Dissertation

submitted to the

Combined Faculty of Natural Sciences and Mathematics

of the Ruperto Carola University Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Presented by

M.Sc. Alexander Constantin Sommerkamp (né Jarke)

born in: Bonn

Oral examination: 17.09.2020

Molecular comparison, preclinical modeling and improved diagnostics of pediatric low-grade gliomas

Referees:

Prof. Dr. Peter Angel

Prof. Dr. Stefan Pfister

Für Pia

*"Science makes people reach selflessly for truth and objectivity;
it teaches people to accept reality, with wonder and admiration"*

Lise Meitner

Summary

Project 1: Comparison of human & murine PA/PXA characteristics

Pediatric low-grade gliomas (pLGGs) are the most common brain tumors in this age group. Despite their favorable prognosis, the young patients often have to suffer from long-term sequalae of the (repeated) therapy or the tumor itself for the rest of their lives. Recent large-scale sequencing studies have advanced the molecular characterization of pLGGs to a new level and are driving the development of a tumor classification based on molecular features rather than primarily histology. However, this is hampered by the enormous heterogeneity of pLGGs, and even the separation of established tumor types is typically not acknowledged, for example in clinical trial stratification. Pilocytic astrocytoma (PA; WHO grade I) and pleomorphic xanthoastrocytoma (PXA; WHO grade II) are two of these pLGG types. They can be difficult to distinguish based on histology alone, but PXA tumors show a clearly worse clinical course than PA tumors and are in many cases lethal. Nevertheless, in clinical trials, both are often simply referred to as "pLGG" and thus treated in the same way. In the first part of this dissertation, I therefore investigated and compared the molecular characteristics of PA and PXA in more detail.

Using molecular data from 89 human pediatric tumor samples, I identified considerable differences between the methylome and transcriptome profiles of PA and PXA. The differentially expressed genes between both tumor types were highly enriched for cell cycle and developmental processes. My results confirm the distinctively more proliferative nature of PXA and suggest fundamental differences in the regulatory circuits that are implicated in tumor development and growth. While most PXA samples in the analysis cohort harbored the typical BRAF V600E mutation, I also identified *NTRK* fusions as a previously underappreciated genetic driver event for this tumor type and, in a separate case, discovered an *EGFR:BRAF* fusion for the first time. In addition, some tumors apparently retained expression of the *CDKN2A/B* locus, which is typically deleted in PXA. Given the important role of the tumor microenvironment, I extracted information on the immune cell content from the bulk sequencing data of PA and PXA. While both tumor types displayed similarly strong signs of overall immune cell infiltration (about 20%), I found a distinctive upregulation of a CD8 T cell gene signature in PXA tumors, which might represent a vulnerability for immunotherapeutic intervention.

In addition, I further developed a genetic mouse model for PXA tumors on the basis of previous work in the group. It relies on viral introduction of the human gene sequence for the BRAF V600E oncogene into neural progenitor cells of *Cdkn2a* knockout mice and nicely complements an existing PA mouse model. I demonstrate that tumors from both models faithfully resemble their human counterparts at the levels of growth behavior, histology, gene

expression and immune cell infiltration. Strikingly, in contrast to the PA model, the murine PXA-like tumors were lethal for the mice, again indicating that the designation of this tumor type as "low-grade" might be inappropriate and should be reconsidered. In addition, these tumors showed the same characteristic CD8 T cell signature that I had observed in human PXA. My work thus paves the way for further molecular and preclinical research on these pLGG entities and suggests that the potential use of immunotherapy should not be rejected prematurely. (Sommerkamp et al., in preparation)

Project 2: KIAA1549:BRAF fusion detection from RNA-Seq data

The most common genetic alteration observed in human PA is the *KIAA1549:BRAF* fusion. It shows high specificity for this tumor type and is therefore of great diagnostic and prognostic value. In addition, it constitutes a vulnerability for targeted therapy. Thus, it is of major importance to reliably detect the *KIAA1549:BRAF* fusion in a clinical setting. RNA sequencing (RNA-Seq) is now widely used in research, and has also become increasingly popular for diagnostic purposes. In addition to providing gene- and isoform-specific expression data, RNA-Seq can be used to identify expressed fusion genes in an unbiased (i.e. not pre-selected) way. The *KIAA1549:BRAF* fusion, however, seems to be expressed at low levels and has previously proven difficult to detect by RNA-Seq. In the second part of this dissertation, I therefore examined the detection reliability of this fusion and developed an optimized workflow for detection from RNA-Seq data.

By analyzing the RNA-Seq data from known fusion-positive PA tumor samples in my analysis cohort, I found that deep sequencing alone is not sufficient to reliably detect the fusion. Instead, the detectability is influenced by different factors, including RNA integrity / library size, tumor cell content, *KIAA1549* expression levels as well as the library preparation protocol. Strikingly, all samples in which the fusion was initially not identified did harbor supporting reads in their raw data that could in principle provide a basis for fusion detection. I identified bottlenecks in standard workflows of the alignment algorithm STAR and the fusion caller Arriba, and show that proper alignment of split reads is a major hurdle for reliable detection of the *KIAA1549:BRAF* fusion. An optimized workflow based on adjusted STAR parameters and a new version of Arriba significantly improved confidence of already detected fusions and allowed identification of fusions that had previously been missed. In an independent diagnostic cohort from Montreal, I demonstrate that this workflow considerably outperforms the previously used standard analysis tools. Moreover, I prove that the higher detection sensitivity is not accompanied by the identification of false positive fusions in an RNA-Seq dataset of > 1000 formalin-fixed, paraffin-embedded (FFPE) samples of diverse origin.

In summary, my work provides a novel workflow that substantially improves the detection of this important fusion from RNA-Seq data, and which will most likely also result in increased fusion detection performance in other tumor contexts. (Sommerkamp et al., 2020)

Zusammenfassung

Projekt 1: Vergleich humaner und muriner PA/PXA Charakteristika

Pädiatrische niedriggradige Gliome (pLGG) sind die häufigsten Hirntumore in dieser Altersgruppe. Trotz ihrer guten Prognose müssen die jungen Patienten oft ihr Leben lang unter den Spätfolgen der (wiederholten) Therapie oder des Tumors leiden. Neue groß angelegte Sequenzierstudien haben die molekulare Charakterisierung von pLGG deutlich verbessert und treiben nun die Entwicklung einer Tumorklassifikation voran, die auf molekularen Eigenschaften statt primär auf Histologie basiert. Dies wird allerdings durch die starke Heterogenität von pLGG erschwert und selbst die Unterscheidung etablierter Tumortypen wird, zum Beispiel bei der Stratifizierung für klinische Studien, typischerweise nicht anerkannt. Das pilozytische Astrozytom (PA; WHO Grad I) und das pleomorphe Xanthoastrozytom (PXA; WHO Grad II) sind zwei dieser pLGG Typen. Beide sind histologisch nicht immer leicht zu unterscheiden, jedoch zeigen PXA Tumore einen deutlich schlechteren klinischen Verlauf als PA Tumore und sind in vielen Fällen tödlich. Dennoch werden beide in klinischen Studien oft einfach als „pLGG" zusammengefasst und somit in der gleichen Art und Weise behandelt. Im ersten Teil dieser Dissertation habe ich daher die molekularen Charakteristika von PA und PXA eingehend untersucht und vergleichen.

Auf der Grundlage molekularer Daten von 89 humanen pädiatrischen Tumorproben habe ich erhebliche Unterschiede in den Methylom- und Transkriptom-Profilen von PA und PXA identifiziert. Die differenziell exprimierten Gene zwischen beiden Tumortypen zeigten eine Überrepräsentation von Zellzyklus- und Entwicklungs-Prozessen. Meine Ergebnisse bestätigen die unverkennbar stärker proliferative Natur des PXA und deuten auf fundamentale Unterschiede in den regulatorischen Schaltkreisen hin, die an Tumorentstehung und -wachstum beteiligt sind. Während die meisten PXA Proben in der Analysekohorte die typische BRAF V600E Mutation aufwiesen, konnte ich zusätzlich *NTRK* Fusionen als bisher unterschätzte genetische Veränderungen in diesem Tumortyp identifizieren und in einem weiteren Fall erstmals eine *EGFR:BRAF* Fusion nachweisen. Einige Tumore behielten zudem scheinbar die Expression des *CDKN2A/B* Lokus bei, der im PXA typischerweise deletiert wird. Da die Mikroumgebung eine wichtige Rolle in Tumoren spielt, habe ich zudem Informationen zum Immunzellgehalt aus den Sequenzierdaten von PA und PXA extrahiert. Während beide Tumortypen gleich starke Zeichen von allgemeiner Immunzellinfiltration zeigten (ca. 20%), konnte ich eine markante Hochregulation einer CD8 T Zell Gensignatur in PXA Tumoren identifizieren, die möglicherweise eine Anfälligkeit für immuntherapeutische Intervention darstellt.

Darüber hinaus habe ich auf der Grundlage früherer Arbeiten der Gruppe ein genetisches Mausmodell für PXA Tumore weiterentwickelt. Es basiert darauf, die humane Gensequenz für das BRAF V600E Onkogen mittels eines Virus in neurale Vorläuferzellen von *Cdkn2a* Knockout Mäusen einzubringen, und ergänzt so optimal ein bereits existierendes PA Mausmodell. Ich zeige, dass die Tumore beider Modelle ihren humanen Pendants in Bezug auf ihr Wachstumsverhalten, ihre Histologie, ihre Genexpression und ihre Immunzellinfiltration stark ähneln. Dabei fiel besonders auf, dass die murinen PXA-ähnlichen Tumore im Gegensatz zum PA Modell tödlich für die Mäuse waren, was erneut darauf hinweist, dass die Einstufung dieses Tumortyps als „niedriggradig" möglicherweise nicht passend ist und überdacht werden sollte. Zusätzlich zeigten diese Tumore die gleiche charakteristische CD8 T Zell Signatur, die ich auch in humanen PXA Tumoren beobachtet hatte. Meine Arbeit ebnet daher den Weg für weiterführende molekulare und präklinische Forschung an diesen pLGG Entitäten und legt nahe, dass der potenzielle Einsatz von Immuntherapie nicht voreilig verworfen werden sollte. (Sommerkamp et al., in Vorbereitung)

Projekt 2: Detektion der KIAA1549:BRAF Fusion in RNA-Seq Daten

Die häufigste genetische Veränderung, die man im humanen PA beobachten kann, ist die *KIAA1549:BRAF* Fusion. Sie zeigt eine hohe Spezifität für diesen Tumortyp und ist daher von großem diagnostischem und prognostischem Wert. Zudem stellt sie eine Anfälligkeit für zielgerichtete Therapie dar. Aus diesem Grund ist es von wesentlicher Bedeutung, die *KIAA1549:BRAF* Fusion in einem klinischen Umfeld zuverlässig zu detektieren. RNA Sequenzierung (RNA-Seq) ist in der Forschung mittlerweile weit verbreitet und wird auch für diagnostische Zwecke immer häufiger genutzt. RNA-Seq kann dabei neben der Bereitstellung Gen- und Isoform-spezifischer Expressionsdaten auch zur unvoreingenommenen (d.h. nicht vorselektierten) Identifikation exprimierter Fusionsgene genutzt werden. Die *KIAA1549:BRAF* Fusion scheint jedoch nur schwach exprimiert zu sein und es hat sich in der Vergangenheit als schwierig erwiesen, sie in RNA-Seq Daten zu detektieren. Im zweiten Teil dieser Dissertation habe ich daher die Detektionszuverlässigkeit dieser Fusion untersucht und einen optimierten Arbeitsablauf für die Detektion in RNA-Seq Daten entwickelt.

Indem ich RNA-Seq Daten von PA Tumorproben analysiert habe, bei denen eine Fusion bereits bekannt war, konnte ich herausfinden, dass eine tiefe Sequenzierung alleine nicht ausreicht, um die Fusion zuverlässig zu detektieren. Stattdessen wird die Detektierbarkeit von verschiedenen Faktoren einschließlich RNA Integrität / Library Größe, Tumorzellgehalt, *KIAA1549* Expressionslevel sowie dem Protokoll zur Library Präparation beeinflusst. Dabei war auffällig, dass die Rohdaten aller Proben, in denen die Fusion initial nicht identifiziert werden konnte, dennoch unterstützende Reads beinhalteten, die prinzipiell eine Grundlage für die Fusionsdetektion darstellen könnten. Ich habe Engpässe in Standardabläufen des Alignment Algorithmus STAR und des Fusions Callers Arriba identifiziert und zeige, dass ein

ordnungsgemäßes Alignment von geteilten Reads eine große Hürde für die zuverlässige Detektion der *KIAA1549:BRAF* Fusion darstellt. Ein optimierter Arbeitsablauf, der auf angepassten STAR Parametern und einer neuen Version von Arriba beruht, verbesserte die Nachweissicherheit bereits detektierter Fusionen signifikant und ermöglichte die Identifizierung von Fusionen, die zuvor übersehen worden waren. In einer unabhängigen diagnostischen Kohorte aus Montreal demonstriere ich zudem, dass dieser Arbeitsablauf die zuvor genutzten Standard-Analysewerkzeuge deutlich übertrifft. Darüber hinaus weise ich nach, dass die höhere Detektionssensitivität in einem RNA-Seq Datensatz von über 1000 mit Formalin fixierten und in Paraffin eingebetteten (FFPE) Proben diversen Ursprungs nicht mit der Identifikation falsch-positiver Fusionen einhergeht.

Zusammenfassend lässt sich sagen, dass meine Arbeit einen neuen Arbeitsablauf bereitstellt, der die Detektion dieser wichtigen Fusion in RNA-Seq Daten erheblich verbessert und höchstwahrscheinlich auch zu einer besseren Erkennungsleistung von Fusionen in anderen Tumor-Zusammenhängen führen wird. (Sommerkamp et al., 2020)

Declarations

Although I have originally started my doctoral studies in a different research group in May 2016, the work presented in this dissertation was performed from December 2017 until May 2020 at the German Cancer Research Center (DKFZ) in Heidelberg and the Hopp Children's Cancer Center Heidelberg (KiTZ) under the supervision of Dr. David Jones and Prof. Dr. Stefan Pfister.

The results described in the first part of this dissertation ("Comparison of human & murine PA/PXA characteristics") have not been published yet, but will be part of an upcoming manuscript. However, a summary of the results has been submitted and accepted as an abstract for the 19th International Symposium on Pediatric Neuro-Oncology (ISPNO) and will therefore be published by the symposium organizers in an online supplement to the journal *Neuro-Oncology*.

The results described in the second part of this dissertation ("*KIAA1549:BRAF* fusion detection from RNA-Seq data") have been published as a paper in *Acta Neuropathologica* (Sommerkamp et al., 2020). This includes all figures, except for Figure 3.18, Figure 3.25 and Figure 3.26.

Moreover, I have contributed to additional projects, some of which are published (Buhl et al., 2019; Usta et al., 2020), in revision (Hahn, Bode et al.) or submitted (Stichel et al.).

Declarations according to § 8 of the doctoral degree regulations:

1) The thesis I have submitted is my own work.
2) I have only used the sources indicated and have not made unauthorized use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given.
3) I have not yet presented this thesis or parts thereof to a university as part of an examination or degree.

Table of contents

1 Introduction

1.1 Pediatric cancer

Cancer is one of the leading causes of death in children (Siegel et al., 2020). Every year, approximately 400,000 new cases occur worldwide, including many undiagnosed cases in low-income countries (Ward et al., 2019). Significant advances in the understanding and treatment of childhood cancer have substantially reduced the mortality rate for these patients over the last decades. While the overall 5-year survival rate was about 60% in the mid-1970s, it has now exceeded 80% (Siegel et al., 2020), although this strongly depends on the tumor (sub)type. Many cancer survivors, however, have to suffer from treatment-induced long-term sequelae later in life, including organ damage, endocrine deficiency, psycho-social problems and secondary tumors (Bhakta et al., 2017).

Originating from an organism that is still in development, pediatric cancer is biologically very distinct from adult cancer. On the one hand, this applies to the putative cells and tissue types of origin. While adult cancers are mostly carcinomas, i.e. develop from epithelial tissues, pediatric cancers usually originate from mesenchymal or neuroectodermal tissues (Friedman and Gillespie, 2011). On the other hand, this also applies to the underlying genetic alterations. In particular, childhood tumors are characterized by a lower number of somatic mutations, likely reflecting the shorter time span for cells to acquire these mutations (Alexandrov et al., 2013; Kandoth et al., 2013). At the same time, this means that single genetic events can be sufficient to induce cancer in children. Recent large-scale landmark studies on the genetics of various tumor types have demonstrated that pediatric cancer is not only very different from adult cancer but also highly heterogeneous in itself (Gröbner et al., 2018; Ma et al., 2018). The diverse tumor entities harbor a plethora of different mutational signatures, driver genes and structural variants. This strong heterogeneity stresses once more that personalized medicine is the only way to effectively tackle these diseases by individually characterizing each tumor on a molecular level and applying tailored therapy regimens.

Broadly speaking, pediatric cancer can be subclassified into hematological malignancies (i.e. leukemias and lymphomas), tumors of the central nervous system (CNS) and other solid tumors. The following paragraphs will focus on the special characteristics of CNS tumors in children.

1.1.1 Pediatric brain tumors

Brain and other CNS tumors are the most common non-hematological tumor type and the primary cause of cancer-related death in children 0-14 years of age (Ostrom et al., 2019). Like other pediatric cancers, they are characterized by a remarkable heterogeneity. Their classification is defined by the World Health Organization (WHO) and was last updated in 2016 (Louis et al., 2016). While it is still mostly based on traditional histological criteria, this update was a major step forward, as it was the first to incorporate molecular definitions into the classification. As in the previous editions, the different tumor entities are graded from I to IV, reflecting their degree of malignancy.

Naming of the different entities has mainly been based on the suspected cell of origin (or closest lookalike) so far. Among all pediatric CNS tumors, gliomas constitute the largest group with over 50% (Figure 1.1). These tumors are thought to arise from glial progenitor cells, as they display an astrocytic, oligodendroglial, ependymal or mixed neuronal-glial morphology. Grade I and II gliomas are commonly referred to as "low-grade gliomas" (LGGs), while grade III and IV gliomas are grouped as "high-grade gliomas" (HGGs) (Sturm et al., 2017). Accordingly, HGGs have a much worse prognosis than LGGs and are therefore treated much more aggressively (Sturm et al., 2017). The dissertation at hand will focus on pediatric LGGs.
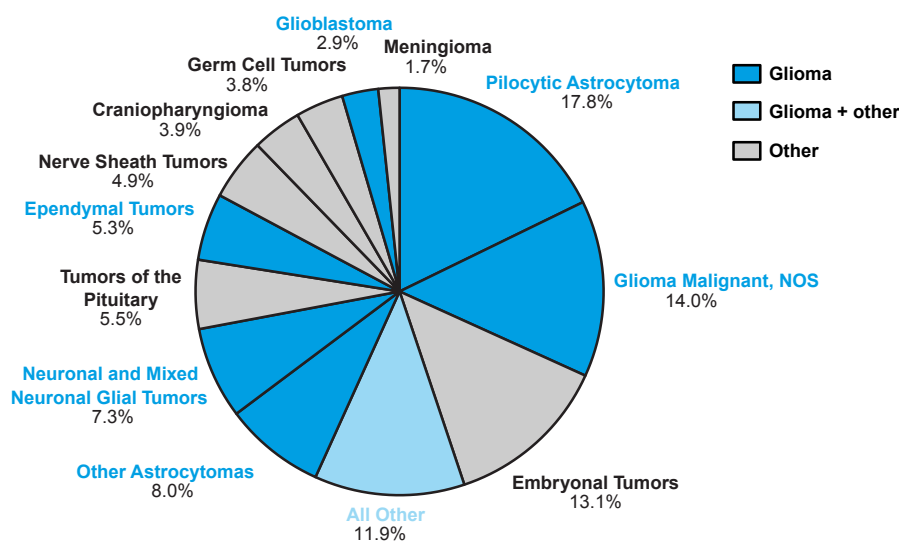
Figure 1.1 | Relative frequency of CNS tumor subtypes by histology in children

Distribution of primary brain or other CNS tumors in children 0-14 years of age based on their histology. The "all other" group includes glioma subtypes like (anaplastic) oligodendroglioma but also non-glioma tumors. NOS = not otherwise specified. Adapted and modified from Ostrom et al. (2019) with permission from Oxford University Press.

1.1.2 Pediatric low-grade gliomas

Pediatric LGGs (pLGGs) account for the largest single subgroup of childhood brain tumors and comprise a heterogenous group of entities, with pilocytic astrocytoma (PA) being most common (Ostrom et al., 2019). In general, pLGGs have an excellent prognosis with a 5-year overall survival of about 90% (Gnekow et al., 2017). Surgery represents the mainstay of treatment for this group of tumors and is usually followed by a "watch-and-wait" approach. This is fundamentally different from pediatric HGGs (pHGGs), where patients directly receive chemo- and radiotherapy (Jones et al., 2019b). With a value of about 50% at 5 years (Gnekow et al., 2017), progression-free survival, however, paints a rather bleaker picture for pLGGs. Albeit not deadly, tumors can recur several times and become a chronic disease, especially if total resection cannot be achieved initially (Wisoff et al., 2011). Upon progression, patients usually receive adjuvant chemotherapy and/or radiation, which again may entail devastating long-term effects. Nevertheless, it is important to note that pediatric LGGs are very distinct from adult (typically *IDH1*-mutated) LGGs, and do not share their propensity for malignant transformation (Jones et al., 2018a).

Pediatric LGGs can arise throughout the CNS and show a wide range of histological features. These features have traditionally been used for tumor classification, and this still holds true when following the current WHO guidelines. Unlike adult LGGs, which have seen the addition of molecular markers for the specification of certain subgroups in the 2016 update of the WHO classification, the definition of pediatric LGG entities is currently still solely based on histology (Louis et al., 2016). While many histological entities are easily distinguishable, there is considerable morphological overlap in other cases (Ryall et al., 2020a). In addition, some very benign entities can show similarities to more malignant glioma types (Collins et al., 2015). This obviously poses challenges for accurate classification and expedient treatment of these tumors.

Recent large-scale sequencing efforts have revolutionized our biological understanding of pLGGs and paved the way for a more molecularly-based classification. Not only are more and more previously unappreciated subgroups being discovered, these studies have also illustrated the discrepancy between histological and genetic features (Jones et al., 2019a). While some genetic marks are enriched in certain histological entities (and vice versa), there are no true one-to-one associations of both (Figure 1.2). Although the molecular classification of pLGGs is still not as well-developed as for other brain tumor types, mounting evidence suggests that it is more accurate and thus superior to a classification based on histology alone (Jones et al., 2019a). Therefore, an adoption into the WHO classification is very likely to happen in the 5th edition, which is currently being written.
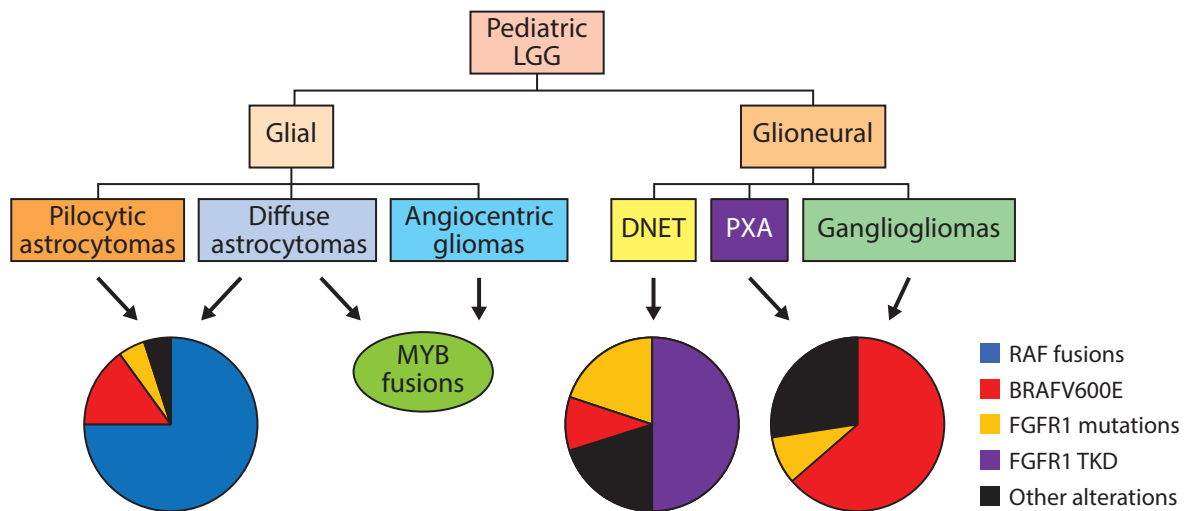
Figure 1.2 | Pediatric LGG subgroups and their molecular alterations

Pediatric LGGs comprise both glial and glioneuronal tumors with diverse histological entities. The most established entities are illustrated and usually lack a clear association with one specific genetic driver. In addition, more and more novel molecular entities are being discovered in genetic profiling studies. Adapted from Jones et al. (2019a) with permission from Annual Reviews, Inc..

Despite the high number of distinct molecular subgroups evolving, pLGGs are considered rather simple genetically. They show a low frequency of somatic mutations and are often driven by a single genetic alteration (Jones et al., 2019a). These oncogenic events can hit different genes (Figure 1.2), but there is one overarching theme that connects virtually all of them: activation of the RAS/mitogen-activated protein kinase (MAPK) pathway (Jones et al., 2013; Zhang et al., 2013). Almost all pLGGs feature an upregulation of this pathway, making them a "single pathway disease". This is especially true as the RAS/MAPK alterations are directly involved in tumor initiation and do not occur as secondary genetic events during tumor progression, as is the case for other tumor types (Jones et al., 2019b).

The first clear indication that RAS/MAPK pathway activation and pLGGs are linked was the observation that about 15% of patients with neurofibromatosis type 1 (NF1) develop optic pathway gliomas (Listernick et al., 1999). NF1 is a tumor predisposition syndrome caused by germline mutations in the *neurofibromin 1* (*NF1*) gene. The product of this gene, neurofibromin, is an inhibitor of RAS, and inactivation therefore leads to aberrant upregulation of RAS/MAPK signaling (Rasmussen and Friedman, 2000). Over the years, the involvement of this pathway in pLGG development has been supported by the discovery of many other drivers (Figure 1.3) (Jones et al., 2019a). Two of these, *KIAA1549:BRAF* and BRAF V600E, will be discussed in the following sections. Notably, recent sequencing efforts also identified some genetic events that do not have a (yet known) direct link to the RAS/MAPK pathway, like alterations of *MYB/MYBL1* (Ramkissoon et al., 2013; Tatevossian et al., 2010; Zhang et al., 2013). However, latest studies suggest that these tumors still show the same level of RAS/MAPK activation, although the exact mechanism remains to be determined (Ryall et al., 2020b).
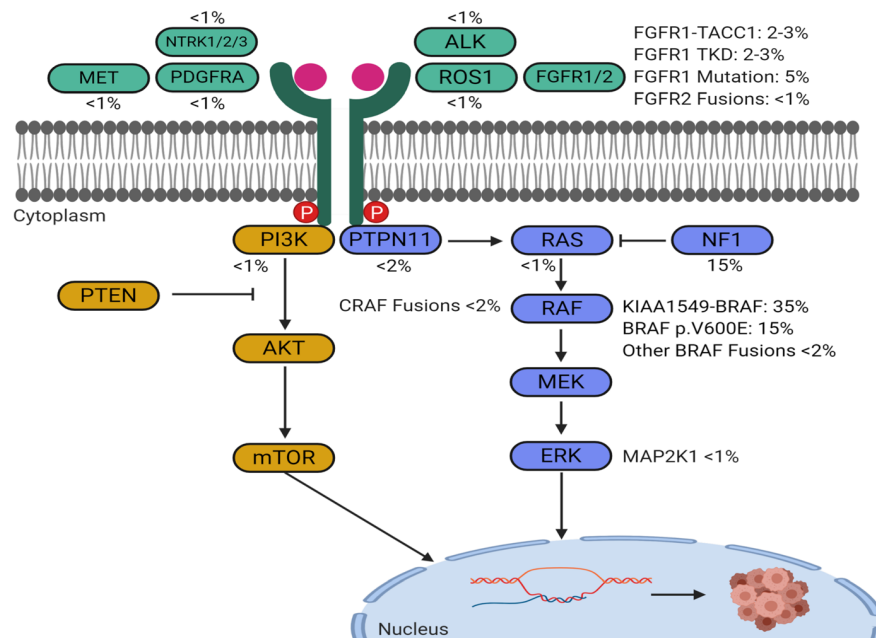
Figure 1.3 | RAS/MAPK alterations in pLGGs

Different genetic alterations can be observed in pLGGs, but almost all affect the RAS/MAPK pathway. Their frequency is shown, with *KIAA1549:BRAF* fusions (35%), BRAF V600E mutations (15%) and *NF1* alterations (15%) being most common. Adapted from Ryall et al. (2020a) under a Creative Commons license.

Essentially being a disease of one single pathway, targeting the RAS/MAPK pathway is an obvious therapeutic strategy for the treatment of pLGGs (Jones et al., 2019b). Clinical trials of BRAF and MEK inhibitors are currently underway, with first results being promising (Banerjee et al., 2017; Fangusaro et al., 2019; Hargrave et al., 2019).

1.1.2.1 *KIAA1549:BRAF* fusion

The *KIAA1549:BRAF* fusion (Figure 1.4) is the most abundant alteration in pLGGs and usually associated with PA (see 1.1.3), particularly when located infratentorially (Jones et al., 2019a). Therefore, its detection is generally indicative of a very good prognosis. Originally, the fusion was indirectly identified as a focal 7q34 gain in genome-wide DNA copy number data of PA, which was shown to result in increased *BRAF* expression (Pfister et al., 2008). Subsequently, it was discovered that the focal gain originated from a tandem duplication, which caused the formation of the previously unknown *KIAA1549:BRAF* fusion (Jones et al., 2008). Depending on the exact combination of exons, different variants of the fusion can occur. While the most common variants are 16:9 and 15:9, in which exon 16 or 15 of *KIAA1549* is fused to exon 9 of *BRAF*, respectively (Jones et al., 2008), more variants have been identified, including 16:11, 18:10, 19:9 (Forshew et al., 2009; Jones et al., 2008; Sievert et al., 2009) and more recently 16:10 (Dahiya et al., 2012), 15:11, 17:10 (Lin et al., 2012), 13:9 and 15(+34 bp of intron 15):10 (Jones et al., 2013). Nevertheless, all fusion variants have in common that the N-terminal regulatory domain of BRAF is lost, while the kinase domain is preserved in frame (Figure 1.4). As a consequence, the BRAF kinase domain is constitutively active and drives downstream MAPK signaling (Jones et al., 2008).
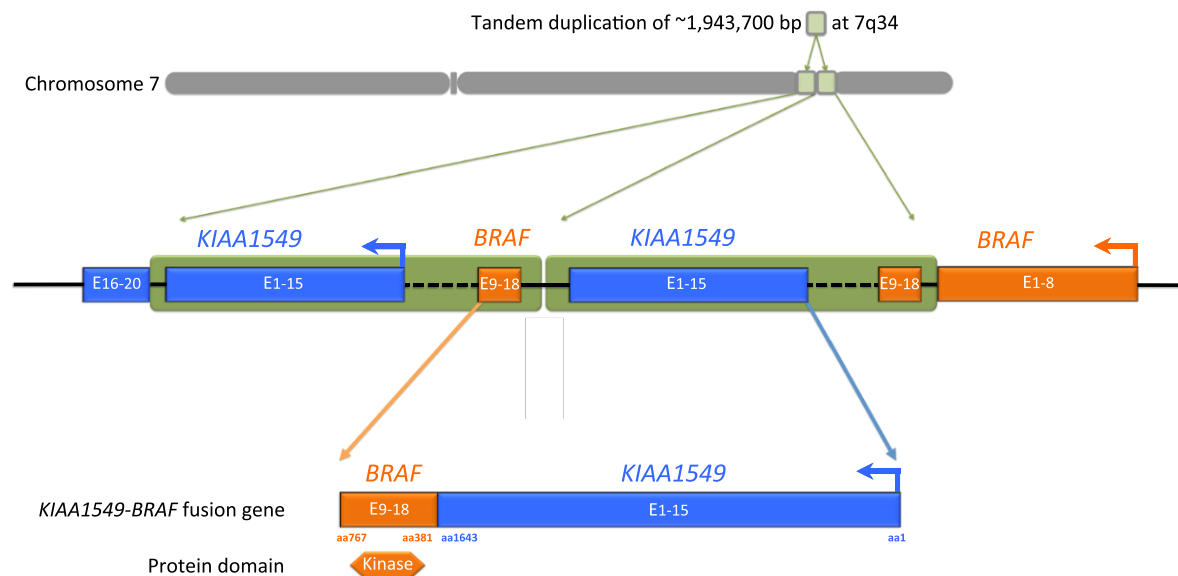
Figure 1.4 | The *KIAA1549:BRAF* fusion

The *KIAA1549:BRAF* fusion gene is generated by a focal tandem duplication at 7q34. It usually comprises exons 1-16 or 1-15 of *KIAA1549* and exons 9-18 of *BRAF*. In the illustration, the exon combination 15:9 is shown. The KIAA1549:BRAF fusion protein retains the kinase domain and ATP binding pocket of BRAF but not its N-terminal regulatory domain including conserved regions 1 and 2 (CR1/CR2). This renders the kinase domain constitutively active and leads to permanent signaling. Adapted and modified from Subbiah et al. (2014) under a Creative Commons license.

In addition, alternative *BRAF/RAF* fusions have been observed in pLGGs that follow the same oncogenic mechanism. Amongst others, these include *SRGAP3:RAF1* (Jones et al., 2009), *FAM131B:BRAF* (Cin et al., 2011), *RNF130:BRAF*, *CLCN6:BRAF*, *MKRN1:BRAF* and *GNAI1:BRAF* (Jones et al., 2013), with new variants still being identified.

1.1.2.2 BRAF V600E mutation

Mutations in *BRAF* constitute another group of genetic events in pLGGs and are also found in a variety of other cancers (Davies et al., 2002). In most cases, a genetic missense mutation results in the substitution of the amino acid valine at position 600 by glutamic acid (V600E). This mutation mimics phosphorylation of BRAF and thus induces constitutive MAPK signaling (Garnett and Marais, 2004; Wan et al., 2004). Again, rare variants of BRAF mutations have been described in pLGGs that harbor single amino acid insertions very close to position 600 (Jones et al., 2009) but share the same mechanism of action.

In contrast to the *KIAA1549:BRAF* fusion, which is only rarely seen outside of PA, *BRAF* mutations can occur in various pLGG entities (see Figure 1.2). They are most commonly encountered in pleomorphic xanthoastrocytoma (PXA; see 1.1.4) and ganglioglioma (GG) but can also be found in PA, dysembryoplastic neuroepithelial tumor (DNET) and many more (Schindler et al., 2011).

1.1.2.3 *CDKN2A/B* deletion

Some tumor entities that formally belong to the pLGG group can also harbor additional genetic alterations that are not related to the RAS/MAPK pathway. One of them is deletion of the *cyclin-dependent kinase inhibitor 2A/B* (*CDKN2A/B)* locus at 9p21, which is recurrently found across pediatric cancer types (Gröbner et al., 2018). The fact that it is usually accompanied by one of the typical RAS/MAPK pathway alterations, like BRAF V600E (Ryall et al., 2020a), and that it shows an enrichment in relapsed tumors (Eleveld et al., 2015), underlines its role as a secondary genetic hit. In the pLGG spectrum, *CDKN2A/B* deletion is mostly found in PXA (Weber et al., 2007).

The *CDKN2A* locus encodes for two tumor suppressors by alternative exon usage, p16INK4a and p14ARF (Figure 1.5), which act as regulators of the cell cycle. While p16INK4A activates the retinoblastoma (Rb) family of proteins via inhibition of cyclin dependent kinases 4 and 6 (CDK4/6) (Liggett and Sidransky, 1998; Ruas and Peters, 1998), p14ARF is a negative regulator of mouse double minute 2 homolog (MDM2) and thus indirectly activates p53 (Ko et al., 2018). The neighboring *CDKN2B* locus encodes for p15INK4B, which has a similar function as p16INK4A (Figure 1.5).
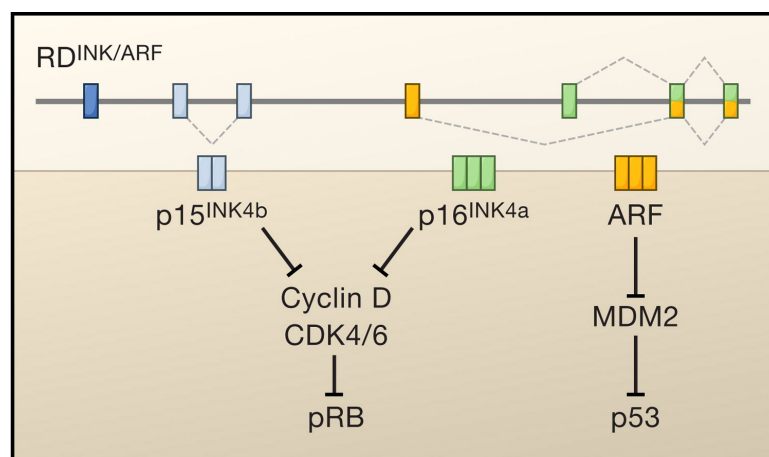
Figure 1.5 | The *CDKN2A/B* locus

The *CDKN2A/B* locus encodes three different tumor suppressors. Both p16INK4A and p14ARF are transcribed from the same part of the locus, using alternative first exons and a different reading frame of exons 2 and 3. In contrast, p15INK4B is transcribed from another part of the locus. All three regulate the cell cycle, p16INK4A and p15INK4B negatively regulate CDK4/6 and thus activate pRB, while p14ARF activates p53 by inhibiting MDM2. Adapted from Kim & Sharpless (2006) with permission from Elsevier.

It has been shown that oncogenic MAPK signaling, e.g. caused by BRAF V600E, activates these products of the *CDKN2A/B* locus as a counter mechanism to induce so-called oncogene-induced senescence (OIS) (Jacob et al., 2011; Michaloglou et al., 2005; Serrano et al., 1997). Deletion of *CDKN2A/B* can thus serve as a mechanism for tumor cells to evade this tumor-suppressive process and promote cell cycle progression.

1.1.3 Pilocytic astrocytoma

PA is the most common childhood brain tumor (see Figure 1.1). It shows benign growth and is therefore classified as WHO grade I (Louis et al., 2016). Malignant progression is rarely observed, and affected pediatric patients accordingly have an excellent prognosis with a 10-year survival rate of more than 95% (Ostrom et al., 2019). PA tumors occur mostly in younger children and are preferentially found in the cerebellum and supratentorial locations (Burkhard et al., 2003).

Histologically, PA is characterized by a biphasic pattern of compact densely fibrillated areas with abundant Rosenthal fibers and loose microcystic areas (Collins et al., 2015). The term "pilocytic" refers to the appearance of tumor cells in the compact areas, which display long hair-like processes. Matching its non-malignant growth pattern, PA shows a very low proportion of cells positive for the proliferation marker Ki-67 (Tu et al., 2018). Histopathological identification of PA is often clear but might be challenging in cases of "non-classical histology". Among others, morphological overlap has been observed with GG, DNET and PXA (Collins et al., 2015; Jones et al., 2019a). Similarly, PA with a diffuse appearance might be confused with high-grade tumors due to the presence of microvascular proliferation, with great implications in terms of prognosis and treatment (Collins et al., 2015).

On a molecular level, PA is a prime example for being a "single pathway disease" with almost all cases featuring activating alterations of the RAS/MAPK pathway (Jones et al., 2012). The most common events are *BRAF* fusions (mainly *KIAA1549:BRAF*), *NF1* alterations and *BRAF* mutations (mainly BRAF V600E). Notably, while PA can occur virtually everywhere in the CNS, the distribution of genetic alterations is not uniform, as it has been observed for other pLGG types as well (Sturm et al., 2017). *KIAA1549:BRAF* fusions, for example, are mostly found in infratentorial PA, whereas BRAF V600E mutations are rather found in supratentorial PA (Jones et al., 2012). Other, less common, affected driver genes include *fibroblast growth factor receptor 1* (*FGFR1*), *neurotrophic receptor tyrosine kinase 2* (*NTRK2*), *KRAS proto-oncogene, GTPase* (*KRAS*) and *protein tyrosine phosphatase non-receptor type 11* (*PTPN11*) (Jones et al., 2013).

