK fusion-driven in vivo model

To assess the tumorigenic potential of the most commonly detected *ALK* gene fusion variant (*PPP1CB:ALK*) in a model system, we attempted to generate an *in vivo*

model using two complementary somatic gene transfer-based methods (RCAS/Ntv-a viral gene transfer and *in utero* electroporation (IUE)) (Figure 5A). When using the RCAS approach with injection of cells producing *PPP1CB:ALK*-containing virus at p0 on a *Cdkn2a* null background, tumor formation was rare (2/19 mice), and only after 300 days. By contrast, *in utero* electroporation at E14.5 with *PPP1CB:ALK* alone was able to generate consistent tumour formation with 100% penetrance, albeit with a relatively long latency of more than 250 days. Although not commonly found in the human disease, when combined with CRISPR/Cas9-mediated knockout of either *Trp53* or *Cdkn2a* for practical purposes, we observed highly efficient tumor formation with a median survival of 32 and 52 days, respectively (Figure 5B). *PPP1CB:ALK* mice +/- Cdkn2a-ko gave rise to tumors which reflected the human setting, including the typical foci of palisading necrosis, mitotic activity, glial cytology and/or clear astrocytic differentiation (Supplementary Figure S7A). All tumors would be classified as high-grade astrocytomas or glioblastomas. Staining for the HA epitope tag included at the C-terminus of the ALK fusion protein in the IUE/*Cdkn2a*-ko setting indicated widespread expression of the fusion protein, with invasion of individual tumor cells into the brain parenchyma (Supplementary Figure S7B).

To test the potential efficacy of targeted ALK inhibition in the context of this tumor model, we first dissociated tissue from a murine tumor into a single-cell suspension for growth in neurosphere (serum-free, non-adherent) conditions. Four different ALK inhibitors were then tested for *in vitro* growth inhibitory effects (crizotinib, ceritinib, alectinib, lorlatinib), representing different generations of inhibitor either approved for clinical use or currently in trials. Whilst all inhibitors showed a significant growth inhibitory effect at nanomolar concentrations (17) (Supplementary Figure S7C), there were differences in potency between the different compounds (Supplementary Table S7).

Due to its clear *in vitro* efficacy and reportedly good blood-brain barrier penetration (an important consideration for clinical translation for brain tumors), lorlatinib was chosen as the primary candidate for *in vivo* testing in our preclinical *ALK* fusion model. For this purpose, adult CD1 mice were allografted with short-term *in vitro*-cultured *PPP1CB:ALK*;*Cdkn2a-/-* cells and monitored for tumor growth using bioluminescence imaging (BLI). At the start of treatment (14 days after injection), mice were stratified into temozolomide (standard chemotherapy), vehicle control or lorlatinib arms, based on consecutive ranking (highest BLI signal to lorlatinib, 2nd highest to control, 3rd highest to temozolomide and so on). Whilst temozolomide was found to slow tumor growth in comparison with vehicle control, all tumors in these two treatment arms continued to grow. In contrast, all but one lorlatinib-treated animal displayed a significant reduction in BLI signal compared with the pre-treatment baseline (Figure 5C,D). This imaging response corresponded with a significant increase in survival in the lorlatinib-treated group compared with the two control arms ($p<0.0001$; although all tumors re-grew after stopping treatment after 28 days, with all mice ultimately needing to be sacrificed due to onset of tumor symptoms) (Figure 5E). No significant difference in body weight was observed between mice on the different treatment arms (data not shown), and the compounds were generally well tolerated. A similar experiment was performed using lorlatinib *versus* temozolomide in mice transplanted with cells from an ALK fusion-only mouse tumor. This also revealed a significant tumour regression (Supplementary Figure S7D) and survival increase with lorlatinib ($p=0.004$, log-rank test), with one animal showing prolonged survival at last follow-up, 8 months post injection (~6 months after end of treatment) (Supplementary Figure S7E). Overall, these findings provide a strong pre-clinical rationale for the potential use of targeted ALK inhibition in a clinical setting. For one of the cases in our study, DKFZ_INF_307, we have been able to demonstrate this directly. Here, a 1 month old boy underwent a left craniotomy with gross total resection, and was diagnosed as glioblastoma (WHO grade IV). He

underwent successive rounds of HIT SKK / ACNS and temozolomide chemotherapy, eventually showing progressive disease after both. He was found to have a *MAD1L1:ALK* fusion and was started on ceritinib, resulting in stable residual disease for nearly two years to date (Figure 5F).

Patient-derived models and clinical experience with NTRK inhibitors

Finally, we explored the utility of treating RTK fusion-positive infant gliomas with targeted inhibitors. We established two primary patient-derived cell cultures from infant glioma specimens with either *TPM3:NTRK1* or *ETV6:NTRK3* fusions (Figure 6A) and compared their *in vitro* sensitivities to three small molecule inhibitors of TrkA/B/C with two fusion-negative paediatric glioma cultures (Figure 6B). *NTRK* fusion-positive cells were more sensitive to entrectinib, crizotinib and milciclib, with differential sensitivities ranging from 2-9 fold over fusion-negative cells ($p=0.0253$, crizotinib; $p=0.0786$, entrectinib; $p=0.0141$, milciclib) (Supplementary Table S7), and reduction in downstream signalling via phospho-Akt and phospho-Erk (Supplementary Figure S7F). The infant glioma models were not tumorigenic after multiple orthotopic implantation experiments in immunodeficient mice, precluding *in vivo* assessment (data not shown).

Clinical treatment with Trk inhibitors was given to two patients in our cohort with *ETV6:NTRK3* fusions. The first case, OPBG_INF_035 was a girl diagnosed with a large frontal mass at 36 weeks' gestation (Figure 6C). It was a large, heterogenous mass with solid, cystic and haemorrhagic components. A biopsy was performed after birth and it was diagnosed histologically as a glioblastoma (WHO grade IV). The child subsequently received chemotherapy (methotrexate, vincristine, etoposide, cyclophosphamide, thiotepa) before undergoing a subtotal resection 3 months later. An *ETV6:NTRK3* fusion was identified in the DNA from both the biopsy and resection specimens, and four months post-surgery, the child was commenced on crizotinib.

An MRI scan performed after 9 months of treatment with crizotinib showed a 56% reduction in the size of the remaining solid component of the tumor compared to the post-surgery MRI scan (RANO criteria size reduction of >50% and stable). After an additional 3 months treatment with larotrectinib, the remaining solid component showed a further reduction in size now reaching 73% (Supplementary Figure S8A). Clinically, the child remains well. The second patient, MSKC_INF_006, presented with a generalized seizure aged 11 months (Figure 6D). An MRI scan revealed a pontine mass with central haemorrhage. The child underwent surgery and a gross total resection was achieved. Histologically the tumor was diagnosed as a low grade neuroepithelial neoplasm. The child developed a recurrence, at which point vincristine and carboplatin were commenced and a complete response was achieved. However, the tumor progressed two years after the original resection; a further gross total resection was achieved and the child treated with larotrectinib after an *ETV6:NTRK3* fusion was identified, with the aim of preventing further recurrence. To date, the child remains well with no evidence of recurrence after 12 months of treatment.

Notably, the patients from whose tumors our primary cell lines were derived have both only received surgery to date, and remain well. QCTB_INF_R077 was diagnosed with a tumor in the left fronto-parietal lobe *in utero* and underwent biopsy and subsequent resection shortly after birth . Histologically, the tumor was reported as a primary neuro-epithelial tumor. The child was not treated with any adjuvant therapy. At 5.5 years old, there has been no progression or relapse and the child has stable disease (Supplementary Figure S8B). The second patient, QCTB_INF_R102 aged 8 months, presented with a tumor in the left temporal lobe aged 8 months. He subsequently received a gross-total resection, with the tumor diagnosed as a ganglioglioma (WHO grade I). He also did not receive any adjuvant therapy post

resection and is currently 4 years old and remains stable under regular surveillance (Supplementary Figure S8C).

In summary, diffuse infant gliomas represent distinct disease entities marked by characteristic clinicopathological profiles and in most cases clinically actionable gene fusions (Figure 7).

Discussion

Malignant glioma presenting in infancy represents a specific clinical challenge, involving diagnostic uncertainty and a hesitancy to aggressively treat given the reported superior outcomes compared with older children, coupled with the high risk of neurocognitive deficits (39). This is compounded by a lack of biological understanding due to the rarity of these tumors. The present international collaborative study brings together the largest collection of tumors originally reported as high grade or diffuse gliomas in this age group, by contrast with another recent multi-institutional study which was predominantly comprised of low grade tumours (17). Our study uniquely includes methylation and gene expression data, and allows for refinement of subgroups within the malignant spectrum of disease with important clinical management implications; we also present experience of clinical responses with targeted agents even after progression on standard chemotherapies.

A first key finding relates to the difficulty of differential diagnoses in these very young children, with ~10% cases unequivocally classifying as other tumor entities on the basis of methylation profiling (40) or the presence of pathognomonic gene fusions (36), even after discounting mis-diagnosed or mis-assigned pilocytic astrocytomas. Often this uncertainty is reflected in the original pathology report, with atypical features highlighted. However, the highly heterogeneous nature of high grade glial tumors provides for a broadly inclusive category in the current WHO classification, which in many cases may result in what is considered to be a relatively uncontroversial histological diagnosis despite widely varying morphologies. Similarly, combined genetic and epigenetic analyses reveal a third of remaining cases to be biologically identical to known high or low grade glioma subtypes, with substantially different prognoses reflective of the known clinical course of the relevant tumor categories. Together, these data make the important points that histopathologic evaluation alone is insufficient to predict outcome, and that high grade gliomas

predominantly occuring in older childhood may also present in the infant population with little survival benefit from standard treatment protocols.

After these exclusions, there remains what we define as an intrinsic set of infant gliomas, which are largely restricted to the cerebral hemispheres and occur in the youngest patients, usually under 12 months old. These patients, despite more than three-quarters unequivocally reported as WHO grade III or IV astrocytoma, have an overall survival more akin to lower grade tumors, yet lack the key molecular features of both HGG and LGG. They appear to form a biological continuum of disease between the recognized MAPK-driven desmoplastic lesions (DIGG/DIA), which may respond clinically to targeted *BRAF* V600E inhibitors, even after previous chemotherapy (41), and a novel assignation of diffuse infant hemispheric glioma. This latter end of the spectrum is strikingly defined by nearly two-thirds of tumors harboring fusions in genes encoding the receptor tyrosine kinases *ALK*, *NRTK1/2/3*, *ROS1* and *MET*. Although structural variants involving these genes within the age group have been described in case reports (27-35) and a recent larger study (n=29) (17), the current report represents a uniquely powerful study of these rare tumors, by accumulating a series of 82 infant cases with RTK fusions with full methylation profiles.

Molecularly, these events included interstitial microdeletions such as those at chromosome 2p23 resulting in the fusion of *CCDC88A or PPP1CB* and *ALK* (17,34) and at 6q21 fusing *ROS1* and *GOPC* (previously known as *FIG*, and originally described in an adult GBM cell line (42)); additional focal DNA copy number losses targeted *MET* at 7q31 (43). There were multiple instances of inter-chromosomal copy number gains fusing *ALK* to a series of novel partners, including *MAD1L1* (7p22), *ZC3H7A* (16p13), *MSI2* (17q22), *SYNDIG1* (20p11) and *SPECC1L* (22q11), as well as the intra-chromosomal *EML4:ALK* fusion that is well-characterized in non-small

cell lung cancer and others (44). The *NTRK* genes had a variety of inter-chromosomal partners, with around half of cases marked by a DNA copy imbalance at either locus. Notably, *NTRK2* fusions (also described in LGG (25,26,45)) were largely found in tumors classifying as other glioma subtypes, as were the previously described *FGFR:TACC* fusions (46).

