Proteogenomics Analysis of Low- and High-Grade Gliomas

Inaugural dissertation

for the attainment of the title of doctor
in the Faculty of Mathematics and Natural Sciences
at the Heinrich Heine University Düsseldorf

presented by

Daniel Joseph Picard

from Sudbury, Canada

Düsseldorf, 2024

Clinic for Pediatric Oncology, Hematology and Clinical Immunology,
University Hospital Düsseldorf

Published by permission of the
Faculty of Mathematics and Natural Sciences at
Heinrich Heine University Düsseldorf

Supervisor: Prof. Dr. Marc Remke

Co-supervisor: Prof. Dr. Sebastian Wesselborg

Date of the oral examination: July 18, 2025

Hereby, I declare on oath that I composed this dissertation independently by myself. I used only the references and resources indicated in this thesis. With the exception of such quotations, the work presented in this thesis is my own. I have accredited all the sources of help. This PhD thesis was never submitted or presented in a similar form to any other institution or examination board. I have not undertaken a doctoral examination without success so far.

Homburg (Saar), 2026

_____

Daniel Joseph Picard

Table of Contents

Acknowledgments

First, I would like to thank Professor Marc Remke for bringing me to Düsseldorf and being my supervisor. I came for a year and stayed because of the great experience you provided.

I would also like to thank Professor Wesselborg for being my second supervisor and helping me in these unusual circumstances.

I would like to thank Nan Qin for being a great colleague and a wonderful friend. You supported me when I needed help both professionally and personally, and I will not forget it.

To the AG Remke, I hope I was able to teach you as much as you taught me. Specifically, I would like to thank Ulvi for being the best student ever. I hope all of you are successful in your respective careers.

Next, I would like to thank the KMT lab. You were always a source for scientific help as well as friendship. I made some amazing friends for which I will always be grateful. Daniel Scholtyssik, Daniel Hein, Sanil, Franzi, Kai, Ciro, Heinz, Melf, Julian, Katie, Pawel and the others I have not mentioned: thank you for all the good times.

To Professor Arndt Borkhardt and Ute Fischer, you let us into your department and your support over the years has been invaluable. From the bottom of my heart, I thank you.

To Professor Guido Reifenberger, although you were Marc's mentor, it seems at times that you were mine as well. I have been very happy for all of the collaborations over the years and I hope that we will still get to work together in the future.

To my dear family: being away from home has been difficult, especially with the birth of the twins, which I missed. However, I always knew you were there supporting me from a distance.

Finally, I would like to thank Stewart, for coming on this wonderful adventure to Germany, for supporting me during the rough patches and for being my husband.

Abbreviations

AKT	AKT serine/threonine kinase
ATP	Adenosine triphosphate
BRAF	B-Raf proto-oncogene, serine/threonine kinase
CLCN6	Chloride voltage-gated channel 6
CNS	Central nervous system
CPTAC	Clinical Proteomic Tumor Analysis Consortium
EGFR	Epidermal growth factor receptor
EMILIN2	Elastin microfibril interfacer 2
ERK	Extracellular signal-regulated kinase
FAM131B	Family with sequence similarity 131 member B
FGFR1	Fibroblast growth factor receptor 1
FGFR2	Fibroblast growth factor receptor 2
GB	Glioblastoma
GPR17	G protein-coupled receptor 17
GSEA	Gene Set Enrichment Analysis
HRAS	HRAS proto-oncogene, GTPase
IDH	Isocitrate dehydrogenase
IPA	Ingenuity Pathway Analysis
IRAG1/MRVI1	Inositol 1,4,5-triphosphate receptor associated 1
JAK1	Janus kinase 1
KIAA1549	KIAA1549
KRAS	KRAS proto-oncogene, GTPase
lncRNA	Long non-coding RNA
MAPK	Mitogen activated protein kinase
MB	Medulloblastoma
MEK	MAPK/ERK kinase
MGMT	O6-methylguanine-DNA methyltransferase
mTOR	Mechanistic target of rapamycin kinase

NF1	Neurofibromin 1
NRAS	NRAS proto-oncogene, GTPase
NTRK2	Neurotrophic receptor tyrosine kinase 2
PA	Pilocytic astrocytoma
PI3K	Phosphatidylinositol 3-kinase
PKG1	Phosphoinositide 3-kinase
PTEN	Phosphatase and tensin homolog
RAF1	Raf-1 proto-oncogene, serine/threonine kinase
RAS	Rat sarcoma virus
RCAS/t-va	Replication-competent avian retrovirus/tumor virus A
SNF	Similarity Network Fusion
SRC	SRC proto-oncogene, non-receptor tyrosine kinase
SYK	Spleen associated tyrosine kinase
TACC1	Transforming acidic coiled-coil containing protein 1
TERT	Telomerase reverse transcriptase
TGM2	Transglutaminase 2
TLR3	Toll like receptor 3
TP53	Tumor protein p53
TRPM8	Transient receptor potential cation channel subfamily M member 8
t-SNE	t-Distributed Stochastic Neighbor Embedding
TUBB	Beta-tubulin
WHO	World Health Organization

Summary

Glial cells provide protection and support of neurons in the central nervous system and tumors arising from these cells or their progenitors are termed gliomas. Gliomas are classified and graded based on histopathological features, molecular characteristics, and clinical behavior according the World Health Organization classification. The glioma types selected for this thesis were pilocytic astrocytoma (PA), the most common brain tumor in children, and glioblastoma, isocitrate dehydrogenase (IDH)-wildtype, the most common malignant brain tumor in adults. The aim of this thesis was to analyze glioma data using novel molecular techniques and bioinformatic methods with a focus on integrative proteogenomic approaches. For this purpose, proteogenomics was applied to integrate RNA sequencing and mass spectrometry-based proteomic profiling data from a cohort of 62 primary PA samples. In addition, a cohort of 70 IDH-wildtype glioblastomas was analyzed focusing on the non-coding genes available from a microarray-based gene expression dataset combined with a subsequent proteogenomic study. Similarity network fusion analysis revealed two PA subgroups that were validated in three non-overlapping cohorts. Interestingly, Group 1 patients were significantly younger than Group 2 patients and displayed a worse progression-free survival. Additional pathway analysis of these two subgroups revealed that Group 1 was enriched for immune response pathways, such as interferon signaling, while Group 2 showed enrichment for action potential and neurotransmitter signaling pathways. When analyzing the glioblastoma gene expression data, *HOTAIRM1* was identified as a candidate long non-coding RNA whose up-regulation was significantly associated with shorter survival of glioblastoma patients. Integrated proteogenomic analyses of glioblastoma cell lines with stable or transient knockdown of *HOTAIRM1* revealed impaired mitochondrial function, and determination of reactive oxygen species (ROS) levels confirmed increased ROS levels upon *HOTAIRM1* knock-down. Finally, *HOTAIRM1* was determined as part of the *HOTAIRM1*/*hsa-miR-17-5p*/*TGM2* axis which was linked to shorter patient survival. In summary, the research work summarized in this thesis conclusively demonstrates that analyzing high-throughput data using integrative approaches focusing on the proteogenome or the lncRNAome provide important novel biological insights into the pathogenesis of the most common low-grade and high-grade glioma types.

Zusammenfassung

Gliale Zellen schützen und unterstützen physiologischer Weise die Nervenzellen im zentralen Nervensystem (ZNS). ZNS-Tumoren, die aus glialen Zellen oder ihren Vorläuferzellen entstehen, werden Gliome genannt. Gliome werden basierend auf ihren histopathologischen Merkmalen, molekularen Biomarkern und ihrem klinischem Verhalten gemäß der Klassifikation der Weltgesundheitsorganisation für ZNS-Tumoren klassifiziert und gradiert. Die in dieser Arbeit untersuchten Gliomtypen waren pilozytische Astrozytome (PA), die häufigsten Hirntumoren im Kindesalter, und Glioblastome, Isocitratdehydrogenase (IDH)-Wildtyp, die häufigsten und zugleich bösartigsten Hirntumoren bei Erwachsenen. Das Ziel dieser Arbeit war es, molekulare Hochdurchsatzdaten mithilfe neuartiger bioinformatischer Methoden und einem Schwerpunkt auf integrativen proteogenomischen Ansätzen zu analysieren, um daraus neue Erkenntnisse zur Pathogenese dieser häufigen Gliome abzuleiten. Zu diesem Zweck wurden proteogenomische Verfahren angewendet, um RNA-Sequenzierungsdaten und Massenspektrometrie-basierte proteomische Profilierungsdaten einer Kohorte von 62 primären PA-Proben zu integrieren. Darüber hinaus wurde eine Kohorte von 70 IDH-Wildtyp-Glioblastomen analysiert, wobei der Schwerpunkt auf Expresssionsdaten zu nicht-kodierenden RNAs lag, die aus einem Microarray-basierten Datensatz stammten und in Kombination mit Proteomdaten in einer proteogenomischen Analyse untersucht wurden. Die integrative bioinformatische Analyse mittels *Similarity Network Fusion* ergab zwei PA-Untergruppen, die in drei nicht überlappenden Patientenkohorten validiert wurden. Interessanterweise waren die Patienten der Gruppe 1 signifikant jünger als die Patienten der Gruppe 2 und zeigten ein kürzeres progressionsfreies Überleben. Eine Signalweganalyse dieser beiden PA-Untergruppen ergab, dass Gruppe 1 eine Anreicherung von Signaturen der Immunantwort wie dem Interferon-Signalweg aufwies, während Gruppe 2 eine Anreicherung von Aktionspotential- und Neurotransmitter-Signalwegen zeigte. Bei der Analyse der Glioblastom-Genexpressionsdaten wurde *HOTAIRM1* als Kandidat für eine lange nicht-kodierende RNA identifiziert, deren Hochregulierung signifikant mit einem kürzeren Überleben von Glioblastompatienten assoziiert war. Proteogenomische Analysen von Glioblastomzelllinien mit stabilem oder transientem Knock-down von *HOTAIRM1* zeigten eine beeinträchtigte Mitochondrienfunktion in den Knoch-down-Zellen und die

Bestimmung der Spiegel reaktiver Sauerstoffspezies (ROS) bestätigte erhöhte ROS-Spiegel nach dem Knock-down von *HOTAIRM1*. Letztendlich konnte *HOTAIRM1* als Teil der *HOTAIRM1/hsa-miR-17-5p/TGM2*-Achse identifiziert werden, deren Aktivierung sich als prognostisch ungünstiger Faktor in Glioblastompatienten herausstellte. Zusammenfassend zeigen die in dieser Dissertation erbrachten Forschungsarbeiten, dass die Analyse von Hochdurchsatzdaten mithilfe integrativer Ansätze, die sich auf das Proteogenom oder das lncRNAom konzentrieren, wichtige neue biologische Erkenntnisse über die Pathogenese der häufigsten niedrig- und hochgradigen Gliomtypen liefern konnte.

1. Introduction

1.1. Introduction to gliomas

Glial cells provide protection and support for neurons in the central nervous system (CNS). A diverse group of tumors originating from glial cells are called gliomas, and account for approximately 80% of malignant brain tumors in adults (1). Behavior, aggressiveness and prognosis vary widely in these tumors and make their classification and management complex and challenging.

The World Health Organization (WHO) classification of CNS tumors categorizes gliomas according to histopathological features, molecular characteristics, and clinical behavior. In the most recent WHO classification, there is an emphasis on the addition of molecular markers to supplement histology of these tumors for a more accurate diagnosis (2). Although surgical resection is the primary curative option, higher grade tumors tend to be infiltrative, making surgical resection difficult and leading to high rates of recurrences (3).

Although gliomas are a heterogeneous group of tumors, encompassing diffuse astrocytic and oligodendroglial tumors, other astrocytic tumors, ependymal tumors, and other gliomas (2), this thesis will focus on the extremes of the glioma spectrum, specifically pilocytic astrocytoma (PA) and glioblastoma (GB) isocitrate dehydrogenase (IDH)-wildtype. These gliomas were selected based on the fact that PA is the most common brain tumor in pediatrics, whereas GB IDH-wildtype is the most common malignant brain tumor in adults (4, 5).

1.2. Pilocytic astrocytoma

1.2.1 Incidence

Pilocytic astrocytoma, previously known as cystic cerebellar astrocytoma or juvenile pilocytic astrocytoma (6), is a CNS WHO grade 1, astrocytic tumor (4, 7). This tumor entity constitutes the most common solid brain tumor in children accounting for 18.7% of all tumors in this age group (4). The incidence rate of PAs in the United States of America is 0.95 per 100,000 for pediatric patients. The incidence rate peaks between ages of one to four years of life and decreases with advancing age where it decreases to only 0.08 per 100,000 (4, 5, 7).

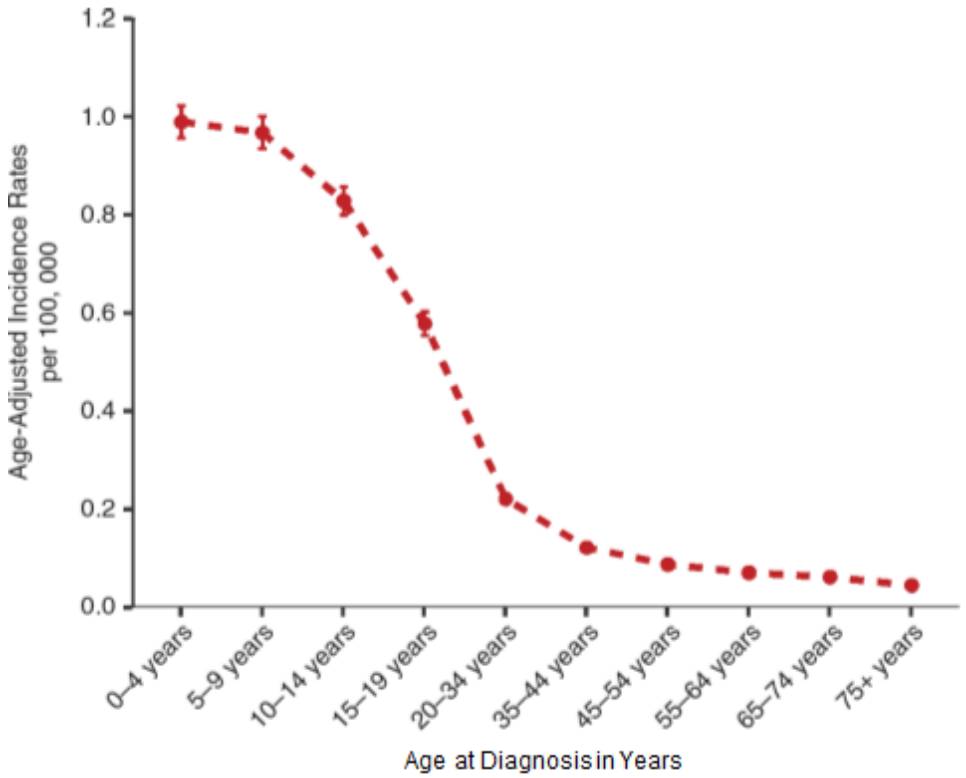


Figure 1.2.1: Incidence rate of pilocytic astrocytomas.
Pilocytic astrocytomas are observed from infancy to adulthood. The incidence is highest from birth until four years of age and is generally much higher in childhood (<19 years of age) compared to adults and becomes uncommon with increased age (8).

1.2.2 Survival

The five-year overall survival is greater than 96% and the 10-year overall survival exceeds 95% for patients under 19 years of age (5). Unfortunately, the overall survival is only 70% for infants diagnosed with PA under <1 year of age (5). In addition, it has been shown in multiple studies that outcome of adult patients is inferior compared to the pediatric cohorts (9-12).

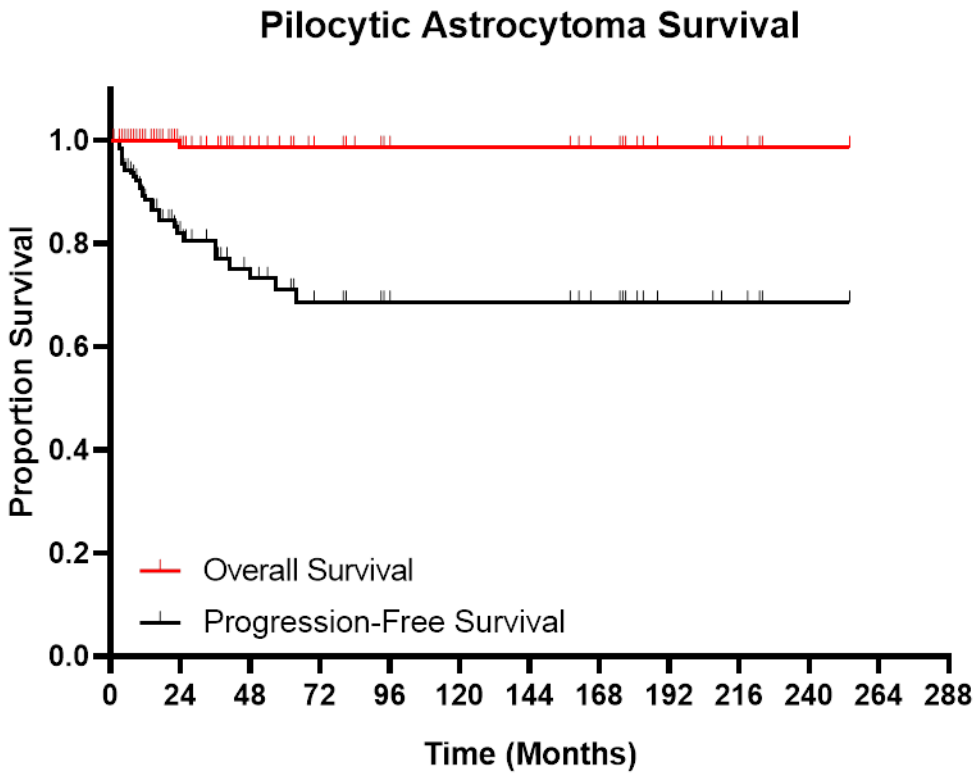


Figure 1.2.2: Survival rates of pilocytic astrocytomas.
Pediatric PA patients have an excellent overall survival, which is estimated to be greater than 95% at then years. However, the ten year progression-free survival is estimated to be less than 70% (13, 14).

On the other hand, incompletely resected PAs tend to recur and progression-free survival is more dismal compared to overall survival. In published cohorts, the progression-free survival ranges between 70% (13, 14) and 44% (8). Management of recurrent disease involves long-term treatment approaches or sequential treatment strategies (8, 15).

1.2.3 Location

Pilocytic astrocytomas can be located across the CNS from the optic nerve to the brainstem and spinal cord (16, 17). Predominantly, PA can be found in optic pathway, thalamus and basal ganglia, cerebral hemispheres, fourth ventricle, cerebellum, brainstem and spine (7, 18). Importantly, an association between age and tumor location was established. Specifically, PAs tend to be located in the cerebellum, spine and optic pathway in younger patients (18). Interestingly, Younes *et al* (2020) revealed that the cell-of-origin for cerebellar PA (19), which is the most common tumor location (6, 20), is found within the ventricular zone using an *in silico* analysis. Their findings suggest that a developmental pathway has been hijacked for PA

formation (19). In a different study, deconvolution at the single cell level showed that PA cells follow a developmental differentiation hierarchy from oligodendrocyte precursor-like cells to mature astrocyte-like cells (21). These findings were confirmed by a study that combined single-cell RNA (scRNA) and bulk RNA sequencing, which implicated oligodendrocyte progenitor cells as the cell-of-origin (22). In addition, they revealed that PA and oligodendrocyte progenitor cells have overexpression of cell surface genes in common such as *GPR17* and *TRPM8.* Taken together, these findings conclusively showed that tumors are formed from progenitor cells and not mature astrocytes (8).

1.2.4 Characteristics

In general, PA is a slowly growing, well-circumscribed tumor, which often presents as cystic lesion with one or more mural tumor nodules (6, 16, 20). The term pilocytic refers to cells that have hair-like, bipolar processes (6, 16, 20).