Importantly, all of these genetic alterations are usually mutually exclusive (Figure 1.6), again underlining that one is sufficient to induce PA. This is also supported by the absence of any secondary genetic hits that are independent of RAS/MAPK signaling.
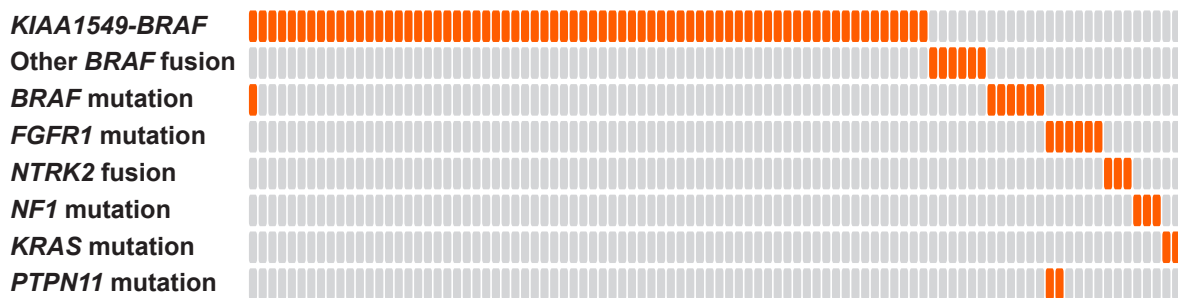
Figure 1.6 | Genetic alterations in PA

Overview of genetic alterations identified in 96 PA tumors by DNA and RNA sequencing. The events (colored boxes) all affect the RAS/MAPK pathway and are, with rare exceptions, mutually exclusive. By far the most frequent alteration is the *KIAA1549:BRAF* fusion. Adapted and modified from Jones et al. (2013) with permission from Springer Nature.

Further research has demonstrated that the aberrant RAS/MAPK activation, despite initially driving tumor growth, eventually leads to OIS in PA (Jacob et al., 2011; Raabe et al., 2011). This "rescue mechanism" is mediated by the senescence-associated secretory phenotype (SASP) (Buhl et al., 2019) and is thought to be one of the main reasons for the benign phenotype of PA (Jones et al., 2012).

1.1.4 Pleomorphic xanthoastrocytoma

PXA is a rare neoplasm that mostly affects children and adolescents. It is almost exclusively found in supratentorial regions of the brain, mostly in the temporal lobe (Ida et al., 2015). In contrast to PA, it is associated with inferior prognosis and a 10-year overall survival rate of around 67% (Perkins et al., 2012). This means that one in three patients will die within 10 years after diagnosis. PXA is therefore classified as WHO grade II (Louis et al., 2016). During the last years, it has been acknowledged that PXA can also display anaplastic features, which are atypical for low-grade tumors, and that these tumors have a worse clinical outcome (Ida et al., 2015). As a result, "anaplastic PXA" was added to the 2016 WHO classification as a separate entity and assigned to WHO grade III (Louis et al., 2016). The usefulness of this split, however, is debated, as it is primarily based on the histopathological assessment of proliferative activity and separates tumors with a similar molecular profile and largely overlapping histological appearance into low- and high-grade (Jones et al., 2019a), resulting in completely different clinical approaches (see 1.1.2). For this reason, a group of intermediate grade might be of use in the future, which contains tumor types like PXA that show a worse clinical outcome than PA but still more favorable than glioblastoma (GBM).

In general, PXA is histologically characterized by cellular pleomorphism of spindle cells and giant multi-nucleated cells with xanthomatous cytoplasm and nuclear atypia (Giannini et al., 1999). In addition, it often shows a dense network of reticulin surrounding the tumor cells, particularly in BRAF V600E-mutated cases (Koelsche et al., 2014), something that is rarely observed in other tumor types. Nevertheless, morphological overlap can be found with GG

(Aisner et al., 2014) as well as epithelioid glioblastoma (eGBM), a recently defined GBM variant that might be related to PXA (Alexandrescu et al., 2016; Furuta et al., 2018; Korshunov et al., 2018).

Genetically, PXA tumors mainly harbor the BRAF V600E mutation, which drives RAS/MAPK signaling (Schindler et al., 2011). Unlike PA, they are typically characterized by at least one additional secondary mutation. In most cases, this is a deletion of the *CDKN2A/B* locus (Weber et al., 2007; Zhang et al., 2013), which is clearly associated with a worse prognosis (Lassaletta et al., 2017; Mistry et al., 2015). Most likely, this also reflects the potential of PXA for malignant transformation.

1.1.5 Tumor microenvironment

Solid tumors are complex "organs" composed of tumor cells and non-neoplastic cells. A mounting body of evidence indicates that these other normal cells are not only bystanders but heavily communicate with the tumor at various levels and significantly influence its development, maintenance and evolution (Broekman et al., 2018). Several cell types can contribute to this microenvironment in which the tumor develops, including stromal cells as well as immune cells. For gliomas, these are mainly tumor-associated microglia or macrophages (TAMs), which can constitute up to 30-50% of the whole tumor mass in some cases (Hambardzumyan et al., 2016). A role of great importance has also been demonstrated for other cell types and structures including normal brain cells (like neurons, astrocytes and oligodendrocytes), blood vessels, the extracellular matrix (ECM), other cells of the innate immune system (like neutrophils and mast cells) and cells of the adaptive immune system (like T cells) (Figure 1.7). The degree of infiltration as well as the composition of the different cell types in the microenvironment strongly varies between tumor types. For example, LGGs generally show a lower extent of immune infiltration compared with HGGs (Yang et al., 2011). In addition, high-grade tumors are particularly capable of hijacking the cells in their environment for their own benefit. This involves shifting the polarization of TAMs towards an anti-inflammatory phenotype, in which they release signaling molecules that actively support tumor growth, and inactivating the effector function of T cells, which sets them in a state of so-called exhaustion (Broekman et al., 2018).

Monocytes and macrophages
- Recruited to tumour by cytokine and chemokine gradient
+ Support ECM remodelling and induce angiogenesis

Microglia
+ Support tumour progression by ECM remodelling and induction of angiogenesis

Mast cells
- Recruited to tumour by SDF1
- Activated by tumour-derived stimuli
+ Produce growth-enhancing and angiogenic factors

T cells
± Subpopulations with both protumour and antitumour potency depending on the complex array of immune-regulatory mechanisms

Oligodendrocytes
– Inhibit tumour growth through paracrine signalling via WIF1

ECM
+ Stiffness of ECM can affect glioma progression
+ Tenascin C in ECM modulates glioma secretome

Blood vessels
+ Delivery of oxygen and nutrients
+ Support tumour cell dissemination
+ Provision of a barrier that results in compartmentalization within the brain

Neurons
+ Affect glioma growth and invasion by neuronal activity, including secretion of neurotransmitters

Neutrophils
+ Source of cytokines to increase the infiltration of monocytes

Astrocytes
+ Transform into neoplastic glioma cells

Figure 1.7 | Microenvironment of brain tumors

Different cell types and structures can be part of the microenvironment of brain tumors, particularly in high-grade tumors like glioblastoma. They closely interact and communicate with the tumor and can both support (+) or inhibit (–) its growth. The cellular part of the microenvironment encompasses normal brain cells as well as immune cells. The latter are either resident (like microglia) or infiltrate the brain via the blood. The extracellular matrix (ECM) and blood vessels represent the non-cellular part of the microenvironment and also substantially influence tumor growth. Adapted from Broekman et al. (2018) with permission from Springer Nature.

As introduced before, pediatric tumors are very distinct from adult tumors. Their lower mutational burden (Alexandrov et al., 2013; Kandoth et al., 2013) is thought to result in the expression of fewer neoantigens, which could be recognized by the adaptive immune system. This would also explain why pediatric tumors are less immunogenic compared with adult cancer (Majzner et al., 2017). A notable exception are mismatch repair-deficient tumors, which possess the highest number of somatic mutations across all human cancer types due to the lack of appropriate DNA repair mechanisms (Shlien et al., 2015). In addition to the genetic differences of tumor cells, the innate and adaptive immune system of children is still immature and thus substantially different from adults (Simon et al., 2015). This even refers to differences between neonates and older children, as their immune system is strongly evolving during this phase of life. It appears likely that these differences in development also functionally affect the immune response to tumors.

A limited number of studies has been conducted on the immune microenvironment of pediatric brain tumors, but their exact composition and functional role remain largely unknown (Jones et al., 2019b). Surprisingly, some studies have indicated that pediatric low-grade tumors are more strongly infiltrated by immune cells compared with pediatric high-grade lesions. For example, PA showed a higher number of infiltrating myeloid cells and lymphocytes as well as a more robust activation of the myeloid cell compartment in comparison to GBM (Griesinger et al., 2013). Another recent study confirmed that pLGGs have a higher number of infiltrating T cells than pHGGs and that the presence of certain T cell subsets correlated with mutational load (Plant et al., 2018). Immune profiling of a pHGG trial cohort revealed that the highest CD8 T cell infiltration was encountered in PXA-like, i.e. intermediate grade, tumors, besides hypermutator cases (Mackay et al., 2018).

Thus, pediatric low-grade tumors seem to harbor a substantial immune infiltrate despite their predicted paucity of neoantigens and display significant differences to their adult counterparts. Nevertheless, further studies on the exact composition in different pLGG groups are urgently needed.

1.1.5.1 Immunotherapy

Given the high number of infiltrating immune cells in many tumors and the intimate molecular crosstalk between both, immunotherapy aims at harnessing the immune system for therapeutic purposes and redirecting it against the tumor (Mellman et al., 2011). Different immunotherapeutic strategies have been developed, including immune checkpoint inhibition, chimeric antigen receptor (CAR) T cells and vaccine approaches. Immune checkpoint inhibition refers to the blockade of costimulatory immune checkpoint molecules like PD-1 or PD-L1 that modulate and suppress T cell immune responses (Pardoll, 2012). In comparison, CAR T cell therapy is a form of adoptive cell therapy and involves the modification of patient-derived T cells to express a CAR hybrid receptor directed against a tumor antigen (Sadelain et al., 2013). These cells are subsequently reinfused into the patient and become activated upon antigen recognition. Therapeutic cancer vaccines, on the other hand, can be based on different designs and are thus sub-classified into cellular vaccines, protein/peptide vaccines and genetic vaccines (Guo et al., 2013). Despite the large variety of approaches, all of them are aimed at enhancing the patient's immune response against the tumor.

Some of these immunotherapeutic strategies have recently shown tremendous success in clinical studies of adult patients (Maude et al., 2018; Motzer et al., 2018; Wolchok et al., 2017). Immunotherapy thus has the potential to expand the therapeutic toolbox, especially in difficult-to-treat cancers. This applies in particular also to pediatric cancer, in which alternatives to chemo- and radiotherapy would be welcome additions due to the numerous long-term sequalae of these established treatments.

Unfortunately, the special characteristics of childhood brain tumors pose difficulties for the use of immunotherapy in these entities and age groups. In adult patients, tumor mutational load has been identified as one of the key determinants of therapy success (Havel et al., 2019; Samstein et al., 2019). However, as already alluded to above, pediatric cancers are characterized by a low number of somatic mutations and this also holds true for brain tumors (Gröbner et al., 2018). In addition, most childhood brain tumors occur in the posterior fossa, a region of the brain that contains very sensitive structures and where immunotherapy-induced inflammation could lead to serious clinical issues (Plant and Hwang, 2018).

So far, immune checkpoint inhibition for pediatric patients has only been successful for mismatch repair-deficient tumors (Bouffet et al., 2016). Most of the other clinical studies have largely failed to live up to expectations (Davis et al., 2020; Geoerger et al., 2020a; 2020b). However, this might to a large extent be due to a poor understanding of the immune microenvironment of many pediatric tumors and the lack of predictive biomarkers, even for the very small subset of patients that do respond. Further research is hence needed to uncover those cancer types or patient populations that are susceptible to immune checkpoint inhibition.

1.2   Molecular diagnostics

For many years, morphology-based histopathological examination has been considered the gold standard for the classification and diagnosis of brain tumors, and most diagnostic labs are equipped with the necessary tools to perform this type of analysis. However, brain tumors are highly complex, and tumors with distinct underlying biology can possess similar histological characteristics (see above). In addition to this morphological overlap, histopathological examination itself is prone to substantial inter-observer variability (Capper et al., 2018a; van den Bent, 2010). Naturally, this hampers an accurate classification of the tumor type and thus a reliable prediction of the clinical course. The field is therefore seeing a strong shift towards the additional inclusion of molecular diagnostic assessment. This is further being fueled by a growing body of evidence showing that a molecular classification represents the landscape of tumor types more accurately than one that is based on histopathology alone (Jones et al., 2019a).

Nowadays, multiple profiling platforms are at the disposal of diagnostic institutions for the molecular classification of tumor types as well as the identification of potentially targetable genetic alterations. These include gene panel sequencing, whole exome sequencing (WES), low-coverage or full-coverage whole genome sequencing (WGS), RNA sequencing (RNA-Seq), DNA methylation arrays and gene expression arrays (Figure 1.8). While some of these platforms are to a limited extent interchangeable, most provide different types of information that allow different genetic features to be evaluated. In parallel with increasing accessibility and

affordability of these profiling platforms, several precision medicine programs have been launched around the world (Jones et al., 2019b). Some of these, including the Individualized Therapy for Relapsed Malignancies in Childhood (INFORM) study in Germany (Worst et al., 2016), make extensive use of the variety of profiling platforms (Figure 1.8) to provide a basis for an informed choice of targeted or conventional therapies with the aim to ultimately improve clinical outcome for pediatric brain tumor patients. In addition, the data collected in these programs will help to refine the molecular assessment of tumors in the future and to identify novel therapeutic targets as well as prognostic and predictive biomarkers.
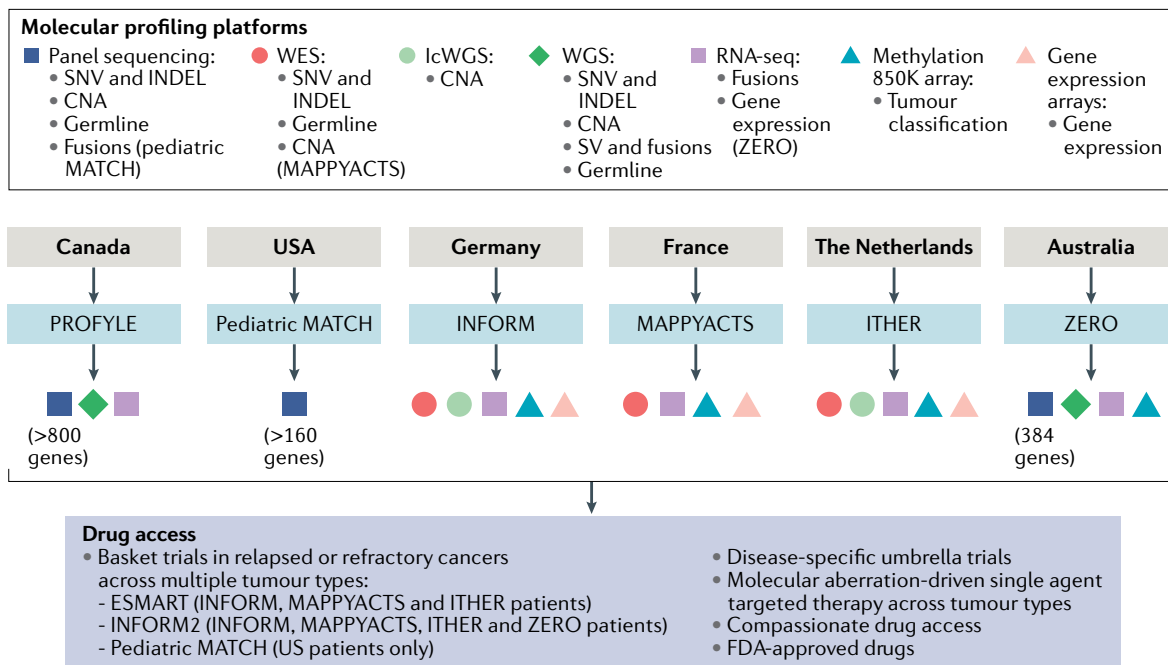
Figure 1.8 | Overview of molecular profiling platforms and pediatric precision medicine programs

Several molecular profiling platforms are nowadays available for clinical diagnostics. These include gene panel sequencing, whole exome sequencing (WES), low-coverage whole genome sequencing (lcWGS), regular whole genome sequencing (WGS), RNA sequencing (RNA-Seq), DNA methylation arrays and gene expression arrays. Each platform can only be used to identify specific types of genetic alterations. Precision medicine programs for pediatric oncology therefore typically use a combination of these methods. Adapted and modified from Jones et al. (2019b) with permission from Springer Nature.

In the following, two commonly used molecular profiling platforms will be described in more detail, DNA methylation arrays and RNA-Seq.

1.2.1 DNA methylation array

Methylation arrays have emerged as one of the primary tools for the molecularly-based classification of brain tumors (Capper et al., 2018a) and are starting to be utilized for the classification of other cancers like sarcomas (Koelsche et al., 2018; Renner et al., 2013; Wu et al., 2017). Methylation arrays can determine genome-wide DNA methylation patterns, i.e. the so-called methylome, by quantifying CpG methylation at more than 850,000 sites (Pidsley et al., 2016). This is achieved by targeted quantitative "genotyping" of DNA methylation. First, the genomic DNA undergoes bisulfite treatment, which converts non-methylated cytosines to uracil, while methylated cytosines remain unchanged. The DNA is subsequently amplified, and specific probes on a microarray are used to detect the presence of a cytosine or uracil at pre-defined CpG sites throughout the genome in a highly multiplexed manner (Bibikova et al., 2011). This information can then be used to infer the methylation status at these sites.

It is believed that changes in methylation patterns are tightly linked to cell lineage and differentiation. As a consequence, the methylome of cancer cells is likely reminiscent of their cell of origin, with an additional component shaped by the acquired somatic mutations that underlie the altered transcriptional programs (Fernandez et al., 2012; Hovestadt et al., 2014). Importantly, this molecular "fingerprint" seems to be stable during tumor evolution (Pajtler et al., 2015). This is underlined by the fact that methylome data can be used to uncover the origin of metastases for which the primary tumor is not known (Moran et al., 2016) and that patient-derived cell lines retain their methylation pattern (at least initially) despite *in vitro* culture and manipulation (Selt et al., 2017). Furthermore, DNA methylation arrays may give robust results also from samples of small size or low quality, like formalin-fixed, paraffin-embedded (FFPE) tissue (Hovestadt et al., 2013).

Due to these advantages, methylation arrays have been widely adopted and used to split tumor entities into molecular subgroups that had previously been considered homogeneous (Capper et al., 2018a; Pajtler et al., 2015; Sturm et al., 2016; 2012). In addition, the first comprehensive methylation-based tumor classifier was published in 2018 (Capper et al., 2018a). It allows mapping of newly diagnosed brain tumors to an existing reference set (Figure 1.9) and is now actively being used to aid diagnosis (Capper et al., 2018b; Pickles et al., 2020).
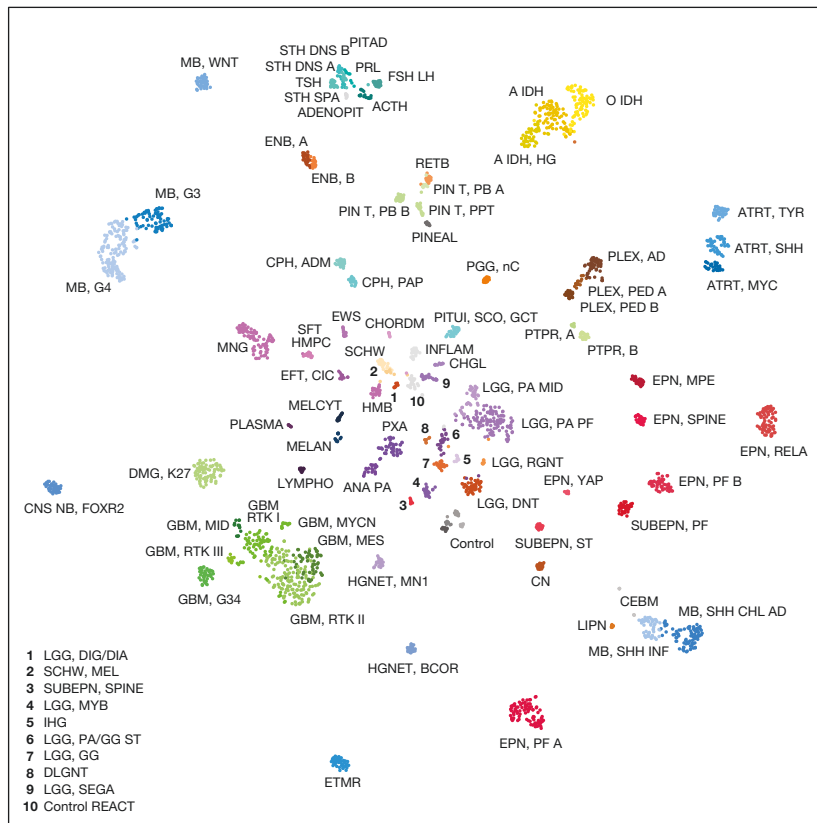
Figure 1.9 | DNA methylation-based CNS tumor reference cohort

The reference cohort of the brain tumor classifier was established on the basis of DNA methylation profiles from 2801 tissue samples. Unsupervised clustering shows the 91 distinct methylation classes, including 82 tumor classes and 9 control tissue classes. The tumor methylation classes are color-coded based on their closest histological equivalent. With the help of a machine-learning algorithm, diagnostic tumor samples can be assigned to one of these classes. Adapted from Capper et al. (2018a) with permission from Springer Nature.

1.2.1.1 Copy number variation calling

Notably, the data generated by methylation arrays can also be used to identify copy number variations (CNVs). For every CpG site included in the array, the intensity of a methylated and an unmethylated probe is measured. The sum of both can therefore be used to determine the copy number in relation to a normal reference profile, showing very similar results compared with approaches based on single nucleotide polymorphism (SNP) profiling (Sturm et al., 2012), albeit at a lower resolution. The resulting genome-wide copy number plots allow identification of large-scale as well as focal copy number variations (Figure 1.10).
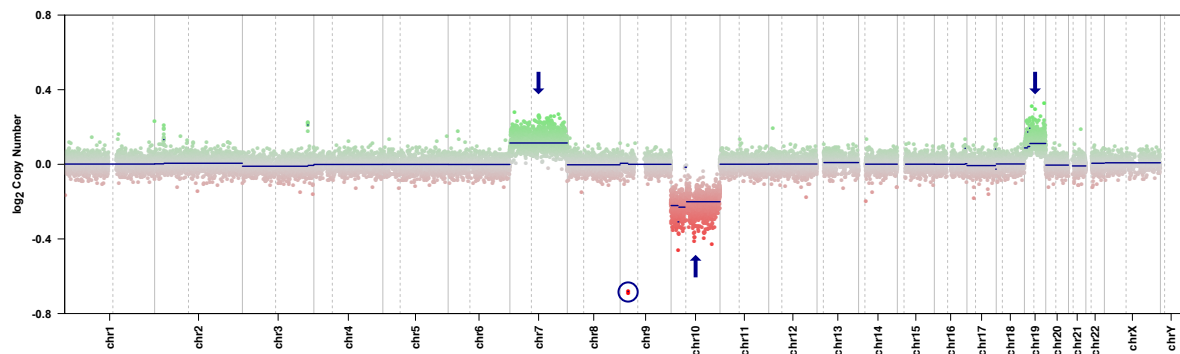
Figure 1.10 | Typical genome-wide copy number plot

DNA methylation array data can be used to generate copy number plots, and a typical example is shown. Large-scale variations, i.e. gain of chromosomes 7 and 19 and loss of chromosome 10 (arrows), as well as focal variations, i.e. loss of the *CDKN2A/B* locus at 9p21 (circle), can be identified.

1.2.2 RNA sequencing

In contrast to DNA methylation arrays, RNA-Seq provides information on the expressed genes, i.e. the so-called transcriptome. It therefore uses extracted RNA as starting material, which is differently processed depending on the type of RNA-Seq. For the standard short-read workflow, the RNA is subsequently enriched for mRNA or freed from ribosomal RNA, fragmented, transcribed into cDNA and ligated to adaptors to eventually generate a sequencing library (Stark et al., 2019). This library is then sequenced on high-throughput machines to the desired read depth, usually employing a paired-end sequencing approach. The base-called sequencing reads are saved as FASTQ files (Cock et al., 2010). To make use of the data, the raw sequencing reads have to be aligned to the reference genome using tools like STAR (Dobin et al., 2013), TopHat (Kim et al., 2013) or HISAT (Kim et al., 2015). The alignment allows the reads to be assigned to the genes and transcripts they originated from. This information can then be used to determine the expression levels of individual genes as well as specific transcript isoforms in an unbiased manner.

1.2.2.1 Fusion calling

Importantly, RNA-Seq data can also serve as a basis for identifying expressed gene fusions (Maher et al., 2009), which play a major role in cancer (Mertens et al., 2015). This is accomplished by identifying split reads, i.e. chimeric reads that contain sequences of two different genes, as well as discordant mates, i.e. paired-end reads from the same library fragment that map to two different genes. Several algorithms have been developed that build on the aligned data to identify fusion-supporting reads (Latysheva and Babu, 2016), like Arriba, FusionCatcher and STAR-Fusion. Notably, these tools do not rely on prior knowledge and thus allow the discovery of novel fusions. In (pediatric) oncology, the identification of fusion transcripts is of pivotal importance. On the one hand, fusions are promising drug targets, particularly when involving a kinase like BRAF, NTRK or anaplastic lymphoma kinase (ALK) (Clarke et al., 2020; Jones et al., 2019b). On the other hand, some fusions show high specificity

for a certain tumor type and are therefore of great diagnostic and prognostic value. For example, detection of the *KIAA1549:BRAF* fusion is under normal circumstances a strong indicator for the diagnosis of pediatric PA and thus of a favorable clinical outcome (Jones et al., 2019a), although it can also occur in some other contexts. Accurate and reliable tools for the detection of gene fusions from RNA-Seq data are thus of major clinical relevance.

1.3   Preclinical modeling

Up to now, the vast majority of drugs used in early phase clinical trials for pediatric cancer have originally been developed for the treatment of adults. Given the manifold differences between children and adults in terms of their developmental status and physiology as well as the biology of their cancers, it is not surprising that many of these drugs show poor efficacy (Waligora et al., 2018). Therefore, preclinical modeling is more than ever a key resource not only to gain a more detailed understanding of tumor initiation and progression but also to aid drug development and patient stratification. Robust preclinical data generated in models of pediatric cancer can help to decide which oncogenic alterations are promising therapeutic vulnerabilities and which drugs for adult cancer are likely to be efficacious in a pediatric setting as well.

Over time, several model systems for oncology have emerged, both *in vitro* and *in vivo*. These include cell lines, organoids, genetically engineered mouse models (GEMMs), humanized mouse models and patient-derived xenografts (PDXs), all of which have individual advantages and drawbacks (Jones et al., 2019b). PDX models have gained a lot of popularity in the last years and are now available also for various pediatric tumor types (Brabetz et al., 2018; Stewart et al., 2017). Depending on the preclinical question, however, PDX models have two major disadvantages: 1) Benign tumors like PA usually fail to engraft *in vivo* due to the lack of proliferative activity and the onset of senescence (Selt et al., 2017). 2) PDX models rely on immunocompromised mice and are thus not suitable for studies of the immune microenvironment, which is an integral component of solid tumors (see 1.1.5), or immunotherapeutic agents like immune checkpoint inhibitors (ICIs) (Hermans and Hulleman, 2020).

For these questions, GEMMs are the method of choice. By introducing defined genetic driver alterations into the relevant cell type, GEMMs can be used to study the biology of both slow- and fast-growing tumors and allow them to originate in their native and immunocompetent environment (Dobson and Gopalakrishnan, 2018). One method that uses GEMMs for tumor induction is the RCAS/TVA system.

1.3.1 RCAS/TVA system

The RCAS/TVA system is based on the replication-competent avian sarcoma-leukosis virus (ASLV) long terminal repeat (LTR) with splice acceptor (RCAS) vector family. These vectors were derived from Rous sarcoma virus A (RSV-A), a retrovirus, by modifying it to express genes of interest instead of the viral oncogene *v-src* (Hughes et al., 1987). Like all retroviruses, RCAS viruses express a glycoprotein as part of their envelope, which needs to bind to its target receptor for cell entry, and this interaction thus mediates cellular specificity. For the most commonly used RCAS virus, this receptor is tumor virus A (TVA), which is exclusively expressed in avian cells (Bates et al., 1993; Young et al., 1993). As a consequence, RCAS viruses can infect and propagate in avian cells, like the DF-1 cell line (Himly et al., 1998; Schaefer-Klein et al., 1998), but not mammalian cells. This can be exploited by equipping certain cell types with the TVA receptor, which thereby acquire exclusive susceptibility to RCAS infection (Bates et al., 1993; Young et al., 1993). Upon cell entry, the virus integrates into the host genome and drives expression of the experimentally introduced gene of interest.

In oncology research, the RCAS/TVA system is used to deliver oncogenes to certain cell types *in vivo* and thus mimic somatic alterations in these cells of origin. This is facilitated by the availability of many different transgenic mouse lines that have been designed to express *tv-a*, the gene that encodes for the TVA receptor, in specific cell types (Ahronian and Lewis, 2014). The transgenic Nestin-*tv-a* (Ntv-a) line, for example, expresses TVA in Nestin-positive cells (Holland et al., 1998). After embryonic development, Nestin expression in the brain is primarily detected in neural stem cells (NSCs) of the subventricular zone (SVZ) (Bernal and Arranz, 2018). The Ntv-a line has therefore widely been used for modeling gliomas (Hambardzumyan et al., 2009).

In 2011, our lab has published an RCAS/TVA-based mouse model for pediatric PA (Gronych et al., 2011). It is based on introducing the kinase domain of the human *BRAF* gene harboring the oncogenic V600E mutation into NSCs using the Ntv-a mouse line. The resulting tumors show strong similarities to human PA and corroborate that a single genetic alteration can be sufficient to drive tumorigenesis *in vivo*.

2 Aim of the thesis

Despite recent progress in the molecular characterization and classification of pLGGs, the separation of specific tumor types is still not fully established. This is also true for PA and PXA, which can be difficult to distinguish based on histology alone (Collins et al., 2015). Although PXA shows a much more aggressive growth pattern and a higher frequency of recurrence than PA (Ida et al., 2015), both are often simply grouped as "pLGG" in clinical trials without further stratification (Ater et al., 2016; 2012; Chintagumpala et al., 2015; Gnekow et al., 2017; Lassaletta et al., 2016). As a consequence, PA and PXA are treated similarly despite clearly different clinical courses. In the first part of this dissertation, I therefore aimed to uncover how exactly these distinct molecular subgroups of pLGG differ by analyzing their methylome and transcriptome. Given the significant role of the immune microenvironment in solid tumors and the recent advent of immune checkpoint inhibition in the pediatric setting, another aim of this thesis was to characterize the immune infiltration of both tumor types to elucidate their susceptibility to immunotherapy. While mouse models of pediatric cancer are urgently needed for preclinical studies, a GEMM for PXA has not yet been developed. Hence, I aimed to develop such a model using the RCAS/TVA system as a complement to the already existing PA model (Gronych et al., 2011). Detailed characterization of both models, also in comparison to their human counterparts, was aimed at supporting their validity.

The *KIAA1549:BRAF* fusion is the most abundant genetic alteration in human PA. Its presence is largely specific for this tumor type and is therefore of immense diagnostic and prognostic value. In addition, it is a biomarker for targeted therapy. Thus, it is critical to reliably detect the *KIAA1549:BRAF* fusion in molecular diagnostics. RNA-Seq has become increasingly popular for diagnostic purposes (Byron et al., 2016) and is now even applied to FFPE material (Stichel et al., 2019). However, the *KIAA1549:BRAF* fusion seems to be expressed at low levels and has proven difficult to reliably detect by RNA-Seq (Lin et al., 2012; Stichel et al., 2019; Tomić et al., 2017). In the second part of this dissertation, I therefore aimed to examine the detection reliability of the *KIAA1549:BRAF* fusion from RNA-Seq data and to optimize the fusion calling workflow to improve detectability.

3   Results

3.1   Comparison of human & murine PA/PXA characteristics

To unravel the molecular differences between PA and PXA, we assembled a cohort of 89 human pediatric fresh-frozen tumor samples that were subjected to different types of analysis (see 3.1.1). In a second step, we generated a PXA mouse model and compared it with an existing PA mouse model as well as with human tumors (see 3.1.2).

3.1.1   Human data analysis

3.1.1.1   PA and PXA show distinct methylome profiles

DNA methylation analysis has emerged as the primary tool for molecular tumor classification and is routinely used in our division to generate a tumor "fingerprint" and determine its identity. The t-distributed stochastic neighbor embedding (t-SNE)-based clustering of an existing set of 64,097 diverse tumor samples shows separation of the different tumor types and allows allocation of newly profiled samples (Figure 3.1a). We generated genome-wide DNA methylation profiles for all 89 PA/PXA tumor samples and were able to confirm their identity, as they clustered with the PA and PXA samples of the reference set (Figure 3.1b). Importantly, the PA and PXA clusters are clearly separated, suggesting that the underlying methylation profiles are consistently distinct from each other.
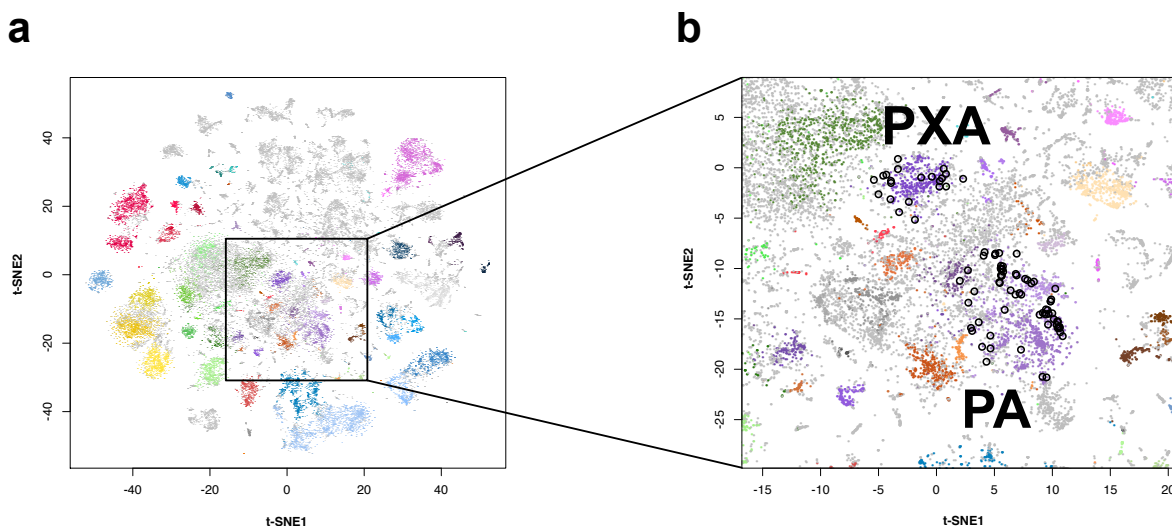
Figure 3.1 | t-SNE-based methylome analysis in comparison to the reference cohort

a) t-SNE analysis of genome-wide DNA methylation profiles of 64,097 tumor samples from various tissues. b) Enlargement of the area in a) containing the PA and PXA clusters. The encircled dots represent the 89 samples of the analysis cohort.

Individual t-SNE-based clustering of the analysis cohort without the reference set similarly revealed two distinct clusters, further supporting our results (Figure 3.2).
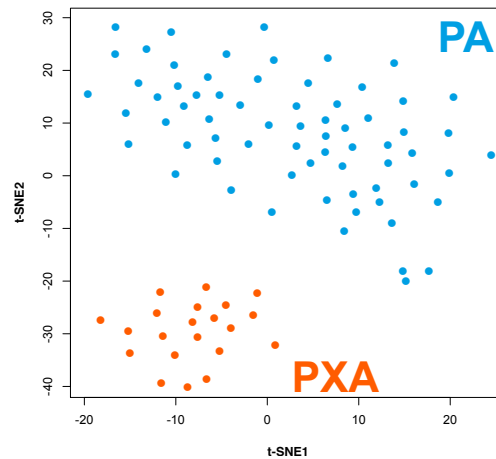
Figure 3.2 | Individual t-SNE-based methylome analysis of the PA/PXA cohort

t-SNE analysis of genome-wide DNA methylation profiles of the 89 PA/PXA samples constituting the analysis cohort.

3.1.1.2   PA and PXA feature different sets of genetic alterations

Subsequently, we performed bulk RNA sequencing for all 89 PA/PXA samples to further investigate the molecular differences of both groups. We also used this data to identify the underlying genetic alterations of all tumors, if not already known from previous studies (International Cancer Genome Consortium PedBrain Tumor Project, 2016; Jones et al., 2013), and found a mix of the different mutations, fusions and deletions usually observed in PA and PXA (Table 3.1). This data was supplemented by methylation array-derived copy number analysis to identify focal deletions, e.g. of the *CDKN2A/B* locus, and Sanger sequencing to confirm BRAF mutations. Furthermore, we added RNA-Seq data of two human brain tumor cell lines that served as controls for parts of the analysis. The BT-66 cell line is derived from a PA and features the typical *KIAA1549:BRAF* fusion (Selt et al., 2017), whereas the BT-40 cell line is derived from a PXA and is characterized by a BRAF V600E mutation and a loss of *CDKN2A/B* (Bid et al., 2013).

Table 3.1 | Cohort overview

	Group		#	Alteration details (#)
PA	*BRAF* Mut	**V600E, ins598/599T**	9	BRAF V600E (6) BRAF ins598T (2) BRAF ins599T (1)
	BRAF Fus	*KIAA1549* **Cereb**	10	*KIAA1549:BRAF* 16:9 (4) *KIAA1549:BRAF* 15:9 (6)
		KIAA1549 **Non-cereb**	12	*KIAA1549:BRAF* 16:9 (7) *KIAA1549:BRAF* 15:9 (5)
		Non-*KIAA1549*	5	*FAM131B:BRAF* (2) *GNAI1:BRAF* (2) *MKRN1:BRAF* (1)
	Non-*BRAF*	*NF1*	21	
		Others	11	*FGFR1* ITD (3) FGFR1 N546K (1) *SRCIN1:FGFR1* (1) *NACC2:NTRK2* (2) *QKI:NTRK2* (1) KRAS E63K (1) KRAS E63K R73M (1) *SRGAP3:RAF1* (1)
PXA	BRAF V600E	*CDKN2A/B* Del	14	
	BRAF WT	*CDKN2A/B* Del	5	*ETV6:NTRK2* (1) *ETV6:NTRK3* (1) *SPECC1L:NTRK2* (1) Putative *NTRK2* Fus (1) *EGFR:BRAF* (1)
	BRAF WT	*CDKN2A/B* WT	2	*FGFR2:CLIP2* (1) Unknown (1)
	Control	**BT-66 (PA), BT-40 (PXA)**	2	
			91	

Overview of the 89 PA/PXA samples of the analysis cohort and the two cell lines used as controls. The different PA and PXA subgroups, their genetic alterations as well as the number of samples in each group are indicated. Mut = mutation, Fus = fusion, Del = deletion, WT = wildtype, Cereb = cerebellar, ITD = internal tandem duplication.

For the PA samples, the identified alterations could be grouped into *BRAF* mutations, *BRAF* fusions and alterations not involving *BRAF* (Table 3.1). The mutations in *BRAF* mainly caused the common V600E amino acid substitution, while some samples featured the rarer ins598T and ins599T insertions. As expected, the vast majority of *BRAF* fusions had *KIAA1549* as the 5' partner (16:9 or 15:9 exon combination), both in the cerebellum and elsewhere. In addition, there were some tumors with a *FAM131B:BRAF*, *GNAI1:BRAF* or *MKRN1:BRAF* fusion. The non-BRAF PA samples harbored *NF1* mutations, *FGFR1* internal tandem duplications (ITDs), *FGFR1* mutations (FGFR1 N546K) and *FGFR1* fusions (*SRCIN1:FGFR1*), *NTRK2* fusions

(*NACC2:NTRK2, QKI:NTRK2*) as well as *KRAS* mutations (KRAS E63K, R73M) and one *RAF1* fusion (*SRGAP3:RAF1*).