Histopathologically, within the context of HGGs, certain common features of the intrinsic infant hemispheric gliomas could be recognized. Cases tended to have a relatively uniform architecture, with marked pleomorphism. There was an enrichment of gemistocytic-like cells, as has been reported for a case with *ZCCHC8:ROS1* fusion (29); a predominance of spindle cell differentiation, reminiscent of mesenchymal tumors with *NTRK* fusions (47), and also described in an *ETV6:NTRK3* infant glioma (35). Our *NTRK* fusion cases in the present study clustered distinctly from *ETV:NTRK3*-positive infantile fibrosarcoma and congenital mesoblastic nephroma, however, suggesting a distinct origin. Several cases also had ependymal differentiation, consistent with two cases with *ALK* fusions (*KTN1:ALK* and *CCDC88A*:*ALK)* reported as not easily fitting the established WHO brain tumor entities (34). Notably, *CCDC88A*:*ALK* cases have been reported clinically as both low- and high- grade glioma, however the same study found tumours generated by overexpressing the fusion in xenografted immortalized human astrocytes to have a high proliferative index, glial marker expression and pseudopallisading necrosis (17), suggestive of high grade lesions in common with our *in utero* electroporation modelling approach. A further case report described a *KIF5B:ALK* fusion in an infant with microglial proliferation, spindle cells with scattered mitotic figures, and a mixed inflammatory infiltrate of scattered lymphocytes, plasma cells and eosinophils, indicating potential microglioma or gliofibroma (31). The recognition of tumors in this series that biologically resemble DIGG/DIA (WHO grade I) is compatible with their histology, in that some cases have been described as presenting with a poorly

differentiated component (39). The case with *ZCCHC8:ROS1* fusion was also described to display a cellular element within a fusocellular desmoplastic component (29), and we noted focal ganglion cells in our series. However, despite these differences, it is still not possible at the present time to define clear histology-only criteria which can reliably distinguish between these molecularly-defined intrinsic infant tumors and other glioma subtypes in the same age group.

The presence of recurrent *ALK/NTRK/ROS1/MET* fusions represent clearly targetable alterations, in common with subgroups of adult epithelial tumors (48,49), and their identification through screening approaches and routine diagnostic sequencing panels (50-53) makes them amenable to selection for clinical trials despite their rarity. The distinct morphological variants, the restricted spatial and temporal patterns of presentation, and the specificity of oncogenic events largely in the absence of other mutations or large-scale chromosomal rearrangements suggests an exquisite developmental susceptibility for transformation which would account for this rare subgroup of tumors.

Multiple *ALK* partners are associated with synapse formation and activity (*CCDC88A, HIP1, SYNDIG1*), neuronal cytoskeletal reorganisation (*CCDC88A, SPECC1L*) and microtubule assembly (*MAP2, PRKAR2A, EML4*), as well as PI3K/MAPK signalling (*PPP1CB, CCDC88A, SPECC1L*) and cell cycle progression (*MAD1L1*) (54-62). Thus in addition to the activated kinase activity of the ALK receptor itself, these fusions likely disrupt key regulatory processes in neurodevelopment, as exemplified by the differential methylation of genes controlling these processes we observed. The most common *ALK* fusion, *PPP1CB:ALK,* was found to be tumorigenic when introduced in prenatal, though largely not postnatal mice, further demonstrating the importance of developmental context associated with the oncogenicity of these alterations.

*AL*K fusion-positive tumors were found to be sensitive to targeted ALK inhibition *in vitro* and *in vivo*, resulting in tumor shrinkage and extension of survival in the latter in contrast to the standard chemotherapeutic agent temozolomide. Excitingly, this experience was mirrored in the clinic, whereby a child diagnosed at 1 month old experienced stable disease for nearly two years on targeted therapy after progressing on two successive chemotherapy protocols, including temozolomide. Critically, *NTRK* fusion cases were also found to respond to targeted inhibitors in patient-derived models *in vitro* as well as in children treated clinically, in common with isolated reported cases (35), whereby for example a 3-year-old girl who had failed multiple therapies including chemotherapy and radiotherapy showed near total resolution of primary and metastatic lesions after treatment with larotrectinib. If validated in larger trials, such agents may represent attractive options in order to spare the long-term sequalae of chemotherapy and radiotherapy, whilst maintaining the generally good prognosis of these patients (27,30,33).

Despite the frequency of alterations identified, not all of the intrinsic infant gliomas were found to harbor RTK fusions. These fusion-negative cases (at least on the basis of the platforms used in this study) had a lower proliferation index compared to *NTRK*-positive cases, but a worse prognosis under standard treatment. Although we could identify no apparent recurrent genetic driver of this subgroup, even with whole genome sequencing of a subset of cases, there were clear epigenetic differences compared to fusion-positive cases, with dysregulated gene networks associated with the regulation of stem cell pluripotency, plausibly suggesting an immature progenitor cell phenotype for these genetically bland lesions. By contrast, *NTRK*-fusion cases were associated with an embryonic, neuronal developmental programme, and *ALK*-fusion cases with later AMPA-receptor synaptic plasticity signatures.

Further work is needed to explore all intrinsic infant glioma subgroups, in particular the fusion-negative cases. However, it is clear that these tumors harbor unique biology with associated clinicopathological differences, and should no longer be diagnosed or treated in the same way as their older counterparts. Maximal safe surgical resection remains the aim of treatment, regardless of subtype (17). However, our study has shown that RTK fusions can be found across all subgroups (although more frequently seen in the IHG group) and so screening (initially via copy number profiling with subsequent validation) will help to identify patients who may be eligible for targeted therapy or clinical trials.

Author contributions

MCl, AMac, BI, DWE, DTWJ, TSJ and CJ designed the study, analysed data, and wrote the manuscript. LHi, JP, RGT, TBa, KMK, CR, MM, SCr, JK, US, UK, DK, MSi, LBr, AJM, LD, SA-S, CCa, BZ, CCh, MSn, IJD, SWG, TM, CH, ISl, RP, H-KN, SCa, BFK, MP, SG, MW, LBa, TKO, PB, MF, JCr, MCa, MKa, MKR, JS, ME, WD, KK, MvB, TP, FA, PHD, AK, DS, MZ, OW, DC, TBl, CM, EM, FDC, GSC, SB, ASM, TEGH, SL, MAK, CMK, MT, MJC, DSZ, KA, DHa, FC, LM, FS, AvD, SMP, SJB, AMas, AC and MV provided cases, data and/or clinical annotation. SN, ISt, BCW, EI, ST, DAS, DMC, VM, AB, LHo, DS, AV, ARF, AA, JCh, MKr, MD, KH, JD, WO, JW MH, RN, JKRB, SPR, DHu, PP, and SP carried out experiments and data analysis. All authors reviewed and approved the final manuscript.

Methods

Cases

All patient samples included were classified as gliomas (WHO grade II, III or IV) aged <4 years old (including congenital cases) from all CNS locations (including spinal tumors). Cases were excluded if they had been diagnosed as a pilocytic astrocytoma with a known *BRAF* fusion or mutation. Ependymal, embryonal, mesenchymal and germ cell tumors were also excluded. Samples were received from national (Great Ormond Street Hospital, London, n=33; King's College Hospital, London, n=21; University Hospitals Bristol, n=9; Newcastle Royal Infirmary, n=6; St George's Hospital, London, n=4) and international collaborators (German Cancer Research Center (DKFZ), n=86; Ospedale Pediatrico Bambino Gesù, n=37; St Jude Children's Research Hospital, Memphis, n=17; Memorial Sloan Kettering Cancer Center, New York, n=6; Queensland Children's Tumor Bank, Brisbane, n=5; Universitätsklinikum Hamburg-Eppendorf, n=5; Children's Cancer Institute, Sydney, n=2; Children's Hospital of Wisconsin, n=2; Emory University Hospital, Atlanta, n=1; St. Petersburg Hospital No. 6, n=1; Wake Forest School of Medicine, Winston-Salem, n=1; The Chinese University of Hong Kong, n=1; Children's National Medical Centre, Washington DC, n=1; Chaim Sheba Medical Center, Tel Aviv, n=1; Oregon Health & Science University, Portland, n=1; University of Ljubljana, n=1). Where possible an H&E slide, 10 unstained sections, formalin-fixed paraffin-embedded (FFPE) tissue rolls, or frozen tissue was provided for each case. In some cases, data alone was provided. A total of 241 cases were entered into the study. Eight cases from King's College and St George's Hospital London (8), and ten cases from St Jude Children's Hospital Memphis (9) have been previously published. All patient samples were collected under full Research Ethics Committee approval at each participating centre.

Nucleic acid extraction

DNA was extracted from frozen tissue by homogenisation prior to following the DNeasy Blood & Tissue kit protocol (QIAGEN, Hilden, Germany). DNA was extracted from formalin-fixed, paraffin-embedded (FFPE) pathology blocks after manual macrodissection using the QIAamp DNA FFPE tissue kit protocol (QIAGEN). Concentrations were measured using a Qubit fluorometer (Life Technologies, Paisley, UK). RNA was extracted by following the RNeasy Mini Kit protocol (QIAGEN), and quantified using a Nanodrop 2000 Spectrophotomer (Thermo Scientific).

Methylation profiling

The quantity and quality of DNA varied between cases with FFPE samples yielding less (range for FFPE: 11.0 – 2960.0ng, range for fresh frozen: 211.0 - 5358.0ng). Methylation analysis was performed when >150ng of DNA was extracted, using either Illumina 450K or EPIC BeadArrays at DKFZ (Heidelberg), University College London (UCL) Great Ormond Street Institute of Child Health or St. Jude Children's Research Hospital. Data from Illumina 450k or EPIC arrays was pre-processed using the minfi package in R (v11b4). DNA copy number was recovered from combined intensities using the conumee package. The Heidelberg brain tumor classifier (molecularneuropathology.org) (40) was used to assign a calibrated score to each case, associating it with one of the 91 tumor entities which feature within the current classifier (v4). Clustering of beta values from methylation arrays was performed based upon correlation distance using a ward algorithm. DNA copy number was derived from combined $\log_2$ intensity data based upon an internal median processed using the R packages minfi and conumee to call copy number in 15,431 bins across the genome. Gene ontology analysis of differentially methylated regions was carried out using methylGSA (rdrr.io/bioc/methylGSA/), adjusting the number of CpGs for each gene by weighted resampling and Wallenius non-

central hypergeometric approximation in methylgometh (63). Ontology networks were constructed using ShinyGO (bioinformatics.sdstate.edu/go/).

Fusion panel

A custom fusion panel consisting of 22 genes associated with fusions in paediatric brain tumors (*ALK, BCOR, BRAF, c11orf95, C19MC, CIC, ETV6, FGFR1-3, FOXR2, KIAA1549, MET, MN1, MYB, MYBL1, NTRK1-3, RAF, RELA, TPM3* and *YAP1*) was designed with a library of probes to ensure adequate coverage of the specified regions (Roche Sequencing Solutions) (64). Where available, 100-200ng of DNA was used for library preparation using KAPA Hyper and HyperPlus Kit (Kapa Biosystems) and SeqCap EZ adaptors (Roche). Following fragmentation, DNA was end-repaired, A-tailed and indexed adaptors ligated. DNA was amplified, multiplexed and hybridized using 1ug of the total pre-capture library DNA. After hybridisation, capture libraries were amplified and sequencing was performed on a MiSeq and NextSeq (Illumina). Quality control (QC), variant annotation, deduplication and metrics were generated for each sample. The raw list of candidates provided by Manta (https://github.com/Illumina/manta) were filtered for more than 2 reads covering both genes, common false positive base pairs (bp) positions/fusions outside of the capture set at both ends, common breakpoint/false positives within 10 bp, common false positive gene pairs, fusions within the same gene and homologous sequences greater than 10bp. Breakdancer was used to confirm all the breakpoints in all samples. Sequences either side of the break points were annotated to look for repetitive elements. A BLAT score was obtained to remove loci which were not uniquely mapped. Integrative Genomics Viewer (IGV) was used to view the fusions.