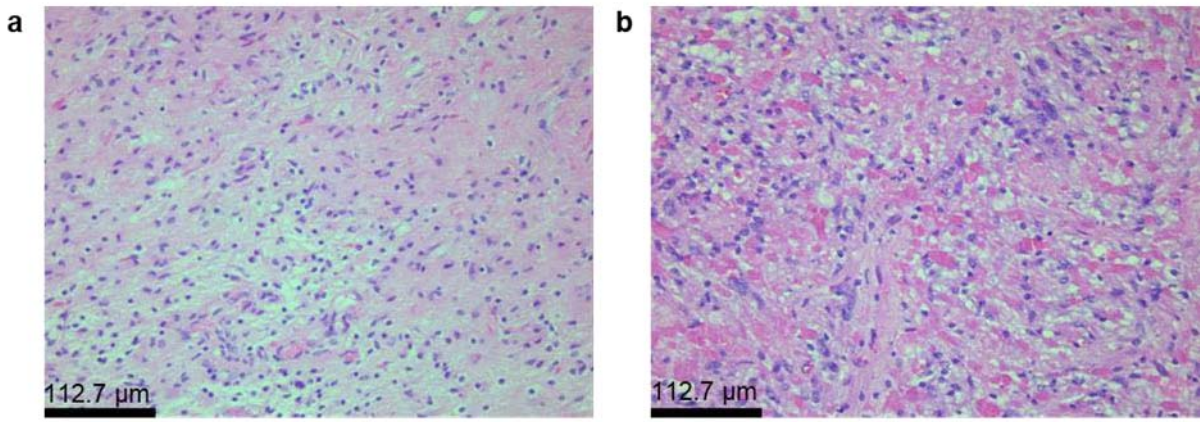


Figure 1.2.3: Pilocytic astrocytoma histology.
Representative images of pediatric pilocytic astrocytoma (PA) histochemistry stained with hematoxylin and eosin to show classical biphasic histology of PA with fibrillar and microcystic areas (a) as well as formation of eosinophilic processes called Rosenthal fibers (b) (images kindly provided by Dr. David Pauck).

The tumors tend to be discrete and displace the surrounding brain instead of being invasive. In contrast, PAs affecting the optic nerve are known to show a more invasive growth pattern (17). Microscopically, PAs constitute a tumor with low to moderate cellularity with some typical features such as Rosenthal fibers, eosinophilic granular bodies and a relatively high degree of vascularization (2, 7, 8, 16, 17). However, presence of Rosenthal fibers is not required for the neuropathological diagnosis of PAs also because these Rosenthal fibers are not exclusively observed in

this entity (2, 7). As a slow growing tumor, PAs tend to have a low mitotic index, around 4%, and, if much higher, then an alternate diagnosis should be considered (16). The diagnosis is relatively straightforward when PAs present with classical features (16).

To improve the diagnostic workup for tumors of the CNS, Capper and colleagues have devised a bioinformatic tool, which takes advantage of high-resolution DNA methylation profiling (23). This algorithm can accurately distinguish molecular entities and discriminates PAs on the basis of location (supratentorial, posterior fossa and midline, version 11b4). When Métais et al (2023) performed their study looking into spinal PAs, they generated a t-distributed stochastic neighbor embedding (t-SNE) plot of astrocytic tumor and control tissues (24).

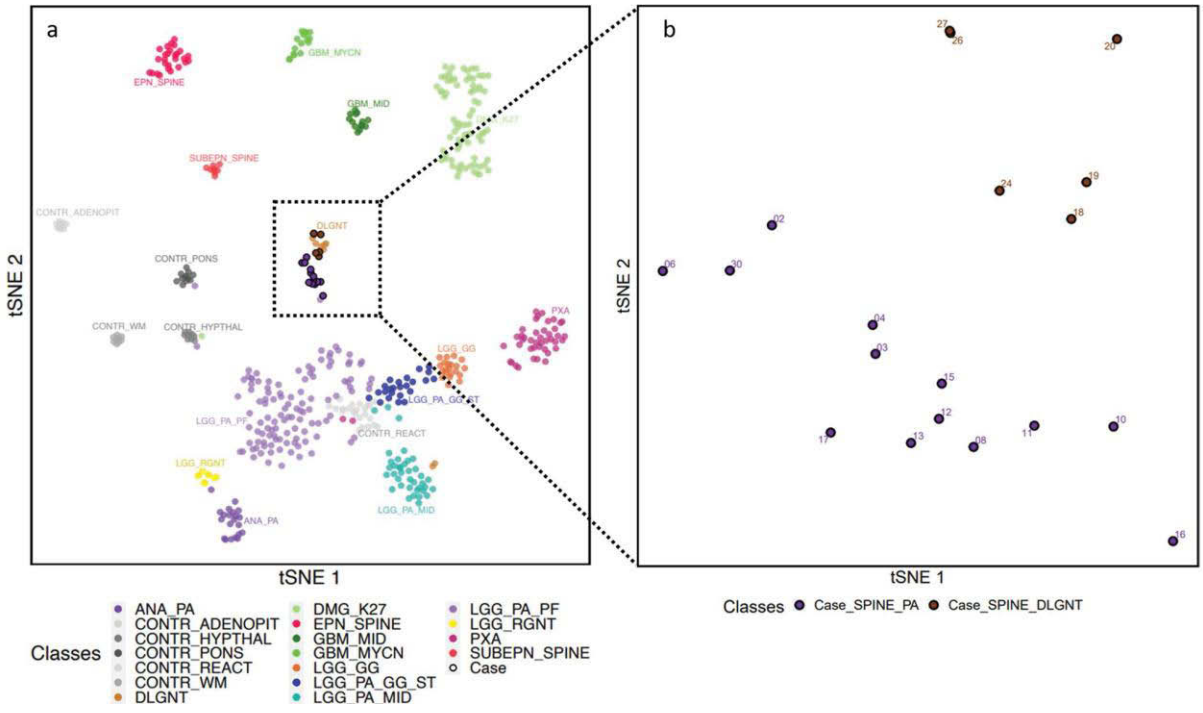


Figure 1.2.4: DNA methylation based classification of glial tumors.
Métais *et al* (2023) aimed to compare spinal pilocytic astrocytomas (PA) to PAs in other locations with epigenetics using the algorithm from MolecularNeuropathology.org. **a** tSNE plot distinguishes astrocytic tumors and control tissues based on the epigenetic profile. Pilocytic astrocytomas (purple and dark blue) cluster closely with control reactive tissue and spinal PA cluster with diffuse leptomeningeal glioneuronal tumours (DLGNT). **b** Spinal PAs cluster closely to DLGNT spinal tumors, but form a distinct cluster in the tSNE plot.

Interestingly and not pertinent to their results, PAs from the posterior fossa, supratentorial and midline location all clustered around the “control tissue, reactive brain”. According to the documentation (https://www.molecularneuropathology.org/mnp/classifiers/1), “control tissue, reactive brain” is a recurrently observed methylation profile of unclear status which constitutes histopathologically assigned low-grade tumors such as gangliogliomas or PAs possibly due to low tumor content. This might suggest that the algorithm cannot appropriately segregate PAs into distinct subgroups and that epigenetics alone may not be sufficient to fully appreciate the biological heterogeneity of PAs.

1.2.5 Stardard-of-care

The standard-of-care for PAs is a watch-and-wait strategy or surgery alone in progressive or symptomatic tumors, due to the fact that tumors are well-circumscribed in which case gross total resection is often achievable for PAs, and, therefore, additional therapy is not required (17, 20, 25). To relieve symptoms such as epileptic seizures, hydrocephalus and raised intracranial pressure surgery is performed (20). However, some tumors, particularly those located in the optic nerve or brainstem, cannot be fully resected and tend to recur (16, 18). The average recurrence rate is double in adult patients compared to children (12). Importantly, the extent of resection is significantly correlated with progression of PAs (26, 27). In some cases, PAs could be considered a chronic lifelong disease with multiple potential recurrences, specifically in children (8, 21). When gross total resection is not possible, chemotherapy and radiotherapy (depending on age) can be administered to the patient (8, 16, 17). For chemotherapy, standards include use of carboplatin and vincristine as combination therapy or vinblastine monotherapy (8, 20, 28), and temozolomide being the most common drug employed in adults (20).

1.3. Genetics of pilocytic astrocytoma

1.3.1 MAPK signaling in pilocytic astrocytomas

Pilocytic astrocytoma is slow growing and is classified as a WHO 1 tumor. As such, it is not surprising that, in a pan-cancer study, PA was found to have the lowest overall mutation rate compared to other tumors (29). However, nearly 100% of alterations

observed in PA are centered on the mitogen activated protein kinase (MAPK) pathway (8, 14, 17, 21). These alterations will have one of two functions: 1) de-repress RAS signaling or 2) cause hyper-activation of the extracellular signal-regulated kinase (ERK) and the MAPK/ERK kinase (MEK) signaling which leads to cell growth and survival (8, 30). The most common alterations are a fusion between *KIAA1549* and B-Raf proto-oncogene, serine/threonine kinase (*BRAF*), followed by mutations in *BRAF* or Neurofibromin 1 (*NF1*) (8, 15, 16, 30). Interestingly, Reitman and colleagues were able to show that MAPK signaling is heterogeneously activated across the tumor suggesting that combination therapy might be required to treat PAs effectively (21). This is important since MAPK activation can lead to oncogene-induced senescence, a process where growth arrest occurs due to a tumor-suppressive mechanism in response to oncogene activation (13, 30).

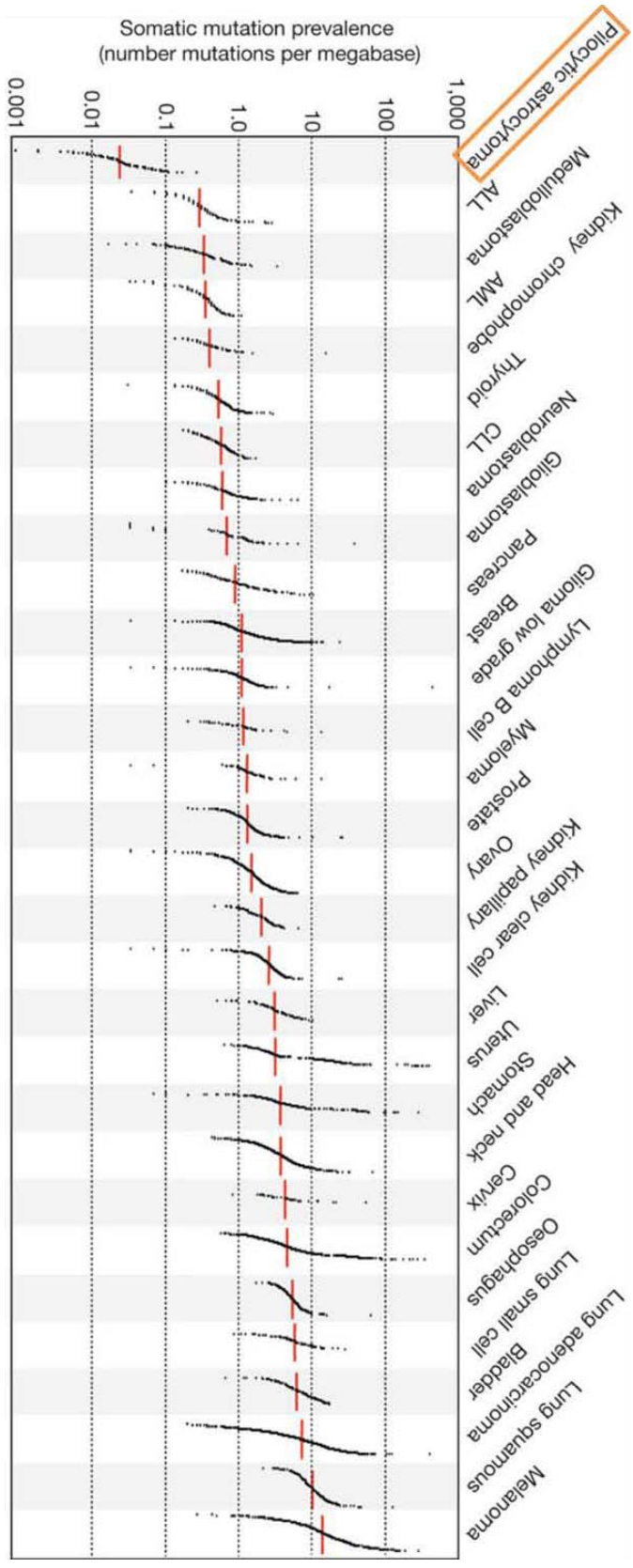


Figure 1.3.1: Pan-cancer mutational landscape.
Number of somatic mutations per megabase is shown across a pan-tumor panel. Mean number of mutations is highlighted with red bar. Tumors are arranged with the smallest number of mutations on the left (pilocytic astrocytoma is highlighted in orange) and highest number on the right (adapted from Alexandrov *et al* (29)).

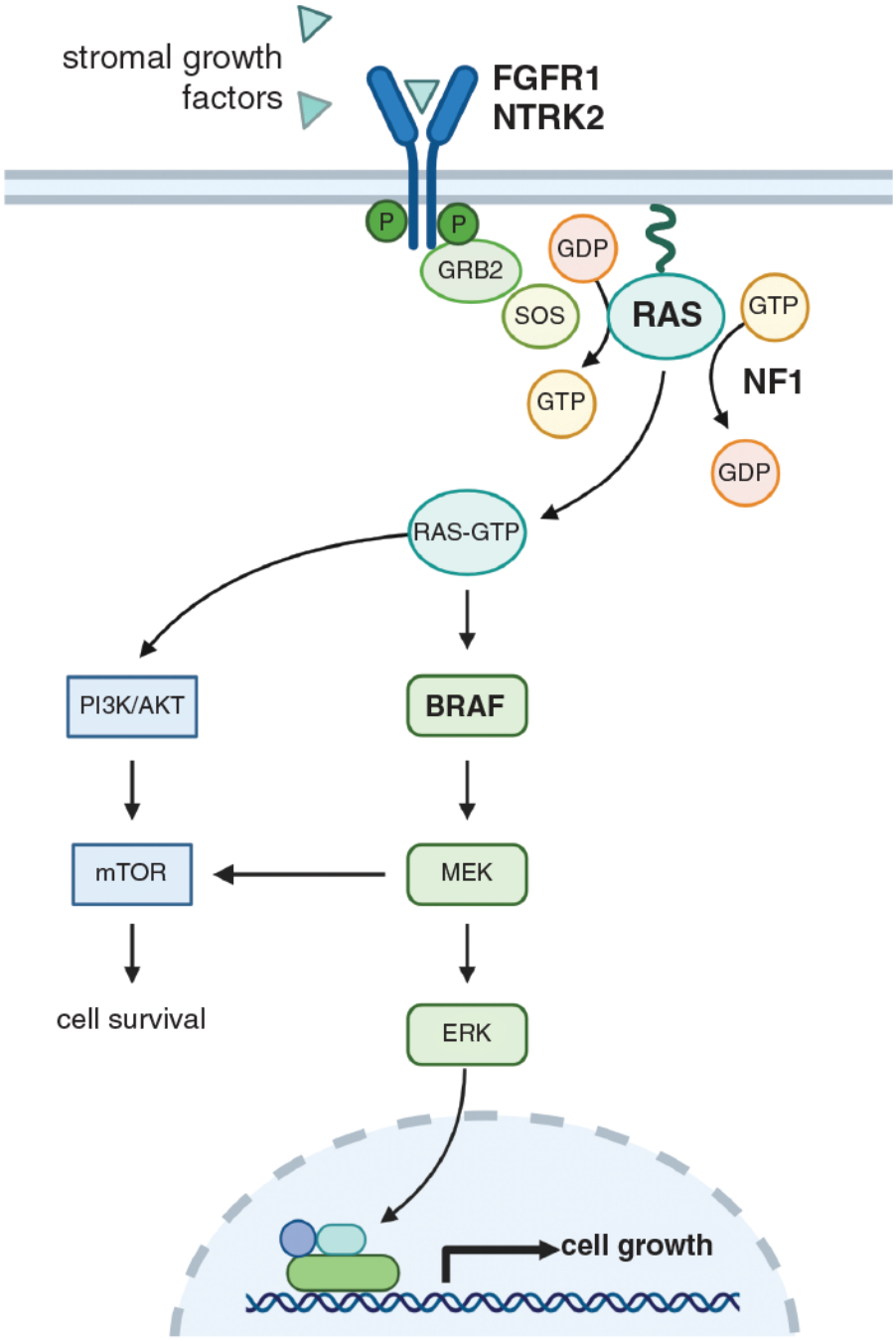


Figure 1.3.2: Schematic overview of mitogen activated protein kinase pathway.
The mitogen activated protein kinase (MAPK) pathway is schematically illustrated. Receptor tyrosine kinases such as FGFR1 or NTRK2 cause pathway activation through RAS, leading to an activation of MEK/ERK through BRAF or of mTOR though PI3K/AKT signaling. The NF1 protein represses the pathway by binding to RAS (8).

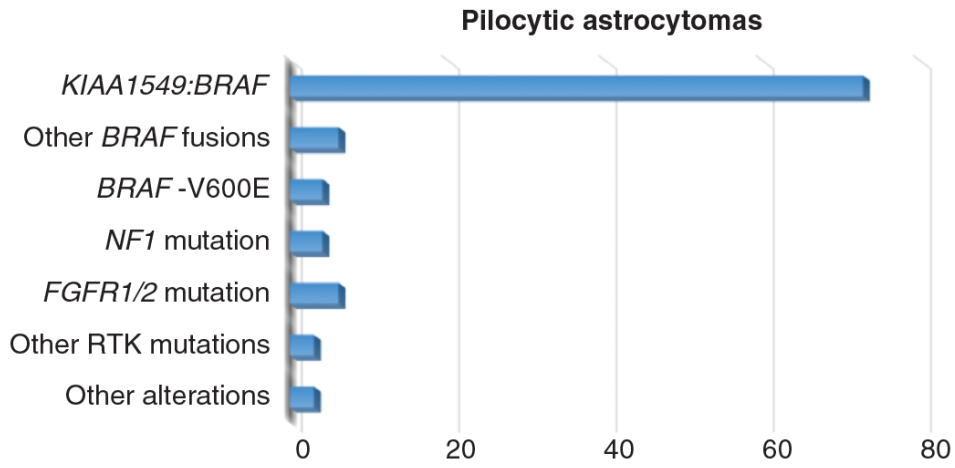


Figure 1.3.3: Genetic alteration spectrum in pilocytic astrocytomas.
Distribution of alterations in pilocytic astrocytomas. X-axis shows the frequency of genetic alterations in percent of tumors identified in this specific cohort. Notably, *BRAF* fusions account for approximately 80% of all alterations (8).

1.3.2 The relevance of *KIAA1549::BRAF* fusion in pilocytic astrocytoma

Over the years, PA studies focused on genomic alterations identified that very few alterations could be observed in most cases (31-33). In fact, most copy number alterations in PA are centered on the q arm of chromosome 7, specifically 7q34, which occurs in up to 70% of all PAs (8, 15, 16). This region, approximately 1.9-2 Mb in size, is often part of a tandem duplication containing *KIAA1549* and *BRAF* at either ends. *BRAF* is a serine/threonine–specific protein kinase in the MAPK pathway and *KIAA1549* is an uncharacterized gene (34, 35). When recombined, a *KIAA1549::BRAF* fusion gene is created. The most common fusions include exons 15::09, 16::09 and 15::11 of *KIAA1549* and *BRAF* (35, 36). In addition, *BRAF* has been known to fuse to other partners such as *FAM131B*, *CLCN6*, and *NTRK2* (14-16). In all fusions observed to date, the kinase domain of *BRAF* is retained and is constitutively active (30, 35). Despite the increasing knowledge about *BRAF* fusions, there is a need for quick and efficient molecular diagnostic tests to detect the full spectrum of *BRAF* fusions (37).