For the PXA samples, two thirds showed the typical combination of BRAF V600E and a focal deletion of the *CDKN2A/B* locus (Table 3.1). Five other samples also had a *CDKN2A/B* deletion but lacked a *BRAF* mutation. Instead, they were mostly driven by *NTRK2* or *NTRK3* fusions (*ETV6:NTRK2, ETV6:NTRK3, SPECC1L:NTRK2*). Interestingly, one sample harbored a so far unknown in-frame fusion of *epidermal growth factor receptor* (*EGFR*) and *BRAF* (Figure 3.3). As with the *KIAA1549:BRAF* fusion, the kinase domain of *BRAF*, which drives MAPK signaling, is retained in this fusion. In contrast, the tyrosine kinase domain of *EGFR* is not fully included in the expressed fusion protein. The same tumor also expressed an in-frame *BRAF:EGFR* fusion, suggesting that both fusions were generated by the same genetic event.
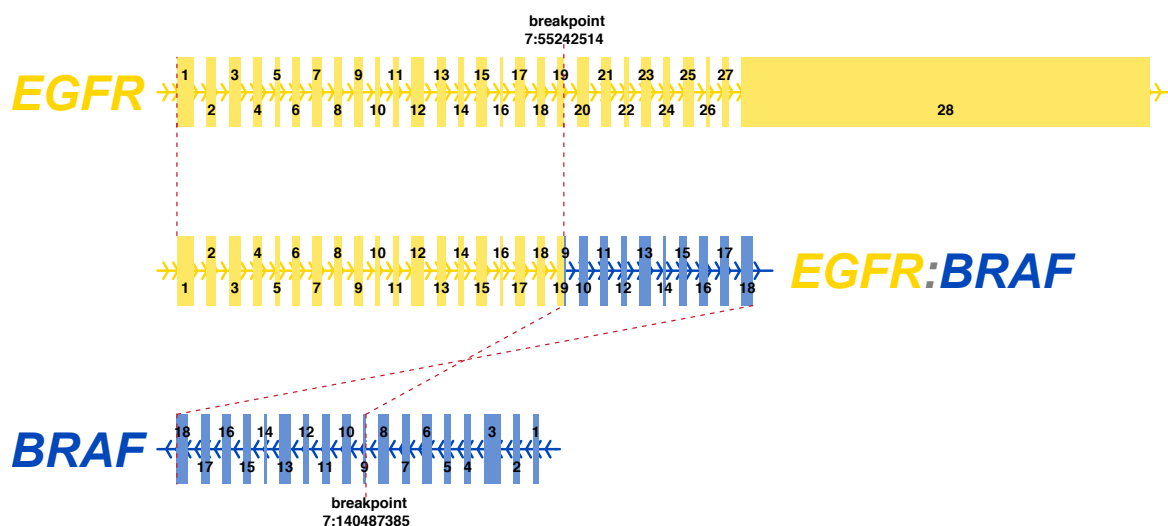
Figure 3.3 | *EGFR:BRAF* fusion

Illustration of the *EGFR:BRAF* fusion identified in the RNA-Seq data of one PXA tumor. The breakpoints and the exons being retained in the fusion transcript are depicted. For *EGFR*, the breakpoint is between exons 19 and 20. For *BRAF*, the breakpoint is between exons 8 and 9. The resulting *EGFR:BRAF* fusion transcript thus has a 19:9 exon combination.

Furthermore, two samples in the PXA group stood out as they lacked the typical deletion of the *CDKN2A/B* locus. Analysis of the RNA-Seq data revealed that these samples indeed expressed both *CDKN2A* and *CDKN2B*, whereas the *CDKN2A/B*-deleted PXA samples showed the expected reduction in expression (Figure 3.4). In terms of activating RAS/MAPK alterations, one of the two samples harbored a *FGFR2:CLIP2* fusion. The driving alteration for the other sample could not be identified.
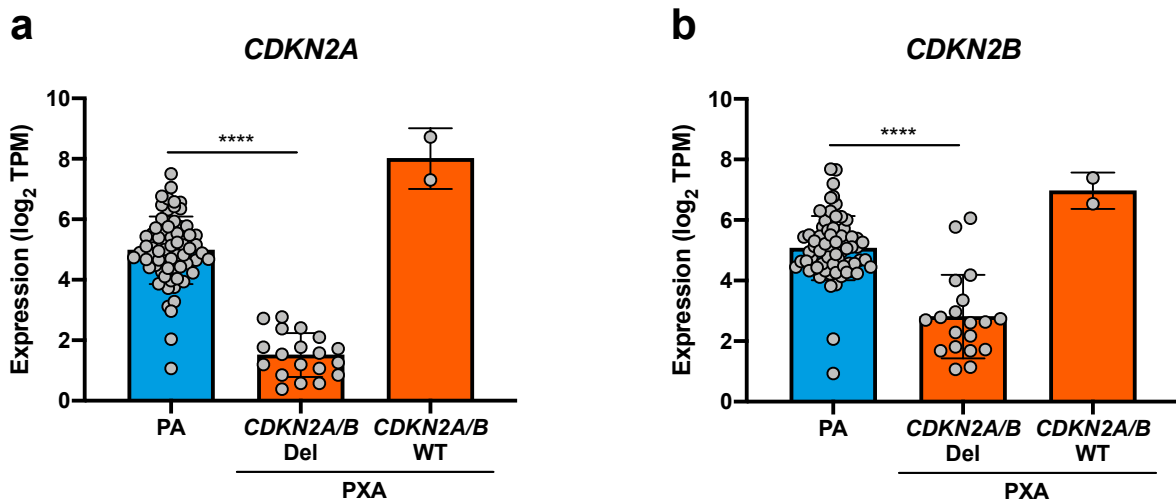
Figure 3.4 | *CDKN2A/B* expression in PA and PXA

Expression of a) *CDKN2A* and b) *CDKN2B* in PA as well as PXA with and without a loss of the *CDKN2A/B* locus. Mean ± SD. One-way ANOVA followed by Tukey multiple comparisons test.

3.1.1.3 PA and PXA activate MAPK signaling to a different degree

Having identified the underlying genetic alterations in our analysis cohort, we wondered if the different PA and PXA subgroups are differently capable of driving MAPK signaling. To this end, we calculated the MAPK Pathway Activity Score (MPAS), which is based on the expression of 10 MAPK target genes (Wagle et al., 2018). The PXA tumors showed a significantly higher MPAS compared with the PA group (Figure 3.5). Interestingly, the two *CDKN2A/B* wildtype PXA samples showed an MPAS more similar to the PA samples, albeit this has to be verified with more samples. Of note, PA tumors driven by a *BRAF* fusion seem to have a lower degree of MAPK activation compared with the other PA subgroups, potentially suggesting that *BRAF* fusions are moderately less potent inducers of MAPK signaling.
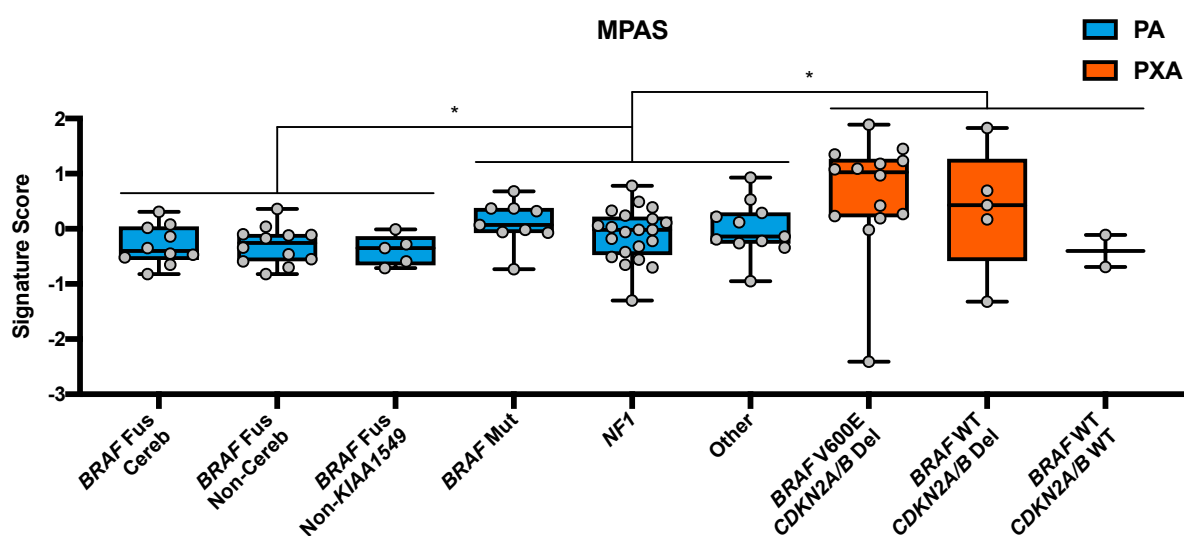
Figure 3.5 | MAPK pathway activity in different subgroups of PA and PXA

Estimation of the MAPK pathway activity in the previously presented subgroups of PA and PXA by calculating their MAPK Pathway Activity Score (MPAS). For statistical testing, the indicated groups of three were combined, respectively. Box with median and interquartile range and whiskers from minimum to maximum. Kruskal-Wallis test followed by Dunn's multiple comparisons test.

3.1.1.4 PA and PXA show distinct transcriptome profiles

Switching to a more global view of the transcriptome, we performed t-SNE analysis using the generated RNA-Seq data and again found PA and PXA to separate (Figure 3.6). Even though this difference was not as pronounced as for the methylome analysis, it clearly indicates that distinct transcriptional programs are active in these two tumor types.
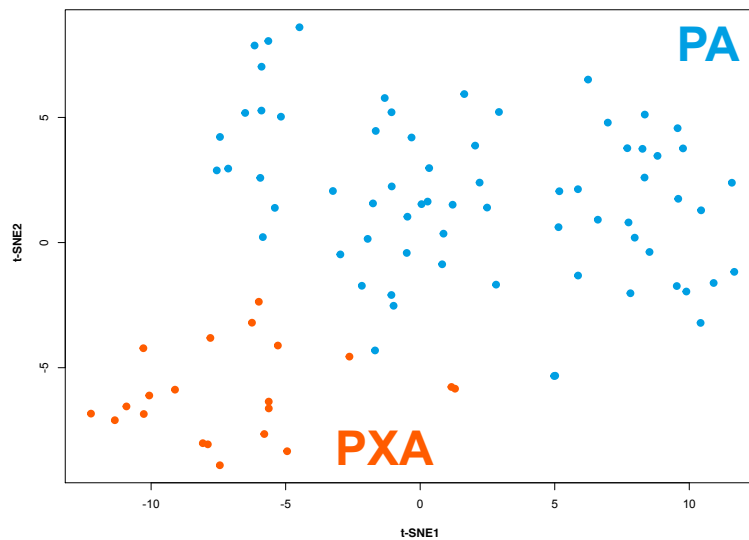
Figure 3.6 | t-SNE-based transcriptome analysis of the PA/PXA cohort

t-SNE analysis of RNA-Seq data of all 89 PA/PXA samples in the analysis cohort using the 1000 most variably expressed genes.

To investigate this in more detail, we performed differential gene expression analysis. A basic gene ontology (GO) analysis of the 1000 most differentially expressed genes showed that cell cycle- and development-associated genes were more frequently found in this list than would be expected by chance (Table 3.2). Signal transduction genes were borderline significant as well (p = 0.05).

Table 3.2 | Basic ontology analysis

Group	In set	Total	%	p-value
All	610	9536	6.4	1.00
DNA repair	16	207	7.7	0.43
Apoptosis	35	623	5.6	0.43
Cell cycle	69	483	14.3	**1.4e-12**
Development	145	1582	9.2	**6.8e-06**
Differentiation	50	670	7.5	0.26
Drug target	62	1142	5.4	0.18
Kinase	43	640	6.7	0.74
Membrane	292	4819	6.1	0.34
Signal transduction	164	2977	5.5	0.05
Transcription factor	54	809	6.7	0.75

Basic ontology analysis of the 1000 most differentially expressed genes between PA and PXA using their annotation to one or more ontology groups. The number of genes in each group is depicted for the 1000 genes in this set as well as in total for the entire genome. The resulting percentages as well as the p-values are shown as well. Chi-square test.

A more advanced GO analysis confirmed that the difference in gene expression between PA and PXA is mainly caused by these processes. The most significant GO term hit was "nervous system development" (p = 1.2e-24), which showed a very clear separation of PA and PXA (Figure 3.7a). This includes genes like *orthodenticle homeobox 1* (*OTX1*), *carbonic anhydrase 10* (*CA10*) and *GDNF family receptor alpha 1* (*GFRA1*) (Figure 3.7b-d). These results suggest a key difference in the regulatory circuits involved in the growth of PA and PXA.
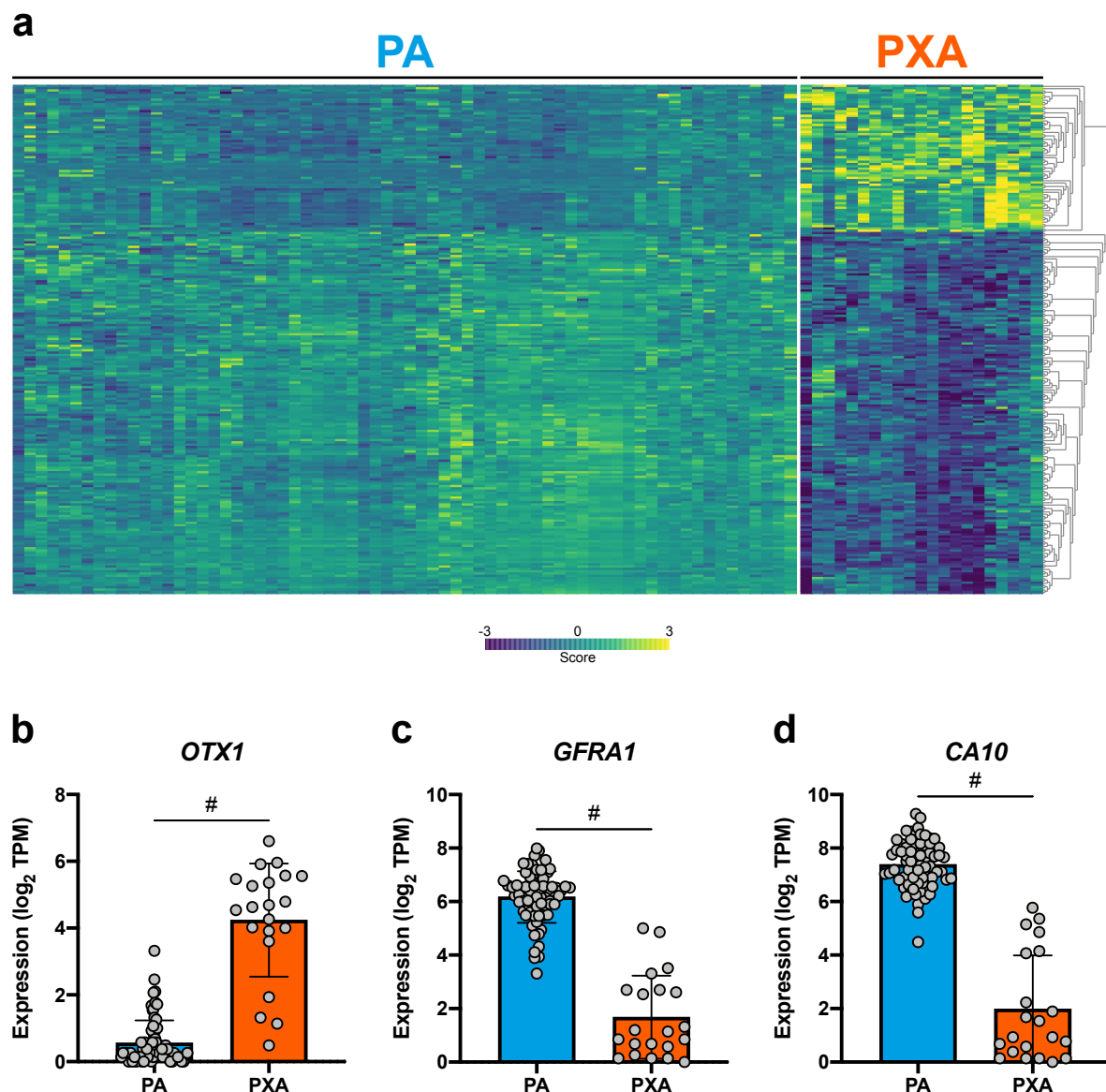
Figure 3.7 | Differentially expressed genes related to nervous system development

a) Heatmap of the 223 differentially expressed genes between PA and PXA related to nervous system development. b)-d) Expression of three exemplary genes from this list, *OTX1*, *GFRA1* and *CA10*. Mean ± SD. # = significant according to differential gene expression analysis.

Again mimicking the basic ontology analysis, cell cycle-related GO terms like "mitotic cell cycle" (p = 1.1e-17), "mitotic sister chromatid segregation" (p = 4.9e-17) or "cell division" (p = 9.4e-14) were among the most differentially regulated processes. As expected from its more proliferative nature, the majority of mitotic cell cycle-associated genes had a higher expression

in PXA (Figure 3.8a), like *E2F transcription factor 7* (*E2F7*) or *cyclin O* (*CCNO*) (Figure 3.8b-c). However, some genes like *phosphotyrosine interaction domain containing 1* (*PID1*; Figure 3.8d) were downregulated in PXA.
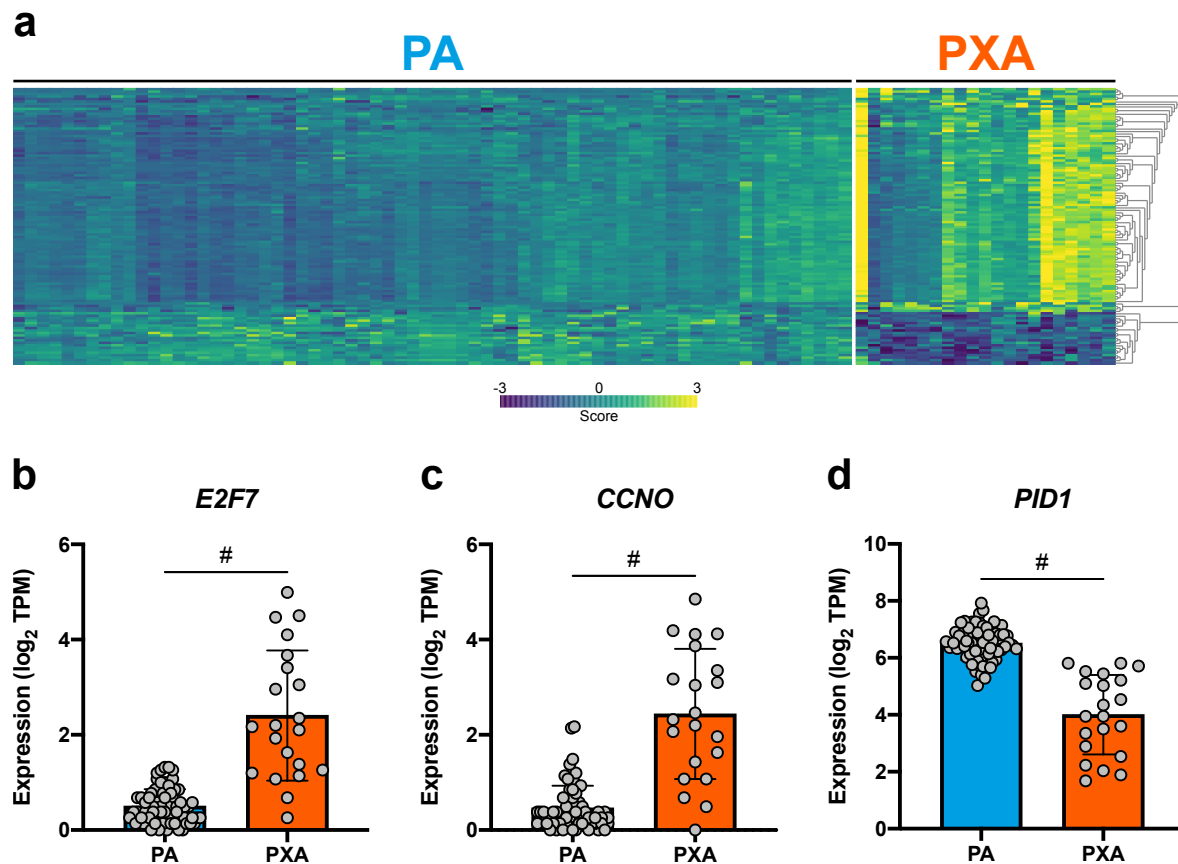
Figure 3.8 | Differentially expressed genes related to mitotic cell cycle

a) Heatmap of the 112 differentially expressed genes between PA and PXA related to mitotic cell cycle. b)-d) Expression of three exemplary genes from this list, *E2F7*, *CCNO* and *PID1*. Mean ± SD. # = significant according to differential gene expression analysis.

3.1.1.5 PA and PXA show qualitative differences in immune infiltration

3.1.1.5.1 Quantitative comparison

Given the important role of the tumor microenvironment, we next extracted information on the infiltrating non-neoplastic cells from the generated bulk RNA-Seq data. We used the ESTIMATE algorithm to deconvolute the data and get a first evaluation of the immune and stromal cell content of the tumors. We included the data of the two human cell lines, BT-66 and BT-40, in this analysis, as they only consist of tumor cells and are hence free of stromal and immune cells. The means of their calculated ESTIMATE scores were used to set the baselines for a zero proportion of immune cells and stromal cells, respectively. Both PA and PXA showed a strong signal of immune infiltration with immune scores of about 1000, but with no significant quantitative difference between the groups (Figure 3.9a). The estimation of the stromal cell content yielded very similar results (Figure 3.9b).

a ESTIMATE Immune
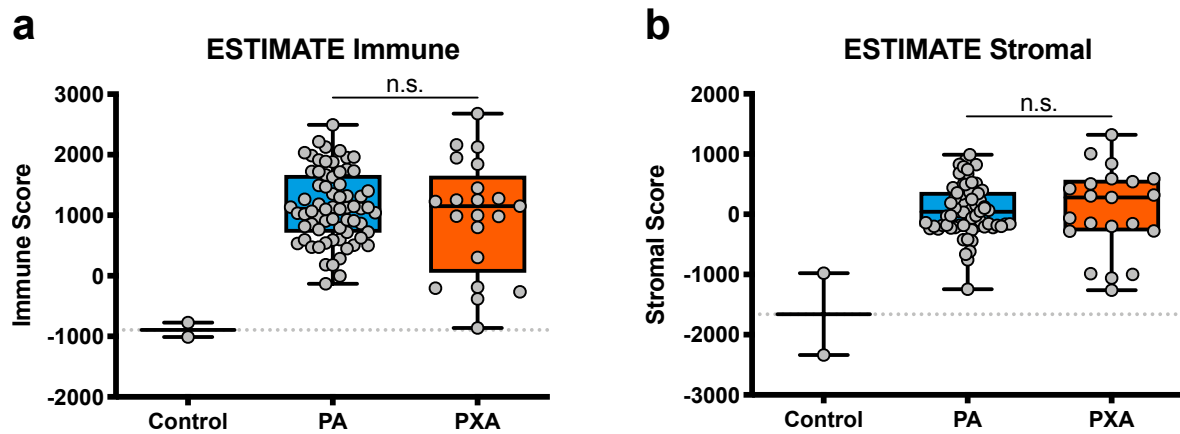
b ESTIMATE Stromal

Figure 3.9 | ESTIMATE analysis of immune and stromal cells

ESTIMATE analysis of human PA and PXA to get an evaluation of the a) immune and b) stromal cell content. Two brain tumor cell lines were used as negative controls to set the baseline. Box with median and interquartile range and whiskers from minimum to maximum. One-way ANOVA followed by Tukey multiple comparisons test.

3.1.1.5.2   Qualitative comparison

Having identified a similarly high degree of immune infiltration in PA and PXA, we continued by investigating if there was a qualitative difference between both groups. As some immune cell types, like microglia, are naturally present in healthy brain tissue, we switched to a largely overlapping microarray (rather than RNA-Seq) dataset of human PA and PXA, as this allowed inclusion of expression data from normal human cerebellum and cerebral cortex as controls for PA and PXA, respectively. Using gene signatures that are characteristic for different immune cell populations, we were able to detect significant qualitative differences not only between tumors and controls but also between tumor types (Figure 3.10). Both PA and PXA showed a stronger signature of infiltrating microglia/macrophages relative to normal brain (Figure 3.10a). Although the mean signature score for PXA was higher than for PA, this difference was not statistically significant. For activated CD4 T cells, both tumor types again showed higher signature scores compared with non-cancerous tissue (Figure 3.10b). Interestingly, the mean score was significantly higher in PA compared with PXA. In contrast, only PXA presented a significantly elevated signature of activated CD8 T cells, while PA was indistinguishable from normal brain (Figure 3.10c).
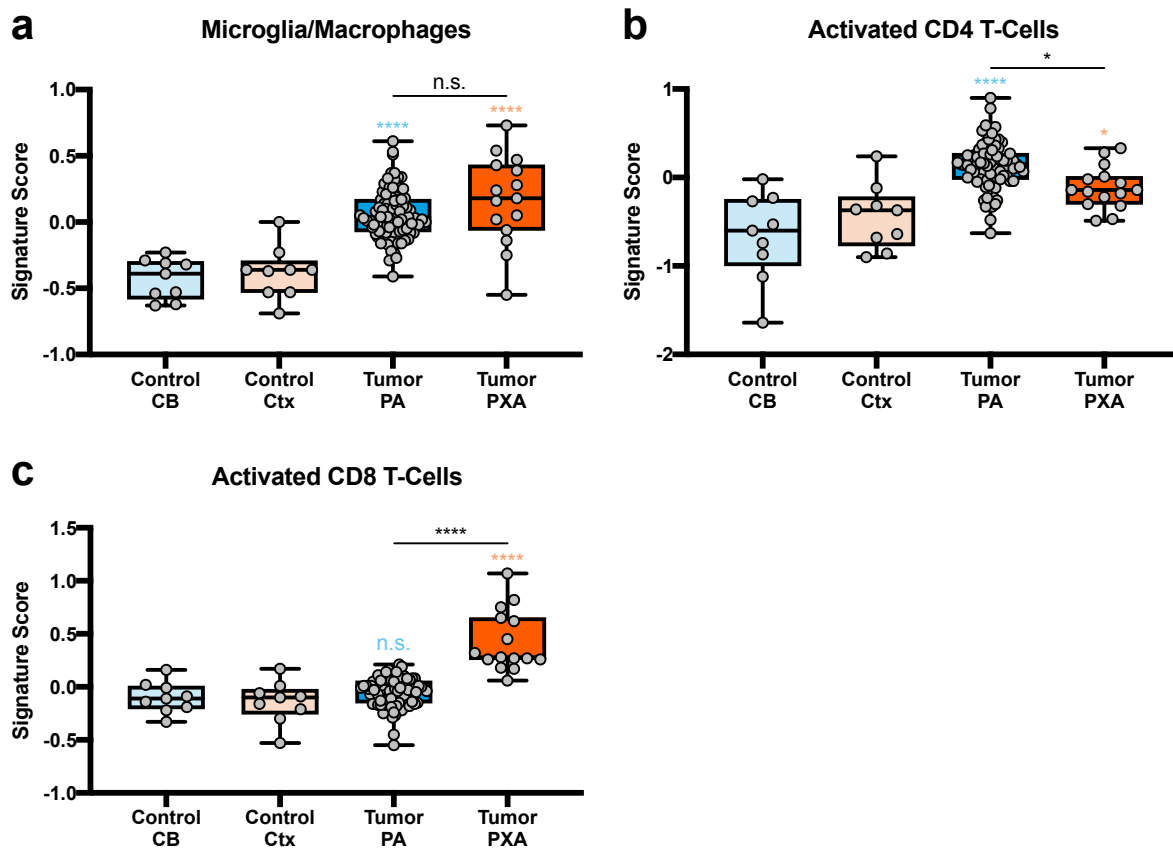
Figure 3.10 | Gene signature analysis of different immune cell subsets

Gene signature analysis of human PA and PXA to get an evaluation of the a) microglia/macrophage, b) activated CD4 T cell and c) activated CD8 T cell content. Healthy cerebellum and cerebral cortex are shown in comparison. Box with median and interquartile range and whiskers from minimum to maximum. One-way ANOVA followed by Tukey multiple comparisons test. Colored significance levels refer to respective control.

3.1.2 Mouse model development and characterization

3.1.2.1 Expansion of a PA mouse model by *Cdkn2a* KO establishes a PXA model

Mouse models closely recapitulating key characteristics of the corresponding human primary tumors are a vital tool for preclinical research. They can not only be used to get a more detailed understanding of the biological processes involved in tumor initiation and progression but also for preclinical studies of novel or established therapeutic compounds. One example is the PA mouse model that has been developed by our lab (Gronych et al., 2011). It is based on the RCAS-mediated introduction of a human sequence coding for the V600E-mutated BRAF kinase domain into newborn Ntv-a mice. To complement the existing PA model and pave the way for further research on PXA, we generated an additional model by adding a second genetic hit. As described before, PXA tumors, in addition to expressing BRAF V600E, usually show a deletion of the *CDKN2A/B* locus (see Table 3.1). For the PXA model, we therefore introduced the oncogenic *BRAF* sequence into Ntv-a *Cdkn2a*$^{-/-}$ mice (Figure 3.11). These mice indeed developed tumors, which were subsequently analyzed in more detail.
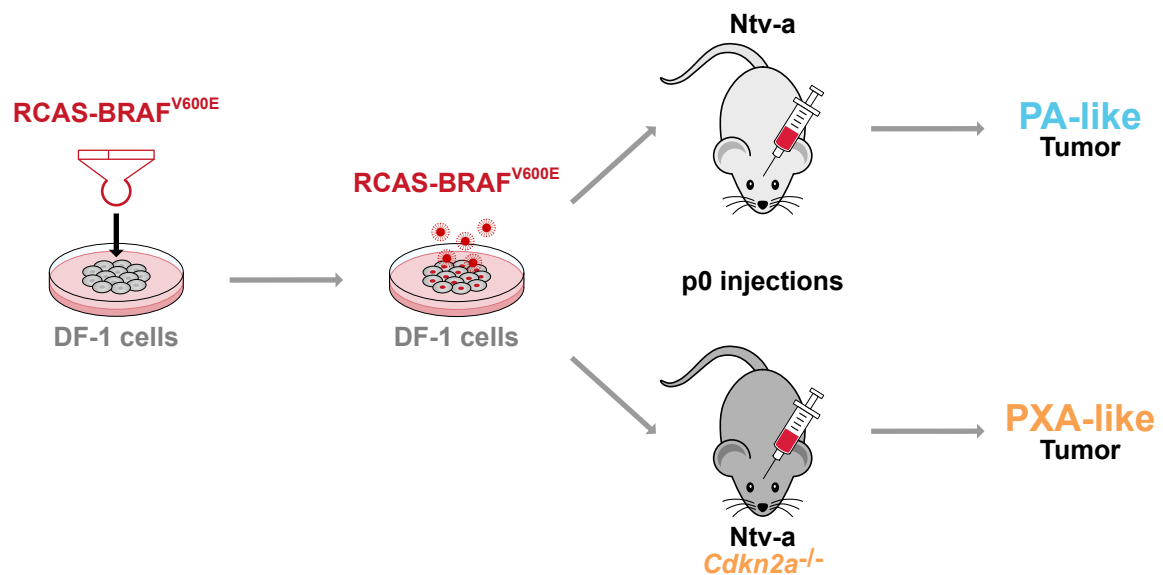
Figure 3.11 | Overview of PA and PXA mouse models

Illustration of the generation of PA- and PXA-like tumors in the respective mouse models. DF-1 cells, which allow propagation of RCAS viruses, are transfected with the RCAS vector containing the oncogenic human sequence coding for the BRAF V600E kinase domain. The DF-1 cells thereupon start producing the RCAS virus harboring the BRAF sequence. The virus-expressing DF-1 cells are subsequently injected into the brain of newborn (p0) mice. Injection into Ntv-a mice generates PA-like tumors, whereas injection into Ntv-a $Cdkn2a^{-/-}$ mice generates PXA-like tumors.

3.1.2.2 Both mouse models resemble human growth patterns

Comparing the growth characteristics of both models revealed striking differences. PA-like tumors started to grow initially but then ceased proliferation. Their benign growth was not lethal for the mice. In contrast, the newly generated PXA model proved more aggressive (Figure 3.12a). Mice developed symptoms after six to eight weeks and had to be sacrificed thereupon.
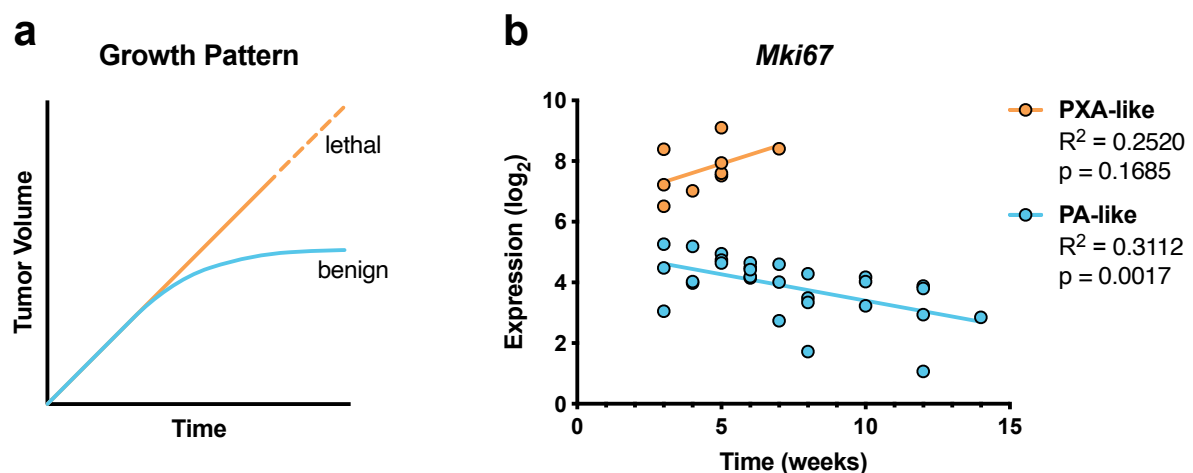
Figure 3.12 | Growth characteristics of the PA-like and PXA-like model

a) Schematic observation-based growth curve of both mouse models over time. b) Expression of $Mki67$ extracted from microarray data of both models over time. Each dot indicates one mouse that was analyzed at the respective time point. The goodness of fit (R^2) as well as the p-values indicating if the slope of the linear regression lines significantly deviate from zero are shown.

This was also reflected by the expression of the proliferation marker gene *marker of proliferation Ki-67* (*Mki67*; Figure 3.12b). Its level was generally lower in the PA model and decreased over time. In contrast, the PXA-like tumors showed higher levels of *Mki67*, which further increased during tumor progression. Both models therefore faithfully resemble growth patterns and proliferation characteristics of their human counterparts.

3.1.2.3 Both mouse models show typical histological features of human tumors

Next, we investigated the histology of the murine tumors. Similar to the published morphological concordance of the PA model and human PA (Gronych et al., 2011), the PXA model closely resembled human PXA. The PXA-like tumors grew close to the ventricle, often ventral of the hippocampus, and were characterized by the typical pleomorphic combination of giant and spindle cells. Notably, they also showed a strong staining for reticulin fibers, which surrounded the individual tumor cells (Figure 3.13). This staining pattern is very typical for human PXA and rarely observed in other brain tumors, let alone normal CNS tissue, again proving a close resemblance to the human counterpart.
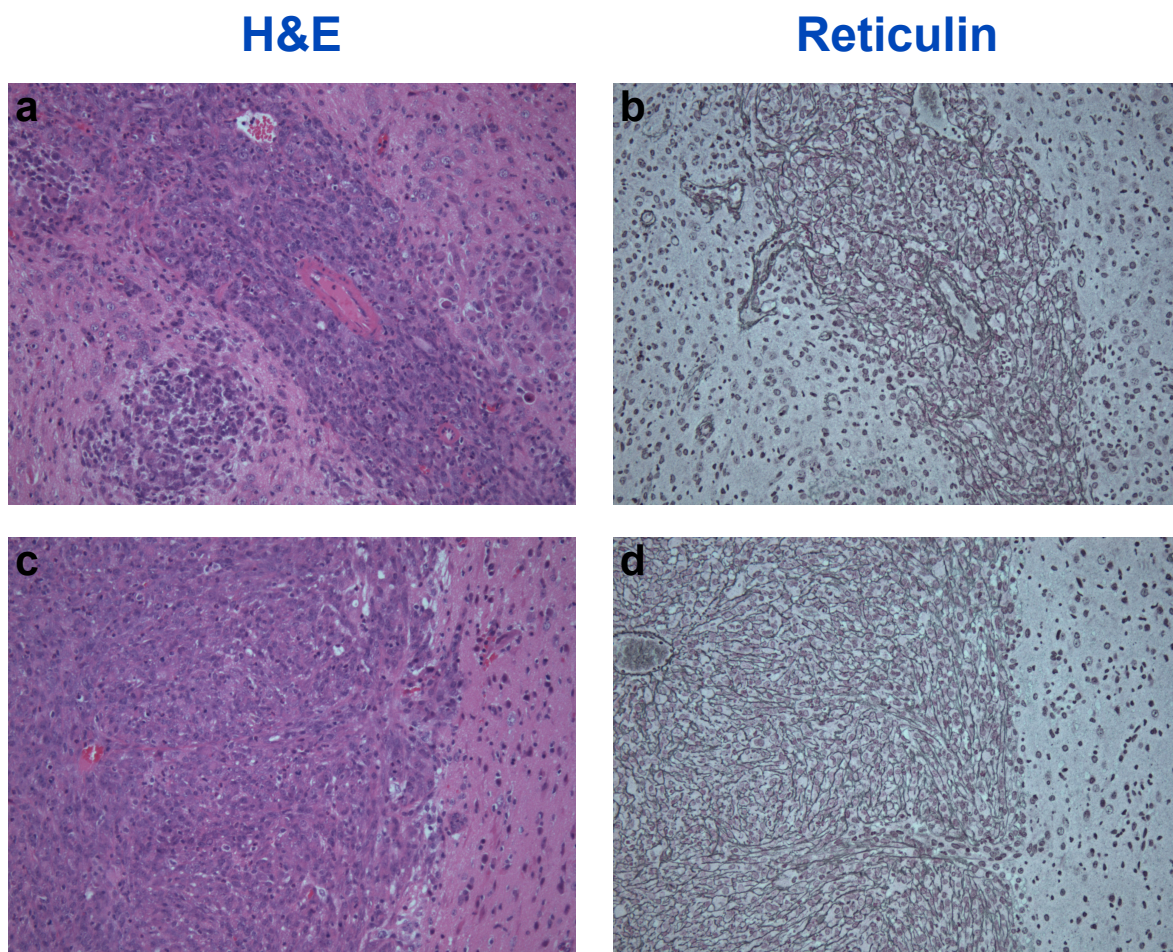
Figure 3.13 | H&E and reticulin staining of murine PXA-like tumors

a),c) Hematoxylin and eosin (H&E)- and b),d) reticulin-stained slides from two different areas of a PXA-like tumor showing diffuse tumor regions and a superficial node, respectively. 100x magnification.

3.1.2.4 Both mouse models show similar expression patterns as human tumors

3.1.2.4.1 Individual genes of interest

Going one step further, we performed microarray expression analysis of the murine tumors and normal brain tissue from non-injected age-matched animals. Using this data, we had a look at the expression of individual genes of interest (Figure 3.14). In addition to *Mki67* (see Figure 3.12b), the PXA-like tumors had a higher expression of another typical proliferation marker gene, *DNA topoisomerase II alpha* (*Top2a*), compared with the PA-like tumors (Figure 3.14a). This was also true for the TWEAK receptor gene *TNF receptor superfamily member 12A* (*Tnfrsf12a*, also known as *Fn14*) and *periostin* (*Postn*) (Figure 3.14c,e), both of which are known to be upregulated in HGGs and to be associated with more malignant tumor growth (Mikheev et al., 2015; Perez et al., 2016; Zhou et al., 2015). Given the success of immune checkpoint inhibition in some tumor entities, we also checked expression of the suppressive immune checkpoint gene *Cd274*, also known as *Pd-l1*, which is expressed on tumor cells. The expression of *Pd-l1* was significantly upregulated in PXA-like tumors compared with the PA model (Figure 3.14g).