DNA and RNA sequencing

DNA was sequenced either as whole genome or captured using Agilent SureSelect whole exome v6 or a custom panel of 329 genes known to present in an unselected series of paediatric high grade glioma (8). Library preparation was performed using 50-200 ng of genomic DNA. Following fragmentation, DNA was end-repaired, A-tailed and indexed adapters ligated. DNA was amplified, multiplexed and hybridized using 1 µg of total pre-capture library. After hybridization, capture libraries were amplified and sequencing was performed on a NextSeq500 (Illumina) with 2 x 150bp, paired-end reads following manufacturer's instructions. Ribosomal RNA was depleted from 500-2000 ng of total RNA from FF and FFPE using NEBNext rRNA Depletion Kit. Following First strand synthesis and directional second strand synthesis resulting cDNAs were used for library preparation using NEBNext Ultra II Directional RNA library prep kit for Illumina performed as per the manufacturers recommendations. Exome capture reads were aligned to the hg19 build of the human genome using bwa v0.7.12 (bio-bwa.sourceforge.net), and PCR duplicates removed with PicardTools 1.94 (pcard.sourceforge.net). Single nucleotide variants were called using the Genome Analysis Tool Kit v3.4-46 based upon current Best Practices using local re-alignment around InDels, downsampling and base recalibration with variants called by the Unified Genotyper (broadinstitute.org/gatk/). Variants were annotated using the Ensembl Variant Effect Predictor v74 (ensembl.org/info/docs/variation/vep) incorporating SIFT (sift.jcvi.org) and PolyPhen (genetics.bwh.harvard.edu/pph2) predictions, COSMIC v64 (sanger.ac.uk/genetics/CGP/cosmic/), dbSNP build 137 (ncbi.nlm.nih.gov/sites/SNP), ExAc and ANNOVAR annotations. RNA sequences were aligned to hg19 and organized into de-novo spliced alignments using bowtie2 and TopHat version 2.1.0 (ccb.jhu.edu/software/tophat). Fusion transcripts were detected using chimerascan version 0.4.5a filtered to remove common false positives.

PCR / Sanger sequencing validation

PCR to validate fusion breakpoints was carried out using primers obtained from Integrated DNA Technologies (Illinois, USA). PCR products were cleaned using the ExoProStar S 20 (Sigma-Aldrich) and were sent for Sanger sequencing (DNA Sequencing and Services, University of Dundee, UK). Sequences were analysed manually with 4Peaks (Nucleobytes, Aalsmeer, Netherlands).

NanoString gene expression analysis

The top 30 genes with the most differentially methylated regions between *ALK*-fusion, *NTRK*-fusion and fusion negative cases were selected for an mRNA expression analysis using a custom nCounter platform and nDesign (NanoString, Seattle, WA, USA). Specimen RNA was mixed in hybridization buffer with CodeSets and hybridized overnight at 65°C. Samples wash reagents and imaging cartridge were processed on the nCounter Prep Station and imaged on the nCounter Digital Analyzer according to the manufacturer's instructions. Data were normalised with NanostringNorm v1.2.1 using variance stabilizing normalization (VSN). Heatmaps were made by clustering the median centred expression values or a correlation matrix based on Euclidean distance using a Ward D2 algorithm.

Immunofluorescence

Paraffin-embedded tissue sections were deparaffinized in three changes of xylene and ethanol. Heat-mediated antigen retrieval was performed (Dako S1699, pH 6.0) and tissue slides were permeabilized with 0.5% Triton X-100 solution for 10 min at room temperature and then blocked with appropriate serum according to the species of secondary antibody for 1 h at room temperature. For STAT1 staining (AHO0832, Invitrogen, 1:800), Alexa Fluor 488 Tyramide Super Boost Kit was used (B40941, Invitrogen) and antibody was incubated at 37 °C for 30 min. For WNT5A (MA5-

15502, Invitrogen, 1:800) and TP63 (39692, Cell signalling, 1:900) staining, samples were incubated at 37 °C for 30 min. Sample slides were then washed in PBS three times and incubated with DyLight 649 (DI-2649, Vector, 1:100) and Alexa Fluor 555 (A31572, Invitrogen, 1:300) -conjugated secondary antibodies for an hour at room temperature. Nuclei were counterstained with DAPI and samples mounted with Vectashield (H1000, Vector Laboratories) and examined using Zeiss Axio Scan.Z1 automated Fluorescence slide scanner.

Histology and immunohistochemistry

Histological review was undertaken according to the WHO Classification of Tumors of the Central Nervous System (2016) (65). Each case was reviewed blinded to the molecular features with a predetermined set of criteria to assess for the presence of histological features characteristic of gliomas such as necrosis, mitotic figures, stromal and astrocytic morphology. Any unusual features not previously associated with these tumors, including unusual nuclear morphology was noted. These features were then re-reviewed in the context of any molecular results identified. Immunohistochemistry for Ki67 (M7240, DAKO, 1:100) was carried out using pressure-mediated antigen retrieval and the Envision^{TM} detection system (DAKO K5007). Slides were mounted using Leica CV Ultra mounting medium, imaged using the high throughput-scanning microscope AxioScan Z1 and quantified using Definiens software.

Novel ALK fusion mouse model

A *PPP1CB:ALK* fusion construct was cloned into either an RCAS or a pT2K vector using RNA from a human glioma sample as template. After cDNA synthesis and PCR amplification, the ends of the product were cut with EcoRI and XhoI (for cloning into pT2K) or NotI and ClaI (for RCAS) and ligated into the target vector using the

Takara Ligation mixture (Clontech). Bacterial amplification and QIAprep® Spin Miniprep kit (QIAGEN) were performed according to manufacturer's instructions to isolate the cloned plasmid. The DNA was sequenced using Sanger Sequencing at GATC Biotech (Heidelberg, Germany) and protein expression was confirmed on Western Blot after transfection of DF-1 cells with the vector.

In utero electroporation: After confirming that the expression vector contained the right inserts, embryos of CD1 mice were injected with plasmid into the fourth ventricle and electroporated *in utero* at E14.5. The *PPP1CB:ALK* fusion plasmid was used alone or in combination with CRISPR guide RNAs against Cdkn2a. Due to the incorporated IRES-Luciferase reporter on the pT2K vector, mice with successful integration of the transgene could be assessed at postnatal day 3 using bioluminescence imaging on an IVIS imager (PerkinElmer). Mice were sacrificed upon first signs of tumor-related symptoms according to humane endpoint criteria. H&E and IHC staining was performed according to standard protocols on 3µm sections.

RCAS: Four days before the calculated birth date, early passage DF-1 fibroblasts for virus production were plated at 2-3x10^5 cells/ T25 flask in 5ml DMEM with 10% FCS + 1% Penicillin/Streptomycin (P/S) + 1% Glutamax at 5% CO_2 at 39°C. One day after, the cells were transfected with the RCAS construct as follows: 4µg of the RCAS plasmid was incubated in 200µl of room-temperature Optimem and 10µl FuGene transfection reagent. After a 15-minute incubation time, this mixture was slowly added to the settled DF-1 cells, mixed well by gently moving the flask and placed back in the incubator. An RCAS-GFP plasmid was always run in parallel in a separate flask to check for transfection success. On the day of birth, the transfected DF-1 cells were harvested using 10x Trypsin-EDTA and counted using the automated cell counter TC20™. 4x10^5 cells in 1µl were used for injection into

newborn Ntv-a;Cdkn2a$^{-/-}$;Pten$^{fl/fl}$ pups at p0. The required amount of cells, depending on the size of the litter, was eluted in DMEM culture medium. The pups were taken out of the cage in a sterile hood and injected into the striatum with 1µl of the DF-1 cell solution using a 10µl Hamilton syringe. Mice were sacrificed upon first signs of tumor-related symptoms according to humane endpoint criteria. All animal protocols were approved by the relevant authority (Regierungspräsidium Karlsruhe) under registration numbers G-212/16 and G-168/17.

In vitro culture and compound testing of murine tumor cells

Murine ALK fusion-positive tumors were dissected immediately post mortem, mechanically dissociated and then filtered through a 40µm cell strainer. Cells were then plated *in vitro* in 10cm dishes and grown as spheres in a 1:1 mix of Neurobasal-A and DMEM/F-12 media containing 1% 1M HEPES buffer solution, 1% 100mM sodium pyruvate MEM, 1% 10mM MEM non-essential amino acids solution, 1% GlutaMAX and 1% antibiotic-antimycotic supplemented with 2% B27, 2µg/ml heparin solution, 10ng/ml H-PDGF-AA, 20ng/ml recombinant human bFGF and 20ng/ml recombinant human EGF. For splitting, cells were dissociated with Accumax at 37°C for 5 minutes.

For *in vitro* drug testing, primary sphere culture cells were plated at 1x10^4 cells/well in 80µl growth factor-containing medium/well in 96-well plates. Triplicates per drug concentration (20µl total volume for each) were added 24 hours after seeding the cells. The drug concentrations ranged between 1nM and 30µM. Corresponding DMSO concentrations were plated as controls, to which the treated wells were normalized. The ALK inhibitors crizotinib, alectinib, ceritinib and lorlatinib were used. All compounds were purchased from Selleck Chemicals and initially diluted in DMSO to either a 10mM or 1mM stock, which were stored at -80°C. A CellTiter-Glo assay (Promega) was used as a readout of compound efficacy. This assay was conducted

72 hours after drugs were added to the cells. For this purpose, 50µl of CellTiter-Glo substrate were added to each well using a multichannel pipette, and plates were incubated for 15 minutes whilst shaking in the dark. After that time, the luminescence signal per well was measured using a Mithras LB940 microplate reader. The respective DMSO control value was subtracted from the drug's value to normalize the readout. The GI_{50} curves show the mean ± SD of the triplicates per condition measured. Representative results from duplicate experiments are shown.

Western blot analysis

Cells were incubated in complete media with vehicle or increasing concentrations of Entrectinib (0.1, 1, 10 µM) and protein was collected 4h post-treatment. Samples were lysed by using lysis buffer (CST) containing phosphatase inhibitor cocktail (Sigma, Poole, UK) and protease inhibitor cocktail (Roche Diagnostics, Burgess Hill, UK). Following quantification using Pierce BCA Protein Assay Kit (Thermo Fisher), cell extracts were loaded for Western blot analysis. Membranes were incubated with primary antibody (1:1000) overnight at 4 °C, and horseradish peroxidase secondary antibody (Amersham Bioscience, Amersham, UK) for 1 h at room temperature. Signal was detected with ECL Prime western blotting detection agent (Amersham Biosciences), visualised using Hyperfilm ECL (Amersham Biosciences) and analysed using an X-ray film processor in accordance with standard protocols. Primary antibodies used were phospho-AKT (Ser473) (CST# 4060), phospho-p44/42 (Thr202/Tyr204) (CST#4370), AKT (CST#9272), p44/42 (CST#9102), GAPDH (CST#2118), all Cell Signalling (Danvers, MA, USA).

In vivo compound testing

To test the effectiveness of ALK inhibition *in vivo*, 6 week old CD1 mice were intracranially allografted with 5x 10^5 mouse *PPP1CB:ALK* tumor cells (see above) in order to give a more standardized latency of tumor formation and to ensure avoid

having to administer treatment to very young animals. The chosen inhibitor was lorlatinib based on the *in vitro* results, as well as HCl and temozolomide as vehicle control and standard-of-care, respectively. Dosing and treatment schedules were as previously described (66). Tumor growth was monitored using bioluminescence imaging on an IVIS imager (PerkinElmer). The tumors were allowed to develop for two weeks before animals were stratified into three treatment groups based on their luciferase signal (rank 1, 4, 7 etc. being assigned to lorlatinib, rank 2, 5, 8 etc. to temozolomide, and rank 3, 6, 9 etc. to vehicle control). Animals were monitored daily for symptoms or abnormal behavior and weighed three times a week, and were sacrificed upon first signs of tumor-related symptoms according to humane endpoint criteria.

Novel patient-derived NTRK fusion models

Each cell culture was initiated using the following method; tissue was first minced using a sterile scalpel followed by enzymatic dissociation with LiberaseTL for 10 min at 37°C. Cells were grown under stem cell conditions, as two-dimensional (2D) adherent cultures on laminin and laminin/fibronectin. Cells were cultured in a serum-free medium, Tumor Stem Media (TSM) consisting of 1:1 Neurobasal(-A), and DMEM:F12 supplemented with HEPES, NEAA, Glutamaxx, sodium pyruvate and B27(-A), human bFGF (20ng/mL), human-EGF (20ng/mL), human PDGF-AA (10ng/mL) and PDGF-BB (10ng/mL) and heparin (2ng/mL). Control lines QCTB-R006 (9.5 years, male, frontal lobe GBM, wildtype) and QCTB-R059 (10.4 years, female, thalamic, *H3F3A*K27M mutant) were also grown as adherent cultures (laminin and laminin-fibronectin). Cells were dissociated enzymatically with accutase and counted using a Beckman-Coulter ViCell cell viability analyser.