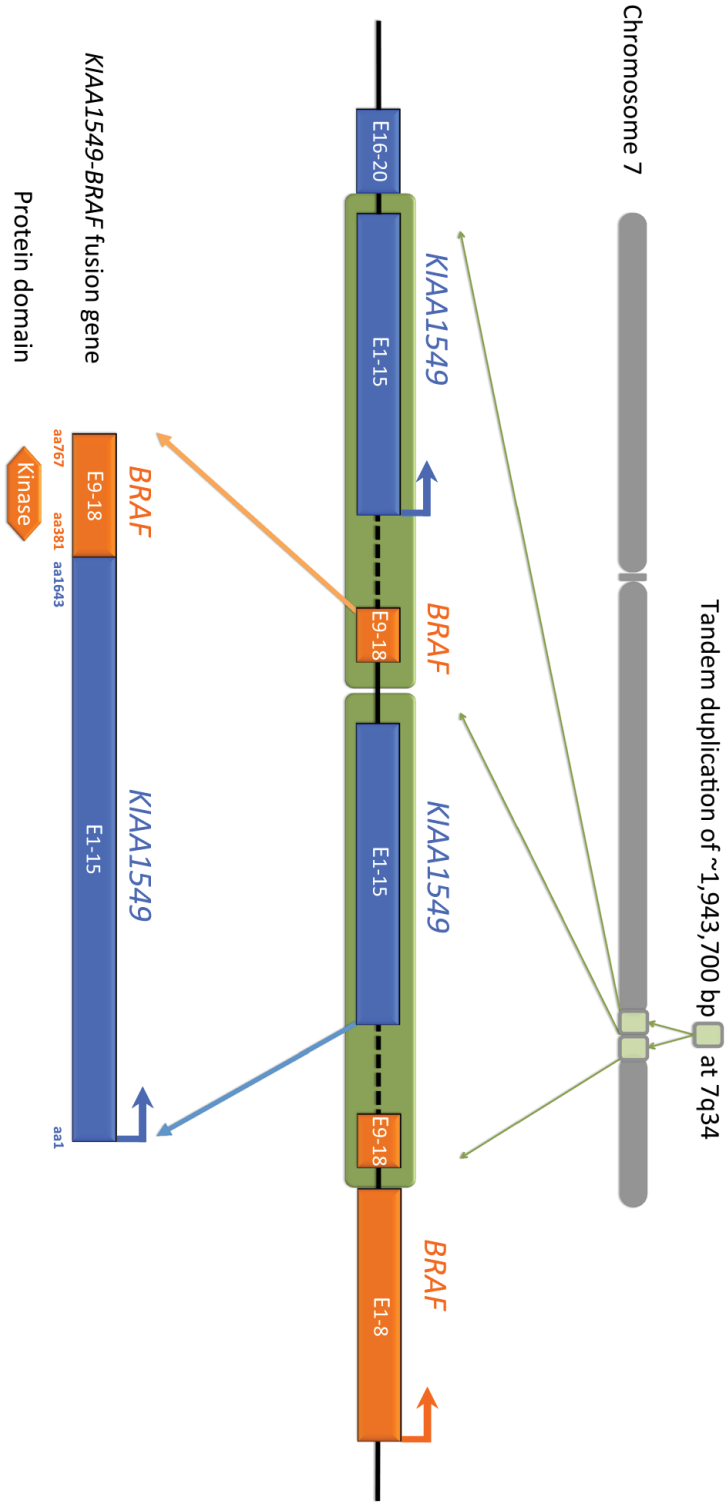


Figure 1.3.4: Schematic overview of the most common *KIAA1549::BRAF* fusion. Example of tandem duplication which creates a fusion protein encompassing the N-terminal end of *KIAA1549* and the C-terminal end of *BRAF*, and retains the kinase domain of *BRAF* (adapted from Subbiah *et al* (38)).

1.3.3 The *BRAF V600E* mutation in pilocytic astrocytomas

One of the most common single nucleotide variations (SNV) in PA is *BRAF V600E*, which is observed in about 5-6% of the patients (8, 14, 33). However, it is important

to note that in melanoma, this alteration is detected in approximately 50% of patients (39, 40). This altered amino acid is located in the kinase domain and is conserved through evolution with the exception of *Drosophila* (34). It is part of a highly conserved motif in a region which has been shown to form a loop that anchors the phosphates of ATP and is important for ATP catalysis (34). Studies have shown that this single mutation results an oncogenic phenotype in cell lines (34), most likely caused by an upregulation of downstream ERK through binding of RAS (35). Finally, it was shown that mutations in *BRAF* could induce PA formation in a RCAS/t-va mouse model (41).

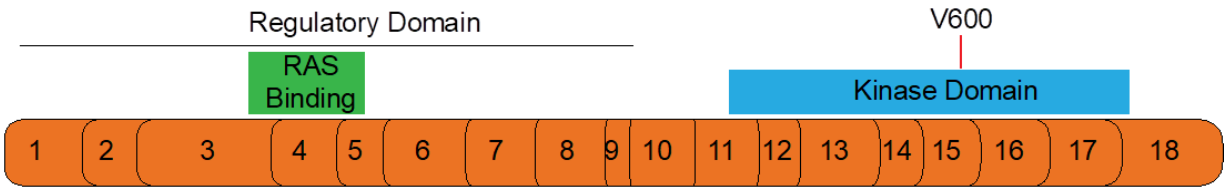


Figure 1.3.5: Schematic overview of the *BRAF* gene.
The *BRAF* gene contains 18 exons, an N-terminal RAS binding domain and a C-terminal kinase domain. The most commonly altered amino acid V600 is highlighted (modified from (37) with information from COSMIC).

1.3.4 The NF1 gene in pilocytic astrocytoma biology

The *NF1* gene contains 60 exons and is ubiquitously expressed, but the highest levels have been found in adult neuronal cells (42). Mutations in the tumor suppressor gene cause neurofibromatosis type 1, an autosomal dominant disorder, which increases the risk of tumors, specifically in the optic pathway amongst other disease predispositions (42-44). In PAs, alterations in the *NF1* gene are observed in 6-10% of patients (8, 30). The NF1 protein is a GTPase that downregulates RAS signaling by binding to RAS (42). Mutations in *NF1* cause a hyper-activation of RAS and subsequent activation of MAPK and mTOR, thereby inducing cell growth (28, 43, 44).

1.3.5 Additional MAPK altered genes

In addition to alterations in *BRAF* and *NF1*, other MAPK pathway alterations have been observed in *KRAS*, *NTRK2*, and *FGFR1/2* (8, 14). To date, three main RAS genes have been characterized: *KRAS*, *HRAS* and *NRAS*. In addition to regulating

metabolism, RAS signals upstream of BRAF and PI3K/AKT to control expression of MEK/ERK and mTOR, respectively (8, 45). Interestingly, no alterations have been observed in HRAS or NRAS, suggesting that KRAS is the key player in PA tumorigenesis (30). NTRK2 is a transmembrane tyrosine kinase and plays an important role in neuronal development (46). Fusions involving the *NTRK2* gene have been found in multiple other cancers, such as neuroblastoma, colon cancer, and melanoma (14, 46). Fusions of *NTRK2* involve the N-terminal of a partner gene and the C-terminal kinase domain (14, 46). The FGFR1 and FGFR2 proteins are fibroblast growth factor receptors and transmembrane tyrosine kinases (15, 47). Previous studies demonstrated that *FGFR1* has an internal tandem duplication (ITD) in the kinase domain or a N-terminal fusion partner (14, 16, 47). One example is *FGFR1-TACC1* where the C-terminal part of *TACC1*, which contains a coiled-coil domain, is retained as part of the fusion (48). This alteration facilitates the localization of the fusion protein to the mitotic spindle, which in turn promotes aneuploidy and tumorigenesis (48). Additionally, this fusion is capable of activating the MAPK pathway (8, 48).

1.3.6 Targeted therapy approaches for pilocytic astrocytoma patients

Due to the fact that MAPK signaling is active in PAs, targeted therapies have focused on this pathway. Initial studies with the BRAF inhibitor PLX4720 showed a decrease in MEK phosphorylation for BRAF V600E expressing cell lines, but an opposite effect in KIAA1549::BRAF expressing lines, where phosphorylation of MEK increased significantly (49). Another example of BRAF as a molecular target is the study by Nobre *et al*., where they retrospectively looked at PA patients with BRAF V600E mutations treated with either dabrafenib or vemurafenib (50). In this case, there was a significant decrease in tumor size with treatment. However, 17 patients that stopped treatment experienced rapid tumor regrowth (50).

There are a number of prospective clinical trials targeting the MAPK pathway, specifically for NF1 or BRAF-altered low-grade gliomas, which include PAs (8). A clinical trial evaluating treatment efficiency of sorafinib, a multikinase inhibitor targeting RAF1 amongst other kinases, was terminated due to the inefficacy of treatment (NCT01338857). Other trials using dabrafenib (BRAF V600 inhibitor) and trametinib (MEK inhibitor) have been more promising. These trials evaluated these targeted agents either as a monotherapy or in combination, or the combination of

therapeutics compared to carboplatin and vincristine standard therapy (NCT01677741 (51), NCT02124772 (52), NCT03975829, NCT02684058). In NCT01677741, dabrafenib was administered to patients with BRAF V600E altered tumors. In total, 13 or 41% of the patients in the trial cohort were diagnosed with PA. Dabrafenib was found to be well tolerated by the patients and showed a manageable safety profile for greater than two years (51). The most common reason in this study for terminating treatment was physician and/or parent decision (42%), likely stemming from the fact that standard treatment is administered for 12-24 months. Adverse effects were only mentioned in 8% of the cases as the reason the termination of treatment (51). In this study, the disease control rate was 88% showing that dabrafenib is well tolerated in pediatric patients, has a distinct clinical benefit and should be evaluated in a larger population (51). In a second prospective study, trametinib monotherapy or trametinib in combination with dabrafenib was tested against BRAF V600E-mutant low-grade gliomas (52). Trametinib and dabrafenib combination therapy has been approved for treatment of BRAF V600E-mutant solid tumors, such as melanoma, in patients ≥ 6 years of age and, in addition, has shown clinical activity in adult gliomas (52). In this study, the mean duration of response was 33.6 months with a progression-free survival of 16.4 months for trametinib treatment and 36.9 months in the combined treatment group (52). However, there were 22% of patients which terminated treatment due to toxicity in the combined therapy compared to 54% in the monotherapy arm. Overall, this study shows that combined BRAF V600E inhibitor and MEK inhibitor treatment is more effective than MEK inhibition alone in BRAF V600E-mutant pediatric low-grade gliomas (52).

Additionally, selumetinib (MEK inhibitor) or randomized trials in comparison to carboplatin and vincristine therapy are currently conducted (NCT01089101 (28), NCT03871257). In NCT01089101, patients were stratified according to the alteration type with stratum 1 containing the BRAF-altered PA patients and stratum 3 NF1-mutated low-grade glioma patients (stratum 2 and 4 were not discussed in this manuscript) (28). Approximately half of patients (56%) in stratum 1 completed the entire treatment regimen with 14/25 patients having tumor progression (28). The two-year progression-free survival was 70%, however, when separated into type of BRAF alteration, BRAF V600E-mutant patients had worse progression-free survival compared to BRAF-KIAA1549 patients (28). Importantly, the toxicity was manageable

with rare serious adverse effects. Patients that achieved 1-49% tumor shrinkage had “stable disease” and this is considered clinically beneficial since the majority of patients do not succumb to PA (28).

Finally, there have been mTOR inhibitor trials using everolimus (also known as RAD001) or temsirolimus in combination with erlotinib (EGFR inhibitor) which have been completed (NCT00112736, NCT01158651, NCT00782626). In one everolimus study (NCT001158651), one patient (4.4%) died, while six patients (26.1%) suffered from serious adverse events or disease-related morbidity such as blindness, anemia and seizures. However, the proportion of participants responding to treatment was 68% where response was considered promising if there was a 25% response rate after 48 weeks. In a second study with everolimus (NCT00782626), although 26% of patients suffered from serious adverse events, 91.3% of patients had stable disease and only 8.7 % of patients had a partial response to treatment. In the third study, where temsirolimus was evaluated in combination with erlotinib (NCT00112736), only 14% of suffered from serious adverse events. However, in this study, only 17% of patients had stable disease, 8% of patients had only a partial response to treatment, but 75% of patients presented with progressive disease.

1.4. Proteogenomics

1.4.1 Proteogenomic approach in cancer research

The term “proteogenomic” first appeared in 2001, according to pubmed (https://pubmed.ncbi.nlm.nih.gov/). This review was focused on the biological principles of chemotherapy (53). The term next appeared in a human research publication three years later (54). In this study, the authors used western blotting, mass spectrometry and genetic profiling on a number of cell lines and serum from patients with prostate cancer with the intent of finding new antigens that have therapeutic, diagnostic or prognostic value. The mass spectrometry was only performed on a single band, which showed differences using western blotting and not the entire proteome. It took an additional ten years before proteogenomics was performed on the entire proteome in cancer, although they only focused on single nucleotide polymorphims (SNPs) of genes and proteins located on chromosome 9 (55).

More recent publications take advantage of multiple layers of –omics data, which includes transcriptomics (RNA and miRNA expression), genomics (mutations and DNA copy-number profiling), epigenetics (DNA methylation and histone modifications), proteomics (phospho-proteomics, glycosylation and acetylation), lipidomics and metabolomics (56-63). Many of these publications were conducted within the context of the Clinical Proteomic Tumor Analysis Consortium (CPTAC). Sample preparation is commonly performed using cryo-pulverization and subsequent sample distribution into the required number of sample tubes for each analytical platform. This approach assures that there will be reduced variation between layers of data. However, even using this method of sample preparation, the mean correlation between protein and RNA expression was only 0.35 (58) which is in-line with other studies (64, 65).In all these examples, proteogenomics was a multi-dimensional analysis which was not integrative, sometimes creating groups for each layer of data analyzed.

1.4.2 Similarity Network Fusion

Data integration processes need to take into account two major key points: First of all, the fact that current studies include few samples in relation to large number of data points (up to 935,000 probes for methylation data). Secondly, there could be differences in scale and noise in each dataset (66). The easiest way to avoid these issues would be to analyze each layer separately and then combine the data, however, this type of analysis often leads to inconsistent conclusions (61, 66). Importantly, many methods to analyze bioinformatic data is to reduce the noise by preselecting a set of genes/probes/proteins before performing such analyses as “consensus clustering” (67) or “K-means clustering” (68). However, pre-selection leads to a biased analysis.

The advantage of using Similarity Network Fusion (SNF) is that this method integrates information at the sample level instead of measurement level (66). Similarity network fusion constructs a sample-similarity network for each layer of data and, then, integrates the networks into a fused similarity network using a non-linear combination method (66). This is a particularly useful tool since the data could consist of a small number of samples, the heterogeneity in data types is limited to a single similarity network layer and, importantly, pre-selection is not used.

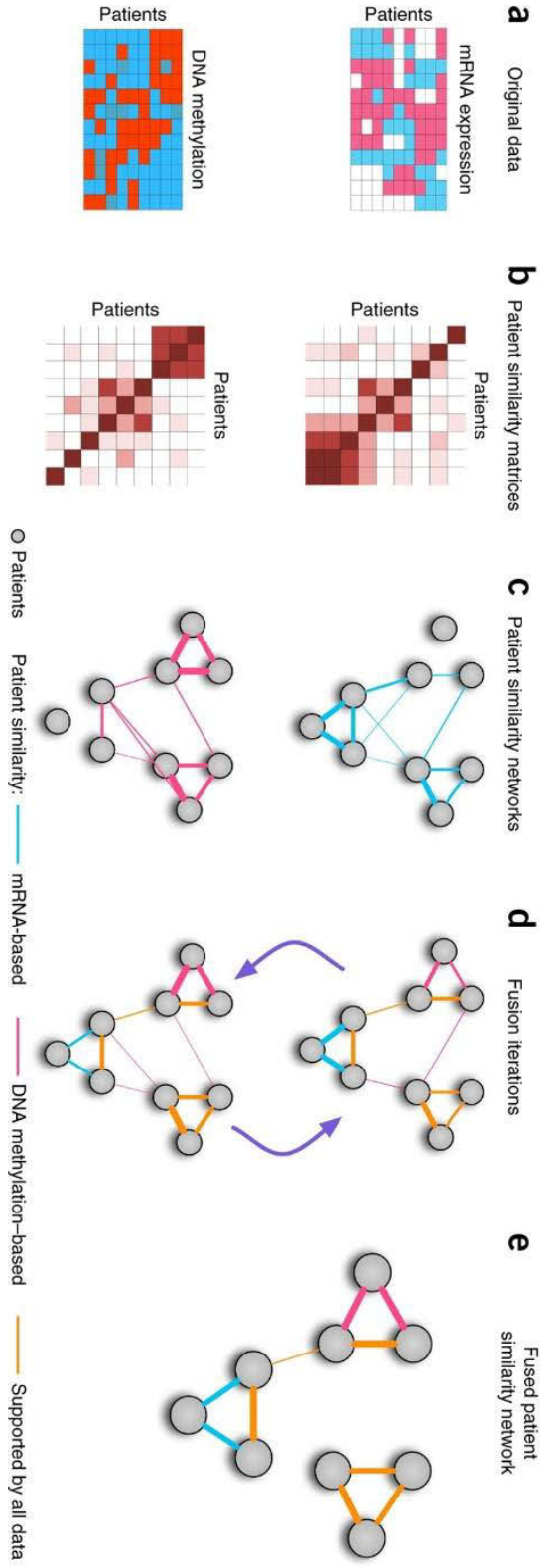


Figure 1.4.1: Application examples of Similarity Network Fusion analyses.
a. Representative images of single layer analyses of mRNA expression and DNA methylation for the same patients (66). **b.** A similarity matrix is constructed for each of the layers. **c.** Conversion of similarity matrix to a similarity network. **d.** Iterations of the fused similarity network are performed to determine the most stable network **e.** Final network visualization constitutes the last step of the approach.

In the first test of SNF, the authors tested previously published data from The Cancer Genome Atlas (glioblastoma, breast invasive carcinoma, kidney renal clear cell carcinoma, lung squamous cell carcinoma and colon adenocarcinoma) (66). In their hands, SNF of mRNA expression, DNA methylation and miRNA expression data identified clinically relevant subgroups for each of the diseases by decreasing experimental noise which is randomly distributed, but the biological signal is consistently recognized when multiple data layers are evaluated.

Others have used this technique which can lead to better understanding of novel pathways or better classification of a disease (62, 69, 70). Interestingly, Forget *et al* were able to consistently classify medulloblastoma into the four subgroups with a relatively small number of samples (62). Their study not only led to the discovery that SRC is a major driver of Group 4 medulloblastomas, but also to the first Group 4 mouse model (62, 71).

1.5. Glioblastoma

1.5.1 Incidence

Glioblastoma IDH-wildtype, referred to as glioblastoma (GB) in this thesis, belongs to the adult-type diffuse glioma group and accounts for 50% of all malignant brain tumors (2, 4, 72). Glioblastoma occurs more frequently in older adults, with a peak incidence between 45 and 75 years of age and a median age of 65 years (4).

1.5.2 Survival

Patients with glioblastoma have a dismal prognosis. The 1-year survival rate is only 42.7%, and there is a dramatic drop to 6.9% and 4.3% for the 5-year and 10-year survival rates, respectively (4, 73).

1.5.3 Location

Glioblastoma is frequently observed in the cerebrum (frontal, temporal, parietal, and occipital lobes) and only rarely located in the cerebellum. Symptomatically, glioblastoma presents with signs of increased intracranial pressure and about one-third of patients experience epileptic seizures. Other neurological symptoms include paralysis, sensory disturbances, visual disturbances, speech impairments, dizziness, memory impairments, or personality changes, depending on the tumor's location (and perilesional edema) (2).

1.5.4 Characteristics

Relevant histopathological features for diagnosis include necrosis, pathological microvascular proliferation and a high mitotic activity, leading to classification in the CNS WHO Grade 4, the highest grade in the WHO classification for central nervous system tumors (2).

The histopathology of glioblastoma is extremely variable, reflected in the old name "glioblastoma multiforme" (7). In the 5th edition of the WHO Classification of CNS Tumors, it is now clearly defined that glioblastoma do not harbor IDH mutations (2). The previously used distinction between "primary glioblastoma" (*de novo* arising and IDH-wildtype) and "secondary glioblastoma" (arising from a less malignant precursor lesion and thus almost always showing an IDH mutation) is now obsolete; the latter group is now defined as astrocytoma, IDH-wildtype (CNS WHO Grade 4). Gliosarcoma and giant cell glioblastoma are rare histological variants of GB (74). For clinical purposes, they are considered variants of IDH-wildtype GB, but genetic data enabling definitive classification are still sparse.

Most GB arise sporadically, although specific SNV have been associated with an increased risk (2, 72). Some of the alterations affect the promoter of *TERT*, the *TP53* and *PTEN* genes and amplifications of the *EGFR* gene (2, 75). Rarely, GB arises within the context of an inherited tumor predisposition syndrome, such as the very rare Li-Fraumeni syndrome or inherited deficiency of the DNA mismatch repair system. Ionizing radiation is the only exogenous risk factor for the development of glioblastoma, such as radiation therapy of the CNS in children with acute leukemia (72).

1.5.5 Standard of care

Current treatment of GB is multifaceted, initially consisting of surgical resection, preferably macroscopically complete ("gross total"), followed by radiotherapy and concurrent and adjuvant chemotherapy with the oral alkylating agent temozolomide (72). Additional administration of lomustine is also possible and offers survival advantages compared to monotherapy with temozolomide for the subgroup of patients whose tumors demonstrate methylation of the O6-methylguanine-DNA methyltransferase (*MGMT*) gene promoter (76).

Despite this multimodal therapeutic approach, the overall survival of GB patients has only marginally improved over the past three decades, and the prognosis remains grim despite advances in treatment. A younger age at diagnosis and good general condition independently represent favorable prognostic factors, as does methylation of the *MGMT* promoter. A problem in GB treatment is that the tumor cells develop resistance to available treatment modalities quickly, resulting in local tumor recurrence (77, 78).