For verification, we checked expression of these genes in the aforementioned microarray dataset of human PA, PXA, cerebellum and cerebral cortex. Although the relative gene expression in tumor and corresponding normal tissue did not reflect the results obtained in the mouse model, all four genes notably were also significantly upregulated in human PXA compared with PA (Figure 3.14b,d,f,h). Analysis of our human RNA-Seq dataset confirmed these differences between PA and PXA (data not shown).
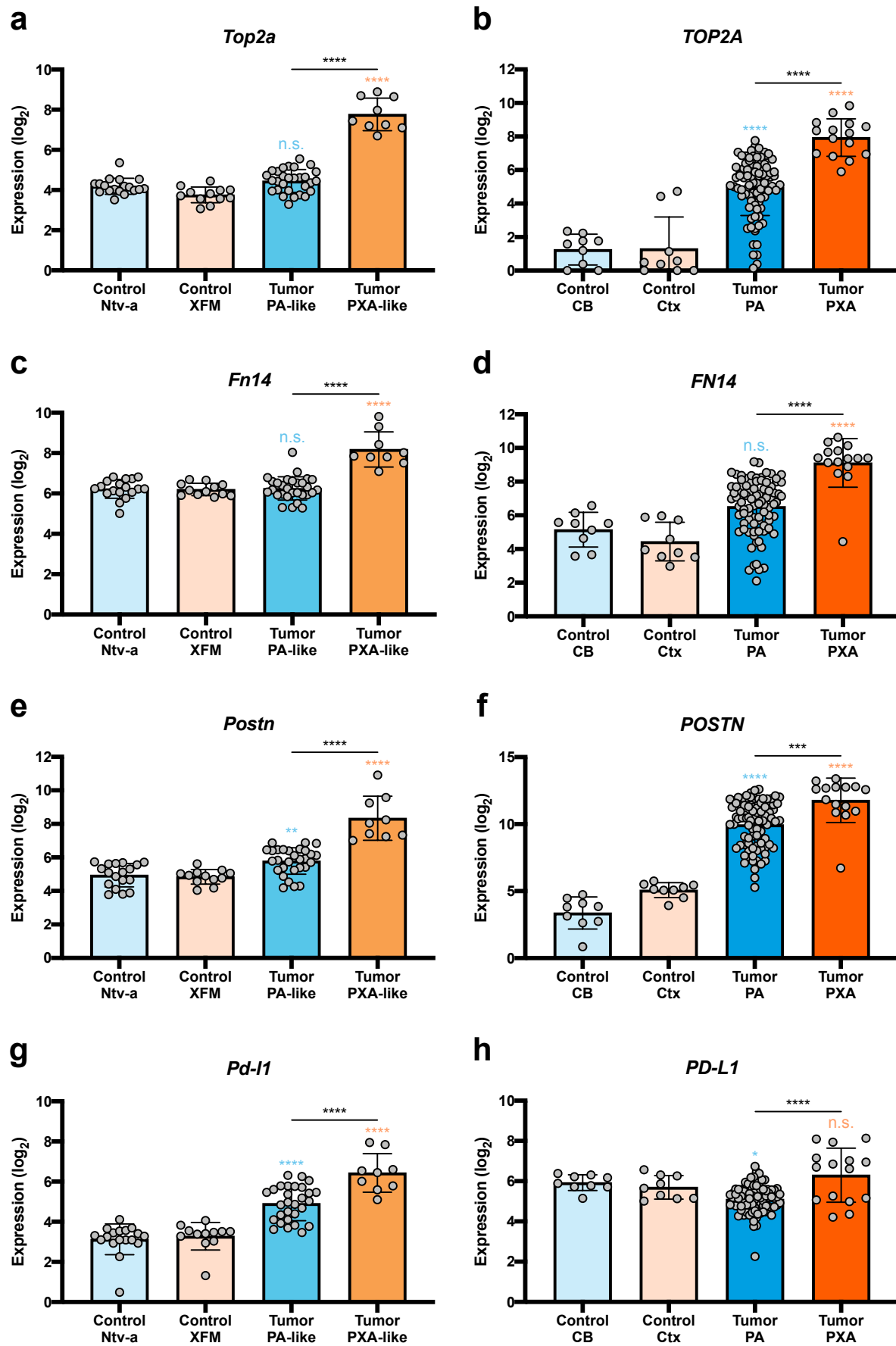
Figure 3.14 | Expression of selected genes in both mouse models and their human counterparts

Expression of a) *Top2a*, c) *Fn14*, e) *Postn* and g) *Pd-l1* in murine PA- and PXA-like tumors as well as normal brain tissue of non-injected Ntv-a and Ntv-a *Cdkn2a*[-/-] mice of the same age based on microarray data. In comparison,

expression of b) TOP2A, d) FN14, f) POSTN and h) PD-L1 in human PA and PXA tumors as well as normal cerebellum and cerebral cortex based on microarray data. *Fn14* and *Pd-l1* are also known as *Tnfrsf12a* and *Cd274*, respectively. The same holds true for the human orthologs. Mean ± SD. One-way ANOVA followed by Tukey multiple comparisons test. Colored significance levels refer to respective control. XFM = Ntv-a *Cdkn2a$^{-/-}$*, CB = cerebellum, Ctx = cerebral cortex.

3.1.2.4.2 MAPK pathway activation

Next, we also wanted to compare the relative degree of MAPK pathway activation between PA and PXA in mouse models and human tumors. Therefore, we evaluated the expression of the 10 marker genes constituting the MPAS. Although the MPAS level in murine PA-like tumors was indistinguishable from normal brain, PXA-like tumors showed a significantly higher score compared with the PA model (Figure 3.15a). This was partly reflected by the analysis of the human microarray data. Similar to the results of the human RNA-Seq dataset (see Figure 3.5), PXA tumors had a higher MPAS compared with PA tumors (Figure 3.15b). However, also the PA tumors showed stronger MAPK activation compared with normal brain.
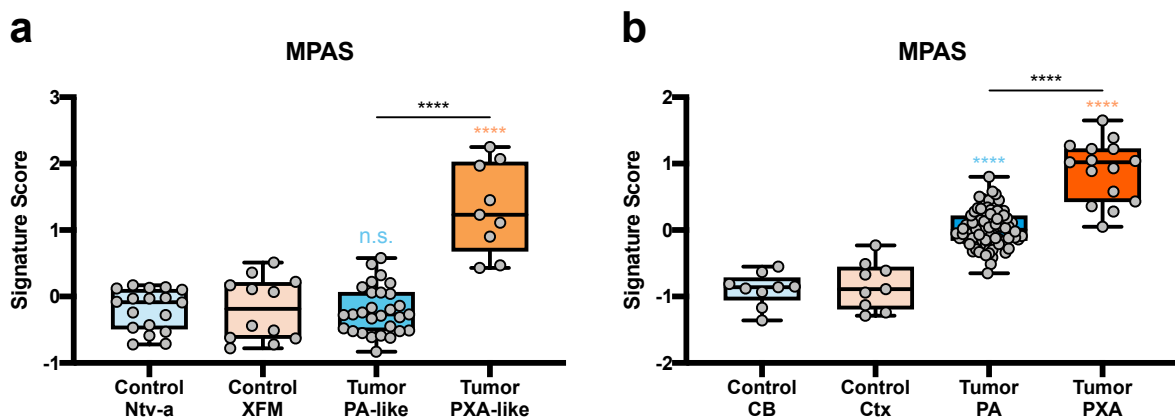
Figure 3.15 | MAPK pathway activity in both mouse models and their human counterparts

a) Estimation of MAPK pathway activity in murine PA- and PXA-like tumors as well as normal brain tissue of non-injected Ntv-a and Ntv-a *Cdkn2a$^{-/-}$* mice of the same age in terms of their MAPK Pathway Activity Score (MPAS) based on microarray expression data. b) In comparison, estimation of the MAPK pathway activity in human PA and PXA tumors as well as normal cerebellum and cerebral cortex by calculating their MPAS based on microarray expression data. Box with median and interquartile range and whiskers from minimum to maximum. One-way ANOVA followed by Tukey multiple comparisons test. Colored significance levels refer to respective control. XFM = Ntv-a *Cdkn2a$^{-/-}$*, CB = cerebellum, Ctx = cerebral cortex.

3.1.2.5 A cross-species analysis validates the authenticity of both mouse models

As a continuation of the analysis of selected genes, we aimed to compare the gene expression of human and murine tumors in a more global and unbiased manner. In collaboration with Konstantin Okonechnikov (DKFZ), we performed a cross-species analysis, which is based on identifying differentially expressed genes and their orthologs in other species. To make this analysis as accurate as possible, we generated RNA-Seq data for four PA-like and five PXA-like mouse tumors and compared it with RNA-Seq data of those human tumors most closely matching the model characteristics. This means that we only used human PA with a *BRAF* mutation and supratentorial location. Similarly, we only used human PXA with a combination of *BRAF* mutation and *CDKN2A/B* deletion as well as supratentorial location. As expected,

principle component analysis (PCA) based on the differentially expressed genes again showed a clear separation of human PA and PXA (Figure 3.16a). Importantly, the murine PA- and PXA-like tumors clustered with their human counterparts, i.e. clustering was based on the overall tumor type and not on the species. This was confirmed by unsupervised hierarchical clustering, which also separated samples by tumor type rather than species (Figure 3.16b). These results convincingly underline that the generated mouse models faithfully resemble human PA and PXA.

a

b
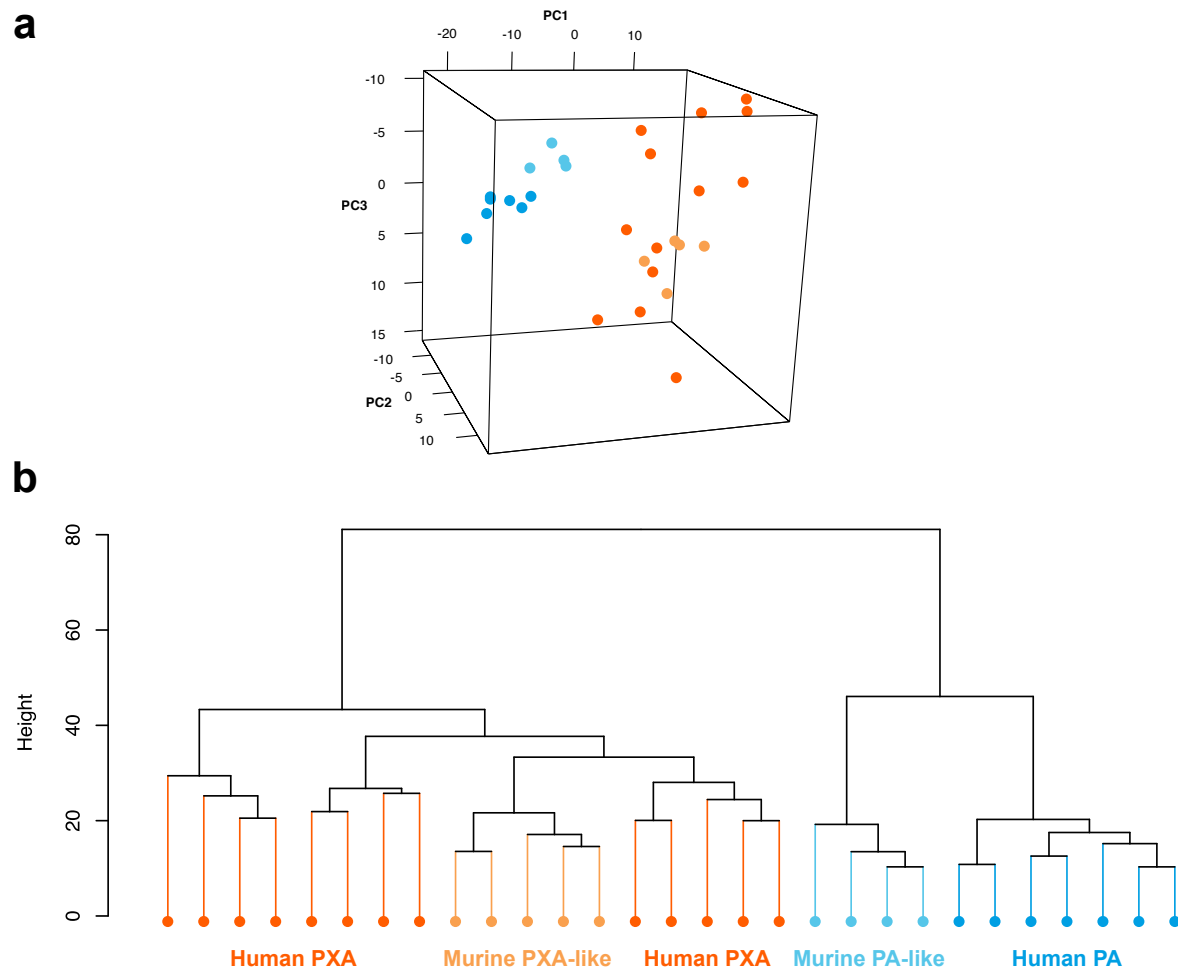
Figure 3.16 | Cross-species comparison of human and murine PA and PXA

Cross-species comparison based on RNA-Seq data of murine PA- and PXA-like and human PA and PXA tumors visualized by a) principle component analysis (PCA) and b) unsupervised hierarchical clustering. All human cases were supratentorial and had a *BRAF* mutation. All human PXA cases were additionally characterized by *CDKN2A/B* deletion. PC = principle component.

3.1.2.6 Both mouse models show similar immune infiltration as human tumors

Given this high degree of similarity, we finally wanted to check if the mouse models also show similar qualitative differences in immune infiltration compared with the human tumors (see Figure 3.10). We again deconvoluted the bulk expression data using gene signatures to identify differences. As in the human analysis, both PA- and PXA-like mouse tumors showed a stronger signature of microglia/macrophages compared with normal brain (Figure 3.17a). However, the score in PXA-like tumors was now significantly higher than in PA-like tumors. In terms of T

cells, PXA-like tumors had a weaker signature of activated CD4 T cells (Figure 3.17b) and a stronger signature of activated CD8 T cells (Figure 3.17c) compared with PA-like tumors, which is similar to the results obtained in the human analysis.
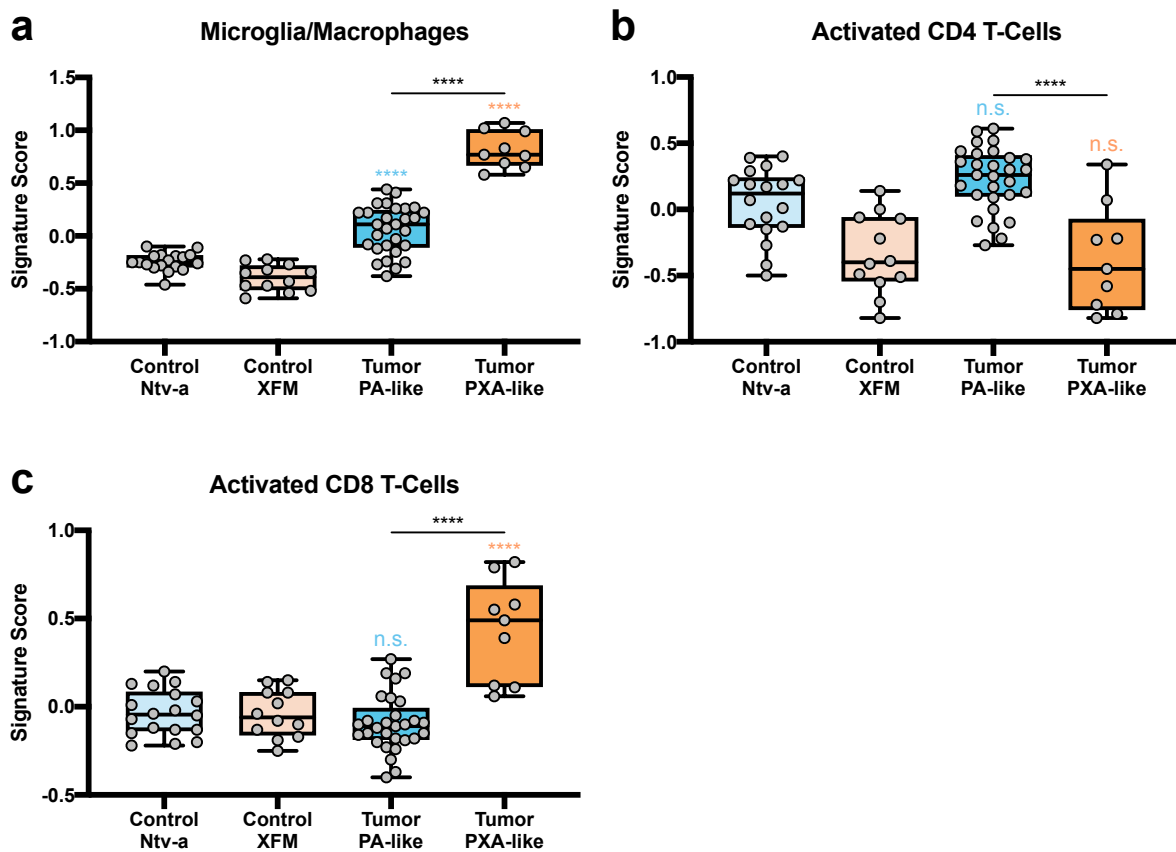
Figure 3.17 | Gene signature analysis of different immune cell subsets in both mouse models

Gene signature analysis of murine PA- and PXA-like tumors as well as normal brain tissue of non-injected Ntv-a and Ntv-a *Cdkn2a*[-/-] mice of the same age to get an evaluation of the a) microglia/macrophage, b) activated CD4 T cell and c) activated CD8 T cell content. Box with median and interquartile range and whiskers from minimum to maximum. One-way ANOVA followed by Tukey multiple comparisons test. Colored significance levels refer to respective control. XFM = Ntv-a *Cdkn2a*[-/-], CB = cerebellum, Ctx = cerebral cortex.

In summary, we identified fundamental molecular differences between PA and PXA on multiple levels that were faithfully recapitulated by our corresponding mouse models.

3.2   *KIAA1549:BRAF* fusion detection from RNA-Seq data

3.2.1   Standard workflow

For the identification of fusion genes in the generated RNA-Seq dataset of human PA and PXA (see 3.1.1.2), we employed our standard workflow involving STAR and Arriba (Figure 3.18). STAR is an algorithm that aligns the raw data to the reference genome, while Arriba identifies fusion genes on the basis of the aligned data. In the course of the analysis, we realized that the *KIAA1549:BRAF* fusion proved difficult to detect with these standard settings, and decided to investigate this further.
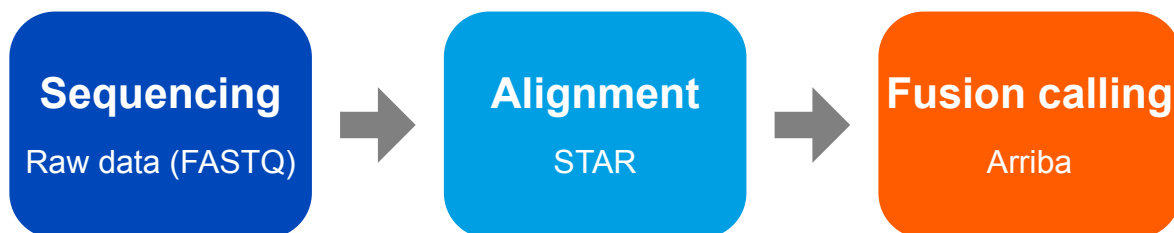
Figure 3.18 | Standard fusion calling workflow

Illustration of our standard analysis workflow to identify gene fusions from RNA-Seq data. Sequencing generates raw FASTQ files, which contain all recorded reads. These reads are subsequently aligned to the reference genome by the STAR algorithm. Eventually, Arriba identifies split reads and discordant mates in the aligned sequencing data to identify expressed fusion genes.

3.2.1.1   Deep RNA sequencing alone is not sufficient for reliable detection

For 22 PA samples in our cohort, a *KIAA1549:BRAF* fusion had previously been found by WGS (Jones et al., 2013). Due to the relatively high number of about 200 million total reads per sample (Figure 3.19a), we expected to identify these fusions in our RNA-Seq data as well. Surprisingly, however, Arriba only reported the fusion in 14/22 samples (Figure 3.19b). In addition to the main output file, Arriba also creates a separate list of potential fusions that are discarded due to certain filtering criteria. Checking these files revealed three additional samples for which the fusion had been discarded due to a low number of supporting reads. For the following analysis, these were still counted as "detected" together with the other reported fusions. However, our standard workflow completely missed the *KIAA1549:BRAF* fusion in five samples. Despite a read count that is already considered to be sufficient for the identification of important driver fusions (S. Uhrig, personal communication), we sequenced those five samples again and merged the data. This considerably increased sequencing depth and generated a mean read count of about 600 million reads per sample (Figure 3.19a). Nevertheless, we were surprised to see that detection of the *KIAA1549:BRAF* fusion was barely improved (Figure 3.19b). One sample now showed the fusion in the discarded file with a single supporting read. However, Arriba still missed the fusion in four of five samples despite extensive sequencing.
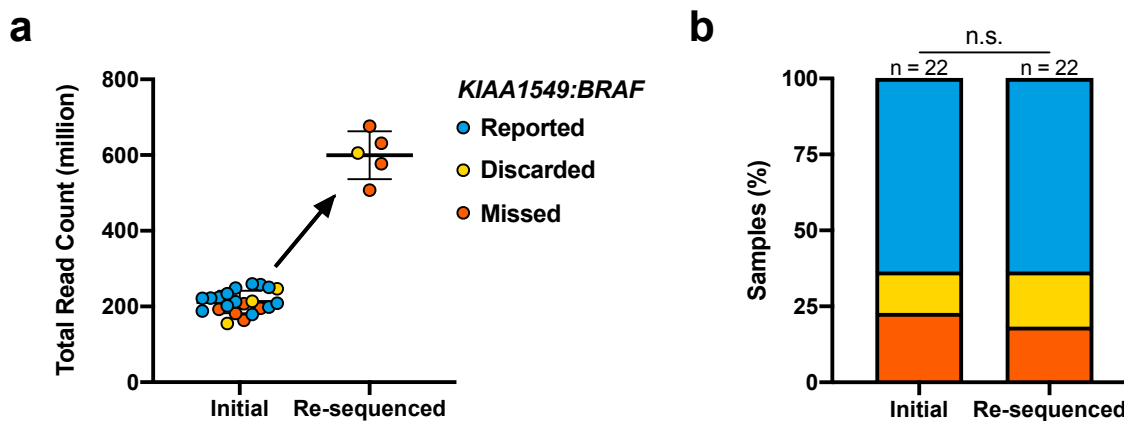
Figure 3.19 | Performance of the standard workflow in detecting *KIAA1549:BRAF* fusions

a) Total read count of the 22 PA samples of our RNA-Seq analysis cohort that were known to harbor a *KIAA1549:BRAF* fusion. In those cases where the fusion was missed in the initial analysis (marked in orange), the respective libraries were re-sequenced, the raw data was merged and the analysis was repeated. Mean ± SD. b) Percentage of *KIAA1549:BRAF* fusions that were reported, discarded or missed by the standard workflow in the initial analysis and after re-sequencing. Chi-square test on the underlying absolute values.

3.2.1.2 Different factors might have a potential influence on detectability

Given these surprising results, we wondered which factors might influence detectability of the *KIAA1549:BRAF* fusion.

3.2.1.2.1   Fusion variant and library size

First, we compared the different exon combinations of *KIAA1549* and *BRAF*. The samples in our cohort had either a 16:9 or a 15:9 fusion. While the detection rate for the 15:9 fusions was slightly higher (10/11 vs. 8/11), this difference was far from significant (Figure 3.20a).

Next, we had a look at the estimated library size, which is a measure of the information content of the library. Samples in which the fusion could be detected showed a trend towards having a higher estimated library size (Figure 3.20b).
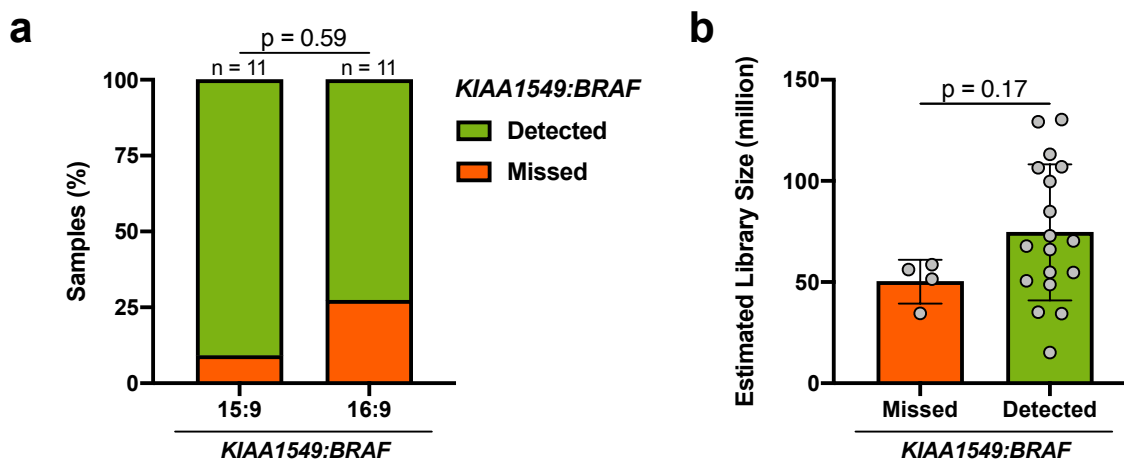
Figure 3.20 | Effect of the exon combination and the estimated library size on detectability

a) Percentage of *KIAA1549:BRAF* fusions that were detected (reported/discarded) or missed depending on their exon combination (16:9 or 15:9). Fisher's exact test on the underlying absolute values. b) Estimated library size of samples in which the *KIAA1549:BRAF* fusion had been detected (reported/discarded) or missed. Mean ± SD. Unpaired t test.

3.2.1.2.2 Tumor cell content

Tumor cell content is an obvious potential factor that might influence the reliability of detection. To get an estimation of the tumor cell content, we used two different approaches.

First, we compared the ESTIMATE immune scores of both groups as already calculated before (see 3.1.1.5). Samples in which the fusion could be detected showed a significantly lower immune score, which is synonymous with a higher tumor cell content (Figure 3.21a).
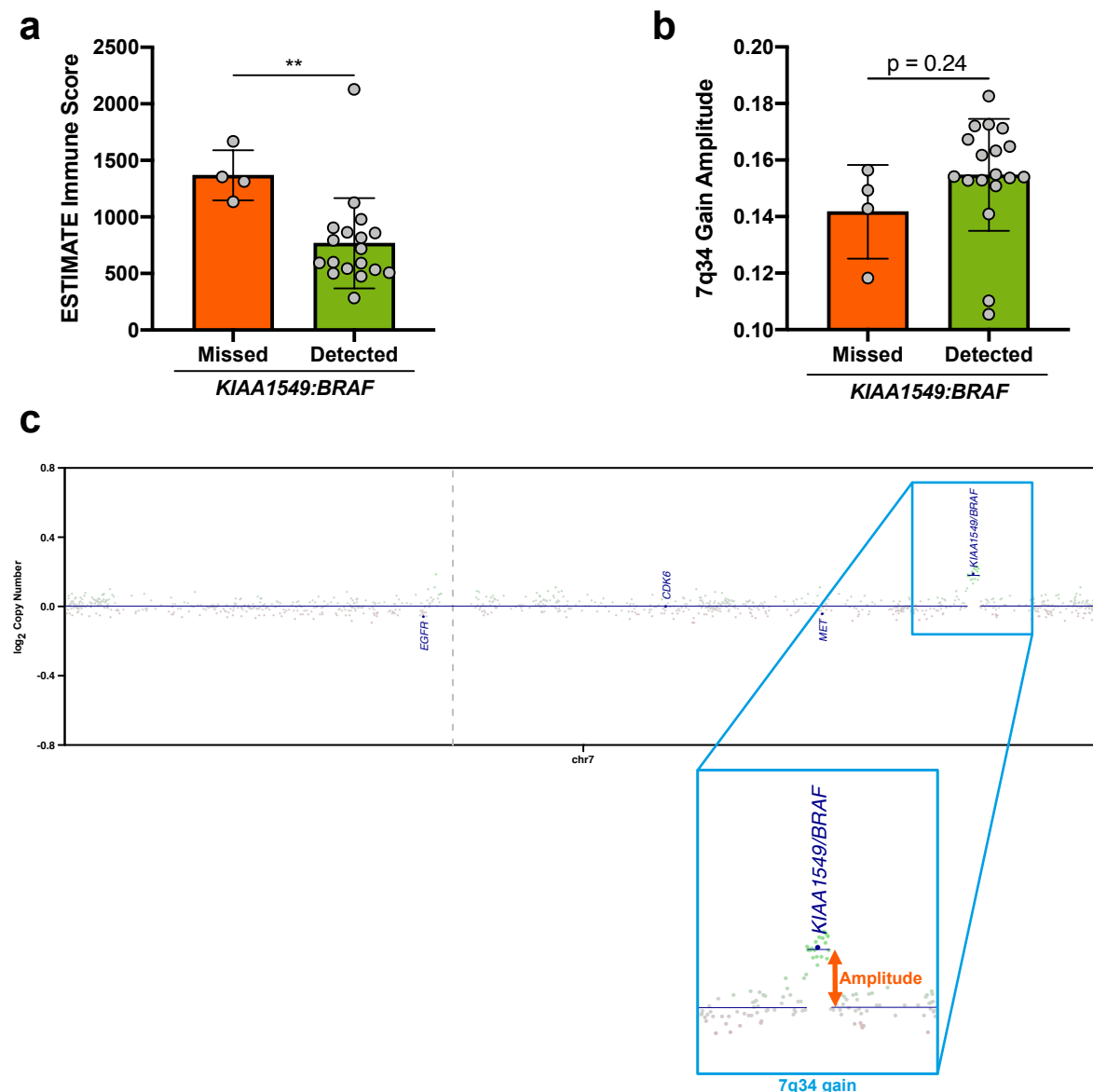
Figure 3.21 | Effect of the tumor cell content on detectability

a) ESTIMATE-based evaluation of the immune cell content of samples in which the *KIAA1549:BRAF* fusion had been detected (reported/discarded) or missed. Mean ± SD. Unpaired t test. b) Amplitude of the genomic 7q34 gain, which is characteristic for the *KIAA1549:BRAF* fusion and correlates with tumor purity, in samples in which the *KIAA1549:BRAF* fusion had been detected (reported/discarded) or missed. Mean ± SD. Unpaired t test. c) Typical copy number plot derived from methylation array data showing the typical focal gain at 7q34. The amplitude was calculated as the difference between gain and baseline.

Secondly, we used methylation array-derived copy number data as the basis for calculating the amplitude of the genomic 7q34 gain, which is characteristic for the *KIAA1549:BRAF* fusion

(Figure 3.21c). This is an alternative measure of tumor cell content, as the amplitude is expected to correlate with tumor purity. Although not significant, the results showed a similar trend (Figure 3.21b). Samples in which the fusion was identified on average showed a higher 7q34 gain amplitude, i.e. a higher tumor cell content, compared with the ones in which the fusion was missed.

3.2.1.2.3 Fusion expression level

As the expression level of the *KIAA1549:BRAF* fusion itself is difficult to measure, we subsequently extracted the expression values of both fusion partner genes from our RNA-Seq data. The expression levels of *KIAA1549*, which drives expression of the fusion, as well as *BRAF* were both significantly elevated in cases where the fusion was detected (Figure 3.22a,b), suggesting that a stronger expression of the fusion facilitates detection.

For comparison, we also had a closer look at seven PA samples harboring *BRAF* fusions with a different fusion partner. Strikingly, these fusions could be detected reliably, and their alternative 5' partners (*FAM131B*, *GNAI1*, *MKRN1* or *RNF130*) were all expressed at higher levels than *KIAA1549* (Figure 3.22a). This again suggests that the fusion expression level is a key determinant of detectability.
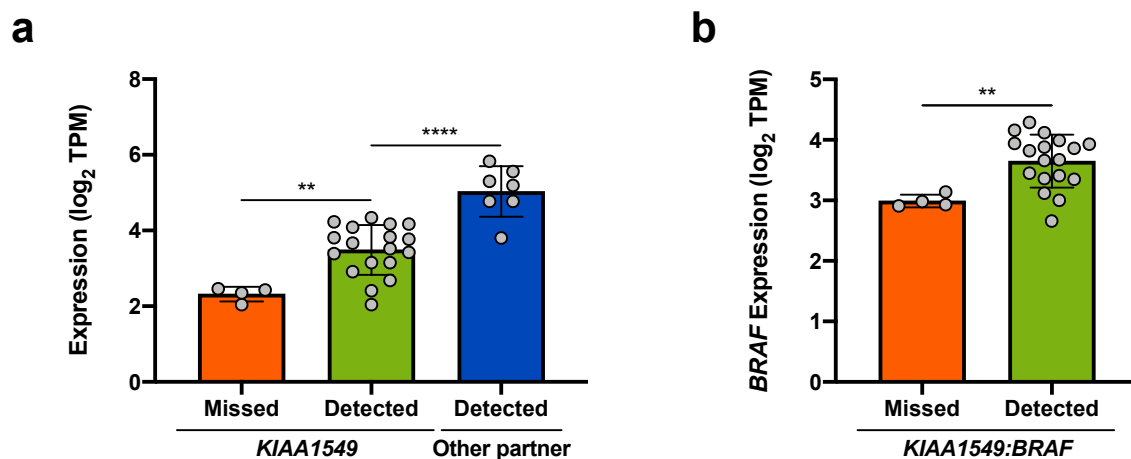
Figure 3.22 | Effect of the expression of both fusion partners on detectability

a) Expression of *KIAA1549* in samples in which the *KIAA1549:BRAF* fusion had been detected (reported/discarded) or missed as well as expression of the 5' partners (*FAM131B*, *GNAI1*, *MKRN1* and *RNF130*) in PA samples with an alternative *BRAF* fusion. Mean ± SD. One-way ANOVA followed by Tukey multiple comparisons test. b) Expression of *BRAF* in samples in which the *KIAA1549:BRAF* fusion had been detected (reported/discarded) or missed. Mean ± SD. Unpaired t test.

3.2.1.2.4 Library preparation protocol

As part of the analysis, we also ran our standard workflow for fusion detection on an older RNA-Seq cohort including 54 PA samples that were known to have a *KIAA1549:BRAF* fusion. The only obvious difference between both cohorts was that the old cohort had been generated using a ribosome-depleted total RNA (RiboZero) library preparation protocol, while our new analysis cohort was generated using a polyA enrichment approach. Surprisingly, detection of

the *KIAA1549:BRAF* fusion was significantly more sensitive in the old cohort (51/54) compared with the new one (17/22) (Figure 3.23). This indicates that the library preparation protocol might be another factor influencing fusion detectability.
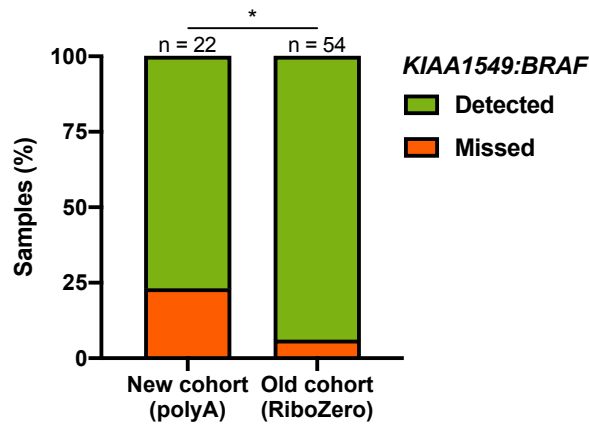
Figure 3.23 | Effect of the library preparation protocol on detectability

Percentage of *KIAA1549:BRAF* fusions that were detected (reported/discarded) or missed in the analysis cohort in comparison to RNA-Seq data of an older cohort, which had been generated using a different library preparation protocol (RiboZero vs. polyA capture). Fisher's exact test on the underlying absolute values.

3.2.1.3 Tumor cell content and fusion expression level limit detectability most

In an attempt to explain why the *KIAA1549:BRAF* fusion was missed in the four difficult samples, we ranked all samples by their "performance" regarding the identified factors that might affect detection reliability. *KIAA1549* expression as well as tumor cell content (evaluated by ESTIMATE) emerged as the most decisive factors for detectability of the *KIAA1549:BRAF* fusion in our cohort. Those samples in which the fusion was missed ranked significantly worse for both of these factors in comparison to the samples in which the fusion was reported (Figure 3.24). Fittingly, the cases where the fusion was identified but then discarded had intermediate ranks.
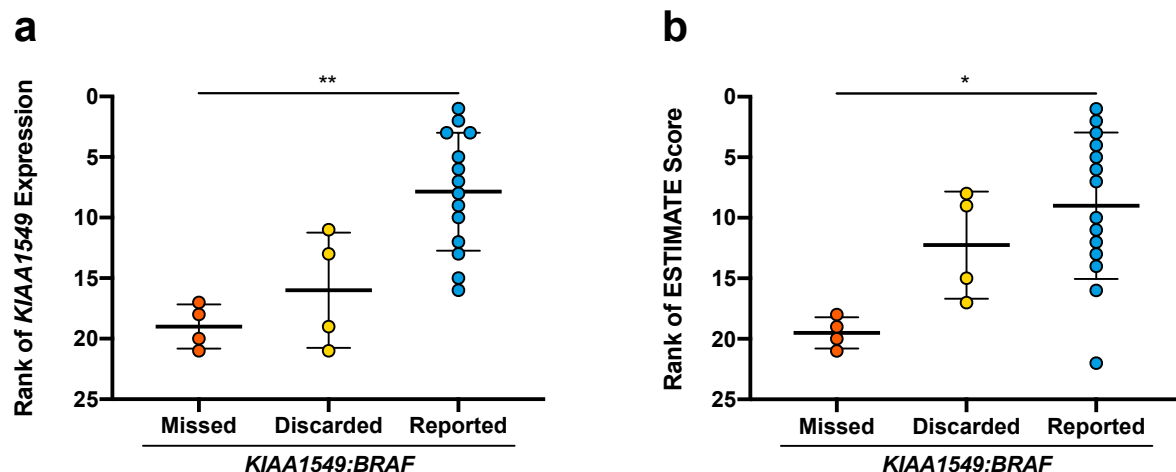
Figure 3.24 | Ranking of samples by *KIAA1549* expression and ESTIMATE immune score

a) Ranking of samples, in which the *KIAA1549:BRAF* fusion had been reported, discarded or missed, by their *KIAA1549* expression (see Figure 3.22a). The sample with the highest expression was assigned the highest rank. Mean ± SD. Kruskal-Wallis test followed by Dunn's multiple comparisons test. b) Ranking of samples, in which the

KIAA1549:BRAF fusion had been reported, discarded or missed, by their ESTIMATE immune score (see Figure 3.21a). The sample with the lowest immune score was assigned the highest rank. Mean ± SD. Kruskal-Wallis test followed by Dunn's multiple comparisons test.

3.2.1.4 Using a different fusion calling algorithm does not improve detectability

Next, we ran a different fusion calling algorithm on our analysis cohort to probe whether that would improve fusion detection. We chose FusionCatcher, as it does not only rely on STAR but uses the output of several aligners as a basis for fusion detection. The overall detection sensitivity of FusionCatcher was very similar to that of our standard workflow (16/22, vs. 18/22 for Arriba), but, notably, the set of samples in which the fusion could be identified was different (Table 3.3). For example, FusionCatcher detected the *KIAA1549:BRAF* fusion in one sample (#3) that had been missed by Arriba. Thus, we hypothesized that the raw data of this sample actually contains fusion-derived reads that are differently aligned and detected by the distinct workflows and that this might also apply to other samples.

Table 3.3 | Detection result comparison of the standard workflow and FusionCatcher

ID	Expected Fusion	Standard Arriba	FusionCatcher
1	*KIAA1549:BRAF* 16:9		
2	*KIAA1549:BRAF* 16:9		
3	*KIAA1549:BRAF* 15:9		✓
4	*KIAA1549:BRAF* 16:9		
5	*KIAA1549:BRAF* 15:9	✓	✓
6	*KIAA1549:BRAF* 15:9	✓	
7	*KIAA1549:BRAF* 15:9	✓	✓
8	*KIAA1549:BRAF* 16:9	✓	✓
9	*KIAA1549:BRAF* 15:9	✓	✓
10	*KIAA1549:BRAF* 15:9	✓	✓
11	*KIAA1549:BRAF* 15:9	✓	✓
12	*KIAA1549:BRAF* 15:9	✓	✓
13	*KIAA1549:BRAF* 15:9	✓	✓
14	*KIAA1549:BRAF* 16:9	✓	
15	*KIAA1549:BRAF* 16:9	✓	✓
16	*KIAA1549:BRAF* 16:9	✓	✓
17	*KIAA1549:BRAF* 16:9	✓	✓
18	*KIAA1549:BRAF* 15:9	✓	✓
19	*KIAA1549:BRAF* 16:9	✓	✓
20	*KIAA1549:BRAF* 16:9	✓	✓
21	*KIAA1549:BRAF* 15:9	✓	✓
22	*KIAA1549:BRAF* 16:9	✓	
		18	16

Overview of all 22 PA samples that were known to harbor a *KIAA1549:BRAF* 16:9 or 15:9 fusion and the results of FusionCatcher compared with our standard Arriba workflow in identifying the fusion. ✓ = reported by Arriba, ✓ = detected but discarded by Arriba, ✓ = detected by FusionCatcher.