For intracranial implantation, all experiments were performed in accordance with the local ethical review panel, the UK Home Office Animals (Scientific Procedures) Act

1986, the United Kingdom National Cancer Research Institute guidelines for the welfare of animals in cancer research and the ARRIVE (Animal Research: Reporting *In Vivo* Experiments) guidelines (67,68). Single cell suspensions were obtained immediately prior to implantation in NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ (NSG) mice (Charles River, UK). Animals were anesthetized with intraperitoneal ketamine (100mg/kg)/xylazine(16mg/kg) and maintained under 1% isoflurane (0.5L/min). Animals were depilated at the incision site and Emla cream 5%(lidocaine/prilocaine) was applied on the skin. A subcutaneous injection of buprenorphine (0.03mg/Kg) was given for general analgesia. The cranium was exposed via midline incision under aseptic conditions, and a 31-gauge burr hole drilled above the injection site. Mice were then placed on a stereotactic apparatus for orthotopic implantation. The coordinates used for the cortex were x=-2.0, z=+1.0, y=-2.5mm from bregma. 300,000 cells in 5µL were stereotactically implanted using a 25-gauge SGE standard fixed needle syringe (SGE™ 005000) at a rate of 2µl/min using a digital pump (HA1100, Pico Plus Elite, Harvard Apparatus, Holliston, MA, USA). At the completion of infusion, the syringe needle was allowed to remain in place for at least 3 minutes, and then manually withdrawn slowly to minimize backflow of the injected cell suspension. An intraperitoneal (IP) injection of the reversing agent atipamezole (1mg/kg) diluted in Hartmann's solution for rehydration was administered. Mice were monitored until fully recovered from surgery and given Carprofen (analgesia) in a gel diet for 48 hours post-surgery. Mice were weighed twice a week and imaged by ^{1}H magnetic resonance imaging (MRI) on a horizontal bore Bruker Biospec 70/20 system (Ettlingen, Germany) equipped with physiological monitoring equipment (SA Instruments, Stony Brook, NY, USA) using a 2cm x 2cm mouse brain array coil. Anaesthesia was induced using 3% isoflurane delivered in oxygen (1l/min) and maintained at 1-2%. Core body temperature was maintained using a thermo-regulated water-heated blanket.

In vitro compound testing of patient-derived cells

Cells were seeded (3000-5000 cells per well) into laminin or laminin-fibronectin coated 96-well plates and treated with different Trk inhibitors at concentration ranging from 0 to 20uM for 8 days. The drugs used were entrectinib (RXDX-101, Selleckchem), crizotinib (PF-02341066, Selleckchem) and milciclib (PHA-848125, Selleckchem). Each assay was performed in three independent biological replicates of three technical replicates each. Cell viability was assessed with Cell Titer-Glo using a FLUOstar Omega plate reader (BMG, LABTECH). Data was analysed and IC50 values were calculated using GraphPad Prism software.

Statistics

Statistical analysis was carried out using R 3.5.0 (www.r-project.org) and GraphPad Prism 7. Categorical comparisons of counts were carried out using Fishers exact test, comparisons between groups of continuous variables employed Student's t-test or ANOVA. Univariate differences in survival were analysed by the Kaplan-Meier method and significance determined by the log-rank test. All tests were two-sided and a *p* value of less than 0.05 was considered significant.

Data availability

All newly generated data have been deposited in the European Genome-phenome Archive (www.ebi.ac.uk/ega) with accession number EGAS00001003532 (sequencing) or ArrayExpress (www.ebi.ac.uk/arrayexpress/) with accession numbers E-MTAB-7802 and E-MTAB-7804 (methylation arrays). Curated gene-level copy number, mutation data and RNAseq data are provided as part of the paediatric-specific implementation of the cBioPortal genomic data visualization portal (pedcbioportal.org).

Acknowledgments

This work was supported by the CRIS Cancer Foundation and the INSTINCT network funded by The Brain Tumour Charity, Great Ormond Street Children's Charity and Children with Cancer UK, Cancer Research UK. The authors acknowledge NHS funding to the National Institute for Health Research Biomedical Research Centre at The Royal Marsden and the ICR, the NIHR Great Ormond Street Hospital Biomedical Research Centre, research nurse funding by the Experimental Cancer Medicines Centre (ECMC) Paediatric Network, as well as CRUK support to the Cancer Imaging Centre at the ICR and Royal Marsden in association with the MRC and Department of Health (England) (C1060/A16464). Further funding support was provided by the German Children's Cancer Foundation (DKKS, project "MNP2.0 – Improving the Diagnostic Accuracy of Pediatric Brain Tumors", and support for the German Brain Tumor Reference Center of the DGNN, grant 2014.17) and the PedBrain Tumour Project contributing to the International Cancer Genome Consortium, funded by German Cancer Aid (109252) and by the German Federal Ministry of Education and Research (BMBF, grant #01KU1201A), and the DKFZ-MOST Cooperation Program. We would like to thank Laura von Soosten (DKFZ) for technical assistance and Richard Buus (ICR) and the Breast Cancer Now NanoString facility for conducting the NanoString gene expression profiling. The authors thank Brain UK for provision of cases and clinical information. The authors thank the Cure Brain Cancer Foundation, Australian Lions Childhood Cancer Research Foundation and Lions Club International Foundation (LCIF). Some of the results are in part based upon data generated by Lions Kids Cancer Genome Project (LKCGP) Partners. The authors thank the German Cancer Research Center (DKFZ) Genomics and Proteomics Core Facility and the Hartwell Center at St. Jude Children's Research Hospital for technical support. The authors acknowledge funding from the American, Lebanese and Syrian-Associated Charities (ALSAC). The Queensland Children's Tumour Bank is funded by the Children's Hospital Foundation

(Queensland). This work was funded in part by the Marie-Josée and Henry R. Kravis Center for Molecular Oncology and the National Cancer Institute Cancer Center Core Grant No. P30-CA008748. We gratefully acknowledge the members of the Memorial Sloan Kettering Molecular Diagnostics Service in the Department of Pathology. MSn acknowledges funding from the Friedberg Charitable Foundation, the Making Headway Foundation and the Sohn Conference Foundation. AK is supported by the Helmholtz Association Research Grant (Germany). MV is a CwCUK Fellow (grant number 16-234). SJB acknowledges funding support from the NIH (CA096832).

References

1. Jones C, Perryman L, Hargrave D. Paediatric and adult malignant glioma: close relatives or distant cousins? Nat Rev Clin Oncol **2012**;9(7):400-13 doi 10.1038/nrclinonc.2012.87.

2. Schwartzentruber J, Korshunov A, Liu XY, Jones DT, Pfaff E, Jacob K*, et al.* Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature **2012**;482(7384):226-31 doi 10.1038/nature10833.

3. 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(3):251-3 doi 10.1038/ng.1102.

4. Jones C, Baker SJ. Unique genetic and epigenetic mechanisms driving paediatric diffuse high-grade glioma. Nat Rev Cancer **2014**;14(10) doi 10.1038/nrc3811.

5. Sturm D, Witt H, Hovestadt V, Khuong-Quang DA, Jones DT, Konermann C*, et al.* Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. Cancer Cell **2012**;22(4):425-37 doi 10.1016/j.ccr.2012.08.024.

6. Bouffet E, Larouche V, Campbell BB, Merico D, de Borja R, Aronson M*, et al.* Immune Checkpoint Inhibition for Hypermutant Glioblastoma Multiforme Resulting From Germline Biallelic Mismatch Repair Deficiency. J Clin Oncol **2016**;34(19):2206-11 doi 10.1200/JCO.2016.66.6552.

7. Shlien A, Campbell BB, de Borja R, Alexandrov LB, Merico D, Wedge D*, et al.* Combined hereditary and somatic mutations of replication error repair genes result in rapid onset of ultra-hypermutated cancers. Nat Genet **2015**;47(3):257-62 doi 10.1038/ng.3202.

8. Mackay A, Burford A, Carvalho D, Izquierdo E, Fazal-Salom J, Taylor KR*, et al.* Integrated Molecular Meta-Analysis of 1,000 Pediatric High-Grade and Diffuse Intrinsic Pontine Glioma. Cancer Cell **2017**;32(4):520-37 e5 doi 10.1016/j.ccell.2017.08.017.

9. Wu G, Diaz AK, Paugh BS, Rankin SL, Ju B, Li Y*, et al.* The genomic landscape of diffuse intrinsic pontine glioma and pediatric non-brainstem high-grade glioma. Nat Genet **2014**;46(5):444-50 doi 10.1038/ng.2938.

10. Duffner PK, Horowitz ME, Krischer JP, Burger PC, Cohen ME, Sanford RA*, et al.* The treatment of malignant brain tumors in infants and very young children: an update of the Pediatric Oncology Group experience. Neuro Oncol **1999**;1(2):152-61 doi 10.1093/neuonc/1.2.152.

11. Lafay-Cousin L, Strother D. Current treatment approaches for infants with malignant central nervous system tumors. Oncologist **2009**;14(4):433-44 doi 10.1634/theoncologist.2008-0193.

12. Isaacs H, Jr. Perinatal (fetal and neonatal) astrocytoma: a review. Childs Nerv Syst **2016**;32(11):2085-96 doi 10.1007/s00381-016-3215-y.

13. Gelabert-Gonzalez M, Serramito-Garcia R, Arcos-Algaba A. Desmoplastic infantile and non-infantile ganglioglioma. Review of the literature. Neurosurg Rev **2010**;34(2):151-8 doi 10.1007/s10143-010-0303-4.

14. Dufour C, Grill J, Lellouch-Tubiana A, Puget S, Chastagner P, Frappaz D*, et al.* High-grade glioma in children under 5 years of age: a chemotherapy only approach with

the BBSFOP protocol. Eur J Cancer **2006**;42(17):2939-45 doi 10.1016/j.ejca.2006.06.021.

15. Geyer JR, Finlay JL, Boyett JM, Wisoff J, Yates A, Mao L*, et al.* Survival of infants with malignant astrocytomas. A Report from the Childrens Cancer Group. Cancer **1995**;75(4):1045-50.

16. Grundy RG, Wilne SH, Robinson KJ, Ironside JW, Cox T, Chong WK*, et al.* Primary postoperative chemotherapy without radiotherapy for treatment of brain tumours other than ependymoma in children under 3 years: results of the first UKCCSG/SIOP CNS 9204 trial. Eur J Cancer **2010**;46(1):120-33 doi 10.1016/j.ejca.2009.09.013.

17. Guerreiro Stucklin AS, Ryall S, Fukuoka K, Zapotocky M, Lassaletta A, Li C*, et al.* Alterations in ALK/ROS1/NTRK/MET drive a group of infantile hemispheric gliomas. Nat Commun **2019**;10(1):4343 doi 10.1038/s41467-019-12187-5.

18. Macy ME, Birks DK, Barton VN, Chan MH, Donson AM, Kleinschmidt-Demasters BK*, et al.* Clinical and molecular characteristics of congenital glioblastoma. Neuro Oncol **2012**;14(7):931-41 doi 10.1093/neuonc/nos125.

19. Takeshima H, Kawahara Y, Hirano H, Obara S, Niiro M, Kuratsu J. Postoperative regression of desmoplastic infantile gangliogliomas: report of two cases. Neurosurgery **2003**;53(4):979-83; discussion 83-4.

20. Haberler C, Slavc I, Czech T, Prayer D, Pirker C, Budka H*, et al.* Malignant predominantly minigemistocytic glioma in two infants: a distinctive glioma variant? Neuropathol Appl Neurobiol **2007**;33(2):169-78 doi 10.1111/j.1365-2990.2006.00823.x.

21. Gielen GH, Gessi M, Buttarelli FR, Baldi C, Hammes J, zur Muehlen A*, et al.* Genetic Analysis of Diffuse High-Grade Astrocytomas in Infancy Defines a Novel Molecular Entity. Brain Pathol **2015**;25(4):409-17 doi 10.1111/bpa.12210.

22. Anestis DM, Tsitsopoulos PP, Ble CA, Tsitouras V, Tsonidis CA. Congenital Glioblastoma Multiforme: An Unusual and Challenging Tumor. Neuropediatrics **2017**;48(6):403-12 doi 10.1055/s-0037-1601858.

23. Wang AC, Jones DTW, Abecassis IJ, Cole BL, Leary SES, Lockwood CM*, et al.* Desmoplastic Infantile Ganglioglioma/Astrocytoma (DIG/DIA) Are Distinct Entities with Frequent BRAFV600 Mutations. Mol Cancer Res **2018**;16(10):1491-8 doi 10.1158/1541-7786.MCR-17-0507.