1.6. Long non-coding RNA

1.6.1 Classification

By definition, long non-coding RNAs are transcripts that are longer than 200 base pairs and do not code for proteins (79). There are seven classes of lncRNAs which include intergenic, intronic, bidirectional, enhancer, promoter-associated, sense and antisense lncRNAs (77). The largest category of lncRNAs consists of long intergenic non-coding RNAs which are situated in the intergenic regions of the genome and do not overlap with any other genes, coding or non-coding. Long non-coding RNAs which are located within a gene, but do not overlap with exonic regions are termed intronic lncRNAs. Bidirectional lncRNAs are transcribed from regions opposite coding genes, typically within 1 kb of promoters, and frequently share the same promoter. Additionally, lncRNAs transcribed from the enhancer or promoter regions of a gene are designated as enhancer and promoter-associated lncRNAs, respectively. Finally, lncRNAs transcribed from the opposite sense or antisense strand of coding genes are termed sense lncRNAs or antisense lncRNAs, respectively (77, 80).

1.6.2 Detection

In the past, microarrays and tiling-arrays have been used to study a limited number of previously identified lncRNAs (81, 82), however RNA sequencing is now considered the gold standard for identifying lncRNAs. RNA sequencing can identify both known and unknown transcripts as approximately 75% of the genome is transcribed of which less than 2% encompasses coding genes (80). The preferred RNA sequencing method is to perform sequencing on ribosome depleted total RNA (83). Although the majority, but not all, lncRNAs contain a poly-A tail, polyadenylated RNA sequencing would not provide a complete analysis of the lncRNAome (84).

1.6.3 Long non-coding RNA in cancer

Due to their significant involvement in gene regulation, lncRNAs are frequently implicated in the pathogenesis of various diseases and noteworthy examples of potential oncogenic lncRNAs are reported in the literature (85). First, HOX transcript antisense RNA (*HOTAIR*) (86) promotes proliferation and migration in cervical cancer (87), contributes to chemoresistance in colorectal cancer (88), and correlates with metastasis and poor survival in breast cancer (89). Another potent oncogenic lncRNA implicated in tumorigenesis is the nuclear paraspeckle assembly transcript 1 (*NEAT1*) (90), which is commonly upregulated and correlated with unfavorable prognosis (91) in multiple cancer types such as hepatocellular carcinoma (92), non-small cell lung cancer (93), cholangiocarcinoma (94), and breast cancer (95, 96). NEAT1 is associated with a number of oncogenic processes including cell proliferation, invasion, tumorigenesis, and metastasis (92-96). Finally, several lncRNAs have been investigated in GB including the following: maternally expressed 3 (*MEG3*) modulates proliferation through interactions with TP53 and MDM2 proteins (97); colorectal neoplasia differentially expressed (*CRNDE*) regulates glioma cell growth via mTOR signaling (98); and HOXA11 antisense RNA (*HOXA11-AS*) which promotes glioma cell growth and metastasis by targeting the miR-130a-5p/high mobility group box 2 (HMGB2) axis (99).

1.7. Aims of this thesis

This thesis aimed to unravel previously understudied molecular questions in the most common brain tumors in children (PA) and adults (GB), namely

1- Is there biological heterogeneity in PA that has, to date, not been observed? and
2- Can a poorly understood molecular data type, specifically the lncRNA transcriptome, be informative in describing the difference between GB patients with shorter *versus* longer survival?

In the first manuscript, the aim was to use proteogenomics to integrate RNA sequencing and proteomic data of overlapping PA tissue samples to potentially observe heterogeneity in this disease. Simply segregating samples was not considered a promising observation as other groups have successfully segregated PA into groups using different techniques. It was important that the grouping was

stable and could be applied to additional datasets. Following segregation of samples, clinical and biological heterogeneity should be studied.

In the second manuscript, the aim was to discern genes which were differentially expressed between short-term and long-term surviving patients with GB. As such analyses had already been performed on the coding genes, the microarray data was re-examined with a focus on long non-coding RNAs as biomarkers to distinguish the two groups. As a biomarker, the candidate must show a difference in survival in several data sets and show that the expression is directly linked with at least one oncogenic phenotype.

Combined, these studies show that there is much value in analyzing data using non-traditional approaches such as integrative analyses of the proteogenome or the lncRNAome.

2. Manuscripts

2.1. Manuscript 1: Integrative multi-omics reveals two biologically distinct groups of pilocytic astrocytoma

Acta Neuropathologica (2023) 146:551–564
https://doi.org/10.1007/s00401-023-02626-5

ORIGINAL PAPER

Integrative multi-omics reveals two biologically distinct groups of pilocytic astrocytoma

Daniel Picard[1,2,3] · Jörg Felsberg[3] · Maike Langini[1,4,5] · Paweł Stachura[1,6] · Nan Qin[1,2,3] · Jadranka Macas[7,8,9] · Yvonne Reiss[7,8,9] · Jasmin Bartl[1,2,3] · Florian Selt[10,11,12] · Romain Sigaud[10,11,12] · Frauke-D. Meyer[1,2,3] · Anja Stefanski[4,5] · Kai Stühler[4,5] · Lucia Roque[13] · Rafael Roque[14] · Aleksandra A. Pandyra[1,15,16] · Triantafyllia Brozou[1] · Christiane Knobbe-Thomsen[3] · Karl H. Plate[7,8,9] · Alexander Roesch[2,17,18] · Till Milde[10,11,12,19] · Guido Reifenberger[2,3] · Gabriel Leprivier[3] · Claudia C. Faria[20,21] · Marc Remke[1,2,3]

Received: 11 May 2023 / Revised: 4 August 2023 / Accepted: 18 August 2023 / Published online: 1 September 2023

Abstract

Pilocytic astrocytoma (PA), the most common pediatric brain tumor, is driven by aberrant mitogen-activated protein kinase signaling most commonly caused by *BRAF* gene fusions or activating mutations. While 5-year overall survival rates exceed 95%, tumor recurrence or progression constitutes a major clinical challenge in incompletely resected tumors. Here, we used similarity network fusion (SNF) analysis in an integrative multi-omics approach employing RNA transcriptomic and mass spectrometry-based proteomic profiling to molecularly characterize PA tissue samples from 62 patients. Thereby, we uncovered that PAs segregated into two molecularly distinct groups, namely, Group 1 and Group 2, which were validated in three non-overlapping cohorts. Patients with Group 1 tumors were significantly younger and showed worse progression-free survival compared to patients with group 2 tumors. Ingenuity pathways analysis (IPA) and gene set enrichment analysis (GSEA) revealed that Group 1 tumors were enriched for immune response pathways, such as interferon signaling, while Group 2 tumors showed enrichment for action potential and neurotransmitter signaling pathways. Analysis of immune cell-related gene signatures showed an enrichment of infiltrating T Cells in Group 1 versus Group 2 tumors. Taken together, integrative multi-omics of PA identified biologically distinct and prognostically relevant tumor groups that may improve risk stratification of this single pathway driven tumor type.

Keywords Pilocytic astrocytoma · Intertumoral heterogeneity · Integrative multi-omics

Introduction

Pilocytic astrocytomas (PAs) are the most common primary brain tumors in children [29]. The majority of tumors develop in the cerebellum, followed by less common locations in other midline structures, such as the optic nerve and chiasm, the hypothalamus and the spinal cord, and by locations in the cerebral cortex [40]. PAs are typically slowly growing, and, if well-circumscribed, can be successfully treated by surgery [28], with 5-year overall survival rates exceeding 95%. Subtotally resected or unresectable tumors due to tumor location, e.g., tumors located in the optic tract and hypothalamus, tend to recur and may require adjuvant therapy by local irradiation or systemic chemotherapy [9]. However, approximately 55% of these tumors progress following current standard of care treatment, and novel treatment strategies are thus urgently needed [24].

Concerning pathogenesis, PAs are considered as a single pathway disease driven by genetic alterations of genes encoding members of the mitogen-activated protein kinase (MAPK) signaling cascade, with the majority of tumors, in particular among the cerebellar PAs, carrying *KIAA1549::* B-Raf proto-oncogene, serine/threonine kinase (*BRAF*) fusions that lead to aberrant MAPK pathway activation [16]. Less common genetic alterations in PAs include activating *BRAF* codon 600 mutations, inactivating neurofibromin 1 (*NF1*) mutations or rarely alterations in the fibroblast growth factor receptor 1/2 (*FGFR1/2*), protein tyrosine phosphatase non-receptor type 11 (*PTPN11*) or neurotrophic receptor tyrosine kinase (*NTRK*)

Gabriel Leprivier, Claudia C. Faria, Marc Remke are Joint senior authors.

Extended author information available on the last page of the article

genes [16]. Based on these molecular findings, individual case observations as well as early clinical trials have been focusing on pharmacological inhibition of MAPK signaling using BRAF and/or MAPK/ERK kinase (MEK) inhibitors as a molecularly guided strategy for targeted therapy (for review see Ruda et al. [37]). Thus far, available data suggest that this targeted strategy may improve outcome of patients with recurrent, refractory or progressive disease. For example, several studies on pediatric patients with *BRAF*- or *NF1*-altered progressive or recurrent low-grade gliomas (LGG), including PAs, reported on promising clinical signals for targeted treatments with BRAF or MEK inhibitors when compared to control patients receiving the standard-of-care [4, 10, 13]. It is also interesting to note that patients who stopped treatment had rapid tumor re-growth [27].

Several molecular profiling studies have been conducted that aimed to stratify PA patients based on large-scale gene expression and/or DNA methylation profiling data [1, 18, 20, 38, 48]. In particular, studies have analyzed epigenetic data to define molecular groups [20, 38]. However, upon clustering of these cohorts, the tumors largely separated according to their anatomical location. Thus, there is no current consensus on subgrouping of PAs based on molecular markers or signatures, indicating a low degree of biological heterogeneity among these tumors. Notably, however, previous molecular profiling studies have been restricted to single layers of molecular data sets, i.e., comprising either gene expression or DNA methylation analyses [20, 32, 51].

In the present study, we employed similarity network fusion (SNF) analysis, which allows for the integration of multiple layers of large-scale molecular data sets [8]. We performed SNF analyses based on RNA sequencing transcriptomic and mass spectrometry (MS)-based proteomic profiling data of a large cohort of PAs to discern the biological heterogeneity of these tumors. As there is a known discordance between mRNA and protein expression [12, 23, 52], it was important to integrate the data to gain a better overview of potential intertumoral biological differences in PAs. Indeed, with pathways such as "Interferon Signaling" and "T Cell Receptor Signaling", these data led us to the discovery that the profile of immune cells, which are part of the tumor microenvironment, may discriminate PAs into two biologically and clinically distinct groups, with Group 1 tumors being more frequently located in the supratentorial compartment, manifesting at younger age and being associated with less favorable progression-free survival.

Materials and methods

Patient samples

Tumor tissue samples from PA patients were obtained from the CNS tumor tissue bank Düsseldorf at the Institute of Neuropathology, University Hospital Düsseldorf, Germany, and from the Hospital de Santa Maria, Centro Hospitalar Universitário Lisboa Norte, in Lisbon, Portugal. Patients or parents provided their written informed consent for the use of the tissue samples for research purposes, in accordance with the requirements of the internal review boards. The study was approved by the Ethics Committee of the Medical Faculty, Heinrich Heine University Düsseldorf (study number: 5604). All samples analyzed in this study were collected from newly diagnosed patients and were flash-frozen directly after surgical resection. Each specimen used for protein and RNA extraction was histologically assessed to assure the presence of cellular tumor tissue with an estimated tumor cell content of > 70%. All tumors were histologically classified as PAs according to the criteria of the World Health Organization (WHO) classification of CNS tumors [21].

Detection of *BRAF* gene alterations

Structural alterations in *BRAF*, i.e., *KIAA1549::BRAF* fusions were demonstrated in the diagnostic setting either by reverse transcription PCR or by in situ hybridization. Briefly, fusions were detected using the primers for the most common fusion products (*KIAA1549::BRAF* exons 15::9 or 16::9) and visualized using gel electrophoresis. Fusions were confirmed using Arriba v2.4.0 algorithm for samples with RNA sequencing [43]. Arriba was run with default settings against the hg38 reference genome with the GENECODE annotation. For *BRAF* V600 missense mutations, droplet digital PCR (ddPCR) was performed as previously published [47].

RNA sequencing

Total RNA was isolated from the fresh frozen PA tissue samples using the Maxwell® RSC simply RNA Tissue Kit (AS1340, Promega, Walldorf, Germany). To prepare the barcoded libraries, 500 ng total RNA was processed using the TruSeq RNA Sample Preparation v2 kit (low-throughput protocol; Illumina, San Diego, CA, USA). Libraries were validated and quantified using either DNA 1000 or high-sensitivity chips on a Bioanalyzer (Agilent, Santa Clara, CA, USA). 7.5 pM denatured libraries were input into cBot (Illumina), followed by deep sequencing using the HiSeq 2500 (Illumina) for 101 cycles, with an additional seven

cycles for index reading. Fastq files were imported into Partek Flow (Partek Incorporated, St. Louis, MO, USA). Quality analysis and quality control were performed on all reads to assess read quality and to determine the amount of trimming required (both ends: 13 bases 5´ and 1 base 3´). Trimmed reads were aligned against the hg38 genome using the STAR v2.4.1d aligner. Unaligned reads were further processed using Bowtie 2 v2.2.5 aligner. Finally, aligned reads were combined before quantifying the expression against the ENSEMBL (release 84) database using the Partek Expectation–Maximization algorithm. Partek Flow default settings were used in all analyses. RNA sequencing data have been deposited in the European Genome–Phenome Archive under the identifier EGAD00001009053 (https://web3.ega-archive.org/).

Mass spectrometry

For MS-based proteome analyses, proteins were extracted from fresh frozen PA tissue. Tissues were homogenized in urea buffer with a TissueLyser (Qiagen, Hilden, Germany) and subsequent sonication. After centrifugation for 15 min at 14,000×*g* and 4 °C, supernatants were collected. Protein concentration was determined via Pierce 660 nm Protein Assay (Thermo Fischer Scientific) and 10 µg protein per sample were desalted through short electrophoretic migration at 50 V for 10 min on a 4–12% Bis–Tris polyacrylamide gel (#EC62352BOX, Novex NuPAGE, Thermo Fischer Scientific). After silver staining, the resulting protein band for each sample was cut out, destained, reduced, alkylated and digested with trypsin before peptide extraction via sonication. Peptides were dissolved and diluted with 0.1% TFA (v/v).

MS-based proteome analyses were performed as previously described [33]. In brief, 15 µL peptide solution per sample were analyzed on a nano-high-performance liquid-chromatography electrospray ionization mass spectrometer. The analytical system was composed of an RSLCnano U3000 HPLC coupled to a QExactive Plus mass spectrometer via a nano-electrospray ion source (Thermo Fischer Scientific). Injected peptides were concentrated and desalted at a flow rate of 6 µL/min using a trapping column (Acclaim PepMao C18, 2 cm × 100 µm × 3 µm particle size, 100 Å pore size, Thermo Fischer Scientific) with 0.1% TFA (v/v) for 10 min. Subsequently, peptides were separated at a constant flowrate of 300 nL/min over a 120 min gradient using an analytical column (Acclaim PepMap RSLC C18, 25 cm × 75 µm × 2 µm particle size, 100 Å pore size, Thermo Fischer Scientific) at 60 °C. Separation was achieved through a gradient from 4% to 40% solvent B [solvent A: 0.1% (v/v) formic acid in water, solvent B: 0.1% (v/v) formic acid, 84% (v/v) acetonitrile in water]. Afterwards, peptides were ionized at a voltage of 1,400 V and introduced into the mass spectrometer operating in positive mode. Mass spectrometry scans were recorded in profile mode at a range from 350 to 2000 *m/z* at a resolution of 70,000, while tandem mass spectra were recorded at a resolution of 17,500. Tandem mass spectra were recorded with a data-dependent Top10 method and 30% normalized collision energy. Dynamic exclusion was activated with a repeat count of 1 for 100 s.

Proteome Discoverer (version 1.4.1.14, Thermo Fisher Scientific) was applied for peptide/protein identification using Mascot (version 2.4, Matrix Science, London, UK) as a search engine employing the UniProt database (human; including isoforms; date 2016–03-01). A false discovery rate of 1% ($p \leq 0.01$) at the peptide level was set as the identification threshold. Proteins were quantified with Progenesis QI for Proteomics (Version 2.0, Nonlinear Dynamics, Waters Corporation, Newcastle upon Tyne, UK). The mass spectrometry proteomics data have been deposited with the ProteomeXchange Consortium via the PRIDE partner repository (https://www.ebi.ac.uk/pride/) with the data set identifier PXD035773.

DNA methylation profiling

Global DNA methylation data of 52 samples presented in this study were generated using tumor DNA extracted either from formalin-fixed paraffin-embedded tissue samples (FFPE, 32 tumors) or from flash-frozen tissue samples (FF, 20 tumors). Tumor DNA was hybridized to Illumina Infinium EPIC Methylation BeadChip Arrays. Methylation profiling was performed according to the manufacturer's instructions at the DKFZ Genomics and Proteomics Core Facility (Heidelberg, Germany). All analyses were performed in Partek Genomic Suite (Partek Incorporated, St. Louis, MO, USA). FFPE and FF samples were processed individually and then combined following beta-value determination. The complete CpG methylation values have been deposited in NCBI's GEO under accession number GSE210353. Normalization and generation of beta values were performed after NOOB background normalization. DNA methylation analysis using the CNS tumor methylation profiling classifier [7] confirmed the diagnosis of PA in 44 patients. In the remaining 8 patients, DNA methylation analysis revealed 4 control tissue samples, 3 samples with no matching methylation class and 1 sonic hedgehog medulloblastoma (histologically a PA with a BRAF-fusion).

Similarity network fusion

This method has been described by Wang et al. [45]. Briefly, patient similarity matrices were constructed for each data type using Euclidean distance on samples that shared collected data for mRNA expression (48 samples, 13,498 features), proteome expression (43 samples, 2457 features)

and methylation beta value (52 samples, 865,860 features) data sets. SNF was performed using 28 samples overlapping between mRNA and protein expression, 25 samples overlapping between mRNA and protein expression and methylation beta values, or individual data sets. SNF was run setting the number of nearest sample neighbors $K = 10$, the hyperparameter alpha = 0.5 and the number of iterations for the diffusion process $T = 10$. To obtain network clusters, spectral clustering was performed on networks representing each of the data types independently, as well as on the fused network to which the SNF process had converged. Analysis was visualized with Cytoscape (www.cytoscape.org) using the minimum number of entries that contained all samples based on the highest degree of relatedness.

Hierarchical clustering and group extension

SNF group extension was conducted using genes with significant differential expression ($p \leq 0.05$ and fold change ± 2) ranked by p value. The top 100, 50 and 25 up- and down-differentially regulated genes and proteins were used to generate signatures (Supplementary Tables 2 and 3). Gene and protein signatures were visualized using hierarchical clustering (HCL) after normalizing mean expression to 0 with a standard deviation of 1 and using Pearson's dissimilarity algorithm and average linkage in Partek Genomics Suite. HCL was first performed with SNF-overlapping samples and then with all samples. Signatures with the lowest number of genes capable of maintaining accurate group associations (50 genes or proteins) were used for subsequent analyses.

Submap analysis

Submap analysis was used to determine similarities between different data sets and was performed using the GenePattern analysis platform (https://cloud.genepattern.org/gp/pages/index.jsf). Data set files were ordered based on groups and class files provided group information. Defaults were used, except for the number of genes used to determine similarities (10,000 markers for transcriptome analysis or 2000 markers for proteome analysis and subclass association (SA) matrix was adjusted using the False Discovery Rate (FDR)).

Calculating correlation and ratio between mRNA and protein expression

For each gene product, we calculated Pearson correlation coefficients between its normalized, centered, and log-transformed transcript and protein levels. The statistical significance of the correlation was assessed by p value. In addition, the divergence of protein and mRNA expression was measured by the protein/mRNA ratio. For all 2107 products, protein/mRNA ratios within each group were computed by dividing the median of the log protein and mRNA levels.

Pathway analysis

Ingenuity pathway analysis (IPA, Qiagen) was conducted using genes with significant differential expression ($p \leq 0.05$ and fold change ± 2). The significance cutoff for IPA was set to $p \leq 0.05$ and an activation z score of ± 1.5. In addition, for upstream regulators, we filtered out biological drugs, all chemicals and miRNA entries.

Gene Set Enrichment Analysis (GSEA) was performed using the t values from the unpaired t tests for both mRNA and protein expression data. Gene sets were comprised of curated pathways from several databases, including GO, Reactome, KEGG (April_01_2019 version; http://download.baderlab.org/EM_Genesets/current_release/ Human/symbol/), and visualized using Cytoscape (www.cytoscape.org; main figure: $p \leq 0.0005$, $q \leq 0.03$, similarity cutoff 0.5; supplementary figure: $p \leq 0.001$, $q \leq 0.05$, similarity cutoff 0.5).