3.2.1.5 All samples harbor supporting reads in their raw data

To check if the raw data of all four samples in which the fusion had been missed by our workflow contained fusion-relevant information, we examined the raw FASTQ files in detail. Using the UNIX utility *grep*, we scanned these files for 20 bp sequences that span the fusion point of *KIAA1549* and *BRAF* in the mRNA of both the 16:9 and the 15:9 variant. Surprisingly, we were able to identify matching reads in all four samples (Figure 3.25). Some samples even had multiple fusion-supporting reads in their raw data.

a

KIAA1549 Ex 16   TCCCTGCAGT GACTTGATTA   BRAF Ex 9
                  AGGGACGTCA CTCAACTAAT

b

KIAA1549 Ex 15   CGGATGCCCA GACTTGATTA   BRAF Ex 9
                  GCCTACGGGA CTCAACTAAT
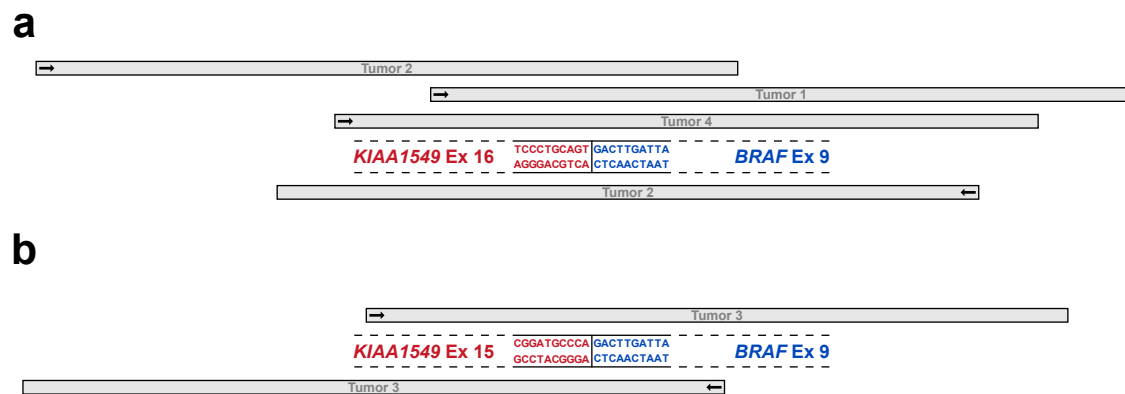
Figure 3.25 | Identification of raw reads comprising fusion-derived sequences

Illustration of the strategy to identify raw reads that are derived from the *KIAA1549:BRAF* fusion and could therefore be used for fusion detection. Using 20 bp sequences spanning the fusion point of a) the 16:9 and b) the 15:9 fusion variant, matching sequences could be determined in the FASTQ files. Shown are the identified unique reads of those samples in which the *KIAA1549:BRAF* fusion had been missed by the standard Arriba workflow (see Table 3.3) and how they align to the input sequence. Arrows indicate read orientation. Ex = exon.

3.2.1.6 Proper alignment of split reads is a major hurdle for reliable detection

Further investigation revealed that not Arriba *per se* but rather STAR was the major bottleneck in our standard workflow. STAR is known to have problems with aligning overlapping paired-end reads as well as split reads with a short overhang, and this turned out to be the problem here as well. The identified fusion-supporting reads were not at all or only partially aligned by STAR. Partially aligned reads are "soft-clipped" and could indeed be identified by manually checking the aligned data (Figure 3.26). Problematically, these reads are not visible to Arriba and thus cannot be used for downstream fusion calling.
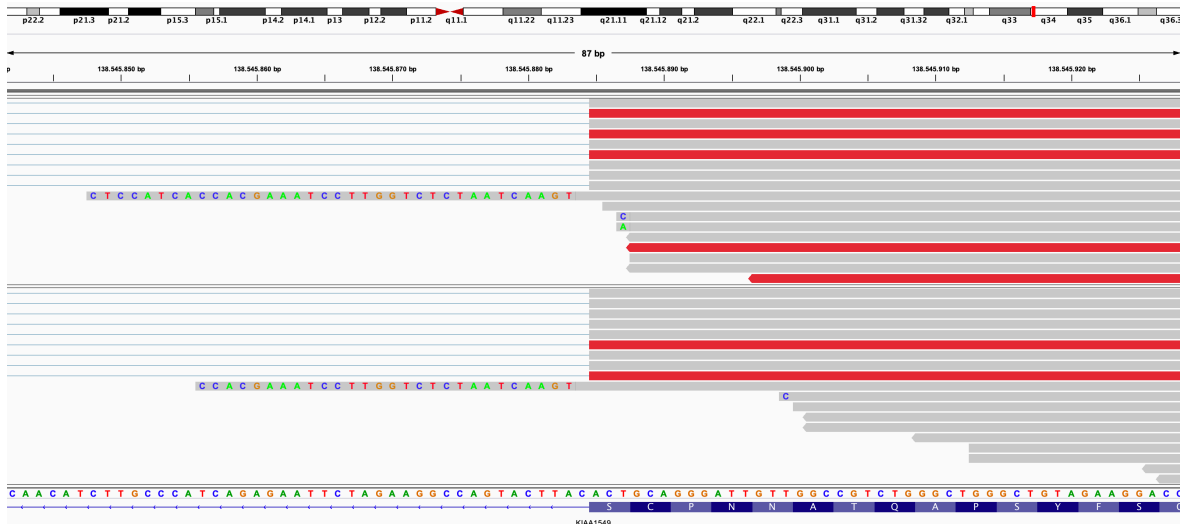
Figure 3.26 | Soft-clipped reads derived from the *KIAA1549:BRAF* fusion

Illustration of exemplary reads that were aligned to exon 16 of the *KIAA1549* gene in two samples. In each sample, one read is only partially aligned and thus displays a soft-clipped part, respectively. These reads originate from the *KIAA1549:BRAF* fusion, and the soft-clipped parts accordingly match to a sequence in exon 9 of the *BRAF* gene.

3.2.2 Optimized workflow

Having demonstrated that the raw data of all samples harbored supporting reads that could in principle provide a basis for fusion identification, one option would be to always manually check the raw data for matching sequences. However, considering the translation into a diagnostic setting, this did not seem practical. Instead, we opted for optimizing the different components of the workflow and worked together with Sebastian Uhrig (DKFZ), the developer of Arriba.

3.2.2.1 The workflow can be optimized by adjusting its individual components

Despite the fact that alignment by STAR was the major hurdle for reliable identification of the *KIAA1549:BRAF* fusion, we adjusted both STAR and Arriba for maximum benefit.

3.2.2.1.1 STAR

As STAR continues to be developed further, new parameters and functions are being added. Serendipitously, some parameters have recently been built into STAR that are intended to address the mentioned issues but are not active by default. We tested and tweaked these parameters and found *--peOverlapNbasesMin 10* and *--alignSplicedMateMapLminOverLmate 0.5* to substantially enhance the alignment of those critical reads that had previously been processed incorrectly.

3.2.2.1.2 Arriba

In addition, we had the objective to reduce the number of fusions that were erroneously being discarded by Arriba, so that all *KIAA1549:BRAF* fusions would be reported in the main output file. As part of the collaboration with Sebastian Uhrig (DKFZ), we created a new version of this

tool, which handles known fusions differently. Now, when providing a curated list of known gene fusions, one supporting read is usually sufficient for Arriba to confidently identify and report the fusion.

3.2.2.2 The optimized workflow improves fusion detection

Subsequently, we applied this optimized workflow of STAR and Arriba to our analysis cohort. The overall detection sensitivity was considerably improved, with 21/22 *KIAA1549:BRAF* fusions now being identified (Figure 3.27a). As intended, all fusions were directly reported in the main output file. Importantly, the optimized workflow also improved the confidence of already detected fusions by supporting them with a significantly higher number of fusion-derived reads (Figure 3.27b).
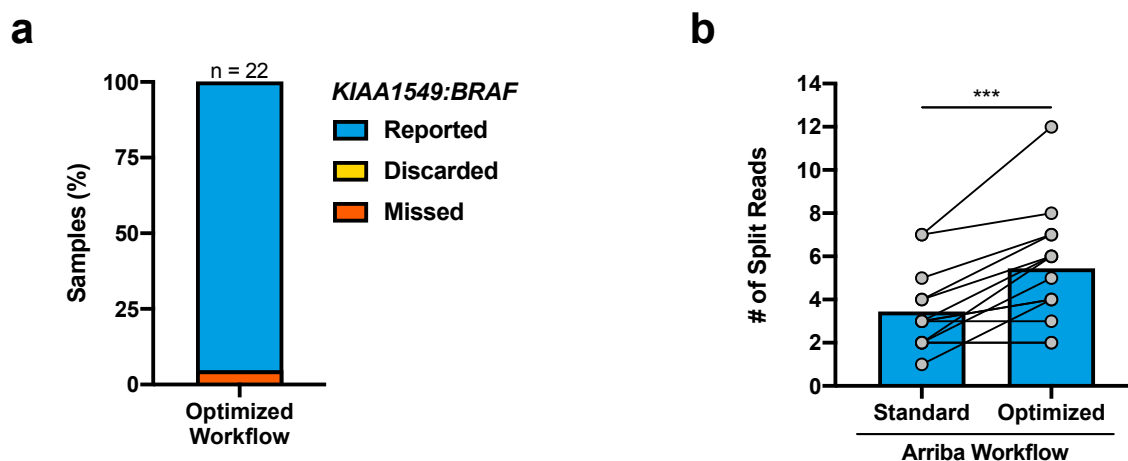
Figure 3.27 | Performance of the optimized workflow in detecting *KIAA1549:BRAF* fusions

a) Percentage of *KIAA1549:BRAF* fusions that were reported, discarded or missed by the optimized workflow. b) Number of fusion-supporting split reads identified before and after workflow optimization in those cases where the *KIAA1549:BRAF* fusion had initially been reported. Mean. Paired t test.

3.2.2.3 The optimized workflow outperforms previously used standard analysis tools

To test its performance in a diagnostic setting, we asked our collaborators in Montreal to use the optimized workflow for the analysis of an independent diagnostic PA cohort. In contrast to the previously used standard analysis tools, FusionCatcher (17/28) and STAR-Fusion (23/28), our workflow clearly performed better and was able to detect the *KIAA1549:BRAF* fusion in all 28 samples (Figure 3.28).
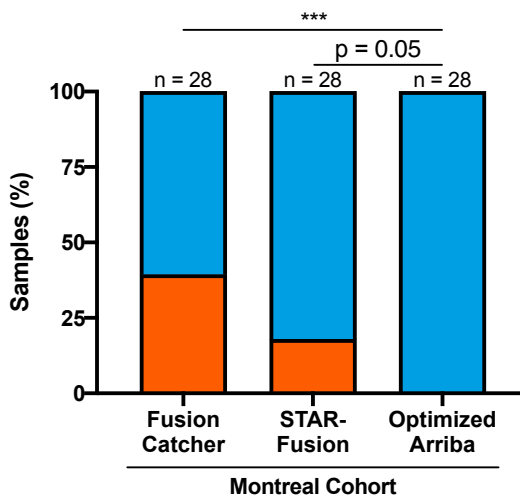
Figure 3.28 | Performance of the optimized workflow in an independent diagnostic cohort

Percentage of *KIAA1549:BRAF* fusions that were reported or missed by the optimized workflow in an independent diagnostic cohort from Montreal compared with the locally used standard analysis tools, FusionCatcher and STAR-Fusion. Fisher's exact test on the underlying absolute values.

3.2.2.4 The higher detection sensitivity is not accompanied by false-positive fusions

Lastly, we wanted to make sure that the higher detection sensitivity is not accompanied by the identification of false-positive fusions. Therefore, we applied our workflow to RNA-Seq data derived from a diagnostic cohort of > 1000 FFPE tissue samples of diverse origin (Stichel et al., 2019). We were able to confirm all previously identified *KIAA1549:BRAF* fusions and, importantly, did not identify any false-positive fusions, neither in other PA samples nor in any other tumor type. The optimized workflow thus had a specificity of 100% in the analyzed cohorts.

4 Discussion

4.1 Comparison of human & murine PA/PXA characteristics

Over the last years, large-scale sequencing studies have massively expanded our understanding of pediatric brain tumors but also made us realize how heterogeneous this group of cancers is (Gröbner et al., 2018). For pLGGs, it has more than ever become clear that these tumors are fundamentally different from adult lower-grade gliomas, which are usually characterized by mutations of *isocitrate dehydrogenase 1/2* (*IDH1/2*) in combination with 1p/19q co-deletion or *tumor protein 53* (*TP53*) mutations (Cancer Genome Atlas Research Network et al., 2015) and are thought to inevitably undergo malignant transformation (Rees et al., 2009). At the same time, pLGGs have shown an astonishing diversity of molecular subgroups and genetic alterations, whose classification thus lags behind that of other childhood brain tumors like medulloblastoma (Northcott et al., 2017; Taylor et al., 2012) or ependymoma (Pajtler et al., 2015).

On the one hand, genetic driver events keep being identified for known histological entities of pLGG, like *protein kinase C alpha* (*PRKCA*) fusions in papillary glioneuronal tumor (PGNT) (Bridge et al., 2013; Hou et al., 2019; Pages et al., 2015) or *FGFR1:TACC1* fusions in extraventricular neurocytoma (EVNCYT) (Sievers et al., 2018). On the other hand, more and more new molecular subgroups are being discovered, like pleomorphic neuroepithelial tumor of the young (PLNTY) characterized by *FGFR2/FGFR3* fusions (Huse et al., 2017), myxoid glioneuronal tumor of the septum pellucidum / lateral ventricle (MYXGNT) characterized by *platelet derived growth factor receptor alpha* (*PDGFRA*) mutations (Solomon et al., 2018) or diffuse glioneuronal tumor with oligodendroglial features and nuclear clusters (DGONC) characterized by chromosome 14 loss (Deng et al., 2019). Integration of histological and molecular information is further complicated by the fact that some molecular subgroups display a wide spectrum of histological characteristics. One example are tumors characterized by alterations of *MYB/MYBL1*, which can resemble angiocentric glioma and diffuse astrocytoma but also many other morphologies (Ramkissoon et al., 2013; Tatevossian et al., 2010; Zhang et al., 2013), including isomorphic diffuse glioma, as recently shown (Wefers et al., 2020).

In addition, the separation of distinct subgroups is not always acknowledged in clinical practice. This refers to PA and PXA, which were the main focus of this thesis. These tumor types can be challenging to distinguish, both in magnetic resonance imaging (She et al., 2018) and histopathological evaluation (Collins et al., 2015), which can pose diagnostic challenges

(Ballester et al., 2018). Besides, both show mutational overlap, as they are characterized by alterations of the RAS/MAPK pathway like BRAF V600E (Schindler et al., 2011). Nevertheless, PXA typically harbors secondary genetic alterations, like *CDKN2A/B* deletion (Weber et al., 2007; Zhang et al., 2013), and, importantly, a clearly worse clinical outcome (Perkins et al., 2012). While this translated into their different grading in the WHO classification (grade I vs. grade II) (Louis et al., 2016), a practical difference in the selected treatment regime is often not noticeable. This can partly be attributed to the fact that clinical practice effectively distinguishes two pediatric glioma groups, LGG and HGG, rather than four WHO grades (or even finer discrimination) for initial treatment decisions (Sturm et al., 2017). Hence, the different grading of PA and PXA is not reflected at the bedside. This is demonstrated by numerous clinical trials that are eligible for any pLGG tumor without further stratification and thus also include PXA cases despite their distinct clinical course (Ater et al., 2012; 2016; Chintagumpala et al., 2015; Gnekow et al., 2017; Lassaletta et al., 2016). Meanwhile, the recognition that PXA with marked anaplastic features can have an even worse prognosis (Ida et al., 2015) led to the creation of a separate tumor entity assigned WHO grade III (Louis et al., 2016). However, the details of this separation leave the impression of being somewhat arbitrary (Jones et al., 2019a). The distinction of anaplastic PXA and "regular" PXA is purely based on histopathology, with tumors displaying five or more mitoses per 10 high-power fields being termed anaplastic. Thus, a single mitotic figure can decide whether a patient receives chemo- and radiotherapy upfront or not.

The discrepancy between these malignant characteristics and the fact that PXA is nevertheless often grouped with extremely benign tumors like PA motivated me to investigate the molecular differences between these two tumor types in more detail. This was fueled by the higher proportion of targetable lesions like BRAF V600E (and potentially *NTRK* fusions) in PXA, which show good response to specific inhibition, e.g. using BRAF/MEK inhibitors. Rather than grouping PXA and other pLGGs together and treating them similarly, it is important to routinely detect these genetic alterations in a clinical setting and apply personalized therapies.

4.1.1 Cohort characteristics

The 68 PA and 21 PXA samples of our analysis cohort faithfully resembled the "fingerprints" of previously analyzed samples of these two tumor types, proving their validity. The fact that PA and PXA form separate clusters in a methylation-based analysis has previously been reported as part of the brain tumor classifier (Capper et al., 2018a) and could indicate that both originate from a different cellular precursor or that the secondary genetic alterations of PXA have an impact on genome-wide DNA methylation. The methylation profile of PA has been shown to separate PA samples by tumor location (Aichmüller et al., 2020; Antonelli et al., 2018; Lambert et al., 2013; Sexton-Oates et al., 2018). In contrast, we have not observed a separation of PXA into sub-clusters so far. This is supported by recent data showing that tumors

histopathologically diagnosed as anaplastic PXA do not possess a unique methylation profile but mostly result in a molecular classification of regular PXA or GBM (Capper et al., 2018a; Nakamura et al., 2019; A. Ebrahimi, personal communication). However, another report suggests that samples of the PXA cluster might still be separated into low- and high-grade tumors with different clinical outcomes based on pathological grade (Fukuoka et al., 2020). Thus, the exact composition of the PXA cluster currently remains unclear.

By stepwise assembling the analysis cohort and identifying the relevant genetic aberrations, we were able to compile a set of tumors with a balanced mixture of different driver events. It should be noted, though, that the ratio of different PA subgroups was intentionally not representative of that statistically observed in patients. While about 70% of PA harbor the characteristic *KIAA1549:BRAF* fusion and are mainly located in the cerebellum (Collins et al., 2015), we intended to generate a dataset with a more equal representation of the different genetic subgroups. Hence, the standard *KIAA1549:BRAF* fusion is strongly underrepresented, while particularly non-*BRAF* events like *NF1* mutations are overrepresented in the cohort. We also made sure to include some non-cerebellar PA tumors with a *KIAA1549:BRAF* fusion to be able to check for differences between tumors with the same genetic profile but different location. Not surprisingly, all *KIAA1549:BRAF* fusions were defined by either a 16:9 or 15:9 combination of exons, as these are the most prevalent fusion events (Jones et al., 2008).

For PXA tumors, we were much more limited by the availability of fresh-frozen tissue samples, and the distribution of genetic alterations is thus more representative. As expected, most samples featured the combination of BRAF V600E mutation and *CDKN2A/B* deletion that is observed in the majority of PXA (Schindler et al., 2011; Weber et al., 2007). A recent study even found biallelic inactivation of *CDKN2A* and *BRAF/RAF1* alterations in all 19 (anaplastic) PXA tumors (Phillips et al., 2019) and another study in 24/27 cases classified by methylation profiling (Fukuoka et al., 2020). Interestingly, *NTRK* fusions emerged as the most common driver event in our cohort when BRAF V600E was not present. While *NTRK* fusions have been observed in several pLGG entities (Jones et al., 2019a), mainly PA (Jones et al., 2013), reports in PXA are rare (International Cancer Genome Consortium PedBrain Tumor Project, 2016; Pehlivan et al., 2020; Zhang et al., 2013). Our results therefore highlight *NTRK* fusions as driver alterations in PXA with potential relevance for treatment options.

Another interesting finding was the presence of an in-frame *EGFR:BRAF* fusion transcript in one of the *CDKN2A/B*-deleted tumors without BRAF V600E. This fusion has not been described in any cancer type before and, to our knowledge, marks the first time that an *EGFR* alteration has been associated with PXA. Generally, genomic events involving *EGFR* are rare in pediatric cancer (Jones et al., 2019b) but have been found in some HGG entities. For example, bithalamic glioma, a pHGG subtype, has recently been described to feature recurrent EGFR alterations (Mondal et al., 2020). The concurrent identification of an in-frame

BRAF:EGFR fusion in our tumor sample indicates that both fusions were generated by a focal inversion on chromosome 7. Looking at the exons of both genes that are retained in the fusion transcript, it becomes clear that the EGFR:BRAF fusion protein comprises the EGF binding domains as well as the transmembrane domain of EGFR. For *BRAF*, the breakpoint is the same as for the *KIAA1549:BRAF* 16:9/15:9 fusions, and its kinase domain is thus preserved in the fusion protein. It therefore seems likely that this fusion joins the ranks of a growing list of *BRAF* fusions in pLGG (Tomić et al., 2017) that have various 5' partners but share the common mechanism of activating BRAF by dropping its inhibitory N-terminal domain (Jones et al., 2008). Nevertheless, it remains unknown whether the EGFR-derived domains tether the fusion protein to the cell membrane or make it responsive to EGF binding. The BRAF:EGFR fusion protein, on the other hand, presumably retains part of the tyrosine kinase domain as well as the autophosphorylation domain of EGFR. While gene fusions with EGFR as the 3' partner are known drivers for example in other gliomas (Stransky et al., 2014) or lung cancer (Konduri et al., 2016), they usually retain the full tyrosine kinase domain. The truncation observed in the BRAF:EGFR fusion raises doubts if this protein is functional, but further studies are needed to answer this question conclusively.

Two PXA tumors without detectable *CDKN2A/B* loss indeed showed unchanged levels of *CDKN2A* and *CDKN2B* transcripts. At this point, however, it cannot be ruled out that *CDKN2A/B* was inactivated by another mechanism. In a recent study, Phillips et al. describe a heterozygous loss of *CDKN2A* in one anaplastic PXA tumor, with the other allele seemingly still intact (Phillips et al., 2019). However, p16INK4A, one of the proteins encoded by the *CDKN2A* locus, was completely absent. Thus, other inactivating mechanisms seem to exist that are not detectable by copy number analysis. Further analysis showed that the two samples in our cohort harbored a mutation of TP53 (R273C) and phosphatase and tensin homolog (PTEN; R130G), respectively. While mutations of these genes have only been described in combination with *CDKN2A* loss in the mentioned study, they might to some extent have the same functional outcome on their own and thus explain why these tumors blend into the PXA cluster for the methylome- and transcriptome-based analysis of this thesis.

Notably, the study of Phillips et al. also suggested a role for *telomerase reverse transcriptase* (*TERT*) promoter mutations in anaplastic PXA (Phillips et al., 2019). In follow-up, we could only detect these mutations in a small fraction of samples and will have to investigate their functional consequences further.

4.1.2 Transcriptional differences

Consistent with their more malignant growth and the avoidance of OIS, the PXA group showed a higher score of MAPK activity compared with PA. However, it was interesting to notice that *BRAF* fusions do not seem to elicit the same degree of MAPK activation as other genetic

alterations in PA. This stands in contrast to the publication that originally presented the discovery of the *KIAA1549:BRAF* fusion (Jones et al., 2008). The authors used a cell-free chemiluminescent kinase assay to demonstrate that the KIAA1549:BRAF fusion protein has a higher kinase activity than BRAF V600E. The opposing results might therefore be explained by the differences in methodology, with the cell-free assay representing an artificial context lacking many regulatory aspects compared with vital tumor cells. It has previously been reported that pLGGs with *KIAA1549:BRAF* have a better prognosis in comparison to BRAF V600E (Hawkins et al., 2011; Ryall et al., 2020b). However, studies like this are often confounded by the fact that the BRAF V600E group includes the more malignant PXA tumors and that the frequency of different genetic events is not evenly distributed throughout the brain. The prognostic value of BRAF V600E is thus highly debated (Jones et al., 2018b; Lassaletta et al., 2017) and requires further investigation.

One of the most important results of this thesis was the distinct transcriptome of PA and PXA, which was mainly caused by genes associated with cell cycle and nervous system development. On the one hand, this demonstrates that the genetic developmental programs in both tumor types are fundamentally distinct. Again, this indicates that the secondary alterations in PXA shape its transcriptional profile and/or that both tumor types arise from different cells of origin. On the other hand, the fact that cell cycle-associated genes clearly separated PA and PXA and were mostly upregulated in the latter supports the more proliferative nature of PXA and renders its low-grade designation somewhat inappropriate. These results thus underline the observation that the loss of *CDKN2A/B* is associated with a worse clinical outcome in pLGG (Lassaletta et al., 2017; Ryall et al., 2020b). They might also explain why tumors characterized by BRAF V600E mutation and *CDKN2A/B* deletion, i.e. molecular "PXA-like" tumors, are often histopathologically diagnosed as pediatric GBM (Korshunov et al., 2015) and can undergo malignant progression to become secondary HGGs (Mistry et al., 2015). In both studies, these PXA-like tumors had a better outcome than molecular GBM but still clearly worse than low-grade lesions. This is also supported by analysis of the HERBY trial (Grill et al., 2018), which revealed a subset of "PXA-like" tumors that had been diagnosed as pHGGs and showed an intermediate outcome (Mackay et al., 2018).

In light of these results, it does not come by surprise that eGBM, a GBM variant that was added to the 2016 WHO classification, shows strong similarities to PXA (Alexandrescu et al., 2016; Furuta et al., 2018) and can even arise from PXA (Tanaka et al., 2014). A recent study confirmed that a subset of eGBM shows molecular characteristics of PXA including the presence of BRAF V600E mutation and *CDKN2A/B* deletion in most samples (Korshunov et al., 2018). Importantly, this study also reported that grade II PXA, grade III PXA and PXA-like eGBM tumors formed one collective cluster in a methylation-based analysis and were evenly distributed. This supports the hypothesis that all three histological groups are very closely related or even represent a spectrum of the same molecular tumor type.

Interestingly, the malignancy-promoting effect of *CDKN2A/B* loss has also become evident by the recent molecular characterization of anaplastic astrocytoma with piloid features (AAP). Besides alterations of the RAS/MAPK pathway and *ATRX*, 80% of these tumors possess a deletion of *CDKN2A/B* and have a much more dismal prognosis than regular PA (Reinhardt et al., 2018).

The results of this thesis together with the described studies therefore support the position that PXA as well as other *CDKN2A/B*-deleted pLGG tumors should be expected to have a more aggressive growth (Jones et al., 2019a) and might thus be treated clinically as intermediate malignancy tumors. Fittingly, a recent study has listed the combination of BRAF V600E mutation and *CDKN2A/B* loss as a high-risk biomarker for pLGG together with the H3.3 K27M mutation, as both induce rapid tumor progression and are often lethal (Ryall et al., 2020b). In any case, a precise molecular diagnosis is essential to identify these more malignant pLGG cases, including PXA. In addition, it helps to uncover those HGG tumors that are more PXA-like and might thus unexpectedly be susceptible to MAPK inhibitor treatment, with partly dramatic responses (Robinson et al., 2014).

4.1.3 Microenvironmental differences

In addition to tumor cell-intrinsic differences between PA and PXA, we also detected a distinct signature of the immune microenvironment, which is known to affect tumor growth on various levels (Broekman et al., 2018). Interestingly, both tumor types showed similarly high levels of a general immune infiltration signature. Simulations have shown that ESTIMATE immune scores of 1000, which corresponds to the levels observed in our analysis, represent an absolute immune infiltration of approximately 20% (P. Sun, personal communication). Thus, while pediatric cancers are thought to be less immunogenic than adult cancers (Majzner et al., 2017) and adult LGGs are often less infiltrated by immune cells than adult HGGs (Yang et al., 2011), our results suggest that this might not be true for pediatric glioma. This is supported by studies showing that pLGGs have a higher relative degree of immune infiltration compared with pHGGs (Griesinger et al., 2013; Plant et al., 2018), albeit subtype-specific variance might exist (Engler et al., 2012). A recent single-cell sequencing study of pediatric PA even reported an immune cell fraction of up to 40% (Reitman et al., 2019), but it remains unclear if this number is biased by the single-cell sequencing method itself. Our qualitative analysis of different immune cell types marked the first direct comparison of PA and PXA in this regard. While both tumor types showed a similarly strong signature of microglia/macrophages in comparison to normal brain, the CD4 T cell signature was enriched in PA and the CD8 T cell signature was enriched in PXA. The phenomenon that microglia/macrophages are recruited to the tumor and accumulate in its microenvironment is commonly observed in gliomas (Hambardzumyan et al., 2016). Microglia also constituted the majority of immune cells in the mentioned single-cell sequencing study of PA (Reitman et al., 2019). The enrichment of CD8 T cells in PXA seems

plausible in the light of a recent analysis of the HERBY trial, which demonstrated that those pHGG tumors that biologically resembled PXA were substantially infiltrated by CD8 T cells (Mackay et al., 2018). This was otherwise only the case for hypermutator cases on the basis of an underlying mismatch repair deficiency.

Given the marked signatures of immune infiltration in both tumor types, the question arises if patients might benefit from immunotherapy. Generally, it is believed that pediatric cancers are less susceptible to immunotherapy due to their lower mutational load compared with adult cancers (Jones et al., 2019b). This would particularly apply to pLGGs, as these tumors have a very low number of somatic mutations, even within the pediatric landscape (Gröbner et al., 2018). However, a recent systematic analysis by Samstein et al. demonstrated that glioma was one of the cancer subtypes in which overall survival after immune checkpoint inhibitor (ICI) treatment was not associated with mutational load (Samstein et al., 2019). It has to be verified if this is true for all age groups and subtypes, but this result opens doors for the use of immunotherapy in pLGGs. In particular, it should be further investigated if PXA tumors are susceptible to ICI treatment given the more extensive CD8 T cell infiltrate. Additionally, using expression data from a large collection of tumors, we found PXA to have the highest *PD-L1* expression across all glioma subtypes (data not shown), which is a potential biomarker for ICI response (Havel et al., 2019). Furthermore, PXA mostly occurs in supratentorial regions, which are not as sensitive to the risk of inflammatory mass effect as infratentorial regions (Plant and Hwang, 2018). If a paucity of neoantigens indeed emerges as a potential problem for ICI treatment, one idea could also be to combine it with epigenetic modifiers (Jones et al., 2019c). A study by Brocks et al. demonstrated that histone deacetylase (HDAC) inhibitors trigger widespread transcription of cryptic non-annotated transcription start sites (TSSs), which often lead to the generation of potential neoantigens (Brocks et al., 2017). This could to some extent compensate the low mutational burden and increase the efficacy of immunotherapy, which hitherto has primarily shown benefit in pediatric patients with hypermutant tumors (Bouffet et al., 2016).

Other immunotherapy approaches besides immune checkpoint inhibition should also not be disregarded for the treatment of pediatric gliomas. For example, CAR T cells directed against the disialoganglioside GD2 have recently shown encouraging efficacy in attacking PDX models of diffuse midline glioma (DMG), an aggressive childhood brain tumor, in mice (Mount et al., 2018). Due to the pre-existence of an immune infiltrate, pLGGs might also benefit from vaccine-based approaches (Griesinger et al., 2014). For example, this has successfully been exploited in adult lower-grade gliomas, where a peptide vaccine for mutant IDH1 induced an effective antitumor response in mouse models (Schumacher et al., 2014) and demonstrated immunogenicity in a first-in-man clinical phase I trial (Platten et al., 2018). It is conceivable that a similar strategy could be exploited for the treatment of PXA using a BRAF V600E vaccine. Previous reports have shown that BRAF V600E peptide vaccines elicit a specific T cell

response and reduce tumor growth in mouse models of BRAF-driven melanoma (Cintolo et al., 2016; Liu et al., 2018). Moreover, it has been demonstrated that T cells from patients respond to BRAF V600E vaccines and exert cytotoxic activity against BRAF V600E-positive melanoma cells (Somasundaram et al., 2006). Thus, although the absolute degree of CD8 T cell infiltration in human PXA should be verified by other methods like immunohistochemistry (IHC), a vaccine-based approach might be promising. The fact that the BRAF V600E mutation is found in the majority of PXA and also other tumor types like GG and extracerebellar PA (Schindler et al., 2011) suggests the possibility of broad applicability in pLGGs.

4.1.4 PXA mouse model

Another achievement described in this dissertation was the generation of a PXA mouse model. Animal models in general are an essential tool for preclinical research. Recent years have seen an advent of PDX models, which mimic the genetic heterogeneity of patient tumors (Rokita et al., 2019). While the concept of single-mouse testing (Gao et al., 2015) might improve the translation of preclinical results into practice and thus be the future for many drug testing studies, it requires a high number of molecularly characterized PDX models (Jones et al., 2019b). This is a problem for less aggressive tumors like pLGGs, which usually fail to engraft (Hermans and Hulleman, 2020; Hidalgo et al., 2014). In a study by Selt et al., the authors report that their efforts to generate a PA PDX model were unsuccessful in all cases (Selt et al., 2017). For PXA, only two PDX models have been published so far (Rokita et al., 2019), one of which is characterized by the typical combination of BRAF V600E mutation and *CDKN2A/B* loss (Kogiso et al., 2017). Since PDX models are also not suitable for studying the immune microenvironment, we decided to follow the GEMM approach, which has been providing valuable insights also for other childhood brain tumors like medulloblastoma (Gibson et al., 2010) or DMG (Larson et al., 2019). Expanding an existing BRAF V600E-driven PA mouse model (Gronych et al., 2011) by *Cdkn2a* knockout allowed us to generate the, to our knowledge, first GEMM for PXA. Also in the light of pLGG being considered a genetically rather simple disease and the vast majority of human PXA harboring both BRAF V600E mutation and *CDKN2A/B* deletion, it seemed an acceptable compromise to forego the genetic heterogeneity that is achieved with PDX models in favor of our GEMM.

The PXA model, as well as the PA model, showed astonishing similarities to human tumors on multiple levels. The growth behavior of both tumors was particularly striking. In the original publication, Gronych et al. demonstrated that the PA-like tumors did not induce any symptoms in mice for at least four months after induction despite the proven presence of tumors (Gronych et al., 2011). We were able to confirm this very benign growth, which mimics the "chronic disease" behavior of human PA. Most probably, this is mediated by the onset of OIS, and we were able to prove an upregulation of SASP factors in the PA model as part of a recent publication (Buhl et al., 2019). In contrast, the PXA-like tumors in our model were lethal 6-8

weeks after induction, underlining the distinctively more aggressive growth of PXA. To put this into context, a similar RCAS-based model for GBM driven by a TFG:MET fusion in combination with *Cdkn2a* knockout was lethal after about three weeks (International Cancer Genome Consortium PedBrain Tumor Project, 2016). This closely matches the observation of human PXA-like HGG tumors having an intermediate outcome between LGG-like tumors and true GBM (Korshunov et al., 2015) and underlines once more that tumors characterized by BRAF V600E mutation and *CDKN2A/B* loss should be expected to show a clinical course that does not deserve a low-grade designation. This is further supported by a related mouse model driven by BRAF V600E mutation and *Cdkn2a* deletion, which developed brain tumors that did not show a PXA-like histology but were also highly malignant (Robinson et al., 2010). The distinct growth patterns of PA- and PXA-like tumors were closely reflected by the expression of proliferation marker genes, which were clearly different three weeks after tumor induction and then drifted even further apart over time.

Histopathological assessment of our PXA model verified its authenticity. This included cellular pleomorphism of giant and spindle cells as well as the presence of a reticulin network, which is very characteristic of human PXA (Koelsche et al., 2014) and thus a strong indicator for its diagnosis.

The analysis of gene expression profiles further strengthened the authenticity of both mouse models. Interestingly, when extracting RNA from the PXA-like tumors, we often noticed a poor integrity. This might be attributed to necrosis within the tumor, which is known to involve RNA degradation (Gallego Romero et al., 2014) and is a feature of high-grade tumors like GBM or anaplastic PXA (Louis et al., 2016). The differences in the expression of selected genes and the MPAS in PA- and PXA-like tumors were consistently matched by their human counterparts. Notably, this was not the case for the difference between tumors and controls. Human cerebellum and cerebral cortex mostly had a lower expression of the analyzed genes as well as a lower MPAS compared with PA and PXA tumors. In contrast, the results obtained from murine brain tissue were mostly indistinguishable from the PA model. On the one hand, this could be due to the different analysis timepoints of both mouse models, which were on average later for the PA-like tumors, and might thus have captured different phases of growth. On the other hand, this might be explained by the relative age of the controls. While age-matched, i.e. young, healthy mice were used as controls for the mouse analysis, the expression profiles of human control tissue were extracted from a dataset of adult post-mortem donors (23-53 years old) (Roth et al., 2006). At a young age, the brain is still actively developing and thus expected to show a higher expression of genes and pathways that are associated with proliferation compared with adulthood. This might be the reason why the murine controls had the same high expression level of these markers as PA-like tumors. The only exception was PD-L1, whose expression pattern was exactly opposite, i.e. low levels in the murine controls but high levels in the human controls. This may also be explained by the

age bias, as PD-L1 is upregulated with age, both in healthy tissue (Luan et al., 2016) as well as tumors (Hamid et al., 2018). Finally, an unbiased assessment of the mouse models was rendered by a cross-species comparison based on differentially expressed genes, which clearly confirmed that both models faithfully resemble their human counterparts on a genetic level. The mouse models and the associated results presented in this thesis therefore pave the way for further research on PA and PXA within the scientific community.

In a last step, we were able to verify similar patterns of immune signatures in our models as compared with human tumors. The clear enrichment of the microglia/macrophage signature in both models matched the results in human PA and PXA. PXA-like tumors showed a higher score than PA-like tumors, which had been observed as a trend in the human analysis as well and might become more evident with more PXA samples. The relative differences in signature gene expression of CD4 T cells in both models similarly mimicked the human analysis. Notably, however, this was the only case where the scores in control mice were not concordant. Potentially, the global *Cdkn2a* knockout in XFM animals could affect the physiology of CD4 T cells or the expression patterns of signature genes. In any case, the results from the tumor comparison should be validated by IHC to exclude that the immune signature analysis is biased by knockout-induced transcriptomic changes. Most importantly, the analysis of a CD8 T cell signature perfectly matched the human results, increasing the value of the PXA model for preclinical proof-of-concept studies involving immunotherapy approaches like immune checkpoint inhibition.

The question might arise if the immune response in the mouse models is directed against the tumor cells or rather caused by the injection-induced tissue injury or the RCAS virus itself. While both factors might indeed play a role at the tumor initiation step, we believe it is unlikely that they have an effect at a later stage. The injection is a single event, and the RCAS virus cannot replicate in mammalian cells (Werder et al., 2012) and should thus disappear from the tumor site shortly after injection. In addition, the qualitative differences between both models cannot be explained by these factors, as the injection procedure and the virus type were identical. Instead, these differences have to originate from tumor-intrinsic characteristics. This allows further analysis of the close interaction between tumor and immune cells in the microenvironment of both PA and PXA, as is successfully being pursued in other mouse models of pLGG, for example driven by *Nf1* alterations (Guo et al., 2019; 2020; Simmons et al., 2011) or *KIAA1549:BRAF* (Chen et al., 2019).

Albeit the immune signatures in the PXA model showed good concordance with the human data, a note of caution should be sounded concerning the global *Cdkn2a* knockout in the model. It simplifies and possibly increases the effectiveness of tumor generation due to the avoidance of a secondary mechanism needed to specifically inactivate *Cdkn2a* in target cells. However, this might at the same time be a disadvantage when using the model to study the

immune microenvironment. It is unclear to what extent the *Cdkn2a* knockout affects immune cell biology and function in our tumor model. On the one hand, the study that originally described the knockout allele with the associated loss of p16INK4A and p19ARF reported features of abnormal extramedullary hematopoiesis in mice carrying a complete knockout (Serrano et al., 1996). In addition, a related mouse model with loss of p16INK4A but retention of p19ARF exhibited thymic hyperplasia as well as higher numbers and enhanced mitogenic responsiveness of T cells (Sharpless et al., 2001). On the other hand, the *Cdkn2a* locus is mostly repressed in young healthy organisms. Expression of *Ink4a* and *Arf* mRNA, for example, is almost undetectable in different tissues of 9-12 weeks old mice (Zindy et al., 1997), a timepoint that is not even reached in our model due to the aggressive growth of PXA-like tumors. Transcript levels subsequently increase with age, and this has also been observed for the different compartments of bone marrow, spleen and lymph node, including T cells (Krishnamurthy et al., 2004).