24. Amatu A, Sartore-Bianchi A, Siena S. NTRK gene fusions as novel targets of cancer therapy across multiple tumour types. ESMO Open **2016**;1(2):e000023 doi 10.1136/esmoopen-2015-000023.

25. Jones DT, Hutter B, Jager N, Korshunov A, Kool M, Warnatz HJ*, et al.* Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. Nat Genet **2013**;45(8):927-32 doi 10.1038/ng.2682.

26. Zhang J, Wu G, Miller CP, Tatevossian RG, Dalton JD, Tang B*, et al.* Whole-genome sequencing identifies genetic alterations in pediatric low-grade gliomas. Nat Genet **2013**;45(6):602-12 doi 10.1038/ng.2611.

27. Aghajan Y, Levy ML, Malicki DM, Crawford JR. Novel PPP1CB-ALK fusion protein in a high-grade glioma of infancy. BMJ Case Rep **2016**;2016 doi 10.1136/bcr-2016-217189.

28. Chmielecki J, Bailey M, He J, Elvin J, Vergilio JA, Ramkissoon S*, et al.* Genomic Profiling of a Large Set of Diverse Pediatric Cancers Identifies Known and Novel

Mutations across Tumor Spectra. Cancer Res **2017**;77(2):509-19 doi 10.1158/0008-5472.CAN-16-1106.

29. Cocce MC, Mardin BR, Bens S, Stutz AM, Lubieniecki F, Vater I*, et al.* Identification of ZCCHC8 as fusion partner of ROS1 in a case of congenital glioblastoma multiforme with a t(6;12)(q21;q24.3). Genes Chromosomes Cancer **2016**;55(9):677-87 doi 10.1002/gcc.22369.

30. Kiehna EN, Arnush MR, Tamrazi B, Cotter JA, Hawes D, Robison NJ*, et al.* Novel GOPC(FIG)-ROS1 fusion in a pediatric high-grade glioma survivor. J Neurosurg Pediatr **2017**;20(1):51-5 doi 10.3171/2017.2.PEDS16679.

31. Maruggi M, Malicki DM, Levy ML, Crawford JR. A novel KIF5B-ALK fusion in a child with an atypical central nervous system inflammatory myofibroblastic tumour. BMJ Case Rep **2018**;2018 doi 10.1136/bcr-2018-226431.

32. Nakano Y, Tomiyama A, Kohno T, Yoshida A, Yamasaki K, Ozawa T*, et al.* Identification of a novel KLC1-ROS1 fusion in a case of pediatric low-grade localized glioma. Brain Tumor Pathol **2019**;36(1):14-9 doi 10.1007/s10014-018-0330-3.

33. Ng A, Levy ML, Malicki DM, Crawford JR. Unusual high-grade and low-grade glioma in an infant with PPP1CB-ALK gene fusion. BMJ Case Rep **2019**;12(2) doi 10.1136/bcr-2018-228248.

34. Olsen TK, Panagopoulos I, Meling TR, Micci F, Gorunova L, Thorsen J*, et al.* Fusion genes with ALK as recurrent partner in ependymoma-like gliomas: a new brain tumor entity? Neuro Oncol **2015**;17(10):1365-73 doi 10.1093/neuonc/nov039.

35. Ziegler DS, Wong M, Mayoh C, Kumar A, Tsoli M, Mould E*, et al.* Brief Report: Potent clinical and radiological response to larotrectinib in TRK fusion-driven high-grade glioma. Br J Cancer **2018**;119(6):693-6 doi 10.1038/s41416-018-0251-2.

36. Sturm D, Orr BA, Toprak UH, Hovestadt V, Jones DT, Capper D*, et al.* New Brain Tumor Entities Emerge from Molecular Classification of CNS-PNETs. Cell **2016**;164(5):1060-72 doi 10.1016/j.cell.2016.01.015.

37. Koelsche C, Mynarek M, Schrimpf D, Bertero L, Serrano J, Sahm F*, et al.* Primary intracranial spindle cell sarcoma with rhabdomyosarcoma-like features share a highly distinct methylation profile and DICER1 mutations. Acta Neuropathol **2018**;136(2):327-37 doi 10.1007/s00401-018-1871-6.

38. Johannessen AL, Torp SH. The clinical value of Ki-67/MIB-1 labeling index in human astrocytomas. Pathol Oncol Res **2006**;12(3):143-7 doi PAOR.2006.12.3.0143.

39. El-Ayadi M, Ansari M, Sturm D, Gielen GH, Warmuth-Metz M, Kramm CM*, et al.* High-grade glioma in very young children: a rare and particular patient population. Oncotarget **2017**;8(38):64564-78 doi 10.18632/oncotarget.18478.

40. Capper D, Jones DTW, Sill M, Hovestadt V, Schrimpf D, Sturm D*, et al.* DNA methylation-based classification of central nervous system tumours. Nature **2018**;555(7697):469-74 doi 10.1038/nature26000.

41. van Tilburg CM, Selt F, Sahm F, Bachli H, Pfister SM, Witt O*, et al.* Response in a child with a BRAF V600E mutated desmoplastic infantile astrocytoma upon retreatment with vemurafenib. Pediatr Blood Cancer **2018**;65(3) doi 10.1002/pbc.26893.

42. Charest A, Lane K, McMahon K, Park J, Preisinger E, Conroy H*, et al.* Fusion of FIG to the receptor tyrosine kinase ROS in a glioblastoma with an interstitial

del(6)(q21q21). Genes Chromosomes Cancer **2003**;37(1):58-71 doi 10.1002/gcc.10207.

43. International Cancer Genome Consortium PedBrain Tumor P. Recurrent MET fusion genes represent a drug target in pediatric glioblastoma. Nat Med **2016**;22(11):1314-20 doi 10.1038/nm.4204.

44. Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, Ishikawa S*, et al.* Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature **2007**;448(7153):561-6 doi 10.1038/nature05945.

45. Qaddoumi I, Orisme W, Wen J, Santiago T, Gupta K, Dalton JD*, et al.* Genetic alterations in uncommon low-grade neuroepithelial tumors: BRAF, FGFR1, and MYB mutations occur at high frequency and align with morphology. Acta Neuropathol **2016**;131(6):833-45 doi 10.1007/s00401-016-1539-z.

46. Singh D, Chan JM, Zoppoli P, Niola F, Sullivan R, Castano A*, et al.* Transforming fusions of FGFR and TACC genes in human glioblastoma. Science **2012**;337(6099):1231-5 doi 10.1126/science.1220834.

47. Davis JL, Lockwood CM, Stohr B, Boecking C, Al-Ibraheemi A, DuBois SG*, et al.* Expanding the Spectrum of Pediatric NTRK-rearranged Mesenchymal Tumors. Am J Surg Pathol **2018** doi 10.1097/PAS.0000000000001203.

48. Farago AF, Azzoli CG. Beyond ALK and ROS1: RET, NTRK, EGFR and BRAF gene rearrangements in non-small cell lung cancer. Transl Lung Cancer Res **2017**;6(5):550-9 doi 10.21037/tlcr.2017.08.02.

49. Pietrantonio F, Di Nicolantonio F, Schrock AB, Lee J, Tejpar S, Sartore-Bianchi A*, et al.* ALK, ROS1, and NTRK Rearrangements in Metastatic Colorectal Cancer. J Natl Cancer Inst **2017**;109(12) doi 10.1093/jnci/djx089.

50. Davare MA, Henderson JJ, Agarwal A, Wagner JP, Iyer SR, Shah N*, et al.* Rare but Recurrent ROS1 Fusions Resulting From Chromosome 6q22 Microdeletions are Targetable Oncogenes in Glioma. Clin Cancer Res **2018**;24(24):6471-82 doi 10.1158/1078-0432.CCR-18-1052.

51. Gatalica Z, Xiu J, Swensen J, Vranic S. Molecular characterization of cancers with NTRK gene fusions. Mod Pathol **2019**;32(1):147-53 doi 10.1038/s41379-018-0118-3.

52. Johnson A, Severson E, Gay L, Vergilio JA, Elvin J, Suh J*, et al.* Comprehensive Genomic Profiling of 282 Pediatric Low- and High-Grade Gliomas Reveals Genomic Drivers, Tumor Mutational Burden, and Hypermutation Signatures. Oncologist **2017**;22(12):1478-90 doi 10.1634/theoncologist.2017-0242.

53. Okamura R, Boichard A, Kato S, Sicklick JK, Bazhenova L, Kurzrock R. Analysis of NTRK Alterations in Pan-Cancer Adult and Pediatric Malignancies: Implications for NTRK-Targeted Therapeutics. JCO Precis Oncol **2018**;2018 doi 10.1200/PO.18.00183.

54. McArthur GA, Laherty CD, Queva C, Hurlin PJ, Loo L, James L*, et al.* The Mad protein family links transcriptional repression to cell differentiation. Cold Spring Harb Symp Quant Biol **1998**;63:423-33.

55. Gripp KW, Aldinger KA, Bennett JT, Baker L, Tusi J, Powell-Hamilton N*, et al.* A novel rasopathy caused by recurrent de novo missense mutations in PPP1CB closely resembles Noonan syndrome with loose anagen hair. Am J Med Genet A **2016**;170(9):2237-47 doi 10.1002/ajmg.a.37781.

56. Dehmelt L, Halpain S. The MAP2/Tau family of microtubule-associated proteins. Genome Biol **2005**;6(1):204 doi 10.1186/gb-2004-6-1-204.

57. Saadi I, Alkuraya FS, Gisselbrecht SS, Goessling W, Cavallesco R, Turbe-Doan A*, et al.* Deficiency of the cytoskeletal protein SPECC1L leads to oblique facial clefting. Am J Hum Genet **2011**;89(1):44-55 doi 10.1016/j.ajhg.2011.05.023.

58. Wilson NR, Olm-Shipman AJ, Acevedo DS, Palaniyandi K, Hall EG, Kosa E*, et al.* SPECC1L deficiency results in increased adherens junction stability and reduced cranial neural crest cell delamination. Sci Rep **2016**;6:17735 doi 10.1038/srep17735.

59. Kalashnikova E, Lorca RA, Kaur I, Barisone GA, Li B, Ishimaru T*, et al.* SynDIG1: an activity-regulated, AMPA- receptor-interacting transmembrane protein that regulates excitatory synapse development. Neuron **2010**;65(1):80-93 doi 10.1016/j.neuron.2009.12.021.

60. Metzler M, Li B, Gan L, Georgiou J, Gutekunst CA, Wang Y*, et al.* Disruption of the endocytic protein HIP1 results in neurological deficits and decreased AMPA receptor trafficking. EMBO J **2003**;22(13):3254-66 doi 10.1093/emboj/cdg334.

61. Nakai T, Nagai T, Tanaka M, Itoh N, Asai N, Enomoto A*, et al.* Girdin phosphorylation is crucial for synaptic plasticity and memory: a potential role in the interaction of BDNF/TrkB/Akt signaling with NMDA receptor. J Neurosci **2014**;34(45):14995-5008 doi 10.1523/JNEUROSCI.2228-14.2014.

62. Enomoto A, Murakami H, Asai N, Morone N, Watanabe T, Kawai K*, et al.* Akt/PKB regulates actin organization and cell motility via Girdin/APE. Dev Cell **2005**;9(3):389-402 doi 10.1016/j.devcel.2005.08.001.

63. Ren X, Kuan PF. methylGSA: a Bioconductor package and Shiny app for DNA methylation data length bias adjustment in gene set testing. Bioinformatics **2019**;35(11):1958-9 doi 10.1093/bioinformatics/bty892.

64. Mackay A, Burford A, Molinari V, Jones DTW, Izquierdo E, Brouwer-Visser J*, et al.* Molecular, Pathological, Radiological, and Immune Profiling of Non-brainstem Pediatric High-Grade Glioma from the HERBY Phase II Randomized Trial. Cancer Cell **2018**;33(5):829-42 e5 doi 10.1016/j.ccell.2018.04.004.

65. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK*, et al.* The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta Neuropathol **2016**;131(6):803-20 doi 10.1007/s00401-016-1545-1.