Deconvolution analyses

ESTIMATE was carried out in R (version 4.0) using default parameters [50]. Briefly, data files were loaded and processed using the estimate package, identifiers were gene symbols and platforms were "illumina" for the RNA sequencing data sets and "affymetrix" for the microarray. Data were then visualized using GraphPad Prism (version 5.0) (https://www.graphpad.com/scientific-software/prism. Single-cell RNA sequencing data signatures were generated by Reitman et al. [34] and imported into CIBERSORT (https://cibersort.stanford.edu/) as a "signature matrix". CIBERSORT was performed for each data set using default settings. Data were visualized using GraphPad Prism.

Multiplex immunofluorescence

FFPE sections of PA patients were stained using Opal Polaris 7 colour kit (NEL861001KT, Akoya Biosciences, Inc.) based on thyramide signal amplification fluorescent immunohistochemistry. The staining targeting anti-human CD4 (1:50, MA5-16,338, Thermo Fisher Scientific), CD8 (1:150, M7103, DAKO), PD-1 (1:300, ab137132, Abcam), FoxP3 (1:200, DIA-FX3, Dianova), Iba-1 (1:450, 019-19741, WAKO) and vWF (1:120, A0082, DAKO) was performed on LabSat™ Research Automated Staining Instrument (Lunaphore Technologies SA). Whole slide multispectral scans were acquired at 0,5 µm/pixel on Vectra Polaris Imaging System using MOTiF™ technology (Akoya Biosciences, Inc.) and analyzed using HALO™ image analysis software (Indica Labs).

Bioinformatic and statistical analyses

For validation purposes, ICGC data of 73 PAs [17] were downloaded from the European Genome–Phenome Archive (https://www.ebi.ac.uk/ega/datasets/EGAD00001000617) and processed in the same way as described for the discovery cohort. The processed and $\log_2$-transformed validation data from a further cohort of 191 PAs published by Kool et al. [5] were downloaded from the R2 Genomics Analysis and Visualization Platform (https://hgserver1.amc.nl/cgi-bin/r2/main.cgi). PA log ratio proteomic profiling data ($n=39$) from PDC000180 was downloaded directly from the CPTAC data portal (https://cptac-data-portal.georgetown.edu/cptacPublic/). Statistical analyses were performed using Partek Genomic Suite or GraphPad Prism. *T* tests or Mann–Whitney tests (non-parametric *t* tests) were used for comparisons between two groups for statistical analysis. χ^2 tests were performed to analyze clinicopathological traits. Differences between groups were considered statistically significant at $p<0.05$. Kaplan Meier progression-free survival analyses were calculated using the log-rank method and multivariate analysis was calculated using the Cox regression method.

Results

Integrative multi-omic analysis of PA tissue samples

Our cohort consisted of a total of 62 flash-frozen primary PA tissue samples that were annotated with various clinical features (see clinical information in Supplementary Table 1). We employed an integrative multi-omics approach and performed DNA methylation, transcriptomic and proteomic measurements on 52, 48 and 43 partially overlapping samples, respectively (Supplementary Fig. 1). For proteomic analysis, we retained only proteins with at least three detected peptide ratios and with no missing values. Together, these stringent criteria led to the unambiguous quantification of 2456 proteins.

To uncover potential intertumoral heterogeneity in the investigated PA samples, we integrated transcriptomic and proteomic data sets (constituting 28 overlapping samples) using SNF [8, 12]. Strikingly, this efficiently identified two distinct tumor clusters that were designated as Group 1 and Group 2 (Fig. 1, Supplementary Fig. 2a). By contrast, separated SNF-based integrative clustering of each of the single omic layers, transcriptomics or proteomics, only poorly segregated groups (Supplementary Fig. 2b, c), highlighting the importance of combining multiple omics data. Moreover, integrating DNA methylation data to transcriptomic and proteomic data (for a total of 25 overlapping samples) did not further refine the SNF-based identification of patient groups (Supplementary Fig. 2d, e). On the contrary, this disrupted the initial group segregation obtained by transcriptomics and proteomics (Supplementary Fig. 2d, e, as compared to Fig. 1 and Supplementary Fig. 2a), arguing against the discriminatory power of DNA methylation to discern PA groups. Together, our integrative analyses of transcriptomics and proteomics revealed two PA groups.

Validation of PA classification and clinical features of the PA groups

To further extend our findings to the non-overlapping samples of our cohort and validate these in independent cohorts, we generated a gene and a protein signature capable of distinguishing between the two identified PA groups, as calculated by SNF clustering. These signatures were determined by extracting the most differentially expressed genes or proteins between Group 1 and Group 2, Resulting

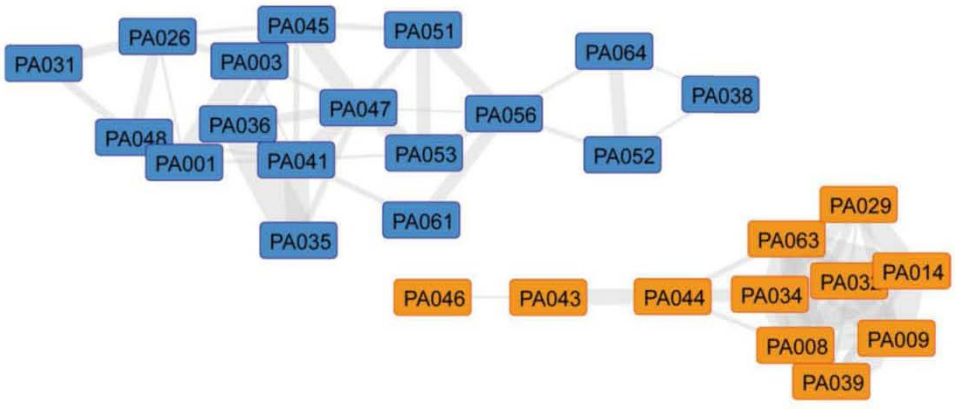
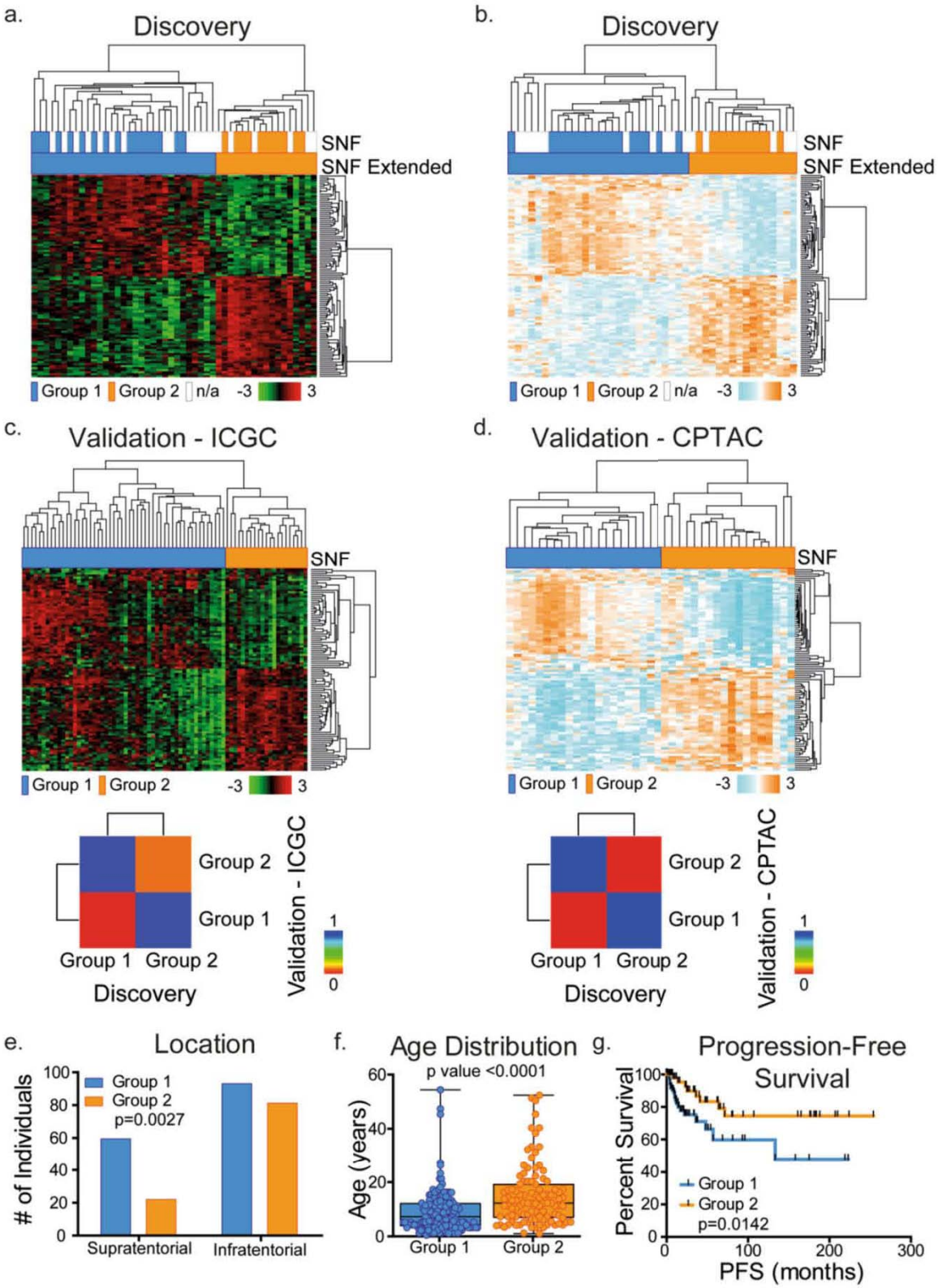


Fig. 1 Similarity network fusion identifies two pilocytic astrocytoma groups using integrative multi-omics data. Similarity network fusion (SNF) representation clearly segregates two PA groups. Shorter edge length and greater thickness between samples (nodes) indicate more similarity


a.
Discovery
SNF
SNF Extended
Group 1 Group 2 n/a -3 3
b.
Discovery
SNF
SNF Extended
Group 1 Group 2 n/a -3 3
c.
Validation - ICGC
SNF
Group 1 Group 2 -3 3
Group 2
Group 1
Group 1 Group 2
Discovery
Validation - ICGC
1
0
d.
Validation - CPTAC
SNF
Group 1 Group 2 -3 3
Group 2
Group 1
Group 1 Group 2
Discovery
Validation - CPTAC
1
0
e.
Location
# of Individuals
Group 1
Group 2
p=0.0027
Supratentorial
Infratentorial
f.
Age Distribution
p value <0.0001
Age (years)
Group 1
Group 2
g.
Progression-Free Survival
Percent Survival
Group 1
Group 2
p=0.0142
PFS (months)

◂**Fig. 2** SNF groups are recapitulated in an extended PA and PA validation cohorts. **a**, **b** Hierarchical clustering of 100 gene/protein signatures based on *p* value allows for the expansion of groups to non-overlapping RNA sequencing (**a**) and mass spectrometry (**b**) samples. **c**, **d** 100-gene/protein signature applied to non-overlapping transcriptomic (ICGC) [17], **c** and proteomic (CPTAC) [31], **d** validation cohorts segregate samples into two groups. Lower panel, submap analyses show close relatedness between discovery and both validation cohorts. **e**–**g** Combined data set analysis of clinical features shows that the majority of younger patients belong to Group 1 and adults belong to Group 2. ($p < 0.0001$, Mann Whitney test, **e**), location shows an enrichment of infratentorial regions for Group 2 ($p = 0.0024$, Fisher's Exact test, **f**), and Kaplan–Meier plot shows patients in Group 1 are more likely to develop recurrent tumors ($p = 0.0142$, log rank method, **g**)

in 100 validated gene and 100 validated protein signatures (see Supplementary Tables 2 and 3). These allowed us to recapitulate our PA classification using only single omics data. We first assessed the validity of the PA classification using 20 additional PA samples that were subjected to RNA sequencing and 15 additional PA samples profiled by proteomic analysis from the same original cohort, but that had not been used in the integrative multi-omics analysis. Semi-unsupervised clustering analysis of transcriptomic data of this extended PA cohort (48 samples total) segregated the same groups as originally identified (Fig. 2a), further reinforcing these PA groups. Applying the same type of analysis to proteomic data of the extended cohort similarly led to the identification of the same PA groups (Fig. 2b), showing that using either a gene or a protein signature alone was sufficient to stratify PAs in an extended patient cohort. We next determined whether the patient stratification could be validated in non-overlapping PA cohorts. To this aim, we used RNA expression data from the International Cancer Genome Consortium (ICGC) ($n = 73$) [17] and Kool et al. ($n = 191$) [5] cohorts and protein expression from the Clinical Proteomic Tumor Analysis Consortium (CPTAC) ($n = 65$) [31]. A semi-unsupervised clustering analysis of these data using the same 100-gene or protein signature was conducted which led to the segregation of two groups in all cohorts (Fig. 2c, d, Supplementary Fig. 3). These were highly similar to the ones originally defined in our discovery cohort as Group 1 and Group 2, as measured by submap analysis (Fig. 2c, d, Supplementary Fig. 3).

After combining all four PA cohorts for a total of 365 PA samples, differences in relevant clinical parameters were analyzed between Group 1 and Group 2. This did not reveal any differences in gender distribution, *BRAF* mutation rate or *KIAA1549::BRAF* fusion occurrence between the two groups (Supplementary Fig. 4a, b). However, tumor location was different between groups, with more tumors in Group 1 being located in the supratentorial compartment, as compared to the preference of infratentorial tumors in Group 2 (*p* value = 0.0027; Fig. 2e). A more detailed analysis revealed that Group 1 tumors were either exclusively observed in optic pathway system and that Group 2 tumors were preferentially observed in the posterior fossa (*p* value = 0.0011; Supplementary Fig. 4c). In addition, the age of patients was significantly different between the two groups, with Group 1 patients being younger than Group 2 patients (mean ages were 7 years versus 12 years, respectively; *p* value < 0.0001; Fig. 2f), potentially corresponding to a younger and an older PA group. Remarkably, Group 1 patients exhibited reduced progression-free survival compared to Group 2 patients, highlighting the clinical significance of the identified groups (*p* value = 0.0142; Fig. 2g). This association was independent of age or tumor location, as determined by multivariate analysis (Supplementary Fig. 4d). Indeed, when we analyzed progression-free survival differences of the two subgroups, we observed no significant survival differences depending on the detailed tumor location. (Supplementary Fig. 5). Altogether, our analyses in additional cohorts confirmed the identification of two PA groups, and highlight that these likely discriminate younger versus older PA patients with distinct progression-free survival.

Posttranscriptional regulation and pathway analysis of PA groups

Given that transcript and protein expression levels are poorly correlated in primary tumors [12, 23, 52], we assessed their level of correlation in PAs and checked whether this level is group-specific. To this end, we computed pairwise Pearson test correlations for 2102 matched mRNA–protein pairs extracted from the original 28 samples of our discovery cohort. In line with previous studies, we uncovered a median Pearson correlation coefficient of 0.168, suggesting the occurrence of posttranscriptional mechanisms regulating gene expression in PA. To further explore such discrepancies between transcript and protein levels in this disease, we calculated the ratio of relative expression of protein and mRNA for each individual pair in each group. We observed that Group 1 and Group 2 displayed remarkably distinct distributions of such a ratio. While the protein/mRNA ratio is mainly distributed towards transcript expression in Group 1, it is the opposite in Group 2, for which protein expression is prevalent (Fig. 3a). This highlights that posttranscriptional regulation is group-related, suggesting different modes of control of gene expression in Group 1 versus Group 2 PA.

To address the biological heterogeneity of PAs, we analyzed biological pathways that are active in each of the two groups. This was achieved by performing two types of complementary analyses: Ingenuity pathways analysis (IPA) and gene set enrichment analysis (GSEA). Using our transcriptomic data, we uncovered that Group 1 and Group 2 are characterized by distinct biological pathways. In particular, IPA showed that "interferon signaling", "dendritic cell

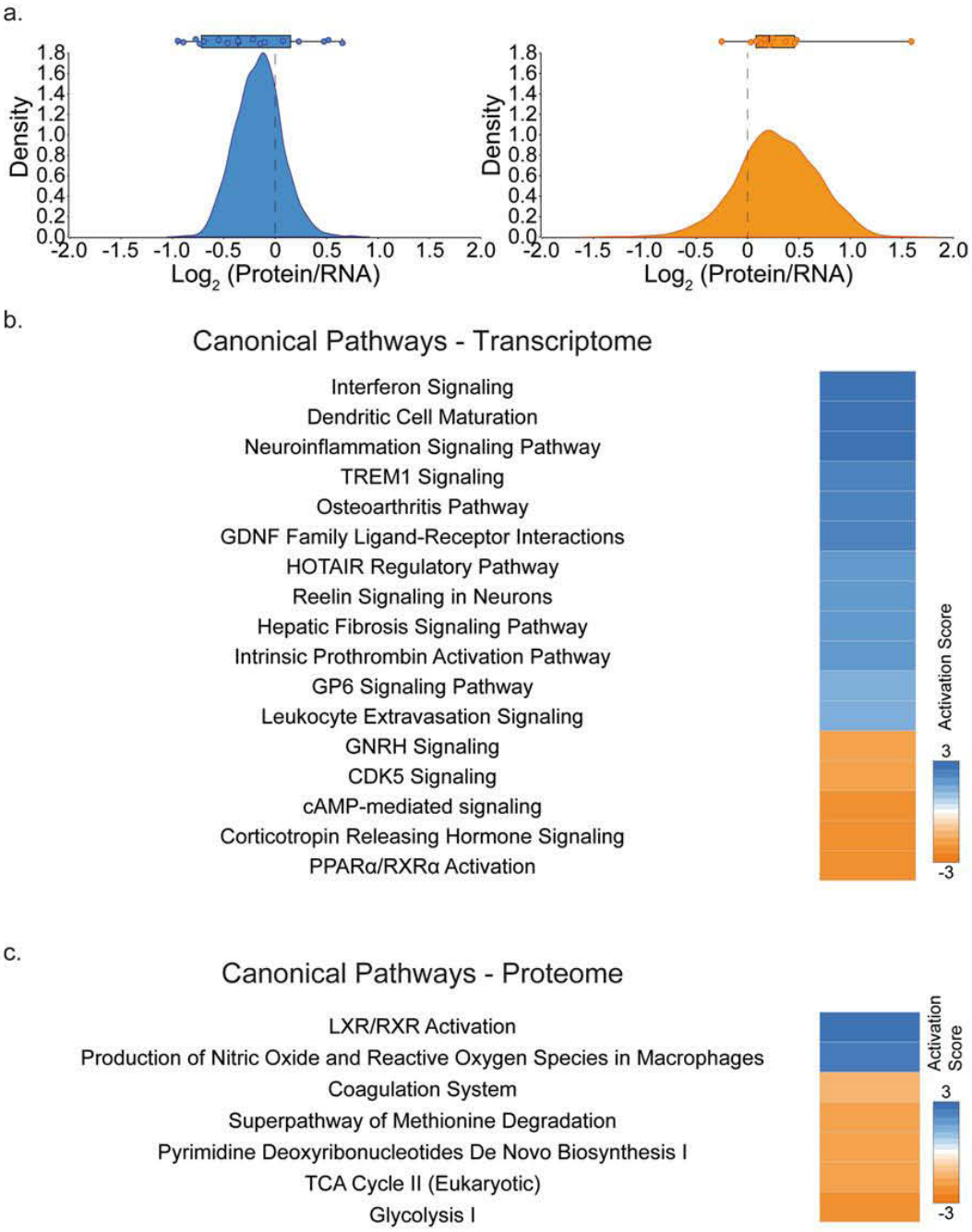


Fig. 3 Ingenuity pathway analysis identifies differential canonical pathway activation between PA groups. **a** Distribution of protein/mRNA ratios in PA groups. Both groups display a non-centric ratio distribution (Mann Whitney test; p value = 0.0144), with an imbalance in favor of mRNA for Group 1 and protein for Group 2. Dotplots show the median protein/mRNA ratios for individual samples. **b**, **c** Group 1 PA were compared to Group 2 PA for both transcriptome (**b**) and proteome (**c**). Significant genes (fold change ± 2 and False Discovery rate q value < 0.05) were processed using IPA and significantly activated canonical pathways are shown. (Activation z-score ± 2, $-\log_{10}$ (p value) > 1.30.)

maturation" and "neuroinflammation signaling pathway" are specifically active in Group 1, while "gonadotropin-releasing hormone (GNRH)" and "cyclin-dependent kinase 5 (CDK5) signaling" are active in Group 2 (Fig. 3b and Supplementary Table 4). In keeping with the IPA, GSEA revealed an overrepresentation of immune response pathways (Fig. 4a and Supplementary Tables 6–7)—in particular of "Interferon Signaling" (top term, Fig. 4b and Supplementary Table 6) and "T Cell Receptor Signaling"—in Group 1. This was confirmed in the ICGC [17] and Kool et al. [5] cohorts (Supplementary Fig. 6), further reinforcing the importance of such pathways in this group. In addition, Group 1 was enriched for cell cycle and RNA processing-related pathways (Fig. 4a), while Group 2 was characterized by enrichment for action potential and neurotransmitter signaling pathways, as determined by GSEA (Fig. 4a). Supporting the discovery analysis, enrichment for action potential was confirmed in Group 2 using the ICGC and Kool et al. cohorts (Supplementary Fig. 6).