Even though it might be hypothesized that the functional consequences of the global *Cdkn2a* knockout on the immune microenvironment of PXA-like tumors are thus negligible, it should be considered to modify the mouse model in a way that allows specific *Cdkn2a* deletion only in tumors cells. One option would be to use a floxed *Cdkn2a* allele, which has already been generated (Krimpenfort et al., 2001), in combination with an RCAS virus carrying both BRAF V600E and Cre. Another option would be to expand the RCAS/TVA system by CRISPR/Cas9 and use transgenic mice that express both Ntv-a and Cas9 in Nestin-positive cells to induce a knockout of *Cdkn2a* in these cells, as recently described (Oldrini et al., 2018).

4.2 *KIAA1549:BRAF* fusion detection from RNA-Seq data

The detection of gene fusions has become a valuable resource for both research and diagnostics. While there are many different sequencing-based approaches (Heyer and Blackburn, 2020), RNA-Seq is most popular and has recently been used to create fusion catalogs across a variety of cancer types (Gao et al., 2018; Hu et al., 2018). The second part of this dissertation focused on the detection of the *KIAA1549:BRAF* fusion, which is largely specific for pediatric PA (Jones et al., 2019a). This is supported by the fact that its frequency is significantly lower in adult PA (Hasselblatt et al., 2011). Notably, there are at least two exceptions in which *KIAA1549:BRAF* fusions can be encountered in other tumor entities, diffuse leptomeningeal glioneuronal tumor (DLGNT) (Deng et al., 2018; Rodriguez et al., 2015) and anaplastic astrocytoma with piloid features (AAP) (Reinhardt et al., 2018). However, besides a different histology, these tumors usually harbor additional genetic alterations like 1p deletion for DLGNT and *CDKN2A/B* and *ATRX* loss for AAP. Comprehensive molecular profiling should thus be able to differentiate those *KIAA1549:BRAF*-positive tumor entities from PA. Independent of the exact tumor type, the identification of a *KIAA1549:BRAF* fusion might qualify tumors for treatment with targeted MAPK pathway inhibitors (Fangusaro et al., 2019; Usta et al., 2020), further increasing its value as a diagnostic marker.

Previous studies had reported that the *KIAA1549:BRAF* fusion is expressed at low levels and thus difficult to detect from RNA-Seq data (Stichel et al., 2019; Tomić et al., 2017). The expression level seems to be significantly lower than that of wildtype *BRAF* (Lin et al., 2012). The fact that we also had issues detecting the *KIAA1549:BRAF* fusion in our PA analysis cohort using Arriba, which has recently proven to be one of the most accurate fusion calling algorithms (Haas et al., 2019; Stichel et al., 2019), motivated me to investigate this further.

4.2.1 Factors affecting detectability

We found out that simply increasing the sequencing depth was not sufficient to reliably detect the *KIAA1549:BRAF* fusion. Subsequently, we provided evidence that the detectability might be influenced by different factors, including library size, tumor cell content, fusion expression level as well as the library preparation protocol. The library size presumably correlates with the integrity of the starting material but was higher than 30 million in both groups, which is usually considered the minimum for reliable fusion detection (unpublished observations). Tumor cell content and fusion expression level, on the other hand, are intrinsically prone to affecting detectability, as they limit the number of fusion-supporting reads, and emerged as the most decisive factors.

The library preparation protocol was the only factor that can actively be influenced. Our conclusion that it might play a role was based on the comparison of two independent PA

cohorts that seemingly only differed in the type of library preparation protocol. We found a higher detection rate in the dataset based on ribosomal RNA depletion compared with our new dataset based on the enrichment for polyA-tailed mRNA. While this association is not necessarily causative, it has also been observed in a transcriptomic profiling study of leukemia, in which a ribosomal RNA depletion approach recovered more gene fusions (Kumar et al., 2017). Particularly for degraded samples, polyA enrichment induces a 3' bias in the data and thus impedes identification of fusions with breakpoints that are distant from the 3' end (Davila et al., 2016). We indeed found signs of 3' bias in our data, but it was difficult to judge the full impact of this on *KIAA1549* and *BRAF* due to their low expression levels. In addition, the fusion point of both is rather close to the 3' end of the transcript (about 1 kb) (Jones et al., 2008), so further investigation is needed to elucidate to what extent the library preparation protocol affects *KIAA1549:BRAF* fusion detection. As RNA-Seq is not only used for fusion detection, the advantages and drawbacks of the different protocols for gene quantification should be considered as well (Zhao et al., 2018).

4.2.2 Analysis bottlenecks and optimization strategies

Driven by the result that FusionCatcher identified the fusion in a sample that was completely missed by Arriba, we checked the raw data and were surprised to find fusion-supporting split reads in all samples. The phenomenon that fusions can only be identified by manually scanning the raw reads has previously been observed for the fusion of *capicua transcriptional repressor* (*CIC*) and *double homeobox 4* (*DUX4*) (Panagopoulos et al., 2014). However, in this case, the reason seems to be that there are numerous copies of *DUX4* within the genome (Gabriëls et al., 1999), and the reads are thus not properly mapped by the alignment algorithms. Recently, Stichel et al. also identified *KIAA1549:BRAF*-supporting reads in the raw data that had been missed by three different fusion callers (Stichel et al., 2019).

We subsequently identified proper alignment of split reads by STAR as the major hurdle for reliable detection of the *KIAA1549:BRAF* fusion. This obviously poses the question why FusionCatcher did not have a better detection rate, as it integrates the information from several alignment algorithms (Nicorici et al., 2014). Our hypothesis is that FusionCatcher uses other or stricter filters that discard these reads again. To our knowledge, it does not create a list of discarded fusions that could be used to scan for *KIAA1549:BRAF* entries, as we successfully did for Arriba. Having identified insufficient alignment as the bottleneck of our workflow, different ideas for optimization emerged:

One idea was to use an assembly-based approach, which assembles reads *de novo* into transcripts and only then aligns them to the genome (Haas et al., 2019). While the alignment of transcripts might indeed be more dependable, these algorithms require a certain number of fusion-supporting reads to be able to build a transcript. However, the low number of reads

originating from the *KIAA1549:BRAF* fusion would likely not be sufficient, and assembly-based approaches thus do not reach the sensitivity of alignment-based approaches (Haas et al., 2019).

Another idea was to replace STAR with a different alignment algorithm. However, this idea failed for two main reasons. First, no method is perfect, and other alignment algorithms likely have other issues that affect fusion detection. While some fusion callers like FusionCatcher try to balance out these weaknesses by utilizing multiple alignment algorithms, this requires a high amount of computing power and might obstruct widespread application. Despite its shortcomings, STAR is most probably still one of the best RNA alignment algorithms. Secondly, there are technical limitations that render an exchange almost impossible. Many alignment algorithms do not support chimeric alignments, which is a prerequisite for the detection of fusion transcripts. Moreover, STAR is, to our knowledge, the only algorithm that reports chimeric reads not in a proprietary but a standardized format. Arriba has been developed to handle this format, and an exchange would thus involve considerable effort to re-develop the software tool.

4.2.3 Optimized workflow

While manually scanning for matching sequences in the raw data might be the most sensitive way to detect *KIAA1549:BRAF* fusions, it is limited to known fusion points, sensitive to mutations and would likely not see widespread implementation either. We therefore decided to optimize our automated fusion detection workflow to supersede the need for a separate complementary analysis. Based on adjusted STAR parameters and a new version of Arriba, the optimized workflow improved confidence of already detected fusions and allowed identification of fusions that were previously missed. In our analysis cohort, there was only one case where the fusion was still not identified. Checking the raw reads revealed that they barely overlapped with the fusion point, making alignment a difficult task.

The significantly higher *KIAA1549:BRAF* detection rate in an independent diagnostic cohort compared with the previously used standard analysis tools underlines that the choice of fusion calling workflow can indeed offer tangible advantages in a clinical setting, with direct impact on diagnosis, prognosis and treatment.

In a final step, we proved that the detection specificity did not suffer from the enhanced sensitivity by confirming the absence of false positive *KIAA1549:BRAF* fusions in a diagnostic RNA-Seq dataset of FFPE samples. This also demonstrated that our workflow is compatible with the modified RNA-Seq protocols that are needed for processing FFPE-derived RNA (Stichel et al., 2019). However, the different type of protocol might be the reason why some *KIAA1549:BRAF* fusions with only one supporting read were again discarded (data not shown),

despite the fact that this should have been prevented with the new version of Arriba. Seemingly, despite a clearly lower number of total reads, FFPE RNA-Seq can generate a high coverage for *KIAA1549* and/or *BRAF*. A high coverage, however, is associated with false positive fusion calls, and Arriba therefore limits the coverage of fusions that will be rescued from the "discarded" file. This has to be kept in mind when analyzing similar data.

In summary, we provide a novel workflow that significantly improves the detection of this important fusion in both research and diagnostics. We expect that the presented modifications will most likely also result in increased fusion detection performance in other contexts. Nevertheless, it should be noted that, even after optimization, it was not possible to extract all evidence of *KIAA1549:BRAF* fusions from the raw data. Therefore, additional optimizations will likely be possible in an ongoing iterative process as STAR, Arriba and other tools continue to be further developed.

It might also be considered to combine RNA-Seq-based fusion detection with other methods to achieve maximum sensitivity. As previously mentioned, the *KIAA1549:BRAF* fusion is generated by a focal tandem duplication at 7q34 (Jones et al., 2008). This characteristic gain is readily visible in methylation array-based copy number plots and has shown significant concordance with sequencing-based fusion identification (Capper et al., 2018b). This also includes the tumor samples presented in this thesis. It is currently not clear if there are cases where the *KIAA1549:BRAF* fusion is only identifiable by RNA-Seq and cannot be deduced from the copy number plot. Either way, available methylation array data should be used to identify those cases that are likely to harbor the fusion and might require a closer look if not directly identified by RNA-Seq.

5 Materials & Methods

5.1 Materials

5.1.1 Human tumor samples

All human tumor samples had been diagnosed as gliomas in pediatric patients (< 18 years of age). They originated from different institutions, and informed consent was obtained from all patients.

Molecular data for some samples in the cohort has been published previously (International Cancer Genome Consortium PedBrain Tumor Project, 2016; Jones et al., 2013).

5.1.2 Transgenic mouse lines

5.1.2.1 Ntv-a

The Ntv-a mouse line harbors a single transgene and expresses the *tv-a* receptor gene under the control of the *Nestin* promoter (Tg(NES-TVA)J12Ech; MGI:2663944) (Holland et al., 1998). Ntv-a mice were maintained on a mixed background.

5.1.2.2 XFM

The XFM mouse line harbors the Ntv-a transgene and two additional genetic changes. First, the *Cdkn2a* gene is knocked out (Cdkn2a^{tm1Rdp}; MGI:1857942), eliminating expression of both p16INK4A and p19ARF (corresponds to p14ARF in humans) (Serrano et al., 1996). Secondly, the *Pten* gene is flanked by loxP sites (Pten$^{tm2.1Ppp}$; MGI:2679886) (Trotman et al., 2003). The latter can be used for Cre-mediated excision of the *Pten* gene but this was not part of our tumor modelling strategy. *Pten* was thus still intact in all mice. XFM mice were maintained on a mixed background.

5.1.3 Cell lines

DF-1 cells had been obtained from the American Type Culture Collection (ATCC) in February 2012 (catalog no. CRL-12203), expanded for two passages and cryopreserved in aliquots.

RNA from BT-66 (passage 15) and BT-40 (passage 16) cell lines was a gift from the lab of Dr. Till Milde at the DKFZ.

5.1.4 Chemicals, reagents and medium

Table 5.1 | Chemicals, reagents and medium

Item	Catalog No.	Manufacturer
Agarose Standard	3810.3	Carl Roth
Aluminium sulfate	1102	Merck
Ammonia solution 25%	1.05432.1000	Merck
Ammonium iron(III) sulphate dodecahydrate	P734.1	Carl Roth
Colorless GoTaq Reaction Buffer (5X)	M792A	Promega
DMEM – high glucose	D5796	Sigma-Aldrich
DNA Loading Dye (6X)	R0611	Thermo Fisher Scientific
dNTP Mix (10 mM each)	R0192	Thermo Fisher Scientific
Dulbecco's Phosphate Buffered Saline (PBS)	14190-094	Thermo Fisher Scientific
Eosin Y solution 0.5% in water	X883.2	Carl Roth
Ethanol absolute	20821.330	VWR Chemicals
Eukitt Mounting Medium	12871	O. Kindler GmbH
Fetal bovine serum (FBS)	S 0115	Biochrom
Formaldehyde solution, non-acidic	P733.1	Carl Roth
Formalin solution, neutral buffered, 10%	HT501128	Sigma-Aldrich
FuGENE HD	E2311	Promega
GeneRuler 1 kb DNA Ladder	SM0311	Thermo Fisher Scientific
GeneRuler 100 bp DNA Ladder	SM0241	Thermo Fisher Scientific
GoTaq DNA Polymerase	M300A	Promega
L-Glutamine (200 mM)	25030081	Thermo Fisher Scientific
Mayer's Hematoxylin solution	254766.1611	AppliChem
Nuclear fast red	5189	Merck
Nuclease-Free Water	AM9937	Ambion
Opti-MEM	31985-062	Thermo Fisher Scientific
Paraffin Histo-Comp	VO-5-1001	Vogel
Penicillin-Streptomycin (10.000 U/mL)	15140148	Thermo Fisher Scientific
peqGREEN DNA/RNA Dye	37-5000	Peqlab
Potassium disulphite	P019.1	Carl Roth
Potassium hydroxide	5033	Merck
Potassium permanganate	8004.1	Carl Roth
Silver nitrate	7908.2	Carl Roth
Sodium thiosulfate pentahydrate	1.06513.2500	Merck
Tetrachloroauric(III) acid trihydrate	3867.1	Carl Roth
Trypan Blue Dye 0.4%	1450013	Bio-Rad
Trypsin – EDTA Solution 0.25%	T4049	Sigma-Aldrich
Xylene (mixture of isomers)	28975.325	VWR Chemicals

5.1.5 Buffer and culture medium composition

Table 5.2 | Buffer and culture medium composition

Buffer or cell culture medium	Composition
50x TAE buffer	50 mM EDTA disodium salt 2 M Tris-base 1 M Glacial/Acetic acid In water
DF-1 culture medium	DMEM – high glucose 10% FBS 1% Penicillin-Streptomycin 1% L-Glutamine
Silver staining solution	15 mL 20% silver nitrate solution 15 mL deionized water 6 mL 10% potassium hydroxide Add ammonia solution dropwise until solution becomes clear. Add 20% silver nitrate solution until the precipitate is only very slightly soluble. Double the volume with deionized water.

5.1.6 Plasmids and primers

Table 5.3 | Plasmids

Plasmid	Source
RCAS-BRAF^{V600E}	Gronych et al. (2011)
RCAS-GFP	Gronych et al. (2011)

Table 5.4 | Primers

Primer	Sequence (5' – 3')
BRAF for	TGCTTGCTCTGATAGGAAAATG
BRAF rev	AGCATCTCAGGGCCAAAAAT

5.1.7 Kits and arrays

Table 5.5 | Kits and arrays

Item	Catalog No.	Manufacturer
Bioanalyzer RNA 6000 Nano Kit	5067-1511	Agilent Technologies
Bioanalyzer RNA 6000 Pico Kit	5067-1513	Agilent Technologies
GeneChip Human Genome U133 Plus 2.0 Array	900467	Affymetrix
GeneChip Mouse Genome 430 2.0 Array	900497	Affymetrix
IDT for Illumina – TruSeq RNA UD Indexes	20022371	Illumina
Infinium MethylationEPIC Kit	WG-317-1003	Illumina
Maxwell RSC simplyRNA Tissue Kit	AS1340	Promega
Maxwell RSC Tissue DNA Kit	AS1610	Promega
QIAquick Gel Extraction Kit	28706	QIAGEN
TruSeq Stranded mRNA kit	20020595	Illumina

5.1.8 Consumables

Table 5.6 | Consumables

Item	Manufacturer
Biopsy embedding cassettes Swingsette	Carl Roth
Centrifuge Tube 50	TPP
Conical Tube 15	Corning
Counting Slides, Dual Chamber for Cell Counter	Bio-Rad
Cover Slips 24 x 60 mm	Thermo Fisher Scientific
Disposable Scalpel	Feather
Microscope Slides Superfrost Ultra Plus	Thermo Fisher Scientific
PCR Foil Seal Strong	4titude
PCR SingleCap 8er-SoftStrips 0.2 mL	Biozym
Petri Dishes 94X16	Greiner Bio-One
Razor blades	Leica
Safe-Lock Tubes	Eppendorf
SafeSeal-Tips professional	Biozym
Stripette 10 mL	Corning
Stripette 25 mL	Corning
Stripette 5 mL	Corning
Stripette 50 mL	Corning
Tissue Culture Flask 25	TPP
twin.tec microbiology PCR Plate 96, skirted, 150 µL, clear	Eppendorf

5.1.9 Equipment

Table 5.7 | Equipment

Equipment	Type	Manufacturer
Agarose gel chamber	Midi 450.000	Labortechnik Armin Baack
Animal hood	BS48	Techniplast
Automated cell counter	TC20	Bio-Rad
Automated DNA/RNA Extractor	Maxwell RSC Instrument	Promega
Bioanalyzer instrument	2100 Bioanalyzer	Agilent Technologies
Cell culture hood	Safe 2020	Thermo Fisher Scientific
Centrifuge	Fresco 17	Heraeus
Centrifuge	Multifuge 3 s	Heraeus
Centrifuge	Biofuge pico	Heraeus
CO_2 incubator	Innova CO-48	New Brunswick Scientific
Drying oven	UT 6060	Heraeus
Electrophoresis power supply	EV231	Consort
Embedding workstation	HistoStar	Thermo Fisher Scientific
Freezer -20°C	Mediline LKexv 3910	Liebherr
Freezer -80°C	Innova U 725-G	New Brunswick Scientific
Fridge 4°C	Mediline LKexv 3910	Liebherr
Gel comb	Diverse	Labortechnik Armin Baack
Gel imager	c200	Azure Biosystems
Gel tray	BAA26900018	Labortechnik Armin Baack

Halogen light source	PL 3000	Photonic
Histology cooling plate	EG1130	Leica
Histology water bath	HI1210	Leica
Homogenizer	Mikro-Dismembrator S	B. Braun Biotech
Injection syringe 10 µL	801 RN (26s gauge)	Hamilton
Inverted microscope	Eclipse TS100	Nikon
Mercury light source	Intensilight C-HGFI	Nikon
Microtome	HistoCore AUTOCUT	Leica
Mini centrifuge	Spectrafuge 3-1810	neoLab
PCR cycler	Mastercycler epgradient S	Eppendorf
Pipette controller	PIPETBOY 2	INTEGRA Biosciences
Pipettes	Pipet-Lite	Rainin
Precision balance	PLJ 1200-3A	Kern
Sequencing system	NovaSeq 6000	Illumina
Sequencing system	HiSeq 4000	Illumina
Spectrophotometer	NanoDrop ND-1000	Thermo Fisher Scientific
Spin Tissue Processor	Microm STP 120	Thermo Fisher Scientific
Stereo microscope	SMZ445	Nikon
Surgical instruments	Diverse	Fine Science Tools
Transilluminator	LB0100	Thermo Fisher Scientific

5.1.10 Software

Table 5.8 | Software

Software	Company/Source
2100 Expert Software	Agilent Technologies
Affinity Designer	Serif
Arriba	https://github.com/suhrig/arriba
BLAT	Kent et al. (2002)
Bowtie	Langmead et al. (2009)
Brain tumor classifier	Capper et al. (2018a)
ComBat	Johnson et al. (2007)
cSeries Capture Software	Azure Biosystems
ESTIMATE	Yoshihara et al. (2013)
Excel 2019 for Mac	Microsoft
FusionCatcher	https://github.com/ndaniel/fusioncatcher
Integrative Genomics Viewer (IGV)	Robinson et al. (2011)
macOS Terminal	Apple
Mastercycler ep CycleManager	Eppendorf
NanoDrop 1000 3.8.1	Thermo Fisher Scientific
Papers 3	Springer Science+Business Media
PowerPoint 2019 for Mac	Microsoft
Prism 8	GraphPad
R	R Core Team
R package *conumee*	http://bioconductor.org/packages/conumee/
RNA-SeQC	DeLuca et al. (2012)
RStudio	RStudio, PBC

SnapGene Viewer	GSL Biotech LLC
STAR	Dobin et al. (2013)
STAR-Fusion	https://github.com/STAR-Fusion/STAR-Fusion
Word 2019 for Mac	Microsoft

5.1.11 Internet resources

Table 5.9 | Internet resources

Resource	Website
COSMIC Cancer Gene Census	https://cancer.sanger.ac.uk/census
ENCODE: Encyclopedia of DNA Elements	https://www.encodeproject.org
Ensembl Genome Browser	https://www.ensembl.org
Gene Ontology Resource	http://geneontology.org
HomoloGene	https://www.ncbi.nlm.nih.gov/homologene
HUGO Gene Nomenclature database	https://www.genenames.org
PubMed	https://www.ncbi.nlm.nih.gov/pubmed
R2: Genomics analysis and visualization platform	https://r2.amc.nl

5.2 Methods

5.2.1 Nucleic acid extraction

Depending on the size, fresh-frozen tissue pieces from human and murine tumors or healthy mouse brain were either first manually crushed or directly homogenized using a Mikro-Dismembrator S. Subsequently, DNA and RNA were extracted in an automated way using a Maxwell RSC instrument according to manufacturer's instructions. Elution was performed in a volume of 50-75 μL. DNA concentration was measured using a Nanodrop spectrophotometer. The integrity and concentration of RNA was evaluated using an Agilent Bioanalyzer.

Isolated DNA was stored at -20°C, RNA was stored at -80°C.

5.2.2 Methylation array

Genome-wide methylation analysis of human tumor DNA was performed by the Genomics and Proteomics Core Facility (GPCF) at the DKFZ using Infinium MethylationEPIC BeadChip arrays, quantifying over 850,000 methylation sites. For some samples, methylation data was already available and had been generated some years ago with Infinium HumanMethylation450 BeadChip arrays, which cover about 450,000 methylation sites.

5.2.2.1 Brain tumor classification

The generated methylation data was fed into version 11b4 of the brain tumor classifier (Capper et al., 2018a). By comparing the data to a reference set of 2,801 samples, each tumor is assigned to one of 91 methylation classes. A score ranging from 0 to 1 indicates how confident the prediction is and should be above 0.9 for a clear result. In addition, the data is added to a t-SNE plot of all tumor samples that have been processed by the classifier. At the time of the analysis, the dataset comprised about 64,000 cases. Determining the position of individual samples within the plot often allows a more accurate classification if the classifier score was low.

An internal tool for methylation data clustering developed by Martin Sill (DKFZ) was used to create individual t-SNE plots for defined sets of tumor samples.

5.2.2.2 Copy number analysis

Genome-wide copy number plots were generated with the help of the Bioconductor package *conumee*. It is based on the combined intensities of methylated and unmethylated probes from the methylation arrays, as previously described (Sturm et al., 2012). Among others, the resulting copy number data was used to identify *CDKN2A/B* deletions and the 7q34 gain that is characteristic for the *KIAA1549:BRAF* fusion.

5.2.2.3 Amplitude of the genomic 7q34 gain

The amplitude of the genomic 7q34 gain was quantified by calculating the average of the $\log_2$ copy number values of features chr7-0773 to chr7-0792 that are reported by *conumee*. These features cover chromosome 7 from position 138,550,001 to 140,400,000 (reference genome hg19) and thus correspond to the region that is duplicated in *KIAA1549:BRAF*-positive tumors.

5.2.3 RNA sequencing

Sequencing of human and murine tumor RNA was performed by the Genomics and Proteomics Core Facility (GPCF) at the DKFZ. Libraries were prepared using the Illumina TruSeq Stranded mRNA kit, which is based on polyA enrichment of total RNA. Subsequently, libraries were sequenced on a NovaSeq 6000 or HiSeq 4000 machine with a 100 bp paired-end configuration. The raw FASTQ files were transferred to the Omics IT and Data Management Core Facility (ODCF) at the DKFZ for storage and further processing, which is handled by the One Touch Pipeline (OTP) (Reisinger et al., 2017).

5.2.3.1 Post processing

Initial post processing involved alignment of reads to the reference genomes hg19 and mm10 for human and mouse data, respectively, using STAR (Dobin et al., 2013) version 2.5.3a with the following parameters:

*--runThreadN 8 *
*--genomeDir "$STAR_INDEX_DIR" --genomeLoad NoSharedMemory *
*--readFilesIn "$READ1" "$READ2" --readFilesCommand zcat *
*--outStd BAM_Unsorted --outSAMtype BAM Unsorted --outBAMcompression 0 *
*--outFilterMultimapNmax 1 --outFilterMismatchNmax 3 *
--chimSegmentMin 10 --chimOutType WithinBAM SoftClip --chimJunctionOverhangMin 10
--chimScoreMin 1 --chimScoreDropMax 30 --chimScoreJunctionNonGTAG 0
--chimScoreSeparation 1 --alignSJstitchMismatchNmax 5 -1 5 5
--chimSegmentReadGapMax 3

The mapped data was saved as BAM files. The total read count as well as the estimated library size were calculated by RNA-SeQC (DeLuca et al., 2012). To make use of the STAR parameters *--peOverlapNbasesMin* and *–alignSplicedMateMapLminOverLmate* as part of the optimized workflow, the data was re-aligned with version 2.7.3a.

For further analysis, transcripts per kilobase million (TPM) values of all genes were uploaded to the genomics analysis and visualization platform R2. As this platform only supports single values per gene but some genes were represented with more than one ENCODE ID, we always chose the lowest ID, as this is supposed to be the canonical isoform.

5.2.4 Fusion detection

5.2.4.1 Fusion calling algorithms

Fusion transcripts were identified by running Arriba version 1.1.0 on the aligned data (BAM files) using standard parameters. A list of known fusions was provided (parameter -k) that was based on the COSMIC Cancer Gene Census (Sondka et al., 2018) and included the *KIAA1549:BRAF* fusion. To turn the more sensitive detection parameters for known fusions as part of the optimized workflow into account, Arriba version 1.2.0 was used instead, which was developed in the course of this project.

Alternatively, FusionCatcher version 1.20 was used to identify fusion transcripts on the basis of the raw FASTQ files. By default, FusionCatcher integrates the data from three different alignment algorithms, Bowtie (Langmead et al., 2009), BLAT (Kent, 2002) and STAR (Dobin et al., 2013).

5.2.4.2 Manual detection

Manual detection of *KIAA1549:BRAF* fusions was performed by scanning the raw FASTQ files for 20 bp sequences spanning the fusion point of *KIAA1549* and *BRAF* using the UNIX utility *grep*. These sequences as well as their reverse complements are listed in the following table for the three most common fusion variants.

Table 5.10 | *KIAA1549:BRAF* fusion point sequences

Fusion variant	Spanning sequence	Reverse complement
16:9	TCCCTGCAGTGACTTGATTA	TAATCAAGTCACTGCAGGGA
15:9	CGGATGCCCAGACTTGATTA	TAATCAAGTCTGGGCATCCG
16:11	TCCCTGCAGTAAAACACTTG	CAAGTGTTTTACTGCAGGGA

Partially aligned fusion-supporting reads were identified by opening BAM files in Integrative Genomics Viewer (IGV) and looking for soft-clipped reads. If such reads in *KIAA1549* could manually be mapped to *BRAF* or vice versa, they had originated from a fusion transcript.

5.2.5 Microarray

Microarray gene expression analysis of RNA from human and murine tumors as well as healthy mouse brain was performed by the Genomics and Proteomics Core Facility (GPCF) at the DKFZ. The GeneChip Human Genome U133 Plus 2.0 Array was used for human RNA, while the GeneChip Mouse Genome 430 2.0 Array was used for murine RNA. For analysis, raw CEL files were uploaded to the R2 platform and MAS5.0 normalized.

The expression data of normal human cerebellum and cerebral cortex was adopted from a publicly available dataset of different tissues from ten post-mortem donors (GEO accession

GSE3526) (Roth et al., 2006). The data had been generated with the same type of microarray and could thus be added to our dataset of human PA and PXA for joint analysis after MAS5.0 normalization.

5.2.6   Sanger sequencing

Sanger sequencing was used to verify *BRAF* mutations on a genomic level. First, the region of interest was amplified from human tumor DNA using the following PCR.

Table 5.11 | *BRAF* PCR mastermix

Component	Volume
5X Colorless GoTaq Reaction Buffer	10 µL
dNTP Mix (10 mM each)	1 µL
Primer *BRAF* for	0.3 µL
Primer *BRAF* rev	0.3 µL
GoTaq DNA Polymerase	0.25 µL
Nuclease-Free Water	36.15 µL
Human DNA	2 µL

Table 5.12 | *BRAF* PCR program

Step	Temperature	Time	Number of cycles
Denaturation	95°C	3 min	
Amplification	95°C	30 sec	35
	62°C	25 sec	
	72°C	30 sec	
Final elongation	72°C	10 min	
Storage	4°C	∞	

Subsequently, the PCR reaction mix was subjected to gel electrophoresis. A suspension of 1.5% agarose in 1x TAE buffer was boiled, and peqGREEN was added (1:20,000). Before complete cooling, the solution was casted into a gel tray, and a comb was inserted. The PCR products were mixed with DNA Loading Dye and loaded onto the gel together with a DNA ladder. A voltage of 120 V was applied to separate the fragments by size. The presence of the expected 230 bp amplicon was validated and documented with a gel imager. Then, the bands were cut out on a transilluminator, and the PCR product was extracted using the QIAquick Gel Extraction Kit according to manufacturer's instructions. In a last step, the DNA was diluted to a concentration of 10-50 ng/µL, and 20 µL were sent to Eurofins Genomics for Sanger sequencing. The presence of *BRAF* mutations in the sequencing results was evaluated using SnapGene Viewer.

5.2.7 Gene expression and ontology analysis

In general, gene expression analysis was based on $\log_2$-transformed expression values and performed using the R2 platform, unless otherwise noted. For RNA-Seq data, only a single TPM value per gene had been uploaded to R2 (see 5.2.3.1). For microarray data, in contrast, gene expression could be defined by more than one probeset per gene. In these cases, the probeset with the highest average present signal was chosen.

Differential gene expression analysis consisted of the identification of the 1,000 most differentially expressed genes between the groups of interest based on p-values calculated by ANOVA and corrected for multiple testing by false discovery rate (FDR).

The basic ontology analysis is a feature of the R2 platform and was used to get a quick but limited overview of the ontologies in the set of differentially expressed genes. It is based on the affiliation of genes to no, one or more than one of 10 gene categories. Overrepresented categories could be identified by comparing the number of genes in each category to the counts that would on average be expected in a similarly sized cohort.

The more advanced ontology analysis was done using the GO term analysis in R2. It works in a similar fashion and tests which GO terms are enriched in the set of differentially expressed genes between the groups. Corresponding heatmaps were based on z-scores of those differentially expressed genes that were associated with the respective GO term.

5.2.8 Gene signature analysis

Unless otherwise noted, gene signature analysis was performed in R2 by evaluating the expression of the signature genes in the samples of interest. For each sample, the signature score was calculated as the mean of z-scores of the individual genes.

5.2.8.1 MPAS

The MPAS was calculated as the signature score from 10 MAPK-specific genes (*PHLDA1, SPRY2, SPRY4, DUSP4, DUSP6, CCND1, EPHA2, EPHA4, ETV4* and *ETV5*) published by Wagle et al. (2018). For mouse data, the murine homologs of these genes were identified in the HomoloGene database and used as a signature.

5.2.8.2 Immune signatures

The signatures for different immune cell subsets were adopted from Engler et al. (2012). Originally, the signature for glioma-infiltrating microglia/macrophages had been published by Murat et al. (2009), whereas the signature for activated CD4/CD8 T cells had been published

by Chambers et al. (2007). For mouse data, the murine homologs of the signature genes were identified in the HomoloGene database.

5.2.8.3 ESTIMATE

ESTIMATE scores for immune and stromal cells were calculated using the original ESTIMATE method (Yoshihara et al., 2013). The underlying algorithm is also based on gene signatures and comprises 141 genes for the immune and for the stromal signature, respectively.

5.2.9 Cross-species analysis

Cross-species analysis of human and murine tumors was performed as previously described (Pajtler et al., 2019; Tao et al., 2019). Briefly, computed read counts from the RNA-Seq data of human PA and PXA were used to identify differentially expressed genes between both tumor types. Subsequently, human-mouse orthologous genes were selected based on this set of genes, and their expression was extracted from the RNA-Seq data of PA- and PXA-like mouse models. Following batch effect adjustment with ComBat (Johnson et al., 2007), these orthologous genes were used as an input for principal component analysis (PCA) as well as unsupervised hierarchical clustering (ward.D2 method).

5.2.10 Mouse lines

Mice were housed at the animal facility of the DKFZ and kept in individually ventilated cages (IVCs) under specific pathogen-free (SPF) conditions. A homozygous status was maintained for both the Ntv-a transgene and the *Cdkn2a* knockout in all animals. After a breeding was set up to generate offspring for tumor induction, the vaginal plug (VP)-positive date was used to estimate the date of p0 injections. Both male and female mice were used for experiments.

5.2.11 Tumor induction in mice

5.2.11.1 DF-1 cell culture and transfection

For each experiment, DF-1 cells were freshly thawed two to three weeks before injection and cultured in medium at 39°C and 5% CO_2 in a humidified incubator. Cells were usually split twice per week. Four days before the expected birth date of Ntv-a or XFM pups, DF-1 cells were seeded into T25 cell culture flasks, with 200,000 cells and 5 mL medium per flask. After 24 hours, the cells were transfected with the RCAS-BRAF^{V600E} vector. To this end, Opti-MEM was equilibrated to room temperature, and 200 μL were added to 4 μg plasmid DNA together with 10 μL FuGENE HD transfection reagent. The mixture was incubated for 15 minutes at room temperature and added dropwise to the DF-1 cells using a 1000 μL pipette. After gentle

but thorough mixing, the cells were put back into the incubator. As a transfection control, one flask received an RCAS-GFP vector instead of RCAS-BRAF^{V600E}.

5.2.11.2 Preparation and p0 injections

On the day of birth, i.e. usually three days after transfection, GFP fluorescence was confirmed in the control cells. The DF-1 cells that had been transfected with RCAS-BRAF^{V600E} were detached using 10x Trypsin-EDTA, centrifuged and resuspended in 1 mL medium. Subsequently, cell number was determined using a TC20 automated cell counter. The cells were centrifuged again, resuspended in DMEM to a concentration of 100,000 cells/µL and kept on ice.

The p0 injections were performed under a sterile hood. A 10 µL Hamilton syringe was filled with the prepared cell suspension, and 1 µL was slowly injected into the striatum of each newborn mouse by puncturing the left cerebral cortex from the lateral side in a 45° angle and inserting the needle to a depth of 2-3 mm. To avoid leakage of the injected suspension, the syringe was slowly retracted from the brain. Subsequently, all pups were put back into the cage and observed until the mother took care of them. All animal procedures were performed according to protocols approved by the German authorities (Regierungspräsidium Karlsruhe; G-69/13, G-212/16, G-38/18, G-229/18, DKFZ342).

5.2.11.3 Follow-up and tissue processing

From that point on, mice were observed daily for symptoms. At an age of three weeks, they were weaned and separated by sex. Mice were sacrificed after different timepoints but latest when first symptoms became visible. Isolated brains were either snap frozen in liquid nitrogen and stored at -80°C until nucleic acid extraction (see 5.2.1) or stored in 10% formalin at 4°C until paraffin embedding (see 5.2.12). As healthy controls for gene expression analysis, corresponding brain regions were isolated from age-matched but non-injected Ntv-a and XFM mice.

5.2.12 Histology

For paraffin embedding, brains were first dehydrated using increasing concentrations of EtOH: 70% EtOH, 2x 80% EtOH and 2x 100% EtOH. Subsequently, they were put twice into xylene and twice into melted paraffin baths. The whole procedure was carried out automatically in a HistoStar embedding workstation and took about 60 hours. Finally, the brains were placed into embedding cassettes, which were then filled with melted paraffin and allowed to cool down.

Paraffin blocks were cut using a microtome. The 3 µm sections were shortly kept in a water bath (42°C) before being mounted onto microscope slides. To ensure complete drying, the slides were placed in a 37°C incubator overnight.

5.2.12.1 H&E staining

Prior to hematoxylin and eosin (H&E) staining, sections were deparaffinized in xylene for 2x 15 minutes and rehydrated with 100% EtOH, 100% EtOH, 95% EtOH, 70% EtOH and deionized water for 5 minutes, respectively. To color nuclei blue, the sections were stained with hematoxylin for 1.5 minutes. Subsequently, they were placed in tap water for up to 5 minutes and rinsed with deionized water to remove excess staining. Eosin was used as a counterstain, as it colors the cytoplasm and some components of the extracellular matrix pink. After staining the sections for 1 minute, they were rinsed with deionized water again to remove excess staining. This was followed by 5 minutes in 70% EtOH, 95% EtOH, 100% EtOH and 100% EtOH, respectively, and 2x 15 minutes in xylene. As a last step, sections were embedded in Eukitt mounting medium, covered with a cover glass and let dry overnight.

5.2.12.2 Reticulin staining

For the reticulin staining, sections were deparaffinized and rehydrated as described for the H&E staining (see 5.2.12.1). Subsequently, sections were placed in 2% potassium permanganate, 2% potassium disulphite, 2% ammonium iron sulphate and the silver staining solution (see Table 5.2) for 1 minute, respectively, and in 4% formaldehyde for 3 minutes. Between each step, the sections were briefly washed in deionized water. To avoid formaldehyde deposits, sections were rinsed with tap water for 3 minutes and briefly washed in deionized water. Then, sections were placed into 0.1% tetrachloroauric acid for 1 minute, briefly washed in deionized water, placed in 2% potassium disulphite and 5% sodium thiosulfate for 1 minute, respectively, and again briefly washed in deionized water. To visualize nuclei, sections were stained with a solution of 5% aluminum sulfate and 0.001% nuclear fast red for 3 minutes. Excess staining was washed away with deionized water. Finally, sections were dehydrated using 70% EtOH, 95% EtOH, 2x 100% EtOH and 2x xylene and embedded in Eukitt mounting medium similar to the H&E staining.

5.2.13 Statistical analysis

Statistical analysis was performed with GraphPad Prism or the R2 platform. The statistical test for each experiment is mentioned in the figure caption. Significance levels are depicted as *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, ****: $p < 0.0001$, n.s.: not significant.