66. Infarinato NR, Park JH, Krytska K, Ryles HT, Sano R, Szigety KM*, et al.* The ALK/ROS1 Inhibitor PF-06463922 Overcomes Primary Resistance to Crizotinib in ALK-Driven Neuroblastoma. Cancer Discov **2016**;6(1):96-107 doi 10.1158/2159-8290.CD-15-1056.

67. Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. PLoS Biol **2010**;8(6):e1000412 doi 10.1371/journal.pbio.1000412.

68. Workman P, Aboagye EO, Balkwill F, Balmain A, Bruder G, Chaplin DJ*, et al.* Guidelines for the welfare and use of animals in cancer research. Br J Cancer **2010**;102(11):1555-77 doi 10.1038/sj.bjc.6605642.

Legends for Figures

Figure 1 – Defining an intrinsic set of infant gliomas. (A) Flow diagram providing an overview of the inclusion and exclusion criteria for the assembled cohort of 241 samples from patients under the age of 4 years. (B) Fusion gene analysis by a variety of means allowed for the identification of 28 fusions marking clearly defined entities that were subsequently excluded from further analysis. (C) Methylation array profiling and analysis by the Heidelberg classifier excluded a further 12 cases closely resembling non-glioma entities or failing quality control (n=9). (D) t-statistic based stochastic neighbor embedding (tSNE) projection of the remaining cases highlighted 61 samples which clustered with previously reported high or low grade glioma subtypes, leaving an intrinsic set of 130 infant gliomas for further characterisation by more histopathological assessment and in-depth sequencing. (E) Anatomical location of infant gliomas after exclusion of pathognomonic fusions and non-glioma entities by methylation profiling (n=130). Left – sagittal section showing internal structures; right – external view highlighting cerebral lobes. Each circle represents a single case and is colored by the glioma subgroup it most closely clusters with, defined by the key below. (F) Kaplan-Meier plot of overall survival of cases separated by methylation subgroups DIGG (desmoplastic infantile ganglioglioma / astrocytoma), IHG (infant hemispheric glioma), LGG (other low grade glioma subgroups) and HGG (other high grade glioma subgroups) (n=102). P value is calculated by the log-rank test (p=0.0566 for HGG *versus* rest). (G) t-statistic based stochastic neighbor embedding (t-SNE) projection of a combined methylation dataset comprising the intrinsic set of the present study (n=130, circled) plus a reference set of glioma subtypes (n=1652). The first two projections are plotted on the x and y axes, with samples represented by dots colored by subtype according to the key provided.

Figure 2 – Mutations in infant gliomas. (A) Oncoprint representation of an integrated annotation of single nucleotide variants, DNA copy number changes and structural variants for infant gliomas excluded as other subgroups (n=24). (B) Oncoprint representation of an integrated annotation of single nucleotide variants, DNA copy number changes and structural variants for infant gliomas in the intrinsic set (n=41). Samples are arranged in columns with genes labelled along rows. Clinicopathological and molecular annotations are provided as bars according to the included key.

Figure 3 - Copy number-associated fusion genes in infant gliomas. (A) Segmented DNA copy number heatmap for *ALK* breakpoint cases, plotted according to chromosomal location. Pink, gain; blue, loss. (B) Segmented DNA copy number heatmap for *ROS1* breakpoint cases, plotted according to chromosomal location. Pink, gain; blue, loss. (C) Segmented DNA copy number heatmap for *MET* breakpoint cases, plotted according to chromosomal location. Pink, gain; blue, loss. (D) *ETV6:NTRK3*. Cartoon representation of the fusion structure, with reads on either side of the breakpoint colored by gene partner and taken from an Integrated Genome Viewer (IGV) snapshot. Below this is a Sanger sequencing trace spanning the breakpoint. Underneath are copy number plots (log2 ratio, y axis) for chromosomal regions spanning the breakpoints (x axis). Points are colored red for copy number gain, blue for loss, and grey for no change. The smoothed values are overlaid by the purple line. (E) *ZC3H7A:ALK*. Cartoon representation of the fusion structure, with reads on either side of the breakpoint colored by gene partner and taken from an Integrated Genome Viewer snapshot. Below this is a Sanger sequencing trace spanning the breakpoint. Underneath are copy number plots (log2 ratio, y axis) for chromosomal regions spanning the breakpoints (x axis). Points are colored red for copy number gain, blue for loss, and grey for no change. The smoothed values are overlaid by the purple line. (F) Circos plot of gene fusions targeting *NTRK1* (light

orange), *NTRK2* (orange) and *NTRK3* (dark orange). Lines link fusion gene partners according to chromosomal location, represented by ideograms arranged around the circle. (G) Circos plot of gene fusions targeting *ALK* (dark blue). Lines link fusion gene partners according to chromosomal location, represented by ideograms arranged around the circle. (H) Kaplan-Meier plot of overall survival of cases separated by fusion event (n=63). P value is calculated by the log-rank test (p=0.085 for any fusion *versus* None).

Figure 4 – Epigenetic alterations in fusion-positive and -negative infant gliomas. (A) Differential methylation-based gene ontology analysis for *ALK*-fusion cases, represented in barplots of -log10 p value for labelled highest scoring categories (top) and aggregated ontology networks (bottom). (B) Differential methylation-based gene ontology analysis for *NTRK*-fusion cases, represented in barplots of -log10 p value for labelled highest scoring categories (top) and aggregated ontology networks (bottom). (C) Differential methylation-based gene ontology analysis for fusion-negative cases, represented in barplots of -log10 p value for labelled highest scoring categories (top) and aggregated ontology networks (bottom). Node size is proportional to the number of genes, shading represents -log10 p value (darker is higher). Thickness of connecting lines reflects the percentage of overlapping genes. (D) Genome browser view of the *WNT5A* locus, with lower methylation, provided as barplots, in selected *ALK*-fusion (blue) cases compared to *NTRK*-fusion (orange) and fusion-negative (grey) cases. (E) Genome browser view of the *STAT1* locus, with lower methylation, provided as barplots, in selected *NTRK*-fusion (orange) cases compared to *ALK*-fusion (blue) and fusion-negative (grey) cases. (F) Genome browser view of the *TP63* locus, with lower methylation, provided as barplots, in selected fusion-negative (grey) cases compared to *ALK*-fusion (blue) and *NTRK*-fusion (orange) cases. Chromosomal ideograms are provided with the red bar indicating the cytoband in which the locus is found. Differentially methylated probes

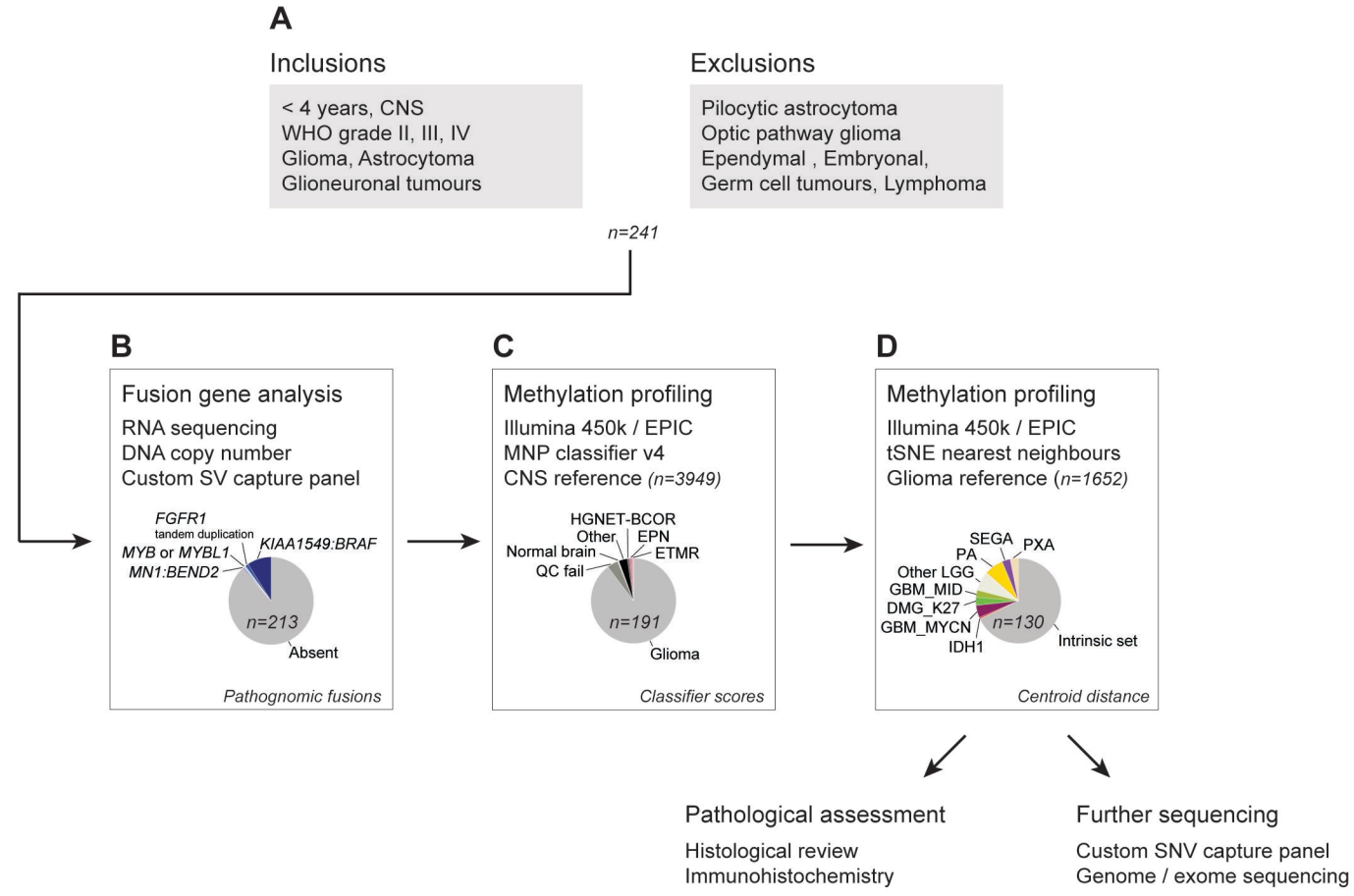
are highlighted by the red box. (G) Immunofluorescent staining of an antibody directed against WNT5A (white) in an *EML4:ALK* fusion infant glioma case, UOLP_INF_001. DAPI is used as a counterstain. Scale bar = 200µM. (H) Immunofluorescent staining of an antibody directed against WNT5A (green) in an *ETV6:NTRK3* fusion infant glioma case, GOSH_INF_007. DAPI is used as a counterstain. Scale bar = 200µM. (I) Heatmap representing gene expression values from a NanoString assay of 30 most differentially methylated genes between *ALK*-fusion (blue), *NTRK*-fusion (orange) and fusion-negative (grey) cases. Expression values are colored according to the scale provided.

Figure 5 - Pre-clinical modelling of *ALK*-fused glioma. (A) Schematic representation of the *in vivo* modelling workflow. IUE, *in utero* electroporation; KD, kinase domain. (B) Kaplan-Meier curve of injected animals using IUE and p0-RCAS method – *PPP1CB:ALK* only IUE, *PPP1CB:ALK* + *Trp53*-ko IUE, *PPP1CB:ALK* + *Cdkn2a*-ko IUE and *PPP1CB:ALK* p0-RCAS only. *, $p<0.05$, **, $p<0.01$ (C, D) Effect of targeted ALK inhibition on growth of allografted *PPP1CB:ALK* + *Cdkn2a*-ko mouse tumor cells *in vivo*. p.i., post injection. (E) Targeted inhibition significantly prolonged the survival of *PPP1CB:ALK* + *Cdkn2a*-ko allografted mice compared with temozolomide or vehicle controls. Two mice in the lorlatinib group were sacrificed due to technical complications with drug delivery, with no tumor being evident upon dissection of the brain. ***, $p<0.001$. (F) Clinical history of DKFZ_INF_307, with confirmed *MAD1L1:ALK* fusion. Timeline of clinical interventions is provided below, with treatment shaded in grey. Axial T2 MRI scans from diagnosis and successive surgeries and chemotherapeutic regimens are provided, in addition to treatment with the ALK inhibitor ceritinib, with tumor circled in red.