Notably, Group 1 was enriched for immune response pathways and, in particular, for "Interferon Signaling", as well as "RNA Processing Pathways" and "Post-Translational Modification" (Fig. 4a) using GSEA of our proteomic data. However, "TCA Cycle II (Eukaryotic)" and "Glycolysis I", among others, were pathways observed in the IPA using the proteomic data of Group 2 samples (Fig. 3c and Supplementary Table 5). These observations were verified using the discovery and validation—CPTAC GSEA. Specifically, cellular respiration (Fig. 4a and Supplementary Table 7) and "Oxidative Phosphorylation", the top gene set (Fig. 4c) in the proteomic analysis, were verified using the validation proteomic data set (Supplementary Fig. 7). Therefore, integration of multiple data layers is required to observe the biological heterogeneity of PA.

Distinct immune cell signatures in PA groups

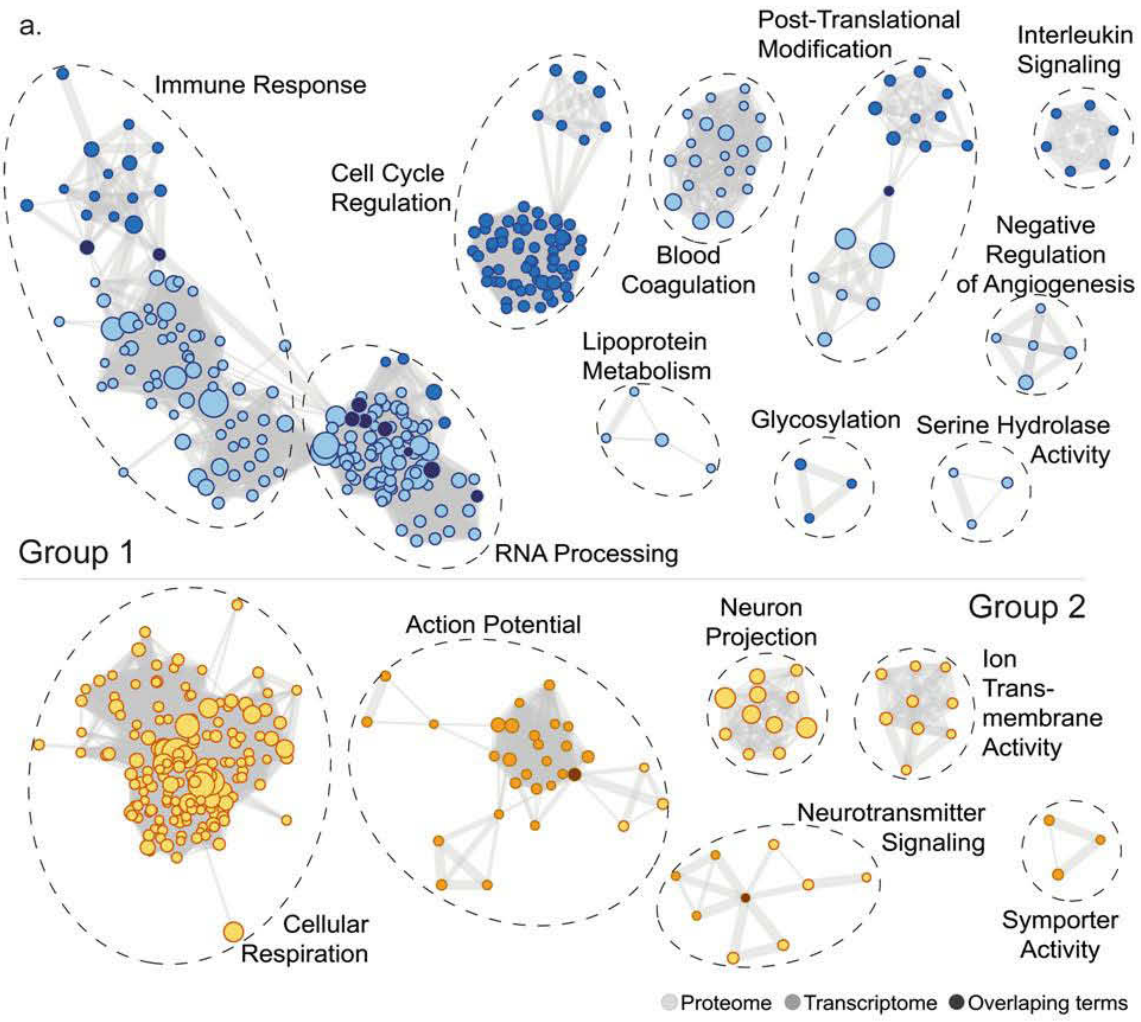
To further investigate our observations of increased immune response gene signature in Group 1, we evaluated the level of immune cell infiltration in Group 1 versus Group 2 samples using the algorithm "Estimation of STromal and Immune cells in MAlignant Tumours using Expression data" (ESTIMATE) [50]. ESTIMATE scores are used to determine tumor purity; however, the ESTIMATE algorithm can also assess the presence of stromal and infiltrating immune cells. Both RNA sequencing data sets (Discovery and Validation—ICGC) showed a significantly higher immune score for Group 1 compared to Group 2, which was not evident in the microarray validation set (Validation-Kool et al. [5]) (Supplementary Fig. 8), possibly reflecting the limited resolution of microarray-based expression profiling compared to RNA-sequencing. Since Reitman et al. [34] showed that microglia, macrophages and T cells are present along with the PA tumor cells using single cell RNA sequencing, we used this information to create a signature file for each of these four cell types. By processing and analyzing our discovery, Validation-ICGC [17] and Validation-Kool et al. [5] data sets using CIBERSORT [26], we uncovered that, in the RNA sequencing data sets (Discovery and Validation—ICGC), T cells were specifically enriched in Group 1 compared to Group 2, while there was no difference in microglia or macrophages (Supplementary Fig. 9).

To confirm the previous results, we performed multiplex immunofluorescence on a subset of cases ($n = 24$). CD4, CD8 or IBA-1 positive cells were similarly represented between PA Groups 1 and 2 (Supplementary Fig. 10a–c). When we analyzed the immunosuppressive T cell populations, although $FOXP3^{+}CD4^{+}$ cells were not different between the two groups, exhausted T cells ($PD1^{+}CD8^{+}$) cells showed a strong trend towards being able to discriminate Group 1 from Group 2, Group 2 having a higher percentage of PD1 positive $CD8^{+}$ T cells (Supplementary Fig. 10d–f).

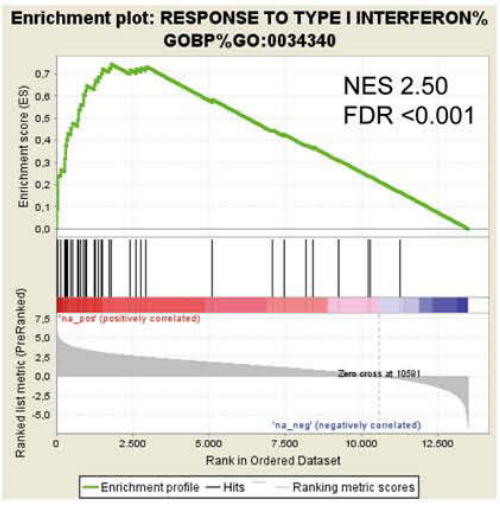
Discussion

Genome-wide profiling and next-generation-based sequencing approaches have provided profound insights into the pathomechanisms underlying PA development. In particular, these methods have revealed that aberrant activation of the MAPK pathway constitutes a hallmark feature of these tumors [6]. Oncogenic activation of MAPK is caused by alterations affecting *BRAF*, *NF1*, *FGFR1* or, rarely, other MAPK pathway genes [17, 24, 32]. However, further biological stratification of this disease has not evolved in the past decade, in contrast to other brain tumor entities in children and adults [8, 30, 41, 44]. Only a limited number of studies have suggested intertumoral heterogeneity of the disease as identified by distinct transcriptomes [20, 39, 48], DNA methylomes [20, 34, 36, 51], ploidy [11] or DNA copy number alterations [14, 32]. Importantly, none of these studies reported reproducible features of the identified subgroups. Furthermore, high-resolution DNA methylation profiles used for the molecular neuropathology classifier [7] are very accurate in dissecting molecular entities, with the important exception of LGG, including PA. At present, the algorithm identifies PA predominantly based on anatomic location. Therefore, currently available approaches fail to differentiate clinically relevant groups of the disease, suggesting either that such PA groups do not exist or that single layer omics approaches insufficiently discriminate the biological heterogeneity of the disease.

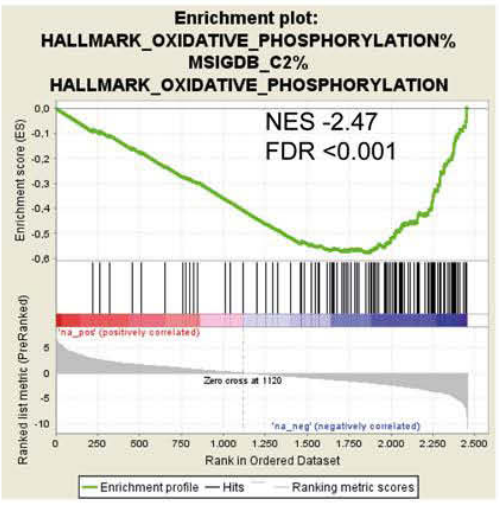
Thus, we decided to apply an innovative, integrative multi-omics approach, which has already provided fundamental insights into the tumor biology of breast cancer [2,


a.
Immune Response
Cell Cycle Regulation
Post-Translational Modification
Interleukin Signaling
Blood Coagulation
Negative Regulation of Angiogenesis
Lipoprotein Metabolism
Glycosylation
Serine Hydrolase Activity
RNA Processing
Group 1
Action Potential
Neuron Projection
Group 2
Ion Trans-membrane Activity
Neurotransmitter Signaling
Cellular Respiration
Symporter Activity
Proteome
Transcriptome
Overlaping terms

b.


Enrichment plot: RESPONSE TO TYPE I INTERFERON% GOBP%GO:0034340
NES 2.50
FDR <0.001
Enrichment score (ES)
Ranked list metric (PreRanked)
'na_pos' (positively correlated)
Zero cross at 10591
'na_neg' (negatively correlated)
Rank in Ordered Dataset
Enrichment profile
Hits
Ranking metric scores

c.


Enrichment plot: HALLMARK_OXIDATIVE_PHOSPHORYLATION% MSIGDB_C2% HALLMARK_OXIDATIVE_PHOSPHORYLATION
NES -2.47
FDR <0.001
Enrichment score (ES)
Ranked list metric (PreRanked)
'na_pos' (positively correlated)
Zero cross at 1120
'na_neg' (negatively correlated)
Rank in Ordered Dataset
Enrichment profile
Hits
Ranking metric scores

◂**Fig. 4** Integrative multi-omics identifies highly divergent pathways. **a** GSEA-based enrichment map representations based on groups (p value = 0; FDR < 0.03). Nodes (circles) represent enriched pathways identified on the basis of proteome or transcriptome expression, which are light or medium colored, respectively, and dark nodes represent overlapping terms between transcriptome and proteome expression. Edges connect pathways/nodes that share at least half of the terms defining them. Nodes grouped according to functional families are indicated on each network. **b**, **c** Enrichment plots of top transcriptomic gene set for Group 1 showing Response to Type I Interferon (**b**) and top proteomic gene set for Group 2 showing Hallmark Oxidative Phosphorylation (**c**)

19], pancreatic ductal adenocarcinoma [6], glioblastoma [46], and medulloblastoma [3, 12], among other entities. Our bioinformatics approach integrates proteomic, transcriptomic, epigenomic and mutational profiles. Subsequently, SNF provided compelling evidence for the existence of two core PA groups using combined proteomic and transcriptomic data in an institutional discovery cohort. Since integrative multi-omics classification requires generating multiple data sets, which is difficult in the clinical setting due to sample quality and/or financial constraints, we established highly accurate classification approaches using only RNA- or protein-based signatures. Notably, by applying this strategy to three independent non-overlapping validation cohorts, we confirmed the existence of the two core PA groups.

Our integrative multi-omics data show highly distinct pathway enrichment according to PA group. Most importantly, immune response and associated pathways, including "Interferon signaling", "Antigen Processing and Presentation", "Cellular Response to Tumor Necrosis Factor", and "T Cell Receptor Signaling Pathway" (Fig. 3, Supplementary Table 6), were significantly overrepresented in Group 1. Notably, this association could be confirmed using both transcriptomic validation cohorts. Furthermore, we uncovered an enrichment for the T cell gene signature in Group 1, likely pointing to a higher infiltration of T cells in Group 1 versus Group 2 PA. This is in agreement with previous reports that detected T cells within PA tumor tissues using histology [35] or single-cell RNA sequencing [34] while being unable to highlight intertumoral heterogeneity for this parameter. PD1 is a marker of T cell exhaustion [25] and our data suggest an exhausted phenotype of $CD8^+$ T cells in Group 2 which is consistent with the lack of observed immune activation at the transcriptomic or proteomic levels in this group. Consistent with our bioinformatics analyses, we did not observe differences in other immunosuppressive populations including regulatory T cells and macrophages.

In our proteomic analysis, we were also able to confirm that Group 2 had greater enrichment of gene sets involved in Cellular Respiration, such as "Oxydative Phosphorylation", "Mitochondrial Respiratory Chain" and "The Citric Acid (TCA) Cycle and Respiratory Electron Transport" using our discovery and validation cohorts. These data are in line with high rate of mitochondrial mutations detected by Leuth et al., where they observed that 53% of the mutations in PA tumors had mutations in genes involved in oxidative phosphorylation pathways [22]. The dysregulated oxidative phosphorylation pathway may enhance reactive oxygen species accumulation and lead to either an increase in proliferation rate or a decrease in apoptotic activity thereby potentially enhancing tumor growth.

Furthermore, "RNA processing" and associated pathways were remarkably divergent between the two PA groups. Notably, we demonstrated that, in general, pathway regulation was predominantly driven by RNA signatures in Group 1, while proteomic-based pathway regulation was significantly increased in Group 2, based on our integrative multi-omic discovery cohort. In addition, "RNA processing" and associated pathways were among the most consistently affected pathways in both transcriptomic validation cohorts.

Finally, we were able to delineate distinct clinical features according to the PA groups using the combined discovery and validation data sets. Group 1 tumors were evenly distributed between the supra- and infratentorial compartments, while Group 2 PAs were more commonly located in the infratentorial region. In addition, age distribution differs significantly, as patients with Group 1 PAs were significantly younger than those with tumors in Group 2. Finally, it has previously been reported that infratentorial tumors are associated with better progression-free survival [38]. While we found a difference in progression-free survival between Group 1 and Group 2, this was independent of tumor location. Importantly, we observed improved progression-free survival in Group 2, which was unexpected, because adults with PA have been reported to have worse prognoses than children [15, 42, 49]. Taken together, our data thus provide novel insights into the biological heterogeneity of PA, which may allow for more accurate biological disease stratification.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00401-023-02626-5.

Acknowledgements We would like to acknowledge Professor Arndt Borkhardt for providing infrastructure for this project and Tatjana Starzetz for technical help. This project funded in part by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project-ID 418179183—KFO 337, RO 3577/7-1 (A.R.), RE 2857/4-1 (M.R.) and supported by Biomed Valley and Day One Therapeutics (T.M.), Fundação Amélia de Melo (C.C.F.) and Fundação Millennium bcp (C.C.F.) Finally, we acknowledge the Biobanco-iMM CAML who enabled the process of tumor specimen collection, processing, and storage.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability The data that support the findings in this study have been deposited in separate repositories. RNA sequencing data has been deposited in the European Genome-phenome Archive under the identifier EGAD00001009053 (https://web3.ega-archive.org/). The mass

spectrometry proteomics data has been deposited with the ProteomeXchange Consortium via the PRIDE partner repository (https://www.ebi.ac.uk/pride/) with the data set identifier PXD035773. The complete CpG methylation values have been deposited in NCBI's GEO under accession number GSE210353.

References

1. Antonelli M, Fadda A, Loi E, Moi L, Zavattari C, Sulas P et al (2018) Integrated DNA methylation analysis identifies topographical and tumoral biomarkers in pilocytic astrocytomas. Oncotarget 9:13807–13821. https://doi.org/10.18632/oncotarget.24480
2. Anurag M, Jaehnig EJ, Krug K, Lei JT, Bergstrom EJ, Kim BJ et al (2022) Proteogenomic markers of chemotherapy resistance and response in triple-negative breast cancer. Cancer Discov 12:2586–2605. https://doi.org/10.1158/2159-8290.CD-22-0200
3. Archer TC, Ehrenberger T, Mundt F, Gold MP, Krug K, Mah CK et al (2018) Proteomics, Post-translational Modifications, and Integrative Analyses Reveal Molecular Heterogeneity within Medulloblastoma Subgroups. Cancer Cell 34:396–410398. https://doi.org/10.1016/j.ccell.2018.08.004
4. Bouffet E, Geoerger B, Moertel C, Whitlock JA, Aerts I, Hargrave D et al (2023) Efficacy and safety of trametinib monotherapy or in combination with dabrafenib in pediatric BRAF V600-mutant low-grade glioma. J Clin Oncol 41:664–674. https://doi.org/10.1200/JCO.22.01000
5. Buhl JL, Selt F, Hielscher T, Guiho R, Ecker J, Sahm F et al (2019) The senescence-associated secretory phenotype mediates oncogene-induced senescence in pediatric pilocytic astrocytoma. Clin Cancer Res 25:1851–1866. https://doi.org/10.1158/1078-0432.CCR-18-1965
6. Cao L, Huang C, Cui Zhou D, Hu Y, Lih TM, Savage SR et al (2021) Proteogenomic characterization of pancreatic ductal adenocarcinoma. Cell 184:5031-5052e5026. https://doi.org/10.1016/j.cell.2021.08.023
7. Capper D, Jones DTW, Sill M, Hovestadt V, Schrimpf D, Sturm D et al (2018) DNA methylation-based classification of central nervous system tumours. Nature 555:469–474. https://doi.org/10.1038/nature26000
8. Cavalli FMG, Remke M, Rampasek L, Peacock J, Shih DJH, Luu B et al (2017) Intertumoral heterogeneity within medulloblastoma subgroups. Cancer Cell 31:737–754736. https://doi.org/10.1016/j.ccell.2017.05.005
9. Collins KL, Pollack IF (2020) Pediatric low-grade gliomas. Cancers (Basel). https://doi.org/10.3390/cancers12051152
10. Fangusaro J, Onar-Thomas A, Young Poussaint T, Wu S, Ligon AH, Lindeman N 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. https://doi.org/10.1016/S1470-2045(19)30277-3
11. Fontebasso AM, Shirinian M, Khuong-Quang DA, Bechet D, Gayden T, Kool M et al (2015) Non-random aneuploidy specifies subgroups of pilocytic astrocytoma and correlates with older age. Oncotarget 6:31844–31856. https://doi.org/10.18632/oncotarget.5571
12. Forget A, Martignetti L, Puget S, Calzone L, Brabetz S, Picard D et al (2018) Aberrant ERBB4-SRC signaling as a hallmark of group 4 medulloblastoma revealed by integrative phosphoproteomic profiling. Cancer Cell 34:379-395e377. https://doi.org/10.1016/j.ccell.2018.08.002
13. Hargrave DR, Bouffet E, Tabori U, Broniscer A, Cohen KJ, Hansford JR 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. https://doi.org/10.1158/1078-0432.CCR-19-2177
14. Jacob K, Albrecht S, Sollier C, Faury D, Sader E, Montpetit A et al (2009) Duplication of 7q34 is specific to juvenile pilocytic astrocytomas and a hallmark of cerebellar and optic pathway tumours. Br J Cancer 101:722–733. https://doi.org/10.1038/sj.bjc.6605179
15. Johnson DR, Brown PD, Galanis E, Hammack JE (2012) Pilocytic astrocytoma survival in adults: analysis of the Surveillance, Epidemiology, and End Results Program of the National Cancer Institute. J Neurooncol 108:187–193. https://doi.org/10.1007/s11060-012-0829-0
16. Jones DT, Gronych J, Lichter P, Witt O, Pfister SM (2012) MAPK pathway activation in pilocytic astrocytoma. Cell Mol Life Sci 69:1799–1811. https://doi.org/10.1007/s00018-011-0898-9
17. Jones DT, Hutter B, Jager N, Korshunov A, Kool M, Warnatz HJ et al (2013) Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. Nat Genet 45:927–932. https://doi.org/10.1038/ng.2682
18. Kim JW, Phi JH, Kim SK, Lee JH, Park SH, Won JK et al (2023) Comparison of the clinical features and treatment outcomes of pilocytic astrocytoma in pediatric and adult patients. Childs Nerv Syst. https://doi.org/10.1007/s00381-023-05839-x
19. Krug K, Jaehnig EJ, Satpathy S, Blumenberg L, Karpova A, Anurag M et al (2020) Proteogenomic landscape of breast cancer tumorigenesis and targeted therapy. Cell 183:1436-1456e1431. https://doi.org/10.1016/j.cell.2020.10.036
20. Lambert SR, Witt H, Hovestadt V, Zucknick M, Kool M, Pearson DM et al (2013) Differential expression and methylation of brain developmental genes define location-specific subsets of pilocytic astrocytoma. Acta Neuropathol 126:291–301. https://doi.org/10.1007/s00401-013-1124-7
21. Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D et al (2021) The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. Neuro Oncol 23:1231–1251. https://doi.org/10.1093/neuonc/noab106
22. Lueth M, Wronski L, Giese A, Kirschner-Schwabe R, Pietsch T, von Deimling A et al (2009) Somatic mitochondrial mutations in pilocytic astrocytoma. Cancer Genet Cytogenet 192:30–35. https://doi.org/10.1016/j.cancergencyto.2009.03.002
23. Mertins P, Mani DR, Ruggles KV, Gillette MA, Clauser KR, Wang P et al (2016) Proteogenomics connects somatic mutations to signalling in breast cancer. Nature 534:55–62. https://doi.org/10.1038/nature18003
24. Milde T, Rodriguez FJ, Barnholtz-Sloan JS, Patil N, Eberhart CG, Gutmann DH (2021) Reimagining pilocytic astrocytomas in the context of pediatric low-grade gliomas. Neuro Oncol 23:1634–1646. https://doi.org/10.1093/neuonc/noab138