A   Appendix

A.1   List of abbreviations

Abbreviation	Explanation
%	percent
°	degree
°C	degree Celsius
AAP	anaplastic astrocytoma with piloid features
ALK	anaplastic lymphoma kinase
ANOVA	analysis of variance
ASLV	avian sarcoma-leukosis virus
ATCC	American Type Culture Collection
ATRX	ATRX chromatin remodeler
bp	base pair
BRAF	B-Raf proto-oncogene, serine/threonine kinase
CA10	carbonic anhydrase 10
CAR	chimeric antigen receptor
Cas9	CRISPR associated protein 9
CB	cerebellum
CCND1	cyclin D1
CCNO	cyclin O
CD	cluster of differentiation
CDK4/6	cyclin dependent kinase 4/6
CDKN2A/B	cyclin dependent kinase inhibitor 2A/B
cDNA	complementary DNA
Cereb	cerebellar
CIC	capicua transcriptional repressor
CLCN6	chloride voltage-gated channel 6
CLIP2	CAP-Gly domain containing linker protein 2
CNS	central nervous system
CNV	copy number variation
CR1/2	conserved region 1/2
CRISPR	clustered regularly interspaced short palindromic repeats
Ctx	cerebral cortex
Del	deletion
DGONC	diffuse glioneuronal tumor with oligodendroglial features and nuclear clusters
DLGNT	diffuse leptomeningeal glioneuronal tumor
DMEM	Dulbecco's Modified Eagle's Medium
DMG	diffuse midline glioma
DNA	deoxyribonucleic acid
DNET	dysembryoplastic neuroepithelial tumor
dNTP	deoxynucleotide triphosphate
DUSP4/6	dual specificity phosphatase 4/6
DUX4	double homeobox 4
E2F7	E2F transcription factor 7
ECM	extracellular matrix
EDTA	ethylenediaminetetraacetic acid

eGBM	epithelioid glioblastoma
EGF	epidermal growth factor
EGFR	epidermal growth factor receptor
EPHA2/4	EPH receptor A2/4
EtOH	ethanol
ETV4/5	ETS variant transcription factor 4/5
ETV6	ETS variant transcription factor 6
EVNCYT	extraventricular neurocytoma
Ex	exon
FAM131B	family with sequence similarity 131 member B
FBS	fetal bovine serum
FDR	false discovery rate
FFPE	formalin-fixed, paraffin-embedded
FGFR1/2/3	fibroblast growth factor receptor 1/2/3
Fus	fusion
GBM	glioblastoma
GEMM	genetically engineered mouse model
GFRA1	GDNF family receptor alpha 1
GG	ganglioglioma
GNAI1	G protein subunit alpha i1
GO	gene ontology
H&E	hematoxylin and eosin
HDAC	histone deacetylase
HGG	high-grade glioma
ICI	immune checkpoint inhibitor
IDH1/2	isocitrate dehydrogenase 1/2
IGV	Integrative Genomics Viewer
IHC	immunohistochemistry
INFORM	Individualized Therapy for Relapsed Malignancies in Childhood
IVC	individually ventilated cage
kb	kilobase
KO	knockout
KRAS	KRAS proto-oncogene, GTPase
lcWGS	low-coverage whole genome sequencing
LGG	low-grade glioma
LTR	long terminal repeat
M	molar
MAPK	mitogen-activated protein kinase
MDM2	MDM2 proto-oncogene
MET	MET proto-oncogene, receptor tyrosine kinase
min	minute
MKI67	marker of proliferation Ki-67
MKRN1	makorin ring finger protein 1
mL	milliliter
mM	millimolar
mm	millimeter
MPAS	MAPK Pathway Activity Score
mRNA	messenger RNA
Mut	mutation
MYB	MYB proto-oncogene, transcription factor
MYBL1	MYB proto-oncogene like 1

MYXGNT	myxoid glioneuronal tumor of the septum pellucidum / lateral ventricle
n.s.	not significant
NACC2	NACC family member 2
NF1	neurofibromatosis type 1
NF1	neurofibromin 1
ng	nanogram
NOS	not otherwise specified
NSC	neural stem cell
NTRK	neurotrophic receptor tyrosine kinase
Ntv-a	Nestin-tv-a
OIS	oncogene-induced senescence
OTP	One Touch Pipeline
OTX1	orthodenticle homeobox 1
PA	pilocytic astrocytoma
PBS	phosphate buffered saline
PC	principle component
PCA	principle component analysis
PDGFRA	platelet derived growth factor receptor alpha
PDX	patient-derived xenograft
PGNT	papillary glioneuronal tumor
pHGG	pediatric high-grade glioma
PHLDA1	pleckstrin homology like domain family A member 1
PID1	phosphotyrosine interaction domain containing 1
pLGG	pediatric low-grade glioma
PLNTY	pleomorphic neuroepithelial tumor of the young
POSTN	periostin
PRKCA	protein kinase C alpha
PTEN	phosphatase and tensin homolog
PTPN11	protein tyrosine phosphatase non-receptor type 11
PXA	pleomorphic xanthoastrocytoma
QKI	QKI, KH domain containing RNA binding
RAF1	Raf-1 proto-oncogene, serine/threonine kinase
Rb	retinoblastoma
RCAS	replication-competent ASLV LTR with splice acceptor
RNA	ribonucleic acid
RNA-Seq	RNA sequencing
RNF130	ring finger protein 130
RSV-A	Rous sarcoma virus A
SASP	senescence-associated secretory phenotype
SD	standard deviation
sec	second
SNP	single nucleotide polymorphism
SPECC1L	sperm antigen with calponin homology and coiled-coil domains 1 like
SPF	specific pathogen-free
SPRY2/4	sprouty RTK signaling antagonist 2/4
SRGAP3	SLIT-ROBO Rho GTPase activating protein 3
SVZ	subventricular zone
t-SNE	t-distributed stochastic neighbor embedding
TACC1	transforming acidic coiled-coil containing protein 1
TAM	tumor-associated microglia or macrophage
TERT	telomerase reverse transcriptase

TFG	trafficking from ER to golgi regulator
TNFRSF12A	TNF receptor superfamily member 12A
TOP2A	DNA topoisomerase II alpha
TP53	tumor protein 53
TPM	transcripts per kilobase million
TSS	transcription start site
TVA	tumor virus A
U	unit
V	volt
VP	vaginal plug
WES	whole exome sequencing
WGS	whole genome sequencing
WHO	World Health Organization
WT	wildtype
XFM	Ntv-a Cdkn2a$^{-/-}$
μg	microgram
μL	microliter
μm	micrometer

A.2 List of figures

A.3 List of tables

A.4   Talks, poster presentations and publications

I have presented some of the results included in this dissertation at various scientific events, either as a talk (see A.4.1) or as a poster (see A.4.2). In addition, I have contributed to the following publications (see A.4.3).

A.4.1   Talks

SIOP-LGG Preclinical Working Group Meeting 2018
13.09. – 14.09.2018 in Cagliari, Italy

Meeting of the Everest Centre for Paediatric Low-Grade Brain Tumour Research
11.03.2019 in London, United Kingdom

PhD Student Forum of the German Academic Scholarship Foundation
14.03. – 17.03.2019 in Mannheim, Germany

SIOP-LGG Preclinical Working Group Meeting 2019
04.04.2019 in Budapest, Hungary

Retreat of the Hopp Children's Cancer Center Heidelberg (KiTZ)
18.07. – 19.07.2019 in Grünstadt, Germany

DFCI – EVEREST – SIOP-E Low-Grade Brain Tumor Meeting
14.01. – 15.01.2020 in Ladenburg, Germany

A.4.2   Poster presentations

Helmholtz International Graduate School for Cancer Research – PhD Retreat
16.07. – 18.07.2018 in Weil der Stadt, Germany

Helmholtz International Graduate School for Cancer Research – Poster Presentation
22.11.2019 in Heidelberg, Germany

A.4.3   Publications

Buhl, J.L., Selt, F., Hielscher, T., Guiho, R., Ecker, J., Sahm, F., Ridinger, J., Riehl, D., Usta, D., Ismer, B., **Sommerkamp, A.C.**, et al. (2019). The Senescence-associated Secretory Phenotype Mediates Oncogene-induced Senescence in Pediatric Pilocytic Astrocytoma. Clin. Cancer Res. 25, 1851–1866.

Usta, D., Sigaud, R., Buhl, J.L., Selt, F., Marquardt, V., Pauck, D., Jansen, J., Pusch, S., Ecker, J., Hielscher, T., Vollmer, J., **Sommerkamp, A.C.**, et al. (2020). A cell-based MAPK reporter assay reveals synergistic MAPK pathway activity suppression by MAPK inhibitor combination in BRAF-driven pediatric low-grade glioma cells. Mol. Cancer Ther.

Sommerkamp, A.C., Uhrig, S., Stichel, D., St-Onge, P., Sun, P., Jäger, N., Deimling, von, A., Sahm, F., Pfister, S.M., Korshunov, A., et al. (2020). An optimized workflow to improve reliability of detection of KIAA1549:BRAF fusions from RNA sequencing data. Acta Neuropathol. 17, 257–3.

B Contributions

The work presented in this dissertation would not have been possible without the help of many people. Some of the individual contributions are recognized in the following.

Andrea Wittmann, Laura von Soosten, Norman Mack, Benjamin Schwalm and Britta Statz provided excellent technical assistance (e.g. tumor isolation and nucleic acid extraction).

RNA sequencing, microarray analysis and methylation analysis was performed by the Genomics and Proteomics Core Facility (GPCF) at the DKFZ. The starting material as well as the processing and sequencing parameters were provided by me.

Import of the methylation array data into the brain tumor classifier was done by Dr. Martin Sill (Division of Pediatric Neurooncology, DKFZ).

Routine processing of the raw RNA sequencing data, i.e. alignment (STAR), quality control (RNA-SeQC) and fusion calling (Arriba), with standard parameters was performed by the Omics IT and Data Management Core Facility (ODCF) at the DKFZ. Analysis with adjusted/optimized parameters was conducted in collaboration with bioinformaticians (see below).

The ESTIMATE analysis was conducted in collaboration with Pengbo Sun (Division of Pediatric Neurooncology, DKFZ).

The cross-species analysis was performed by Dr. Konstantin Okonechnikov (Division of Pediatric Neurooncology, DKFZ) based on the sequencing data that I had generated. Selection of human and murine samples for the analysis was done by me.

Adjustment of STAR parameters and development of a new Arriba version was done in a collaborative effort with Sebastian Uhrig (Division of Applied Bioinformatics, DKFZ). In detail, the strategy for workflow optimization was discussed by Sebastian Uhrig, Dr. David Jones and myself. Implementation as well as analysis of the sequencing cohort with the new parameters was performed by Sebastian Uhrig. Similarly, FusionCatcher was run by Sebastian Uhrig. The results were visualized by me.

Analysis of the validation cohort from Montreal was done by Pascal St-Onge (Division of Hematology-Oncology, Charles-Bruneau Cancer Centre, CHU Sainte-Justine, Montreal, Canada). He also provided previous analysis results of FusionCatcher and STAR-Fusion.

Evaluation of the optimized workflow for the FFPE RNA-Seq dataset was done in collaboration with Dr. Damian Stichel (Department of Neuropathology, University Hospital Heidelberg).

C   Acknowledgments

The fact that I was forced to switch research groups and to work on a new project 1.5 years after starting my doctoral studies made this a special, sometimes frustrating but in the end remarkable and rewarding experience. I was not only able to immerse myself in different areas of today's molecular research and learn so many new things but also to develop as a person. I am proud to have contributed to the progress of translational pediatric cancer research, even if it is only a very small step on the way to a cure for cancer. This would not have been possible without the continuous support of many people.

First of all, I would like to thank **Dr. David Jones** for giving me the chance to work on a new and exciting PhD project. I have always enjoyed our (sometimes longer than expected) scientific discussions and exchange of ideas. Thank you also for being a great supervisor and always having a sympathetic ear for my questions. I very much appreciate that you allowed me to attend various national and international scientific meetings, represent your group and share my results. At the same time, I am grateful that you supported my non-academic activities in a similar way. Thank you for your ongoing support and the trust you put in me.

I would also like to extend my thanks to **Prof. Dr. Stefan Pfister**. I am deeply grateful that you readily accepted me into your division in a difficult situation and thus saved me a lot of additional trouble. This generosity is also part of the wonderful environment that you have created within the division and as part of the KiTZ. I am always impressed by what you have achieved and look forward to visiting the KiTZ building once it is finished. Thank you also for being a great leader. You always treat people with fairness and respect, irrespectively of hierarchies, and enable them to do great science. Your scientific input as well as your pragmatic approach have definitely helped me to progress.

Moreover, I would like to thank the members of my thesis advisory committee, **Prof. Dr. Peter Angel, Prof. Dr. David Capper and Prof. Dr. Stefan Pfister**, for providing invaluable scientific input and guidance for my PhD project and making sure that I am on track.

In addition, I would like to thank **Prof. Dr. Peter Angel, Prof. Dr. Stefan Pfister, Prof. Dr. Frank Lyko and Dr. Darjus Tschaharganeh** for their willingness to be part of my examination commission and for evaluating my thesis.

A big thanks goes out to **Britta I.**, who laid the foundation for the PA/PXA project. I am very grateful that I could build on your initial work and had someone to introduce me to this topic and the different methods. We immediately understood each other well, and this allowed me to integrate into the group very quickly. Thank you also for your positive energy and

enthusiasm. I always remember our joint commitment for the KiTZ club and the fun competitions with pleasure. You are a great friend and I very much enjoyed the time when you were still in the lab.

Furthermore, I would like to express my gratitude to **Sebastian**, who I worked with for the fusion detection project. Your expertise and high working speed allowed us to finish and publish this project so quickly. Thank you also for answering the numerous questions on the processing and analysis of sequencing data I had. Without you, this project would not have been possible.

I would like to thank **all technicians** of B360 and B062 for your experience and invaluable support of all the different projects that are being worked on in the lab. Particularly, I would like to thank **Andrea** not only for the countless DNA/RNA extractions of various tumor samples but also for keeping an overview of where the samples are stored. Thanks for being such a kind person and always being there to help me when needed. In addition, I would like to thank **Norman and Benni** for keeping all the mousework running and being a lot of fun to work with. You guys saved me numerous times by identifying and extracting tiny mouse tumors that simply looked like normal brain tissue to me. Furthermore, I would like to thank **Laura** for your help with genotyping and **Britta S.** for your support with the microtome.

In addition, I would like to extend my thanks to the members of the **DKFZ core facilities**, including the Genomics and Proteomics Core Facility, the Omics IT and Data Management Core Facility and the Central Animal Laboratory. You all are doing a great job and make a decisive contribution to the scientific success of the DKFZ.

Moreover, I am very grateful to our collaborators, including **Dr. Annika Wefers** and **Dr. Damian Stichel** at the Neuropathology department in Heidelberg as well as **Prof. Dr. Nada Jabado** and **Pascal St-Onge** in Montreal. Sciences lives from joint work and collaborative exchange, and your contributions were critical for advancing these projects.

A big thanks goes out to **all present and former members of B360 and B062**. It has been a great pleasure to work with all of you. You are a great group of people that strive to support each other and create an atmosphere in which everyone is willing to share experience and help.

Jasmin and **Ines**, thank you a lot for your help with all kinds of organizational questions regarding working contracts, business travel, seminars and advanced training. Thank you also for understanding the cryptic forms of German bureaucracy and always finding a way to squeeze appointments into already full calendars.

Furthermore, I would like to thank the people in my office, **Agata**, **Anna-Lisa**, **Barbara**, **Dominik**, **Elke**, **Kathrin**, **Max**, **Michaela**, **Mija, Pascal** and **Sebastian**, for all the fun

conversations, the mutual support, the sweets from all over the world, the discussions on the perfect room temperature as well as the "excitement" of finding out what is hiding inside the coffee machine. You are all really great people and I would not have liked to share the room with anyone else.

Huge thanks also to my amazing colleagues and friends **Andrea**, **Annette**, **Felix G.**, **Felix S.**, **Jens**, **Julia**, **Kristin**, **Lisa**, **Luisa**, **Ralph**, **Sarah**, **Sonja** and **Vera**. Thank you for all the great moments and memorable activities, including lunch and coffee breaks, (winter) sports activities, brunch and dinner parties. All of that was great fun and helped me to clear my mind.

Zu guter Letzt möchte ich mich besonders bei meiner wunderbaren Familie bedanken. **Ines**, **Sascha**, **Simon** und **Lina**, vielen Dank, dass man sich bei euch immer wohlfühlen und auf andere Gedanken kommen kann. **Marlies** und **Harald**, vielen Dank, dass ihr als Schwiegereltern mich so herzlich in eure Familie aufgenommen und mir in den letzten 9 Jahren immer Interesse und Unterstützung entgegengebracht habt. Ganz besonderer Dank gilt meinen Eltern, **Anne** und **Axel**. Vielen Dank für eure immerwährende Unterstützung und euren Rückhalt, eure bedingungslose Liebe und euer Vertrauen. Vielen Dank, dass ihr mir alles ermöglicht und mich auf meinem Weg immer bestärkt habt. In guten wie in schweren Zeiten kann ich stets auf euch zählen! Abschließend möchte ich noch meiner Frau, **Pia**, danken. Vielen Dank, dass du mich immer verstehst und unterstützt. Vielen Dank auch, dass du mich bei Rückschlägen immer aufgebaut hast und für mich da warst. Ohne dich wäre ich nicht da, wo ich jetzt bin.

D Bibliography

Ahronian, L.G., and Lewis, B.C. (2014). Using the RCAS-TVA system to model human cancer in mice. Cold Spring Harb Protoc *2014*, 1128–1135.

Aichmüller, C.F., Iskar, M., Jones, D.T.W., Korshunov, A., Radlwimmer, B., Kool, M., Ernst, A., Pfister, S.M., Lichter, P., and Zapatka, M. (2020). Pilocytic Astrocytoma demethylation and transcriptional landscapes link bZIP transcription factors to immune response. Neuro-Oncology.

Aisner, D.L., Newell, K.L., Pollack, A.G., Kleinschmidt-DeMasters, B.K., Steinberg, G.K., Smyth, L.T., and Vogel, H. (2014). Composite pleomorphic xanthoastrocytoma-epithelioid glioneuronal tumor with BRAF V600E mutation - report of three cases. Clin. Neuropathol. *33*, 112–121.

Alexandrescu, S., Korshunov, A., Lai, S.H., Dabiri, S., Patil, S., Li, R., Shih, C.-S., Bonnin, J.M., Baker, J.A., Du, E., et al. (2016). Epithelioid Glioblastomas and Anaplastic Epithelioid Pleomorphic Xanthoastrocytomas--Same Entity or First Cousins? Brain Pathol. *26*, 215–223.

Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A.J.R., Behjati, S., Biankin, A.V., Bignell, G.R., Bolli, N., Borg, A., Børresen-Dale, A.-L., et al. (2013). Signatures of mutational processes in human cancer. Nature *500*, 415–421.

Antonelli, M., Fadda, A., Loi, E., Moi, L., Zavattari, C., Sulas, P., Gentilini, D., Cameli, C., Bacchelli, E., Badiali, M., et al. (2018). Integrated DNA methylation analysis identifies topographical and tumoral biomarkers in pilocytic astrocytomas. Oncotarget *9*, 13807–13821.

Ater, J.L., Xia, C., Mazewski, C.M., Booth, T.N., Freyer, D.R., Packer, R.J., Sposto, R., Vezina, G., and Pollack, I.F. (2016). Nonrandomized comparison of neurofibromatosis type 1 and non-neurofibromatosis type 1 children who received carboplatin and vincristine for progressive low-grade glioma: A report from the Children's Oncology Group. Cancer *122*, 1928–1936.

Ater, J.L., Zhou, T., Holmes, E., Mazewski, C.M., Booth, T.N., Freyer, D.R., Lazarus, K.H., Packer, R.J., Prados, M., Sposto, R., et al. (2012). Randomized study of two chemotherapy regimens for treatment of low-grade glioma in young children: a report from the Children's Oncology Group. J. Clin. Oncol. *30*, 2641–2647.

Ballester, L.Y., Penas-Prado, M., Leeds, N.E., Huse, J.T., and Fuller, G.N. (2018). FGFR1 tyrosine kinase domain duplication in pilocytic astrocytoma with anaplasia. Cold Spring Harb Mol Case Stud *4*.

Banerjee, A., Jakacki, R.I., Onar-Thomas, A., Wu, S., Nicolaides, T., Young Poussaint, T., Fangusaro, J., Phillips, J., Perry, A., Turner, D., et al. (2017). 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-Oncology *19*, 1135–1144.

Bates, P., Young, J.A., and Varmus, H.E. (1993). A receptor for subgroup A Rous sarcoma virus is related to the low density lipoprotein receptor. Cell *74*, 1043–1051.

Bernal, A., and Arranz, L. (2018). Nestin-expressing progenitor cells: function, identity and therapeutic implications. Cell. Mol. Life Sci. *75*, 2177–2195.

Bhakta, N., Liu, Q., Ness, K.K., Baassiri, M., Eissa, H., Yeo, F., Chemaitilly, W., Ehrhardt, M.J., Bass, J., Bishop, M.W., et al. (2017). The cumulative burden of surviving childhood cancer: an initial report from the St Jude Lifetime Cohort Study (SJLIFE). Lancet *390*, 2569–2582.

Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J.M., Delano, D., Zhang, L., Schroth, G.P., Gunderson, K.L., et al. (2011). High density DNA methylation array with single CpG site resolution. Genomics *98*, 288–295.

Bid, H.K., Kibler, A., Phelps, D.A., Manap, S., Xiao, L., Lin, J., Capper, D., Oswald, D., Geier, B., DeWire, M., et al. (2013). Development, characterization, and reversal of acquired resistance to the MEK1 inhibitor selumetinib (AZD6244) in an in vivo model of childhood astrocytoma. Clin. Cancer Res. *19*, 6716–6729.

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.

Brabetz, S., Leary, S.E.S., Gröbner, S.N., Nakamoto, M.W., Şeker-Cin, H., Girard, E.J., Cole, B., Strand, A.D., Bloom, K.L., Hovestadt, V., et al. (2018). A biobank of patient-derived pediatric brain tumor models. Nat. Med. *24*, 1752–1761.

Bridge, J.A., Liu, X.-Q., Sumegi, J., Nelson, M., Reyes, C., Bruch, L.A., Rosenblum, M., Puccioni, M.J., Bowdino, B.S., and McComb, R.D. (2013). Identification of a novel, recurrent SLC44A1-PRKCA fusion in papillary glioneuronal tumor. Brain Pathol. *23*, 121–128.

Brocks, D., Schmidt, C.R., Daskalakis, M., Jang, H.S., Shah, N.M., Li, D., Li, J., Zhang, B., Hou, Y., Laudato, S., et al. (2017). DNMT and HDAC inhibitors induce cryptic transcription start sites encoded in long terminal repeats. Nat. Genet. *49*, 1052–1060.

Broekman, M.L., Maas, S.L.N., Abels, E.R., Mempel, T.R., Krichevsky, A.M., and Breakefield, X.O. (2018). Multidimensional communication in the microenvirons of glioblastoma. Nat Rev Neurol *14*, 482–495.

Buhl, J.L., Selt, F., Hielscher, T., Guiho, R., Ecker, J., Sahm, F., Ridinger, J., Riehl, D., Usta, D., Ismer, B., et al. (2019). The Senescence-associated Secretory Phenotype Mediates Oncogene-induced Senescence in Pediatric Pilocytic Astrocytoma. Clin. Cancer Res. *25*, 1851–1866.

Burkhard, C., Di Patre, P.-L., Schüler, D., Schüler, G., Yaşargil, M.G., Yonekawa, Y., Lütolf, U.M., Kleihues, P., and Ohgaki, H. (2003). A population-based study of the incidence and survival rates in patients with pilocytic astrocytoma. J. Neurosurg. *98*, 1170–1174.

Byron, S.A., Van Keuren-Jensen, K.R., Engelthaler, D.M., Carpten, J.D., and Craig, D.W. (2016). Translating RNA sequencing into clinical diagnostics: opportunities and challenges. Nat. Rev. Genet. *17*, 257–271.

Cancer Genome Atlas Research Network, Brat, D.J., Verhaak, R.G.W., Aldape, K.D., Yung, W.K.A., Salama, S.R., Cooper, L.A.D., Rheinbay, E., Miller, C.R., Vitucci, M., et al. (2015).

Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. N Engl J Med *372*, 2481–2498.

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. (2018a). DNA methylation-based classification of central nervous system tumours. Nature *555*, 469–474.

Capper, D., Stichel, D., Sahm, F., Jones, D.T.W., Schrimpf, D., Sill, M., Schmid, S., Hovestadt, V., Reuss, D.E., Koelsche, C., et al. (2018b). Practical implementation of DNA methylation and copy-number-based CNS tumor diagnostics: the Heidelberg experience. Acta Neuropathol. *136*, 181–210.

Chambers, S.M., Boles, N.C., Lin, K.-Y.K., Tierney, M.P., Bowman, T.V., Bradfute, S.B., Chen, A.J., Merchant, A.A., Sirin, O., Weksberg, D.C., et al. (2007). Hematopoietic fingerprints: an expression database of stem cells and their progeny. Cell Stem Cell *1*, 578–591.

Chen, R., Keoni, C., Waker, C.A., Lober, R.M., Chen, Y.-H., and Gutmann, D.H. (2019). KIAA1549-BRAF Expression Establishes a Permissive Tumor Microenvironment Through NFκB-Mediated CCL2 Production. Neoplasia *21*, 52–60.

Chintagumpala, M., Eckel, S.P., Krailo, M., Morris, M., Adesina, A., Packer, R., Lau, C., and Gajjar, A. (2015). A pilot study using carboplatin, vincristine, and temozolomide in children with progressive/symptomatic low-grade glioma: a Children's Oncology Group study†. Neuro-Oncology *17*, 1132–1138.

Cin, H., Meyer, C., Herr, R., Janzarik, W.G., Lambert, S., Jones, D.T.W., Jacob, K., Benner, A., Witt, H., Remke, M., et al. (2011). Oncogenic FAM131B-BRAF fusion resulting from 7q34 deletion comprises an alternative mechanism of MAPK pathway activation in pilocytic astrocytoma. Acta Neuropathol. *121*, 763–774.

Cintolo, J.A., Datta, J., Xu, S., Gupta, M., Somasundaram, R., and Czerniecki, B.J. (2016). Type I-polarized BRAF-pulsed dendritic cells induce antigen-specific CD8+ T cells that impact BRAF-mutant murine melanoma. Melanoma Res. *26*, 1–11.

Clarke, M., Mackay, A., Ismer, B., Pickles, J.C., Tatevossian, R.G., Newman, S., Bale, T.A., Stoler, I., Izquierdo, E., Temelso, S., et al. (2020). Infant High-Grade Gliomas Comprise Multiple Subgroups Characterized by Novel Targetable Gene Fusions and Favorable Outcomes. Cancer Discov.

Cock, P.J.A., Fields, C.J., Goto, N., Heuer, M.L., and Rice, P.M. (2010). The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. Nucleic Acids Res. *38*, 1767–1771.

Collins, V.P., Jones, D.T.W., and Giannini, C. (2015). Pilocytic astrocytoma: pathology, molecular mechanisms and markers. Acta Neuropathol. *129*, 775–788.

Dahiya, S., Yu, J., Kaul, A., Leonard, J.R., and Gutmann, D.H. (2012). Novel BRAF Alteration in a Sporadic Pilocytic Astrocytoma. Case Rep Med *2012*, 418672.

Davies, H., Bignell, G.R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M.J., Bottomley, W., et al. (2002). Mutations of the BRAF gene in human cancer. Nature *417*, 949–954.

Davila, J.I., Fadra, N.M., Wang, X., McDonald, A.M., Nair, A.A., Crusan, B.R., Wu, X., Blommel, J.H., Jen, J., Rumilla, K.M., et al. (2016). Impact of RNA degradation on fusion detection by RNA-seq. BMC Genomics *17*, 814–819.

Davis, K.L., Fox, E., Merchant, M.S., Reid, J.M., Kudgus, R.A., Liu, X., Minard, C.G., Voss, S., Berg, S.L., Weigel, B.J., et al. (2020). Nivolumab in children and young adults with relapsed or refractory solid tumours or lymphoma (ADVL1412): a multicentre, open-label, single-arm, phase 1-2 trial. Lancet Oncol. *21*, 541–550.

DeLuca, D.S., Levin, J.Z., Sivachenko, A., Fennell, T., Nazaire, M.-D., Williams, C., Reich, M., Winckler, W., and Getz, G. (2012). RNA-SeQC: RNA-seq metrics for quality control and process optimization. Bioinformatics *28*, 1530–1532.

Deng, M.Y., Sill, M., Sturm, D., Stichel, D., Witt, H., Ecker, J., Wittmann, A., Schittenhelm, J., Ebinger, M., Schuhmann, M.U., et al. (2019). Diffuse glioneuronal tumour with oligodendroglioma-like features and nuclear clusters (DGONC) - a molecularly defined glioneuronal CNS tumour class displaying recurrent monosomy 14. Neuropathol. Appl. Neurobiol.

Deng, M.Y., Sill, M., Chiang, J., Schittenhelm, J., Ebinger, M., Schuhmann, M.U., Monoranu, C.-M., Milde, T., Wittmann, A., Hartmann, C., et al. (2018). Molecularly defined diffuse leptomeningeal glioneuronal tumor (DLGNT) comprises two subgroups with distinct clinical and genetic features. Acta Neuropathol. *136*, 239–253.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Dobson, T.H.W., and Gopalakrishnan, V. (2018). Preclinical Models of Pediatric Brain Tumors-Forging Ahead. Bioengineering (Basel) *5*.

Eleveld, T.F., Oldridge, D.A., Bernard, V., Koster, J., Colmet Daage, L., Diskin, S.J., Schild, L., Bentahar, N.B., Bellini, A., Chicard, M., et al. (2015). Relapsed neuroblastomas show frequent RAS-MAPK pathway mutations. Nat. Genet. *47*, 864–871.

Engler, J.R., Robinson, A.E., Smirnov, I., Hodgson, J.G., Berger, M.S., Gupta, N., James, C.D., Molinaro, A., and Phillips, J.J. (2012). Increased microglia/macrophage gene expression in a subset of adult and pediatric astrocytomas. PLoS ONE *7*, e43339.

Fangusaro, J., Onar-Thomas, A., Young Poussaint, T., Wu, S., Ligon, A.H., Lindeman, N., Banerjee, A., Packer, R.J., Kilburn, L.B., Goldman, S., et al. (2019). 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. *20*, 1011–1022.

Fernandez, A.F., Assenov, Y., Martin-Subero, J.I., Balint, B., Siebert, R., Taniguchi, H., Yamamoto, H., Hidalgo, M., Tan, A.C., Galm, O., et al. (2012). A DNA methylation fingerprint of 1628 human samples. Genome Research *22*, 407–419.

Forshew, T., Tatevossian, R.G., Lawson, A.R.J., Ma, J., Neale, G., Ogunkolade, B.W., Jones, T.A., Aarum, J., Dalton, J., Bailey, S., et al. (2009). Activation of the ERK/MAPK pathway: a signature genetic defect in posterior fossa pilocytic astrocytomas. J. Pathol. *218*, 172–181.

Friedman, G.K., and Gillespie, G.Y. (2011). Cancer Stem Cells and Pediatric Solid Tumors. Cancers (Basel) *3*, 298–318.

Fukuoka, K., Mamatjan, Y., Tatevossian, R., Zapotocky, M., Ryall, S., Stucklin, A.G., Bennett, J., Nobre, L.F., Arnoldo, A., Luu, B., et al. (2020). Clinical impact of combined epigenetic and molecular analysis of pediatric low grade gliomas. Neuro-Oncology *19*, 750.

Furuta, T., Miyoshi, H., Komaki, S., Arakawa, F., Morioka, M., Ohshima, K., Nakada, M., and Sugita, Y. (2018). Clinicopathological and genetic association between epithelioid glioblastoma and pleomorphic xanthoastrocytoma. Neuropathology *38*, 218–227.

Gabriëls, J., Beckers, M.C., Ding, H., De Vriese, A., Plaisance, S., van der Maarel, S.M., Padberg, G.W., Frants, R.R., Hewitt, J.E., Collen, D., et al. (1999). Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. Gene *236*, 25–32.

Gallego Romero, I., Pai, A.A., Tung, J., and Gilad, Y. (2014). RNA-seq: impact of RNA degradation on transcript quantification. BMC Biol. *12*, 42.

Gao, H., Korn, J.M., Ferretti, S., Monahan, J.E., Wang, Y., Singh, M., Zhang, C., Schnell, C., Yang, G., Zhang, Y., et al. (2015). High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response. Nat. Med. *21*, 1318–1325.

Gao, Q., Liang, W.-W., Foltz, S.M., Mutharasu, G., Jayasinghe, R.G., Cao, S., Liao, W.-W., Reynolds, S.M., Wyczalkowski, M.A., Yao, L., et al. (2018). Driver Fusions and Their Implications in the Development and Treatment of Human Cancers. Cell Rep *23*, 227–238.e3.

Garnett, M.J., and Marais, R. (2004). Guilty as charged: B-RAF is a human oncogene. Cancer Cell *6*, 313–319.

Geoerger, B., Kang, H.J., Yalon-Oren, M., Marshall, L.V., Vezina, C., Pappo, A., Laetsch, T.W., Petrilli, A.S., Ebinger, M., Toporski, J., et al. (2020a). Pembrolizumab in paediatric patients with advanced melanoma or a PD-L1-positive, advanced, relapsed, or refractory solid tumour or lymphoma (KEYNOTE-051): interim analysis of an open-label, single-arm, phase 1-2 trial. Lancet Oncol. *21*, 121–133.

Geoerger, B., Zwaan, C.M., Marshall, L.V., Michon, J., Bourdeaut, F., Casanova, M., Corradini, N., Rossato, G., Farid-Kapadia, M., Shemesh, C.S., et al. (2020b). Atezolizumab for children and young adults with previously treated solid tumours, non-Hodgkin lymphoma, and Hodgkin lymphoma (iMATRIX): a multicentre phase 1-2 study. Lancet Oncol. *21*, 134–144.

Giannini, C., Scheithauer, B.W., Burger, P.C., Brat, D.J., Wollan, P.C., Lach, B., and O'Neill, B.P. (1999). Pleomorphic xanthoastrocytoma: what do we really know about it? Cancer *85*, 2033–2045.

Gibson, P., Tong, Y., Robinson, G., Thompson, M.C., Currle, D.S., Eden, C., Kranenburg, T.A., Hogg, T., Poppleton, H., Martin, J., et al. (2010). Subtypes of medulloblastoma have distinct developmental origins. Nature *468*, 1095–1099.

Gnekow, A.K., Walker, D.A., Kandels, D., Picton, S., Giorgio Perilongo, Grill, J., Stokland, T., Sandstrom, P.E., Warmuth-Metz, M., Schmidt, R., et al. (2017). A European randomised controlled trial of the addition of etoposide to standard vincristine and carboplatin induction

as part of an 18-month treatment programme for childhood (≤16 years) low grade glioma - A final report. Eur. J. Cancer *81*, 206–225.

Griesinger, A.M., Birks, D.K., Donson, A.M., Amani, V., Hoffman, L.M., Waziri, A., Wang, M., Handler, M.H., and Foreman, N.K. (2013). Characterization of distinct immunophenotypes across pediatric brain tumor types. J. Immunol. *191*, 4880–4888.

Griesinger, A.M., Donson, A.M., and Foreman, N.K. (2014). Immunotherapeutic implications of the immunophenotype of pediatric brain tumors. Oncoimmunology *3*, e27256.

Grill, J., Massimino, M., Bouffet, E., Azizi, A.A., McCowage, G., Cañete, 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.

Gronych, J., Korshunov, A., Bageritz, J., Milde, T., Jugold, M., Hambardzumyan, D., Remke, M., Hartmann, C., Witt, H., Jones, D.T.W., et al. (2011). An activated mutant BRAF kinase domain is sufficient to induce pilocytic astrocytoma in mice. J. Clin. Invest. *121*, 1344–1348.

Gröbner, S.N., Worst, B.C., Weischenfeldt, J., Buchhalter, I., Kleinheinz, K., Rudneva, V.A., Johann, P.D., Balasubramanian, G.P., Segura-Wang, M., Brabetz, S., et al. (2018). The landscape of genomic alterations across childhood cancers. Nature *555*, 321–327.

Guo, C., Manjili, M.H., Subjeck, J.R., Sarkar, D., Fisher, P.B., and Wang, X.-Y. (2013). Therapeutic cancer vaccines: past, present, and future. Adv. Cancer Res. *119*, 421–475.

Guo, X., Pan, Y., and Gutmann, D.H. (2019). Genetic and genomic alterations differentially dictate low-grade glioma growth through cancer stem cell-specific chemokine recruitment of T cells and microglia. Neuro-Oncology *21*, 1250–1262.

Guo, X., Pan, Y., Xiong, M., Sanapala, S., Anastasaki, C., Cobb, O., Dahiya, S., and Gutmann, D.H. (2020). Midkine activation of CD8+ T cells establishes a neuron-immune-cancer axis responsible for low-grade glioma growth. Nat Commun *11*, 2177–15.

Haas, B.J., Dobin, A., Li, B., Stransky, N., Pochet, N., and Regev, A. (2019). Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. Genome Biol. *20*, 213–216.

Hambardzumyan, D., Amankulor, N.M., Helmy, K.Y., Becher, O.J., and Holland, E.C. (2009). Modeling Adult Gliomas Using RCAS/t-va Technology. Transl Oncol *2*, 89–95.

Hambardzumyan, D., Gutmann, D.H., and Kettenmann, H. (2016). The role of microglia and macrophages in glioma maintenance and progression. Nat. Neurosci. *19*, 20–27.

Hamid, O., Vanderwalde, A.M., Szeto, C., Reddy, S., and Pal, S.K. (2018). Differential expression of PD-L1 and immune biomarkers by age: Decreased expression in pediatric/AYA patients with advanced cancer. Annals of Oncology *29*, viii32.

Hargrave, D.R., Bouffet, E., Tabori, U., Broniscer, A., Cohen, K.J., Hansford, J.R., Geoerger, B., Hingorani, P., Dunkel, I.J., Russo, M.W., et al. (2019). Efficacy and Safety of Dabrafenib in Pediatric Patients with BRAF V600 Mutation-Positive Relapsed or Refractory Low-Grade Glioma: Results from a Phase I/IIa Study. Clin. Cancer Res. *25*, 7303–7311.

Hasselblatt, M., Riesmeier, B., Lechtape, B., Brentrup, A., Stummer, W., Albert, F.K., Sepehrnia, A., Ebel, H., Gerß, J., and Paulus, W. (2011). BRAF-KIAA1549 fusion transcripts

are less frequent in pilocytic astrocytomas diagnosed in adults. Neuropathol. Appl. Neurobiol. *37*, 803–806.

Havel, J.J., Chowell, D., and Chan, T.A. (2019). The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. Nat. Rev. Cancer *19*, 133–150.

Hawkins, C., Walker, E., Mohamed, N., Zhang, C., Jacob, K., Shirinian, M., Alon, N., Kahn, D., Fried, I., Scheinemann, K., et al. (2011). BRAF-KIAA1549 fusion predicts better clinical outcome in pediatric low-grade astrocytoma. Clin. Cancer Res. *17*, 4790–4798.

Hermans, E., and Hulleman, E. (2020). Patient-Derived Orthotopic Xenograft Models of Pediatric Brain Tumors: In a Mature Phase or Still in Its Infancy? Front Oncol *9*, 1418.

Heyer, E.E., and Blackburn, J. (2020). Sequencing Strategies for Fusion Gene Detection. Bioessays e2000016.

Hidalgo, M., Amant, F., Biankin, A.V., Budinská, E., Byrne, A.T., Caldas, C., Clarke, R.B., de Jong, S., Jonkers, J., Mælandsmo, G.M., et al. (2014). Patient-derived xenograft models: an emerging platform for translational cancer research. Cancer Discov *4*, 998–1013.

Himly, M., Foster, D.N., Bottoli, I., Iacovoni, J.S., and Vogt, P.K. (1998). The DF-1 chicken fibroblast cell line: transformation induced by diverse oncogenes and cell death resulting from infection by avian leukosis viruses. Virology *248*, 295–304.

Holland, E.C., Hively, W.P., DePinho, R.A., and Varmus, H.E. (1998). A constitutively active epidermal growth factor receptor cooperates with disruption of G1 cell-cycle arrest pathways to induce glioma-like lesions in mice. Genes Dev. *12*, 3675–3685.

Hou, Y., Pinheiro, J., Sahm, F., Reuss, D.E., Schrimpf, D., Stichel, D., Casalini, B., Koelsche, C., Sievers, P., Wefers, A.K., et al. (2019). Papillary glioneuronal tumor (PGNT) exhibits a characteristic methylation profile and fusions involving PRKCA. Acta Neuropathol. *137*, 837–846.

Hovestadt, V., Jones, D.T.W., Picelli, S., Wang, W., Kool, M., Northcott, P.A., Sultan, M., Stachurski, K., Ryzhova, M., Warnatz, H.-J., et al. (2014). Decoding the regulatory landscape of medulloblastoma using DNA methylation sequencing. Nature *510*, 537–541.

Hovestadt, V., Remke, M., Kool, M., Pietsch, T., Northcott, P.A., Fischer, R., Cavalli, F.M.G., Ramaswamy, V., Zapatka, M., Reifenberger, G., et al. (2013). Robust molecular subgrouping and copy-number profiling of medulloblastoma from small amounts of archival tumour material using high-density DNA methylation arrays. Acta Neuropathol. *125*, 913–916.

Hu, X., Wang, Q., Tang, M., Barthel, F., Amin, S., Yoshihara, K., Lang, F.M., Martinez-Ledesma, E., Lee, S.H., Zheng, S., et al. (2018). TumorFusions: an integrative resource for cancer-associated transcript fusions. Nucleic Acids Res. *46*, D1144–D1149.

Hughes, S.H., Greenhouse, J.J., Petropoulos, C.J., and Sutrave, P. (1987). Adaptor plasmids simplify the insertion of foreign DNA into helper-independent retroviral vectors. J. Virol. *61*, 3004–3012.

Huse, J.T., Snuderl, M., Jones, D.T.W., Brathwaite, C.D., Altman, N., Lavi, E., Saffery, R., Sexton-Oates, A., Blümcke, I., Capper, D., et al. (2017). Polymorphous low-grade neuroepithelial tumor of the young (PLNTY): an epileptogenic neoplasm with

oligodendroglioma-like components, aberrant CD34 expression, and genetic alterations involving the MAP kinase pathway. Acta Neuropathol. *133*, 417–429.