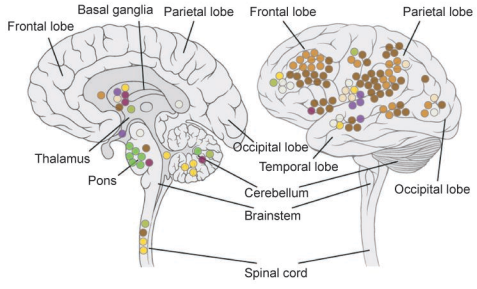
Figure 6 – Preclinical and clinical experience with Trk inhibitors in fusion-positive infant glioma. (A) Light microscopy image of two patient-derived infant glioma cell cultures, harboring either *TPM3:NTRK1* (QCTB-R102, light orange) or *ETV6:NTRK3* (QCTB-R077, dark red) fusions. (B) Concentration-response curves for three Trk inhibitors tested against two *NTRK* fusion-positive infant glioma cell cultures (QCTB-R102, *TPM3:NTRK1,* light orange; QCTB-R077, *ETV6:NTRK3*, dark red) and two fusion-negative glioma cultures (QCTB-R006, light grey; QCTB-R059, dark grey). Concentration of compound is plotted on a log scale (x axis) against cell viability (y axis). Mean plus standard error are plotted from at least n=3 experiments. (C) Clinical history of OPBG_INF_035, with confirmed *ETV6-NTRK3* fusion. Timeline of clinical interventions is provided below, with Trk inhibitor treatment shaded in grey. Diagnosis, post-biopsy, pre/post-surgery, post-crizotinib and post-larotrectinib axial T2 MRI scans are provided, with tumor circled in red. (D) Clinical history of MSKC_INF_006, with confirmed *ETV6:NTRK3* fusion. Timeline of clinical interventions is provided below, with Trk inhibitor treatment shaded in grey. Diagnosis and post-larotrectinib post-contrast axial T1 MRI scans are provided, with tumor circled in red.

Figure 7 – Summary of infant HGG subgroups.