25. Mohme M, Neidert MC (2020) Tumor-specific T cell activation in malignant brain tumors. Front Immunol 11:205. https://doi.org/10.3389/fimmu.2020.00205
26. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y et al (2015) Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 12:453–457. https://doi.org/10.1038/nmeth.3337
27. Nobre L, Zapotocky M, Ramaswamy V, Ryall S, Bennett J, Alderete D et al (2020) Outcomes of BRAF V600E pediatric gliomas treated with targeted BRAF inhibition. JCO Precis Oncol. https://doi.org/10.1200/PO.19.00298
28. Ogiwara H, Bowman RM, Tomita T (2012) Long-term follow-up of pediatric benign cerebellar astrocytomas. Neurosurgery 70:40–47. https://doi.org/10.1227/NEU.0b013e31822ff0ed. (**discussion 47-48**)
29. Ostrom QT, Price M, Ryan K, Edelson J, Neff C, Cioffi G et al (2022) CBTRUS statistical report: pediatric brain tumor foundation childhood and adolescent primary brain and other central nervous system tumors diagnosed in the United States in 2014–2018. Neuro Oncol 24:iii1–iii38. https://doi.org/10.1093/neuonc/noac161
30. Pajtler KW, Witt H, Sill M, Jones DT, Hovestadt V, Kratochwil F et al (2015) Molecular classification of ependymal tumors across all CNS compartments, histopathological grades, and age groups. Cancer Cell 27:728–743. https://doi.org/10.1016/j.ccell.2015.04.002
31. Petralia F, Tignor N, Reva B, Koptyra M, Chowdhury S, Rykunov D et al (2020) Integrated proteogenomic characterization across major histological types of pediatric brain cancer. Cell 183:1962-1985e1931. https://doi.org/10.1016/j.cell.2020.10.044
32. Pfister S, Janzarik WG, Remke M, Ernst A, Werft W, Becker N et al (2008) BRAF gene duplication constitutes a mechanism of MAPK pathway activation in low-grade astrocytomas. J Clin Invest 118:1739–1749. https://doi.org/10.1172/JCI33656
33. Poschmann G, Sitek B, Sipos B, Hamacher M, Vonend O, Meyer HE et al (2009) Cell-based proteome analysis: the first stage in the pipeline for biomarker discovery. Biochim Biophys Acta 1794:1309–1316. https://doi.org/10.1016/j.bbapap.2009.07.001
34. Reitman ZJ, Paolella BR, Bergthold G, Pelton K, Becker S, Jones R et al (2019) Mitogenic and progenitor gene programmes in single pilocytic astrocytoma cells. Nat Commun 10:3731. https://doi.org/10.1038/s41467-019-11493-2
35. Robinson MH, Vasquez J, Kaushal A, MacDonald TJ, Velazquez Vega JE et al (2020) Subtype and grade-dependent spatial heterogeneity of T-cell infiltration in pediatric glioma. J Immunother Cancer. https://doi.org/10.1136/jitc-2020-001066
36. Rodriguez FJ, Giannini C, Asmann YW, Sharma MK, Perry A, Tibbetts KM et al (2008) Gene expression profiling of NF-1-associated and sporadic pilocytic astrocytoma identifies aldehyde dehydrogenase 1 family member L1 (ALDH1L1) as an underexpressed candidate biomarker in aggressive subtypes. J Neuropathol Exp Neurol 67:1194–1204. https://doi.org/10.1097/NEN.0b013e31818fbe1e
37. Ruda R, Capper D, Waldman AD, Pallud J, Minniti G, Kaley TJ et al (2022) EANO-EURACAN-SNO Guidelines on circumscribed astrocytic gliomas, glioneuronal, and neuronal tumors. Neuro Oncol 24:2015–2034. https://doi.org/10.1093/neuonc/noac188
38. Sexton-Oates A, Dodgshun A, Hovestadt V, Jones DTW, Ashley DM, Sullivan M et al (2018) Methylation profiling of paediatric pilocytic astrocytoma reveals variants specifically associated with tumour location and predictive of recurrence. Mol Oncol 12:1219–1232. https://doi.org/10.1002/1878-0261.12062
39. Sharma MK, Mansur DB, Reifenberger G, Perry A, Leonard JR, Aldape KD et al (2007) Distinct genetic signatures among pilocytic astrocytomas relate to their brain region origin. Cancer Res 67:890–900. https://doi.org/10.1158/0008-5472.CAN-06-0973
40. Stokland T, Liu JF, Ironside JW, Ellison DW, Taylor R, Robinson KJ et al (2010) A multivariate analysis of factors determining tumor progression in childhood low-grade glioma: a population-based cohort study (CCLG CNS9702). Neuro Oncol 12:1257–1268. https://doi.org/10.1093/neuonc/noq092
41. Sturm D, Witt H, Hovestadt V, Khuong-Quang DA, Jones DT, Konermann C et al (2012) Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. Cancer Cell 22:425–437. https://doi.org/10.1016/j.ccr.2012.08.024
42. Tomita Y, Hibler EA, Suruga Y, Ishida J, Fujii K, Satomi K et al (2022) Age is a major determinant for poor prognosis in patients with pilocytic astrocytoma: a SEER population study. Clin Exp Med. https://doi.org/10.1007/s10238-022-00882-5
43. Uhrig S, Ellermann J, Walther T, Burkhardt P, Frohlich M, Hutter B et al (2021) Accurate and efficient detection of gene fusions from RNA sequencing data. Genome Res 31:448–460. https://doi.org/10.1101/gr.257246.119
44. Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD et al (2010) Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell 17:98–110. https://doi.org/10.1016/j.ccr.2009.12.020
45. Wang B, Mezlini AM, Demir F, Fiume M, Tu Z, Brudno M et al (2014) Similarity network fusion for aggregating data types on a genomic scale. Nat Methods 11:333–337. https://doi.org/10.1038/nmeth.2810
46. Wang LB, Karpova A, Gritsenko MA, Kyle JE, Cao S, Li Y et al (2021) Proteogenomic and metabolomic characterization of human glioblastoma. Cancer Cell 39:509-528e520. https://doi.org/10.1016/j.ccell.2021.01.006
47. Wolter M, Felsberg J, Malzkorn B, Kaulich K, Reifenberger G (2022) Droplet digital PCR-based analyses for robust, rapid, and sensitive molecular diagnostics of gliomas. Acta Neuropathol Commun 10:42. https://doi.org/10.1186/s40478-022-01335-6
48. Wong KK, Chang YM, Tsang YT, Perlaky L, Su J, Adesina A et al (2005) Expression analysis of juvenile pilocytic astrocytomas by oligonucleotide microarray reveals two potential subgroups. Cancer Res 65:76–84
49. Yang W, Porras JL, Khalafallah AM, Sun Y, Bettegowda A, Mukherjee D (2022) Comparison of adult and pediatric pilocytic astrocytomas using competing risk analysis: A population-based study. Clin Neurol Neurosurg 212:107084. https://doi.org/10.1016/j.clineuro.2021.107084
50. Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W et al (2013) Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun 4:2612. https://doi.org/10.1038/ncomms3612
51. Zakrzewski K, Jarzab M, Pfeifer A, Oczko-Wojciechowska M, Jarzab B, Liberski PP et al (2015) Transcriptional profiles of

pilocytic astrocytoma are related to their three different locations, but not to radiological tumor features. BMC Cancer 15:778. https://doi.org/10.1186/s12885-015-1810-z

52. Zhang H, Liu T, Zhang Z, Payne SH, Zhang B, McDermott JE et al (2016) Integrated Proteogenomic Characterization of Human High-Grade Serous Ovarian Cancer. Cell 166:755–765. https://doi.org/10.1016/j.cell.2016.05.069

Authors and Affiliations

Daniel Picard[1,2,3] · Jörg Felsberg[3] · Maike Langini[1,4,5] · Paweł Stachura[1,6] · Nan Qin[1,2,3] · Jadranka Macas[7,8,9] · Yvonne Reiss[7,8,9] · Jasmin Bartl[1,2,3] · Florian Selt[10,11,12] · Romain Sigaud[10,11,12] · Frauke-D. Meyer[1,2,3] · Anja Stefanski[4,5] · Kai Stühler[4,5] · Lucia Roque[13] · Rafael Roque[14] · Aleksandra A. Pandyra[1,15,16] · Triantafyllia Brozou[1] · Christiane Knobbe-Thomsen[3] · Karl H. Plate[7,8,9] · Alexander Roesch[2,17,18] · Till Milde[10,11,12,19] · Guido Reifenberger[2,3] · Gabriel Leprivier[3] · Claudia C. Faria[20,21] · Marc Remke[1,2,3]

✉ Marc Remke
marc.remke@med.uni-duesseldorf.de

1 Department of Pediatric Oncology, Hematology, and Clinical Immunology, Medical Faculty, and University Hospital Düsseldorf, Heinrich Heine University, Düsseldorf, Germany

2 German Cancer Consortium (DKTK), Partner site Essen/Düsseldorf, Düsseldorf, Germany

3 Institute of Neuropathology, Medical Faculty, and University Hospital Düsseldorf, Heinrich Heine University, Düsseldorf, Germany

4 Molecular Proteomics Laboratory, Biological and Medical Research Center (BMFZ), Heinrich Heine University Düsseldorf, Düsseldorf, Germany

5 Institute for Molecular Medicine I, Heinrich Heine University Medical Faculty, Düsseldorf, Germany

6 Institute for Molecular Medicine II, Heinrich Heine University Medical Faculty, Düsseldorf, Germany

7 Institute of Neurology (Edinger Institute), University Hospital Frankfurt, Frankfurt am Main, Germany

8 German Cancer Consortium (DKTK), Partner site Frankfurt/Mainz, Frankfurt, Germany

9 Frankfurt Cancer Institute, Frankfurt, Germany

10 Hopp Children's Cancer Center Heidelberg (KiTZ), Heidelberg, Germany

11 Clinical Cooperation Unit Pediatric Oncology, German Cancer Research Center (DKFZ) and German Cancer Consortium (DKTK), Heidelberg, Germany

12 KiTZ Clinical Trial Unit (ZIPO), Department of Pediatric Hematology and Oncology, University Hospital Heidelberg, Heidelberg, Germany

13 Portuguese Cancer Institute, Unidade de Investigação em Patobiologia Molecular (UIPM), IPOLFG, Lisbon, Portugal

14 Laboratory of Neuropathology, Neurology Department, Hospital de Santa Maria, Centro Hospitalar Universitário Lisboa Norte (CHULN), Lisbon, Portugal

15 Institute of Clinical Chemistry and Clinical Pharmacology, University Hospital Bonn, Bonn, Germany

16 German Center for Infection Research (DZIF), Partner Site Bonn-Cologne, Bonn, Germany

17 Department of Dermatology, University Hospital Essen, West German Cancer Center, University Duisburg-Essen, Essen, Germany

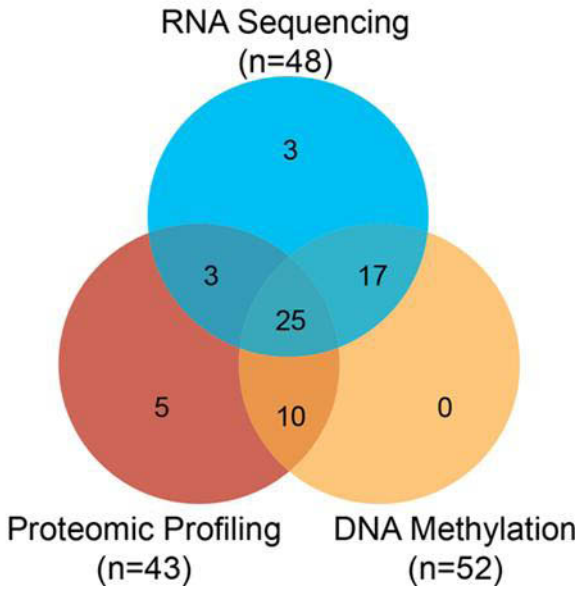
18 Center for Medical Biotechnology (ZMB), University of Duisburg-Essen, Essen, Germany

19 National Center for Tumor Diseases (NCT), Heidelberg, Germany

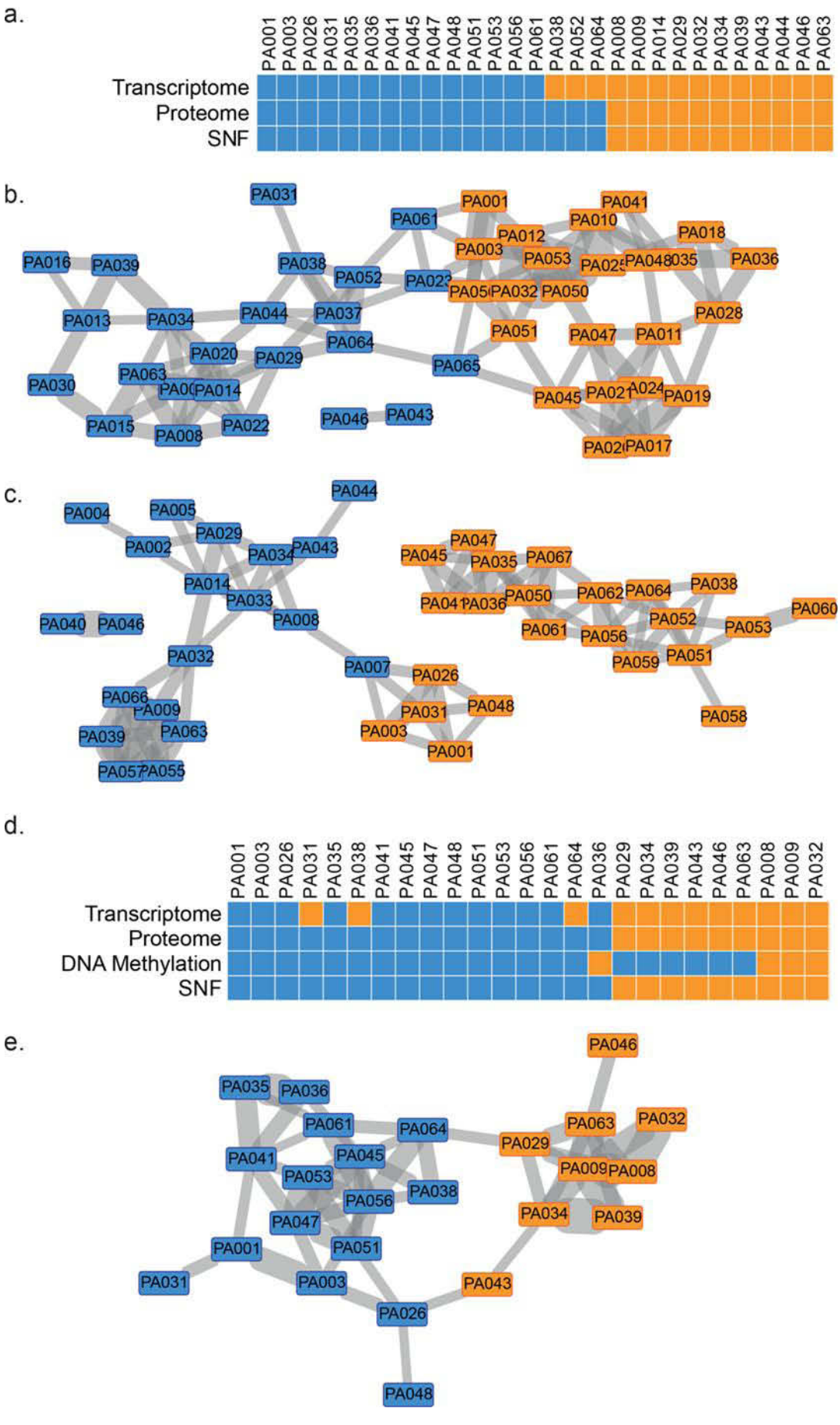
20 Faculdade de Medicina, Instituto de Medicina Molecular João Lobo Antunes, da Universidade de Lisboa, Lisbon, Portugal

21 Department of Neurosurgery, Hospital de Santa Maria, Centro Hospitalar Universitário Lisboa Norte (CHULN), Lisbon, Portugal

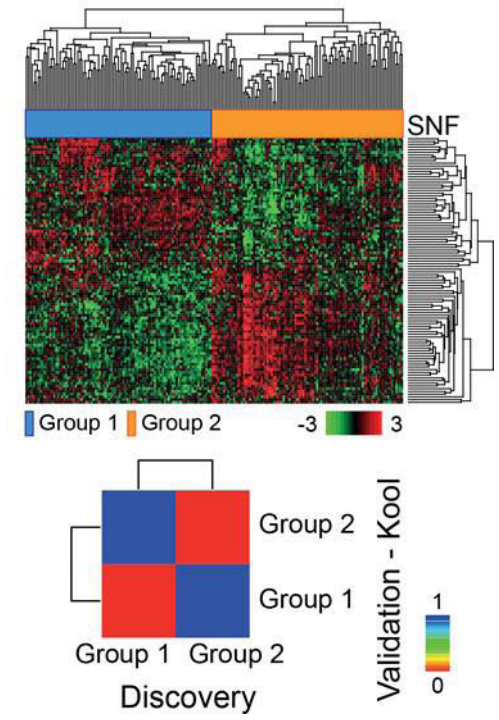
Picard Supplementary Figure 1



Picard Supplementary Figure 2

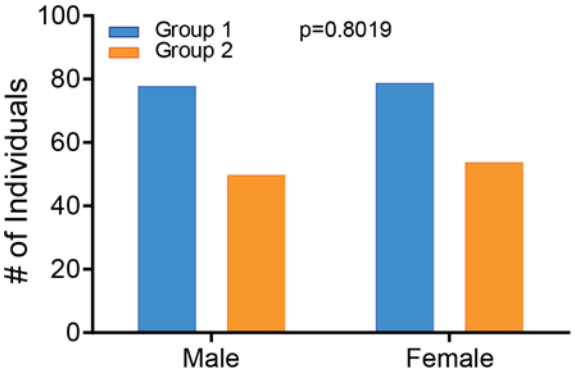


Picard Supplementary Figure 3

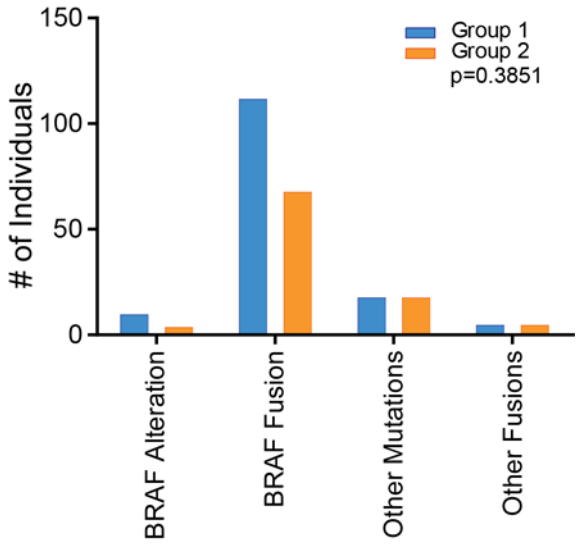


Picard Supplementary Figure 4

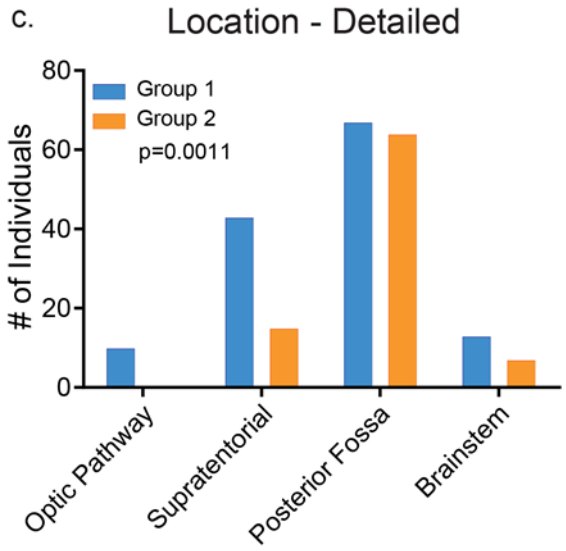
a.



b.



c.



d.