Ida, C.M., Rodriguez, F.J., Burger, P.C., Caron, A.A., Jenkins, S.M., Spears, G.M., Aranguren, D.L., Lachance, D.H., and Giannini, C. (2015). Pleomorphic Xanthoastrocytoma: Natural History and Long-Term Follow-Up. Brain Pathol. *25*, 575–586.

International Cancer Genome Consortium PedBrain Tumor Project (2016). Recurrent MET fusion genes represent a drug target in pediatric glioblastoma. Nat. Med. *22*, 1314–1320.

Jacob, K., Quang-Khuong, D.-A., Jones, D.T.W., Witt, H., Lambert, S., Albrecht, S., Witt, O., Vezina, C., Shirinian, M., Faury, D., et al. (2011). Genetic aberrations leading to MAPK pathway activation mediate oncogene-induced senescence in sporadic pilocytic astrocytomas. Clin. Cancer Res. *17*, 4650–4660.

Johnson, W.E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics *8*, 118–127.

Jones, D.T.W., Kocialkowski, S., Liu, L., Pearson, D.M., Ichimura, K., and Collins, V.P. (2009). Oncogenic RAF1 rearrangement and a novel BRAF mutation as alternatives to KIAA1549:BRAF fusion in activating the MAPK pathway in pilocytic astrocytoma. Oncogene *28*, 2119–2123.

Jones, D.T.W., Bandopadhayay, P., and Jabado, N. (2019a). The Power of Human Cancer Genetics as Revealed by Low-Grade Gliomas. Annu. Rev. Genet. *53*, 483–503.

Jones, D.T.W., Banito, A., Grünewald, T.G.P., Haber, M., Jäger, N., Kool, M., Milde, T., Molenaar, J.J., Nabbi, A., Pugh, T.J., et al. (2019b). Molecular characteristics and therapeutic vulnerabilities across paediatric solid tumours. Nat. Rev. Cancer *19*, 420–438.

Jones, D.T.W., Gronych, J., Lichter, P., Witt, O., and Pfister, S.M. (2012). MAPK pathway activation in pilocytic astrocytoma. Cell. Mol. Life Sci. *69*, 1799–1811.

Jones, D.T.W., Hutter, B., Jäger, N., Korshunov, A., Kool, M., Warnatz, H.-J., Zichner, T., Lambert, S.R., Ryzhova, M., Quang, D.A.K., et al. (2013). Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. Nat. Genet. *45*, 927–932.

Jones, D.T.W., Kieran, M.W., Bouffet, E., Alexandrescu, S., Bandopadhayay, P., Bornhorst, M., Ellison, D., Fangusaro, J., Fisher, M.J., Foreman, N., et al. (2018a). Pediatric low-grade gliomas: next biologically driven steps. Neuro-Oncology *20*, 160–173.

Jones, D.T.W., Kocialkowski, S., Liu, L., Pearson, D.M., Bäcklund, L.M., Ichimura, K., and Collins, V.P. (2008). Tandem duplication producing a novel oncogenic BRAF fusion gene defines the majority of pilocytic astrocytomas. Cancer Res. *68*, 8673–8677.

Jones, D.T.W., Witt, O., and Pfister, S.M. (2018b). BRAF V600E Status Alone Is Not Sufficient as a Prognostic Biomarker in Pediatric Low-Grade Glioma. J. Clin. Oncol. *36*, 96.

Jones, P.A., Ohtani, H., Chakravarthy, A., and De Carvalho, D.D. (2019c). Epigenetic therapy in immune-oncology. Nat. Rev. Cancer *19*, 151–161.

Kandoth, C., McLellan, M.D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J.F., Wyczalkowski, M.A., et al. (2013). Mutational landscape and significance across 12 major cancer types. Nature *502*, 333–339.

Kent, W.J. (2002). BLAT--the BLAST-like alignment tool. Genome Research *12*, 656–664.

Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat. Methods *12*, 357–360.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14,* R36–13.

Kim, W.Y., and Sharpless, N.E. (2006). The regulation of INK4/ARF in cancer and aging. Cell *127*, 265–275.

Ko, A., Han, S.Y., and Song, J. (2018). Regulatory Network of ARF in Cancer Development. Mol. Cells *41*, 381–389.

Koelsche, C., Hartmann, W., Schrimpf, D., Stichel, D., Jabar, S., Ranft, A., Reuss, D.E., Sahm, F., Jones, D.T.W., Bewerunge-Hudler, M., et al. (2018). Array-based DNA-methylation profiling in sarcomas with small blue round cell histology provides valuable diagnostic information. Mod. Pathol. *31*, 1246–1256.

Koelsche, C., Sahm, F., Wöhrer, A., Jeibmann, A., Schittenhelm, J., Kohlhof, P., Preusser, M., Romeike, B., Dohmen-Scheufler, H., Hartmann, C., et al. (2014). BRAF-mutated pleomorphic xanthoastrocytoma is associated with temporal location, reticulin fiber deposition and CD34 expression. Brain Pathol. *24*, 221–229.

Kogiso, M., Qi, L., Lindsay, H., Huang, Y., Zhao, X., Liu, Z., Braun, F.K., Du, Y., Zhang, H., Bae, G., et al. (2017). Xenotransplantation of pediatric low grade gliomas confirms the enrichment of BRAF V600E mutation and preservation of CDKN2A deletion in a novel orthotopic xenograft mouse model of progressive pleomorphic xanthoastrocytoma. Oncotarget *8*, 87455–87471.

Konduri, K., Gallant, J.-N., Chae, Y.K., Giles, F.J., Gitlitz, B.J., Gowen, K., Ichihara, E., Owonikoko, T.K., Peddareddigari, V., Ramalingam, S.S., et al. (2016). EGFR Fusions as Novel Therapeutic Targets in Lung Cancer. Cancer Discov *6*, 601–611.

Korshunov, A., Chavez, L., Sharma, T., Ryzhova, M., Schrimpf, D., Stichel, D., Capper, D., Sturm, D., Kool, M., Habel, A., et al. (2018). Epithelioid glioblastomas stratify into established diagnostic subsets upon integrated molecular analysis. Brain Pathol. *28*, 656–662.

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.

Krimpenfort, P., Quon, K.C., Mooi, W.J., Loonstra, A., and Berns, A. (2001). Loss of p16Ink4a confers susceptibility to metastatic melanoma in mice. Nature *413*, 83–86.

Krishnamurthy, J., Torrice, C., Ramsey, M.R., Kovalev, G.I., Al-Regaiey, K., Su, L., and Sharpless, N.E. (2004). Ink4a/Arf expression is a biomarker of aging. J. Clin. Invest. *114*, 1299–1307.

Kumar, A., Kankainen, M., Parsons, A., Kallioniemi, O., Mattila, P., and Heckman, C.A. (2017). The impact of RNA sequence library construction protocols on transcriptomic profiling of leukemia. BMC Genomics *18*, 629–13.

Lambert, S.R., Witt, H., Hovestadt, V., Zucknick, M., Kool, M., Pearson, D.M., Korshunov, A., Ryzhova, M., Ichimura, K., Jabado, N., et al. (2013). Differential expression and methylation of brain developmental genes define location-specific subsets of pilocytic astrocytoma. Acta Neuropathol. *126*, 291–301.

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

Larson, J.D., Kasper, L.H., Paugh, B.S., Jin, H., Wu, G., Kwon, C.-H., Fan, Y., Shaw, T.I., Silveira, A.B., Qu, C., et al. (2019). Histone H3.3 K27M Accelerates Spontaneous Brainstem Glioma and Drives Restricted Changes in Bivalent Gene Expression. Cancer Cell *35*, 140–155.e147.

Lassaletta, A., Scheinemann, K., Zelcer, S.M., Hukin, J., Wilson, B.A., Jabado, N., Carret, A.S., Lafay-Cousin, L., Larouche, V., Hawkins, C.E., et al. (2016). Phase II Weekly Vinblastine for Chemotherapy-Naïve Children With Progressive Low-Grade Glioma: A Canadian Pediatric Brain Tumor Consortium Study. J. Clin. Oncol. *34*, 3537–3543.

Lassaletta, A., Zapotocky, M., Mistry, M., Ramaswamy, V., Honnorat, M., Krishnatry, R., Guerreiro Stucklin, A., Zhukova, N., Arnoldo, A., Ryall, S., et al. (2017). Therapeutic and Prognostic Implications of BRAF V600E in Pediatric Low-Grade Gliomas. J. Clin. Oncol. *35*, 2934–2941.

Latysheva, N.S., and Babu, M.M. (2016). Discovering and understanding oncogenic gene fusions through data intensive computational approaches. Nucleic Acids Res. *44*, 4487–4503.

Liggett, W.H., and Sidransky, D. (1998). Role of the p16 tumor suppressor gene in cancer. J. Clin. Oncol. *16*, 1197–1206.

Lin, A., Rodriguez, F.J., Karajannis, M.A., Williams, S.C., Legault, G., Zagzag, D., Burger, P.C., Allen, J.C., Eberhart, C.G., and Bar, E.E. (2012). BRAF alterations in primary glial and glioneuronal neoplasms of the central nervous system with identification of 2 novel KIAA1549:BRAF fusion variants. J. Neuropathol. Exp. Neurol. *71*, 66–72.

Listernick, R., Charrow, J., and Gutmann, D.H. (1999). Intracranial gliomas in neurofibromatosis type 1. Am. J. Med. Genet. *89*, 38–44.

Liu, Q., Zhu, H., Liu, Y., Musetti, S., and Huang, L. (2018). BRAF peptide vaccine facilitates therapy of murine BRAF-mutant melanoma. Cancer Immunol. Immunother. *67*, 299–310.

Louis, D.N., Perry, A., Reifenberger, G., Deimling, von, 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.

Luan, D., Willard-Gallo, K., Garaud, S., Duvillier, H., Lodewyckx, J.-N., Solinas, C., De Silva, P., Gu-Trantien, C., Sibille, C., and Bron, D. (2016). Transcription Factors and Checkpoint Inhibitor Expression with Age: Markers of Immunosenescence? Blood *128*, 5983–5983.

Ma, X., Liu, Y., Liu, Y., Alexandrov, L.B., Edmonson, M.N., Gawad, C., Zhou, X., Li, Y., Rusch, M.C., Easton, J., et al. (2018). Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. Nature *555*, 371–376.

Mackay, A., Burford, A., Molinari, V., Jones, D.T.W., Izquierdo, E., Brouwer-Visser, J., Giangaspero, F., Haberler, C., Pietsch, T., Jacques, T.S., et al. (2018). Molecular, Pathological, Radiological, and Immune Profiling of Non-brainstem Pediatric High-Grade Glioma from the HERBY Phase II Randomized Trial. Cancer Cell *33*, 829–842.e5.

Maher, C.A., Kumar-Sinha, C., Cao, X., Kalyana-Sundaram, S., Han, B., Jing, X., Sam, L., Barrette, T., Palanisamy, N., and Chinnaiyan, A.M. (2009). Transcriptome sequencing to detect gene fusions in cancer. Nature *458*, 97–101.

Majzner, R.G., Heitzeneder, S., and Mackall, C.L. (2017). Harnessing the Immunotherapy Revolution for the Treatment of Childhood Cancers. Cancer Cell *31*, 476–485.

Maude, S.L., Laetsch, T.W., Buechner, J., Rives, S., Boyer, M., Bittencourt, H., Bader, P., Verneris, M.R., Stefanski, H.E., Myers, G.D., et al. (2018). Tisagenlecleucel in Children and Young Adults with B-Cell Lymphoblastic Leukemia. N Engl J Med *378*, 439–448.

Mellman, I., Coukos, G., and Dranoff, G. (2011). Cancer immunotherapy comes of age. Nature *480*, 480–489.

Mertens, F., Johansson, B., Fioretos, T., and Mitelman, F. (2015). The emerging complexity of gene fusions in cancer. Nat. Rev. Cancer *15*, 371–381.

Michaloglou, C., Vredeveld, L.C.W., Soengas, M.S., Denoyelle, C., Kuilman, T., van der Horst, C.M.A.M., Majoor, D.M., Shay, J.W., Mooi, W.J., and Peeper, D.S. (2005). BRAFE600-associated senescence-like cell cycle arrest of human naevi. Nature *436*, 720–724.

Mikheev, A.M., Mikheeva, S.A., Trister, A.D., Tokita, M.J., Emerson, S.N., Parada, C.A., Born, D.E., Carnemolla, B., Frankel, S., Kim, D.-H., et al. (2015). Periostin is a novel therapeutic target that predicts and regulates glioma malignancy. Neuro-Oncology *17*, 372–382.

Mistry, M., Zhukova, N., Merico, D., Rakopoulos, P., Krishnatry, R., Shago, M., Stavropoulos, J., Alon, N., Pole, J.D., Ray, P.N., et al. (2015). BRAF mutation and CDKN2A deletion define a clinically distinct subgroup of childhood secondary high-grade glioma. J. Clin. Oncol. *33*, 1015–1022.

Mondal, G., Lee, J.C., Ravindranathan, A., Villanueva-Meyer, J.E., Tran, Q.T., Allen, S.J., Barreto, J., Gupta, R., Doo, P., van Ziffle, J., et al. (2020). Pediatric bithalamic gliomas have a distinct epigenetic signature and frequent EGFR exon 20 insertions resulting in potential sensitivity to targeted kinase inhibition. Acta Neuropathol. *3*, 534–18.

Moran, S., Martínez-Cardús, A., Sayols, S., Musulén, E., Balañá, C., Estival-Gonzalez, A., Moutinho, C., Heyn, H., Diaz-Lagares, A., de Moura, M.C., et al. (2016). Epigenetic profiling to classify cancer of unknown primary: a multicentre, retrospective analysis. Lancet Oncol. *17*, 1386–1395.

Motzer, R.J., Tannir, N.M., McDermott, D.F., Arén Frontera, O., Melichar, B., Choueiri, T.K., Plimack, E.R., Barthélémy, P., Porta, C., George, S., et al. (2018). Nivolumab plus Ipilimumab versus Sunitinib in Advanced Renal-Cell Carcinoma. N Engl J Med *378*, 1277–1290.

Mount, C.W., Majzner, R.G., Sundaresh, S., Arnold, E.P., Kadapakkam, M., Haile, S., Labanieh, L., Hulleman, E., Woo, P.J., Rietberg, S.P., et al. (2018). Potent antitumor efficacy of anti-GD2 CAR T cells in H3-K27M+ diffuse midline gliomas. Nat. Med. *24*, 572–579.

Murat, A., Migliavacca, E., Hussain, S.F., Heimberger, A.B., Desbaillets, I., Hamou, M.-F., Rüegg, C., Stupp, R., Delorenzi, M., and Hegi, M.E. (2009). Modulation of angiogenic and inflammatory response in glioblastoma by hypoxia. PLoS ONE *4*, e5947.

Nakamura, T., Fukuoka, K., Nakano, Y., Yamasaki, K., Matsushita, Y., Yamashita, S., Ikeda, J., Udaka, N., Tanoshima, R., Shiba, N., et al. (2019). Genome-wide DNA methylation profiling shows molecular heterogeneity of anaplastic pleomorphic xanthoastrocytoma. Cancer Sci. *110*, 828–832.

Nicorici, D., Satalan, M., Edgren, H., Kangaspeska, S., Murumagi, A., Kallioniemi, O., Virtanen, S., and Kilkku, O. (2014). FusionCatcher - a tool for finding somatic fusion genes in paired-end RNA-sequencing data. bioRxiv 1–11.

Northcott, P.A., Buchhalter, I., Morrissy, A.S., Hovestadt, V., Weischenfeldt, J., Ehrenberger, T., Gröbner, S., Segura-Wang, M., Zichner, T., Rudneva, V.A., et al. (2017). The whole-genome landscape of medulloblastoma subtypes. Nature *547*, 311–317.

Oldrini, B., Curiel-García, Á., Marques, C., Matia, V., Uluçkan, Ö., Graña-Castro, O., Torres-Ruiz, R., Rodriguez-Perales, S., Huse, J.T., and Squatrito, M. (2018). Somatic genome editing with the RCAS-TVA-CRISPR-Cas9 system for precision tumor modeling. Nat Commun *9*, 1466–16.

Ostrom, Q.T., Cioffi, G., Gittleman, H., Patil, N., Waite, K., Kruchko, C., and Barnholtz-Sloan, J.S. (2019). CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2012-2016. Neuro-Oncology *21*, v1–v100.

Pages, M., Lacroix, L., Tauziede-Espariat, A., Castel, D., Daudigeos-Dubus, E., Ridola, V., Gilles, S., Fina, F., Andreiuolo, F., Polivka, M., et al. (2015). Papillary glioneuronal tumors: histological and molecular characteristics and diagnostic value of SLC44A1-PRKCA fusion. Acta Neuropathol Commun *3*, 85–10.

Pajtler, K.W., Wei, Y., Okonechnikov, K., Silva, P.B.G., Vouri, M., Zhang, L., Brabetz, S., Sieber, L., Gulley, M., Mauermann, M., et al. (2019). YAP1 subgroup supratentorial ependymoma requires TEAD and nuclear factor I-mediated transcriptional programmes for tumorigenesis. Nat Commun *10*, 3914–3916.

Pajtler, K.W., Witt, H., Sill, M., Jones, D.T.W., Hovestadt, V., Kratochwil, F., Wani, K., Tatevossian, R., Punchihewa, C., Johann, P., et al. (2015). Molecular Classification of Ependymal Tumors across All CNS Compartments, Histopathological Grades, and Age Groups. Cancer Cell *27*, 728–743.

Panagopoulos, I., Gorunova, L., Bjerkehagen, B., and Heim, S. (2014). The "grep" command but not FusionMap, FusionFinder or ChimeraScan captures the CIC-DUX4 fusion gene from whole transcriptome sequencing data on a small round cell tumor with t(4;19)(q35;q13). PLoS ONE *9*, e99439.

Pardoll, D.M. (2012). The blockade of immune checkpoints in cancer immunotherapy. Nat. Rev. Cancer *12*, 252–264.

Pehlivan, K.C., Malicki, D.M., Levy, M.L., and Crawford, J.R. (2020). TPM3-NTRK1 fusion in a pleomorphic xanthoastrocytoma presenting with haemorrhage in a child. BMJ Case Rep *13*, e234347.

Perez, J.G., Tran, N.L., Rosenblum, M.G., Schneider, C.S., Connolly, N.P., Kim, A.J., Woodworth, G.F., and Winkles, J.A. (2016). The TWEAK receptor Fn14 is a potential cell surface portal for targeted delivery of glioblastoma therapeutics. Oncogene *35*, 2145–2155.

Perkins, S.M., Mitra, N., Fei, W., and Shinohara, E.T. (2012). Patterns of care and outcomes of patients with pleomorphic xanthoastrocytoma: a SEER analysis. J. Neurooncol. *110*, 99–104.

Pfister, S., Janzarik, W.G., Remke, M., Ernst, A., Werft, W., Becker, N., Toedt, G., Wittmann, A., Kratz, C., Olbrich, H., et al. (2008). BRAF gene duplication constitutes a mechanism of MAPK pathway activation in low-grade astrocytomas. J. Clin. Invest. *118*, 1739–1749.

Phillips, J.J., Gong, H., Chen, K., Joseph, N.M., van Ziffle, J., Bastian, B.C., Grenert, J.P., Kline, C.N., Mueller, S., Banerjee, A., et al. (2019). The genetic landscape of anaplastic pleomorphic xanthoastrocytoma. Brain Pathol. *29*, 85–96.

Pickles, J.C., Fairchild, A.R., Stone, T.J., Brownlee, L., Merve, A., Yasin, S.A., Avery, A., Ahmed, S.W., Ogunbiyi, O., Gonzalez Zapata, J., et al. (2020). DNA methylation-based profiling for paediatric CNS tumour diagnosis and treatment: a population-based study. Lancet Child Adolesc Health *4*, 121–130.

Pidsley, R., Zotenko, E., Peters, T.J., Lawrence, M.G., Risbridger, G.P., Molloy, P., Van Djik, S., Muhlhausler, B., Stirzaker, C., and Clark, S.J. (2016). Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. Genome Biol. *17*, 208–217.

Plant, A.S., and Hwang, E.I. (2018). Immunotherapy and the Immune Infiltrate in Pediatric Brain Tumors: An Illustration and Review of the Unique Challenges Facing Immunotherapy for Pediatric Oncology. Int J Immunol Immunother *5*.

Plant, A.S., Koyama, S., Sinai, C., Solomon, I.H., Griffin, G.K., Ligon, K.L., Bandopadhayay, P., Betensky, R., Emerson, R., Dranoff, G., et al. (2018). Immunophenotyping of pediatric brain tumors: correlating immune infiltrate with histology, mutational load, and survival and assessing clonal T cell response. J. Neurooncol. *137*, 269–278.

Platten, M., Schilling, D., Bunse, L., Wick, A., Bunse, T., Riehl, D., Karapanagiotou-Schenkel, I., Harting, I., Sahm, F., Schmitt, A., et al. (2018). A mutation-specific peptide vaccine targeting IDH1R132H in patients with newly diagnosed malignant astrocytomas: A first-in-man multicenter phase I clinical trial of the German Neurooncology Working Group (NOA-16). Jco *36*, 2001–2001.

Raabe, E.H., Lim, K.S., Kim, J.M., Meeker, A., Mao, X.-G., Nikkhah, G., Maciaczyk, J., Kahlert, U., Jain, D., Bar, E., et al. (2011). BRAF activation induces transformation and then senescence in human neural stem cells: a pilocytic astrocytoma model. Clin. Cancer Res. *17*, 3590–3599.

Ramkissoon, L.A., Horowitz, P.M., Craig, J.M., Ramkissoon, S.H., Rich, B.E., Schumacher, S.E., McKenna, A., Lawrence, M.S., Berghold, G., Brastianos, P.K., et al. (2013). 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. *110*, 8188–8193.

Rasmussen, S.A., and Friedman, J.M. (2000). NF1 gene and neurofibromatosis 1. Am. J. Epidemiol. *151*, 33–40.

Rees, J., Watt, H., Jäger, H.R., Benton, C., Tozer, D., Tofts, P., and Waldman, A. (2009). Volumes and growth rates of untreated adult low-grade gliomas indicate risk of early malignant transformation. Eur J Radiol *72*, 54–64.

Reinhardt, A., Stichel, D., Schrimpf, D., Sahm, F., Korshunov, A., Reuss, D.E., Koelsche, C., Huang, K., Wefers, A.K., Hovestadt, V., et al. (2018). Anaplastic astrocytoma with piloid features, a novel molecular class of IDH wildtype glioma with recurrent MAPK pathway, CDKN2A/B and ATRX alterations. Acta Neuropathol. *136*, 273–291.

Reisinger, E., Genthner, L., Kerssemakers, J., Kensche, P., Borufka, S., Jugold, A., Kling, A., Prinz, M., Scholz, I., Zipprich, G., et al. (2017). OTP: An automatized system for managing and processing NGS data. J. Biotechnol. *261*, 53–62.

Reitman, Z.J., Paolella, B.R., Bergthold, G., Pelton, K., Becker, S., Jones, R., Sinai, C.E., Malkin, H., Huang, Y., Grimmet, L., et al. (2019). Mitogenic and progenitor gene programmes in single pilocytic astrocytoma cells. Nat Commun *10*, 3731–17.

Renner, M., Wolf, T., Meyer, H., Hartmann, W., Penzel, R., Ulrich, A., Lehner, B., Hovestadt, V., Czwan, E., Egerer, G., et al. (2013). Integrative DNA methylation and gene expression analysis in high-grade soft tissue sarcomas. Genome Biol. *14*, r137–26.

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*, 258–5.

Robinson, J.P., VanBrocklin, M.W., Guilbeault, A.R., Signorelli, D.L., Brandner, S., and Holmen, S.L. (2010). Activated BRAF induces gliomas in mice when combined with Ink4a/Arf loss or Akt activation. Oncogene *29*, 335–344.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24–26.

Rodriguez, F.J., Schniederjan, M.J., Nicolaides, T., Tihan, T., Burger, P.C., and Perry, A. (2015). High rate of concurrent BRAF-KIAA1549 gene fusion and 1p deletion in disseminated oligodendroglioma-like leptomeningeal neoplasms (DOLN). Acta Neuropathol. *129*, 609–610.

Rokita, J.L., Rathi, K.S., Cardenas, M.F., Upton, K.A., Jayaseelan, J., Cross, K.L., Pfeil, J., Egolf, L.E., Way, G.P., Farrel, A., et al. (2019). Genomic Profiling of Childhood Tumor Patient-Derived Xenograft Models to Enable Rational Clinical Trial Design. Cell Rep *29*, 1675–1689.e1679.

Roth, R.B., Hevezi, P., Lee, J., Willhite, D., Lechner, S.M., Foster, A.C., and Zlotnik, A. (2006). Gene expression analyses reveal molecular relationships among 20 regions of the human CNS. Neurogenetics *7*, 67–80.

Ruas, M., and Peters, G. (1998). The p16INK4a/CDKN2A tumor suppressor and its relatives. Biochim. Biophys. Acta *1378*, F115–F177.

Ryall, S., Tabori, U., and Hawkins, C. (2020a). Pediatric low-grade glioma in the era of molecular diagnostics. Acta Neuropathol Commun *8*, 30–22.

Ryall, S., Zapotocky, M., Fukuoka, K., Nobre, L., Guerreiro Stucklin, A., Bennett, J., Siddaway, R., Li, C., Pajovic, S., Arnoldo, A., et al. (2020b). Integrated Molecular and Clinical Analysis of 1,000 Pediatric Low-Grade Gliomas. Cancer Cell *37*, 569–583.e5.

Sadelain, M., Brentjens, R., and Rivière, I. (2013). The basic principles of chimeric antigen receptor design. Cancer Discov *3*, 388–398.

Samstein, R.M., Lee, C.-H., Shoushtari, A.N., Hellmann, M.D., Shen, R., Janjigian, Y.Y., Barron, D.A., Zehir, A., Jordan, E.J., Omuro, A., et al. (2019). Tumor mutational load predicts survival after immunotherapy across multiple cancer types. Nat. Genet. *51*, 202–206.

Schaefer-Klein, J., Givol, I., Barsov, E.V., Whitcomb, J.M., VanBrocklin, M., Foster, D.N., Federspiel, M.J., and Hughes, S.H. (1998). The EV-O-derived cell line DF-1 supports the efficient replication of avian leukosis-sarcoma viruses and vectors. Virology *248*, 305–311.

Schindler, G., Capper, D., Meyer, J., Janzarik, W., Omran, H., Herold-Mende, C., Schmieder, K., Wesseling, P., Mawrin, C., Hasselblatt, M., et al. (2011). Analysis of BRAF V600E mutation in 1,320 nervous system tumors reveals high mutation frequencies in pleomorphic xanthoastrocytoma, ganglioglioma and extra-cerebellar pilocytic astrocytoma. Acta Neuropathol. *121*, 397–405.

Schumacher, T., Bunse, L., Pusch, S., Sahm, F., Wiestler, B., Quandt, J., Menn, O., Osswald, M., Oezen, I., Ott, M., et al. (2014). A vaccine targeting mutant IDH1 induces antitumour immunity. Nature *512*, 324–327.

Selt, F., Hohloch, J., Hielscher, T., Sahm, F., Capper, D., Korshunov, A., Usta, D., Brabetz, S., Ridinger, J., Ecker, J., et al. (2017). Establishment and application of a novel patient-derived KIAA1549:BRAF-driven pediatric pilocytic astrocytoma model for preclinical drug testing. Oncotarget *8*, 11460–11479.

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

Serrano, M., Lin, A.W., McCurrach, M.E., Beach, D., and Lowe, S.W. (1997). Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. Cell *88*, 593–602.

Sexton-Oates, A., Dodgshun, A., Hovestadt, V., Jones, D.T.W., Ashley, D.M., Sullivan, M., MacGregor, D., and Saffery, R. (2018). Methylation profiling of paediatric pilocytic astrocytoma reveals variants specifically associated with tumour location and predictive of recurrence. Mol Oncol *12*, 1219–1232.

Sharpless, N.E., Bardeesy, N., Lee, K.H., Carrasco, D., Castrillon, D.H., Aguirre, A.J., Wu, E.A., Horner, J.W., and DePinho, R.A. (2001). Loss of p16Ink4a with retention of p19Arf predisposes mice to tumorigenesis. Nature *413*, 86–91.

She, D., Liu, J., Zeng, Z., Xing, Z., and Cao, D. (2018). Diagnostic accuracy of diffusion weighted imaging for differentiation of supratentorial pilocytic astrocytoma and pleomorphic xanthoastrocytoma. Neuroradiology *60*, 725–733.

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.

Siegel, R.L., Miller, K.D., and Jemal, A. (2020). Cancer statistics, 2020. CA Cancer J Clin *70*, 7–30.

Sievers, P., Stichel, D., Schrimpf, D., Sahm, F., Koelsche, C., Reuss, D.E., Wefers, A.K., Reinhardt, A., Huang, K., Ebrahimi, A., et al. (2018). FGFR1:TACC1 fusion is a frequent event in molecularly defined extraventricular neurocytoma. Acta Neuropathol. *136*, 293–302.

Sievert, A.J., Jackson, E.M., Gai, X., Hakonarson, H., Judkins, A.R., Resnick, A.C., Sutton, L.N., Storm, P.B., Shaikh, T.H., and Biegel, J.A. (2009). Duplication of 7q34 in pediatric low-grade astrocytomas detected by high-density single-nucleotide polymorphism-based genotype arrays results in a novel BRAF fusion gene. Brain Pathol. *19*, 449–458.

Simmons, G.W., Pong, W.W., Emnett, R.J., White, C.R., Gianino, S.M., Rodriguez, F.J., and Gutmann, D.H. (2011). Neurofibromatosis-1 heterozygosity increases microglia in a spatially and temporally restricted pattern relevant to mouse optic glioma formation and growth. J. Neuropathol. Exp. Neurol. *70*, 51–62.

Simon, A.K., Hollander, G.A., and McMichael, A. (2015). Evolution of the immune system in humans from infancy to old age. Proc. Biol. Sci. *282*, 20143085.

Solomon, D.A., Korshunov, A., Sill, M., Jones, D.T.W., Kool, M., Pfister, S.M., Fan, X., Bannykh, S., Hu, J., Danielpour, M., et al. (2018). Myxoid glioneuronal tumor of the septum pellucidum and lateral ventricle is defined by a recurrent PDGFRA p.K385 mutation and DNT-like methylation profile. Acta Neuropathol. *136*, 339–343.

Somasundaram, R., Swoboda, R., Caputo, L., Otvos, L., Weber, B., Volpe, P., van Belle, P., Hotz, S., Elder, D.E., Marincola, F.M., et al. (2006). Human leukocyte antigen-A2-restricted CTL responses to mutated BRAF peptides in melanoma patients. Cancer Res. *66*, 3287–3293.

Sondka, Z., Bamford, S., Cole, C.G., Ward, S.A., Dunham, I., and Forbes, S.A. (2018). The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. Nat. Rev. Cancer *18*, 696–705.

Stark, R., Grzelak, M., and Hadfield, J. (2019). RNA sequencing: the teenage years. Nat. Rev. Genet. *20*, 631–656.

Stewart, E., Federico, S.M., Chen, X., Shelat, A.A., Bradley, C., Gordon, B., Karlstrom, A., Twarog, N.R., Clay, M.R., Bahrami, A., et al. (2017). Orthotopic patient-derived xenografts of paediatric solid tumours. Nature *549*, 96–100.

Stichel, D., Schrimpf, D., Casalini, B., Meyer, J., Wefers, A.K., Sievers, P., Korshunov, A., Koelsche, C., Reuss, D.E., Reinhardt, A., et al. (2019). Routine RNA sequencing of formalin-fixed paraffin-embedded specimens in neuropathology diagnostics identifies diagnostically and therapeutically relevant gene fusions. Acta Neuropathol. *138*, 827–835.

Stransky, N., Cerami, E., Schalm, S., Kim, J.L., and Lengauer, C. (2014). The landscape of kinase fusions in cancer. Nat Commun *5*, 4846–10.

Sturm, D., Orr, B.A., Toprak, U.H., Hovestadt, V., Jones, D.T.W., 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., Pfister, S.M., and Jones, D.T.W. (2017). Pediatric Gliomas: Current Concepts on Diagnosis, Biology, and Clinical Management. J. Clin. Oncol. *35*, 2370–2377.

Sturm, D., Witt, H., Hovestadt, V., Khuong-Quang, D.-A., Jones, D.T.W., Konermann, C., Pfaff, E., Tönjes, 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.

Subbiah, V., Westin, S.N., Wang, K., Araujo, D., Wang, W.-L., Miller, V.A., Ross, J.S., Stephens, P.J., Palmer, G.A., and Ali, S.M. (2014). Targeted therapy by combined inhibition of the RAF and mTOR kinases in malignant spindle cell neoplasm harboring the KIAA1549-BRAF fusion protein. J Hematol Oncol *7*, 8–7.

Tanaka, S., Nakada, M., Nobusawa, S., Suzuki, S.O., Sabit, H., Miyashita, K., and Hayashi, Y. (2014). Epithelioid glioblastoma arising from pleomorphic xanthoastrocytoma with the BRAF V600E mutation. Brain Tumor Pathol *31*, 172–176.

Tao, R., Murad, N., Xu, Z., Zhang, P., Okonechnikov, K., Kool, M., Rivero-Hinojosa, S., Lazarski, C., Zheng, P., Liu, Y., et al. (2019). MYC Drives Group 3 Medulloblastoma through Transformation of Sox2+ Astrocyte Progenitor Cells. Cancer Res. *79*, 1967–1980.

Tatevossian, R.G., Tang, B., Dalton, J., Forshew, T., Lawson, A.R., Ma, J., Neale, G., Shurtleff, S.A., Bailey, S., Gajjar, A., et al. (2010). MYB upregulation and genetic aberrations in a subset of pediatric low-grade gliomas. Acta Neuropathol. *120*, 731–743.

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

Tomić, T.T., Olausson, J., Wilzén, A., Sabel, M., Truvé, K., Sjögren, H., Dósa, S., Tisell, M., Lannering, B., Enlund, F., et al. (2017). A new GTF2I-BRAF fusion mediating MAPK pathway activation in pilocytic astrocytoma. PLoS ONE *12*, e0175638.

Trotman, L.C., Niki, M., Dotan, Z.A., Koutcher, J.A., Di Cristofano, A., Xiao, A., Khoo, A.S., Roy-Burman, P., Greenberg, N.M., Van Dyke, T., et al. (2003). Pten dose dictates cancer progression in the prostate. PLoS Biol. *1*, E59.

Tu, A., Robison, A., Melamed, E., Buchanan, I., Hariri, O., Babu, H., Szymanski, L., and Krieger, M. (2018). Proliferative Index in Pediatric Pilocytic Astrocytoma by Region of Origin and Prediction of Clinical Behavior. Pediatr Neurosurg *53*, 395–400.

Usta, D., Sigaud, R., Buhl, J.L., Selt, F., Marquardt, V., Pauck, D., Jansen, J., Pusch, S., Ecker, J., Hielscher, T., et al. (2020). A cell-based MAPK reporter assay reveals synergistic MAPK pathway activity suppression by MAPK inhibitor combination in BRAF-driven pediatric low-grade glioma cells. Mol. Cancer Ther.

van den Bent, M.J. (2010). Interobserver variation of the histopathological diagnosis in clinical trials on glioma: a clinician's perspective. Acta Neuropathol. *120*, 297–304.

Wagle, M.-C., Kirouac, D., Klijn, C., Liu, B., Mahajan, S., Junttila, M., Moffat, J., Merchant, M., Huw, L., Wongchenko, M., et al. (2018). A transcriptional MAPK Pathway Activity Score (MPAS) is a clinically relevant biomarker in multiple cancer types. NPJ Precis Oncol *2*, 7–12.

Waligora, M., Bala, M.M., Koperny, M., Wasylewski, M.T., Strzebonska, K., Jaeschke, R.R., Wozniak, A., Piasecki, J., Sliwka, A., Mitus, J.W., et al. (2018). Risk and surrogate benefit for pediatric Phase I trials in oncology: A systematic review with meta-analysis. PLoS Med. *15*, e1002505.

Wan, P.T.C., Garnett, M.J., Roe, S.M., Lee, S., Niculescu-Duvaz, D., Good, V.M., Jones, C.M., Marshall, C.J., Springer, C.J., Barford, D., et al. (2004). Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. Cell *116*, 855–867.

Ward, Z.J., Yeh, J.M., Bhakta, N., Frazier, A.L., and Atun, R. (2019). Estimating the total incidence of global childhood cancer: a simulation-based analysis. Lancet Oncol. *20*, 483–493.

Weber, R.G., Hoischen, A., Ehrler, M., Zipper, P., Kaulich, K., Blaschke, B., Becker, A.J., Weber-Mangal, S., Jauch, A., Radlwimmer, B., et al. (2007). Frequent loss of chromosome 9, homozygous CDKN2A/p14(ARF)/CDKN2B deletion and low TSC1 mRNA expression in pleomorphic xanthoastrocytomas. Oncogene *26*, 1088–1097.

Wefers, A.K., Stichel, D., Schrimpf, D., Coras, R., Pages, M., Tauziede-Espariat, A., Varlet, P., Schwarz, D., Söylemezoglu, F., Pohl, U., et al. (2020). 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. *139*, 193–209.

Werder, von, A., Seidler, B., Schmid, R.M., Schneider, G., and Saur, D. (2012). Production of avian retroviruses and tissue-specific somatic retroviral gene transfer in vivo using the RCAS/TVA system. Nat Protoc *7*, 1167–1183.

Wisoff, J.H., Sanford, R.A., Heier, L.A., Sposto, R., Burger, P.C., Yates, A.J., Holmes, E.J., and Kun, L.E. (2011). Primary neurosurgery for pediatric low-grade gliomas: a prospective multi-institutional study from the Children's Oncology Group. Neurosurgery *68*, 1548–54–discussion1554–5.

Wolchok, J.D., Chiarion-Sileni, V., Gonzalez, R., Rutkowski, P., Grob, J.-J., Cowey, C.L., Lao, C.D., Wagstaff, J., Schadendorf, D., Ferrucci, P.F., et al. (2017). Overall Survival with Combined Nivolumab and Ipilimumab in Advanced Melanoma. N Engl J Med *377*, 1345–1356.

Worst, B.C., van Tilburg, C.M., Balasubramanian, G.P., Fiesel, P., Witt, R., Freitag, A., Boudalil, M., Previti, C., Wolf, S., Schmidt, S., et al. (2016). Next-generation personalised medicine for high-risk paediatric cancer patients - The INFORM pilot study. Eur. J. Cancer *65*, 91–101.

Wu, S.P., Cooper, B.T., Bu, F., Bowman, C.J., Killian, J.K., Serrano, J., Wang, S., Jackson, T.M., Gorovets, D., Shukla, N., et al. (2017). DNA Methylation-Based Classifier for Accurate Molecular Diagnosis of Bone Sarcomas. JCO Precis Oncol *2017*, 1–11.

Yang, I., Han, S.J., Sughrue, M.E., Tihan, T., and Parsa, A.T. (2011). Immune cell infiltrate differences in pilocytic astrocytoma and glioblastoma: evidence of distinct immunological microenvironments that reflect tumor biology. J. Neurosurg. *115*, 505–511.

Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-Garcia, W., Treviño, V., Shen, H., Laird, P.W., Levine, D.A., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun *4*, 2612–11.

Young, J.A., Bates, P., and Varmus, H.E. (1993). Isolation of a chicken gene that confers susceptibility to infection by subgroup A avian leukosis and sarcoma viruses. J. Virol. *67*, 1811–1816.

Zhang, J., Wu, G., Miller, C.P., Tatevossian, R.G., Dalton, J.D., Tang, B., Orisme, W., Punchihewa, C., Parker, M., Qaddoumi, I., et al. (2013). Whole-genome sequencing identifies genetic alterations in pediatric low-grade gliomas. Nat. Genet. *45*, 602–612.

Zhao, S., Zhang, Y., Gamini, R., Zhang, B., and Schack, von, D. (2018). Evaluation of two main RNA-seq approaches for gene quantification in clinical RNA sequencing: polyA+ selection versus rRNA depletion. Sci Rep *8*, 4781–12.

Zhou, W., Ke, S.Q., Huang, Z., Flavahan, W., Fang, X., Paul, J., Wu, L., Sloan, A.E., McLendon, R.E., Li, X., et al. (2015). Periostin secreted by glioblastoma stem cells recruits M2 tumour-associated macrophages and promotes malignant growth. Nat. Cell Biol. *17*, 170–182.

Zindy, F., Quelle, D.E., Roussel, M.F., and Sherr, C.J. (1997). Expression of the p16INK4a tumor suppressor versus other INK4 family members during mouse development and aging. Oncogene *15*, 203–211.