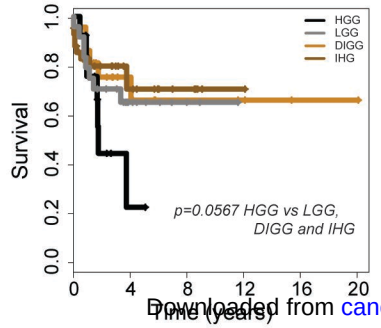
Figure 1



E



F



G

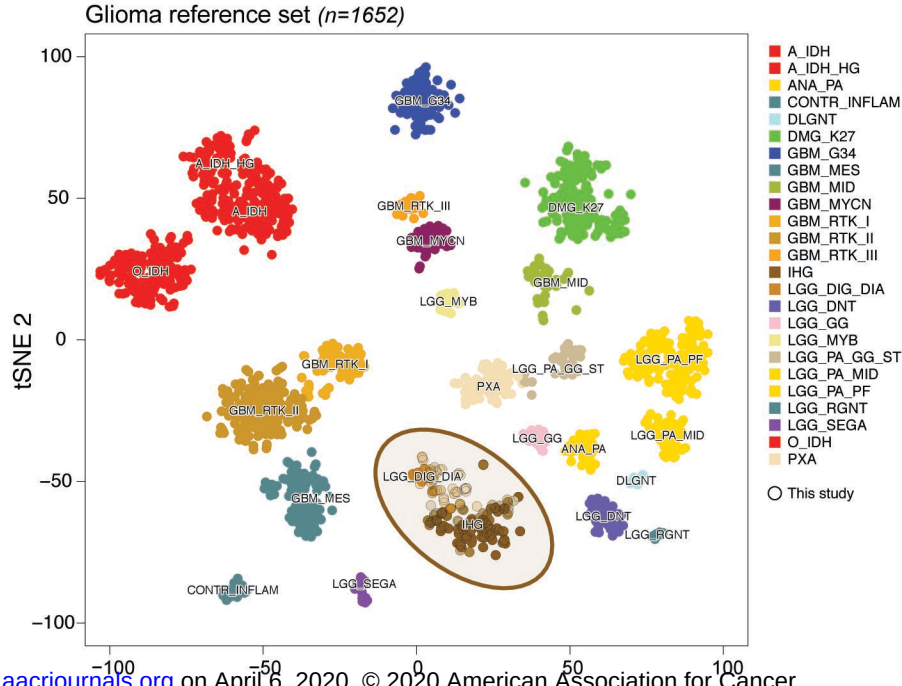


Figure 2

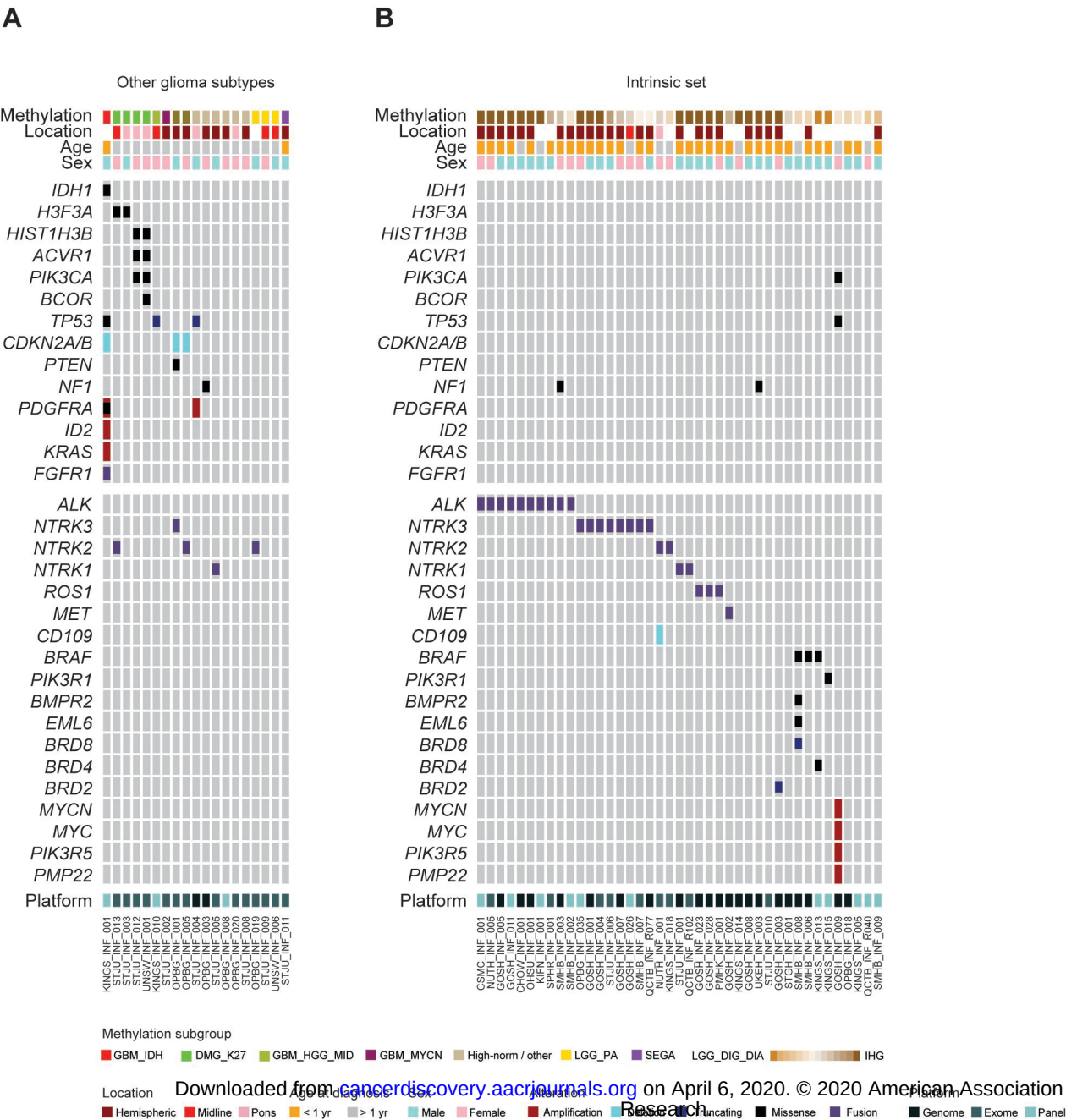


Figure 3

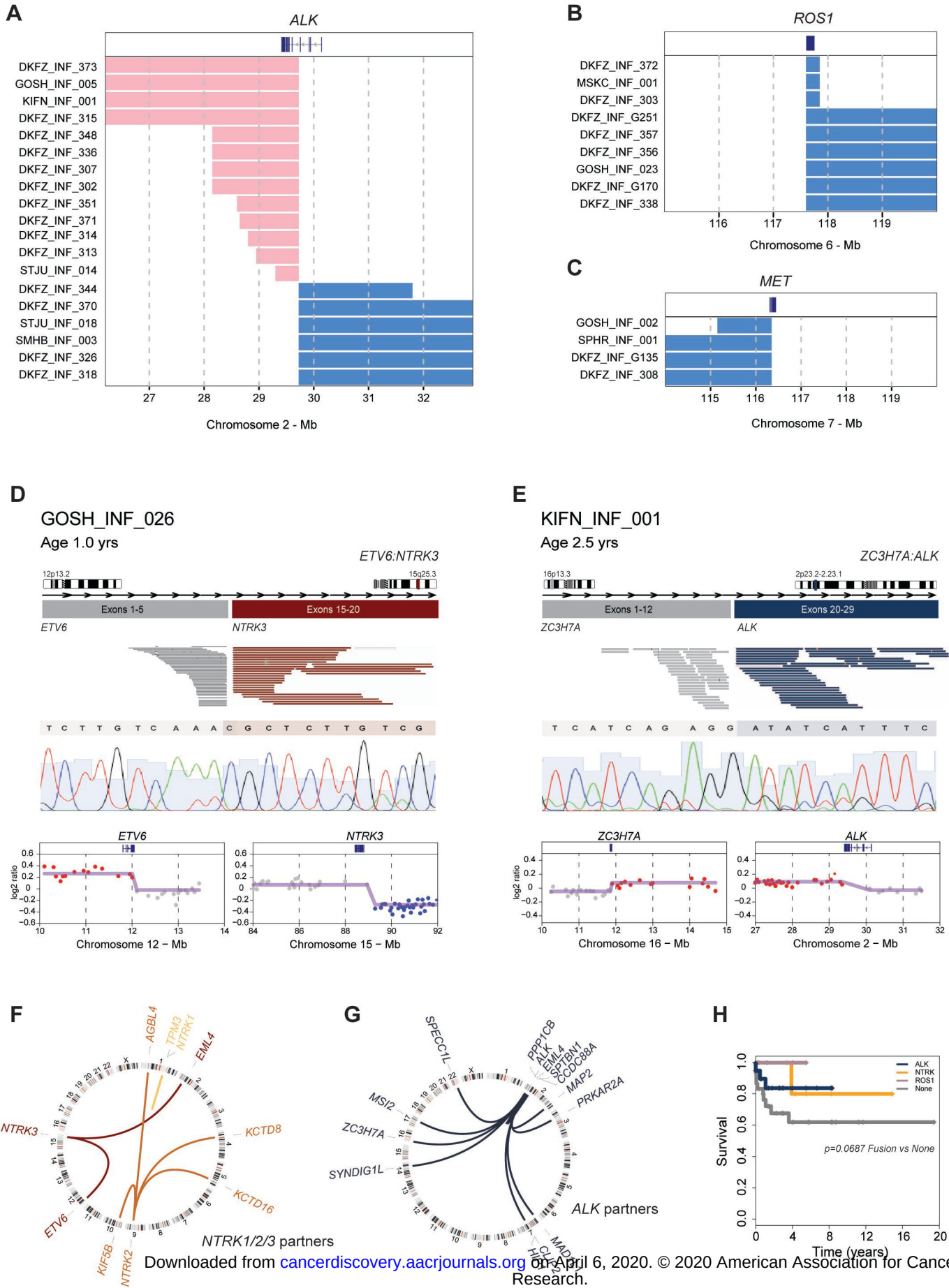


Figure 4

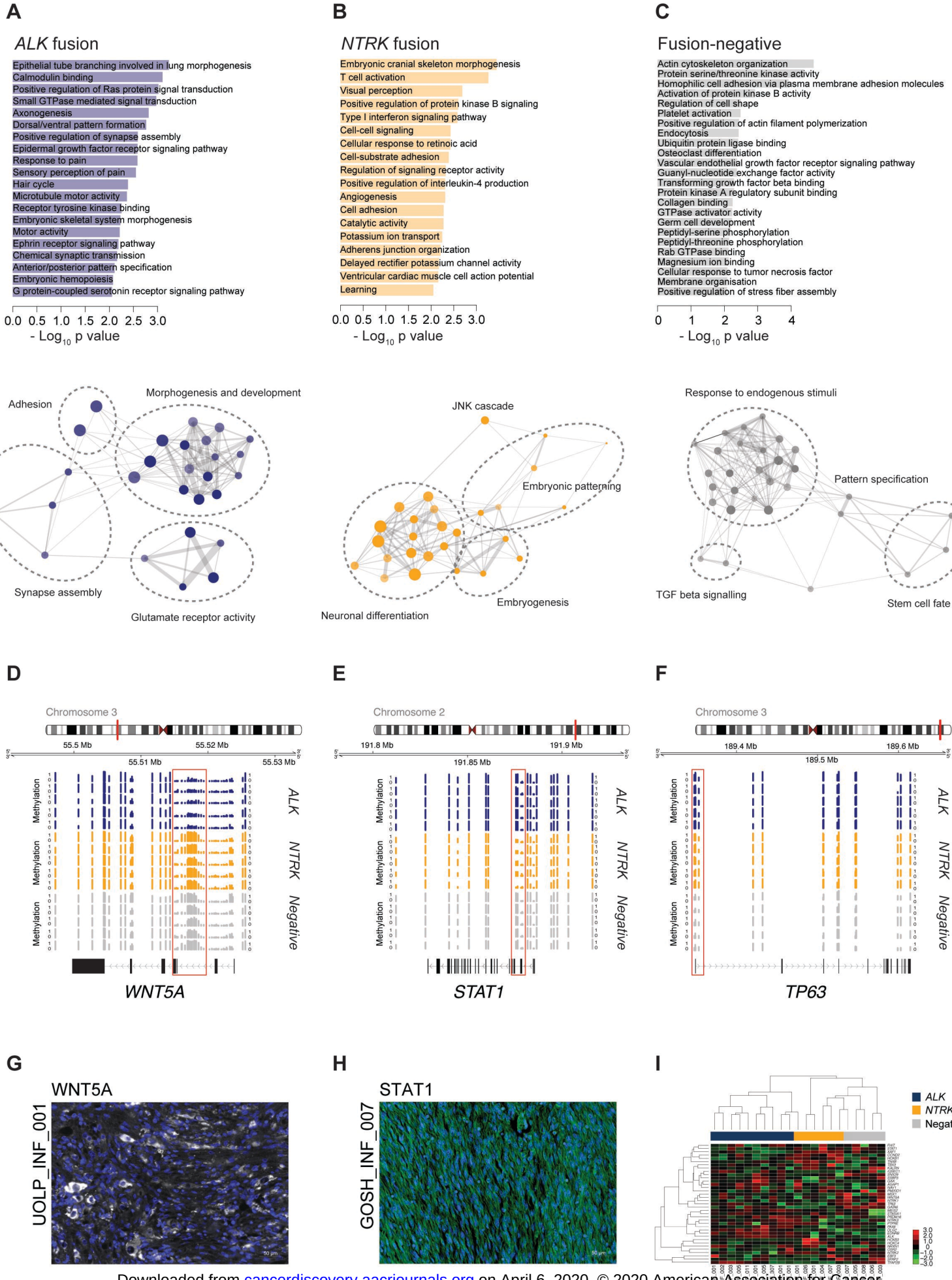


Figure 5

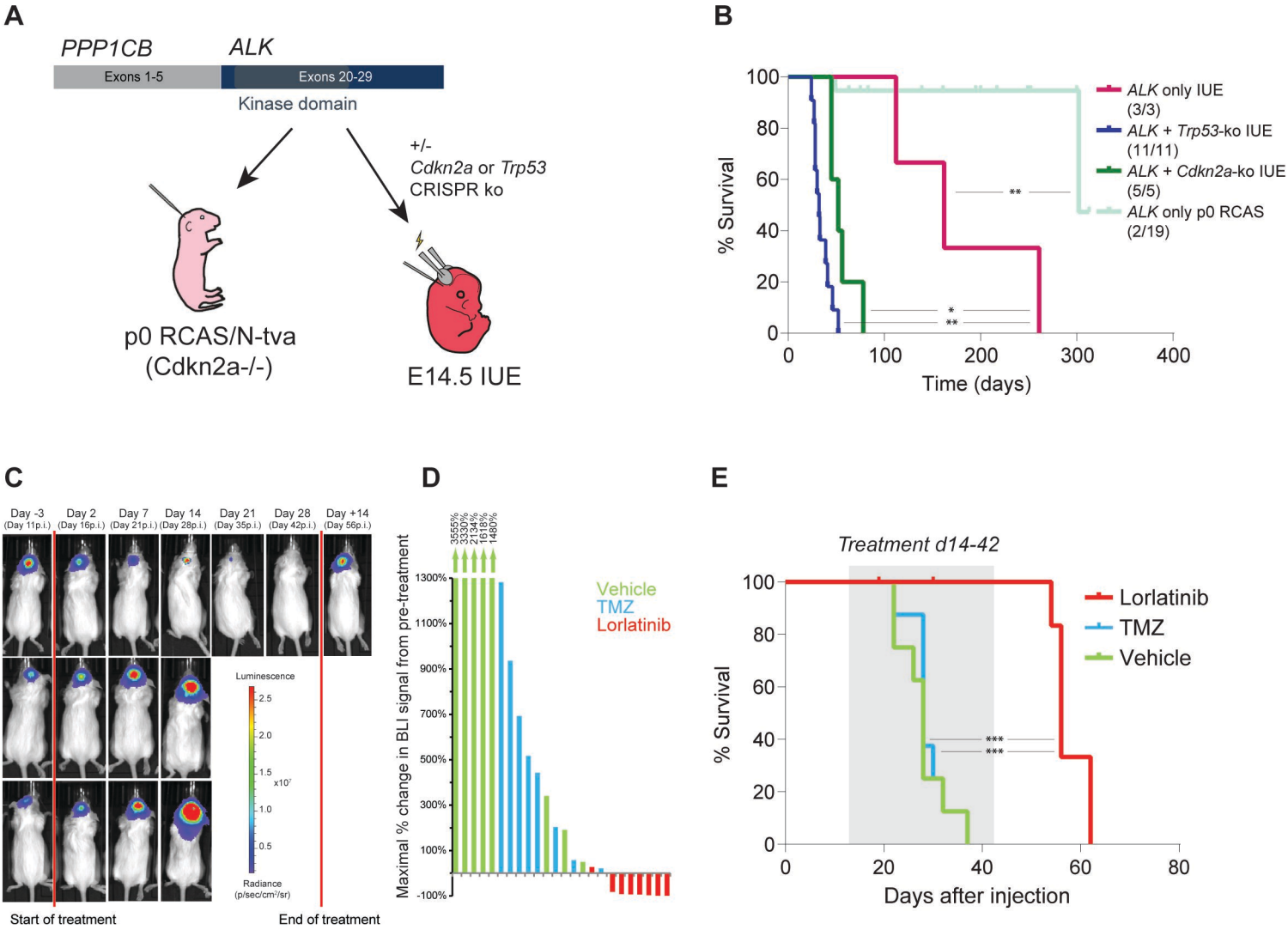


Figure 6

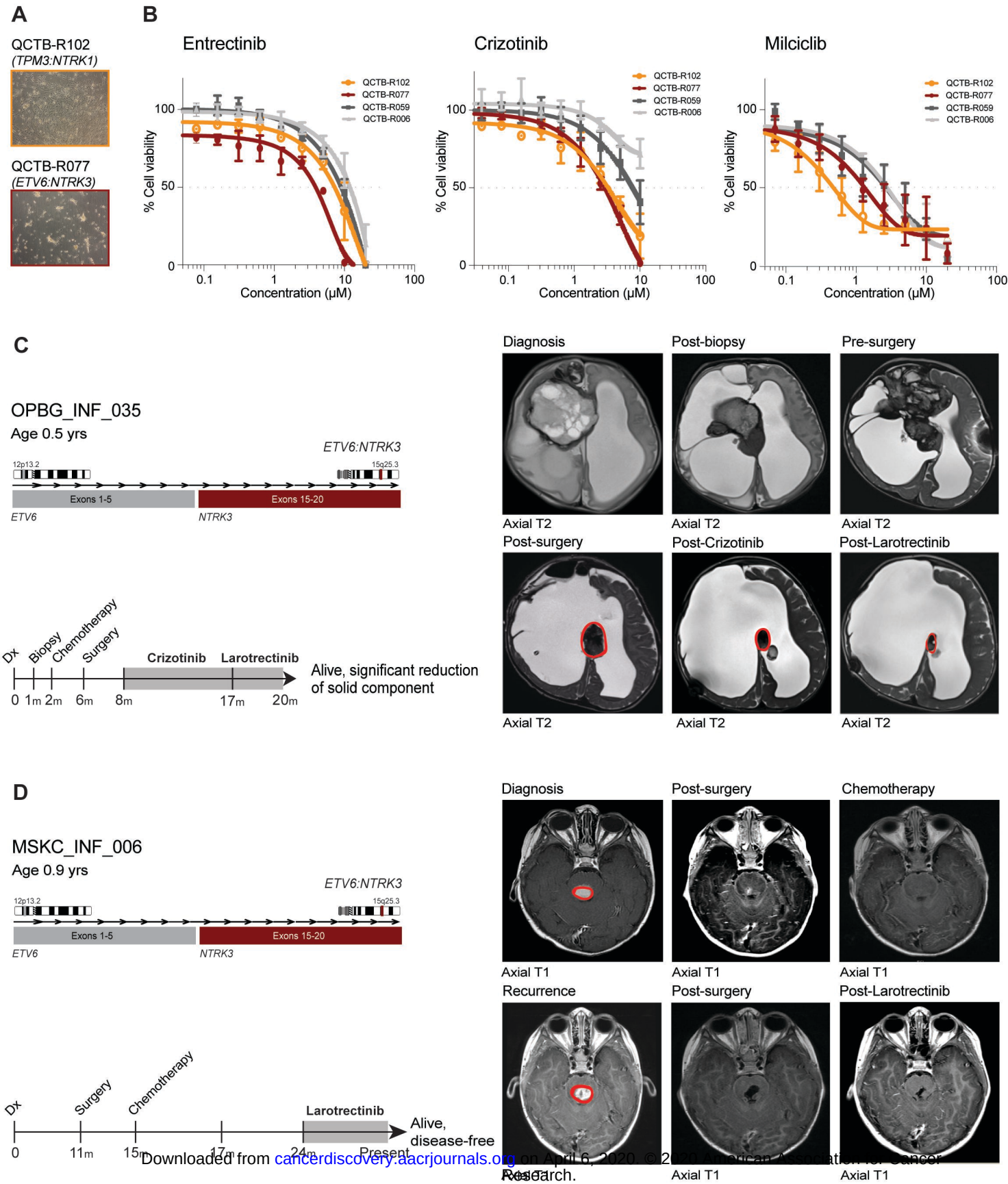
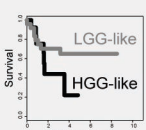
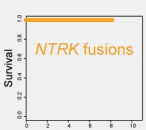
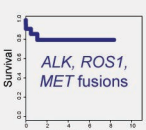
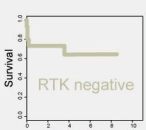


Figure 7

High grade glioma in the infant population

All locations
> 1 year

Hemispheric
< 1 year

Other glioma subgroups	*DIGG/DIA & DIGG/DIA-like*	*Intrinsic set Fusion positive*		*Intrinsic set Fusion negative*
Subgroup-specific additional mutations	*BRAF* V600E Low mutational burden	Few mutations	Few mutations	Few mutations
			ALK	
			ROS1	
		MET		
		NTRK1		
NTRK2				
		NTRK3		
FGFR1				
		Differential methylation *STAT1*, JNK	Differential methylation *WNT5A*, GluRs	Differential methylation *TP63*, TGF beta
		Neurogenesis	Synaptic plasticity	Stem cell pluripotency
				
		Spindle/gemistrocytic Limited mitoses	Spindle/gemistrocytic Limited mitoses	Spindle/gemistocytic