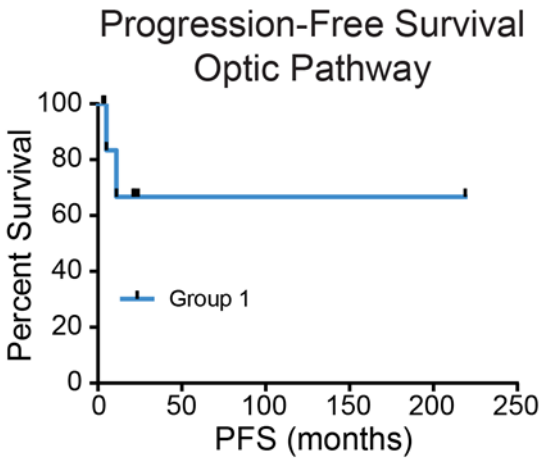
Factor	Hazard Ratio	95% CI	Significance
PA Group	0.365	0.158-0.844	0.018
Age Group	2184.359	0-1.919x10^{64}	0.914
Alteration Type	0.244	0.037-1.635	0.146
Gender	0.923	0.145-5.889	0.933
Location	2779.621	0-3.551x10^{36}	0.838

Picard Supplementary Figure 5

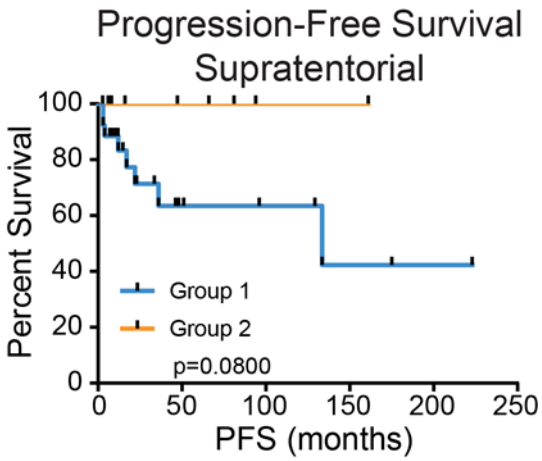
a.

Factor	Hazard Ratio	95% CI	Significance
Location - Detailed	0.949	0.282-3.201	0.933

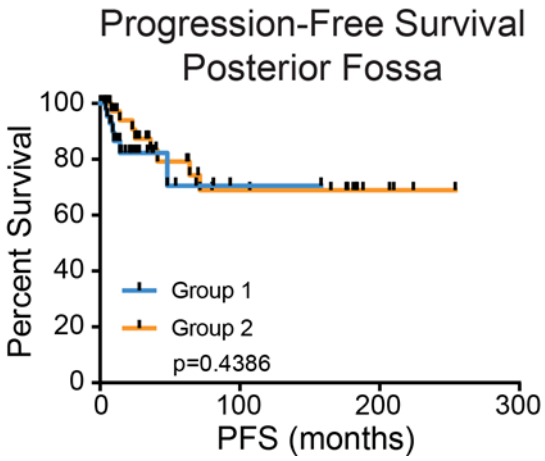
b.



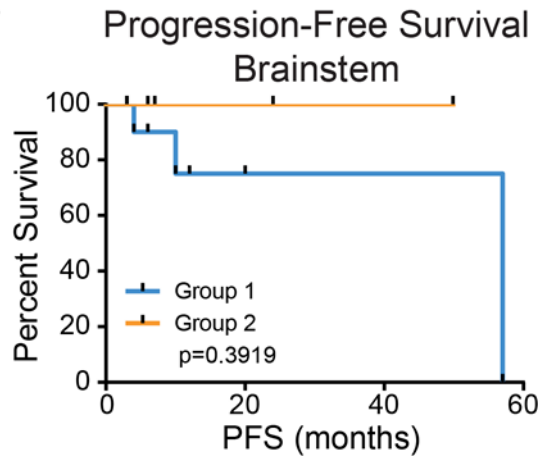
c.



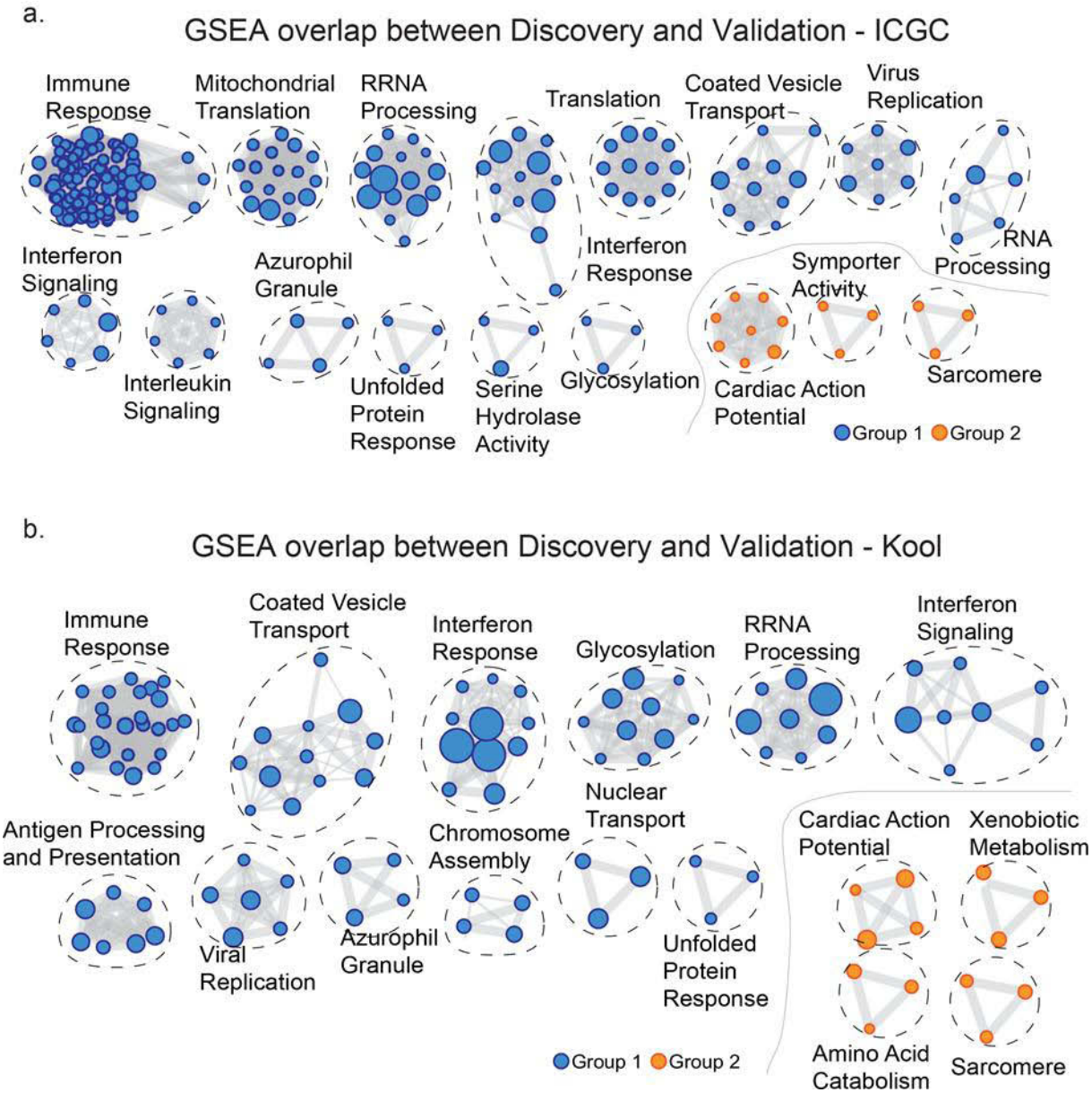
d.



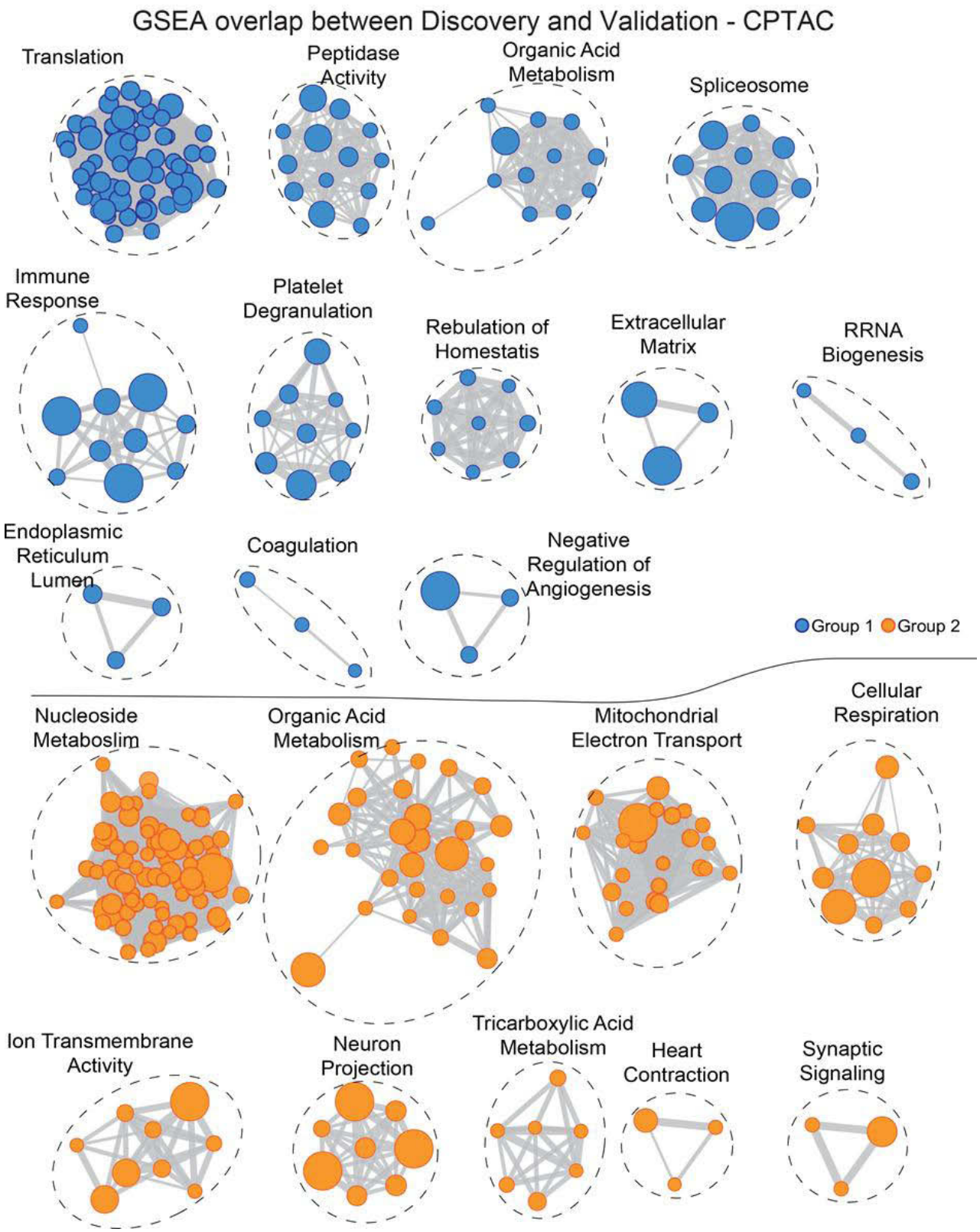
e.



Picard Supplementary Figure 6

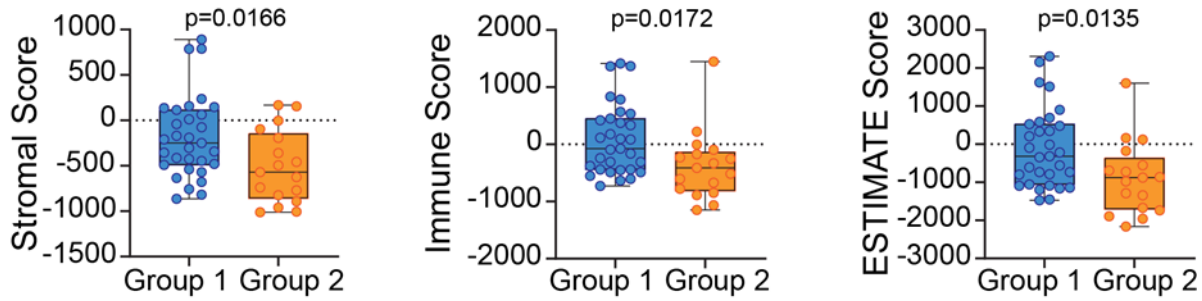


Picard Supplementary Figure 7

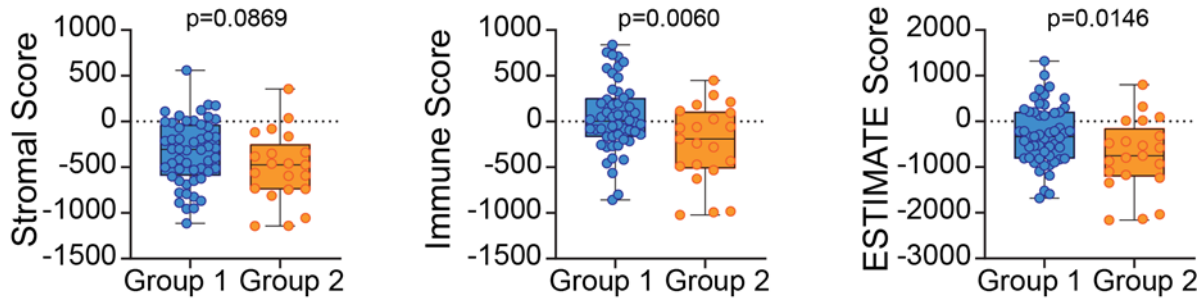


Picard Supplementary Figure 8

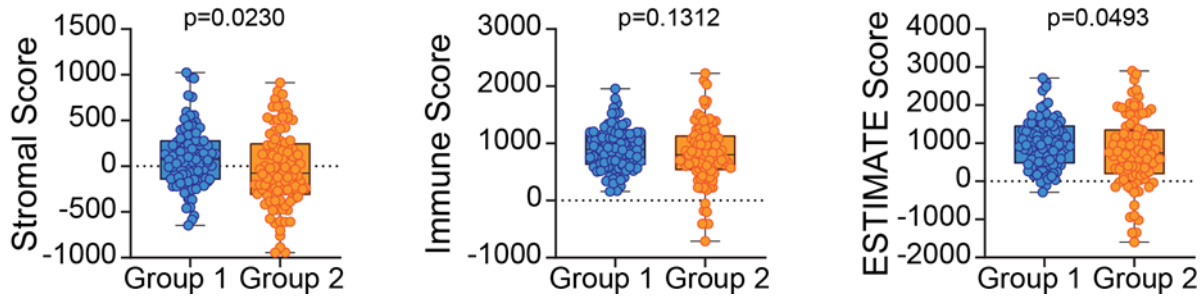
a.



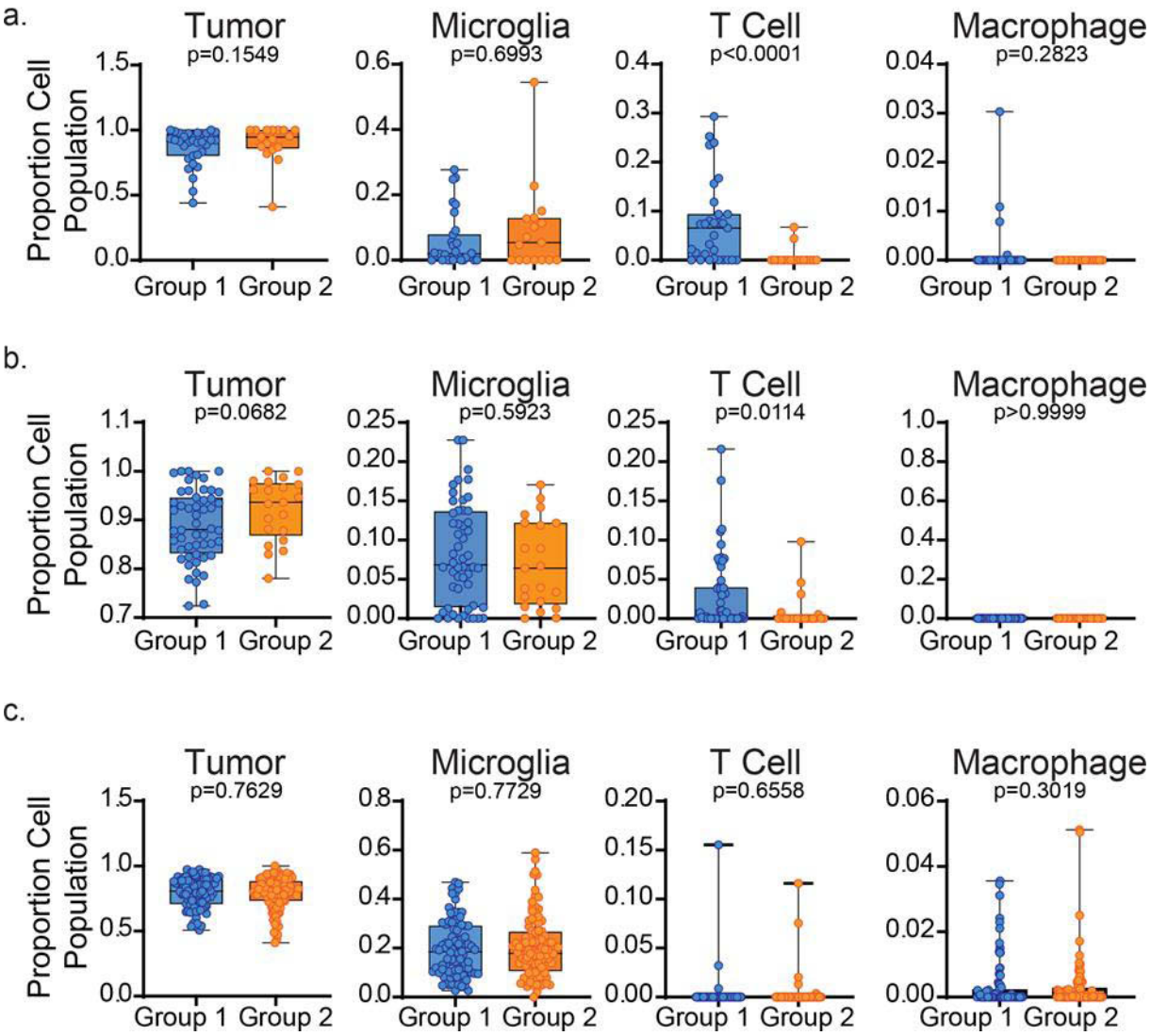
b.



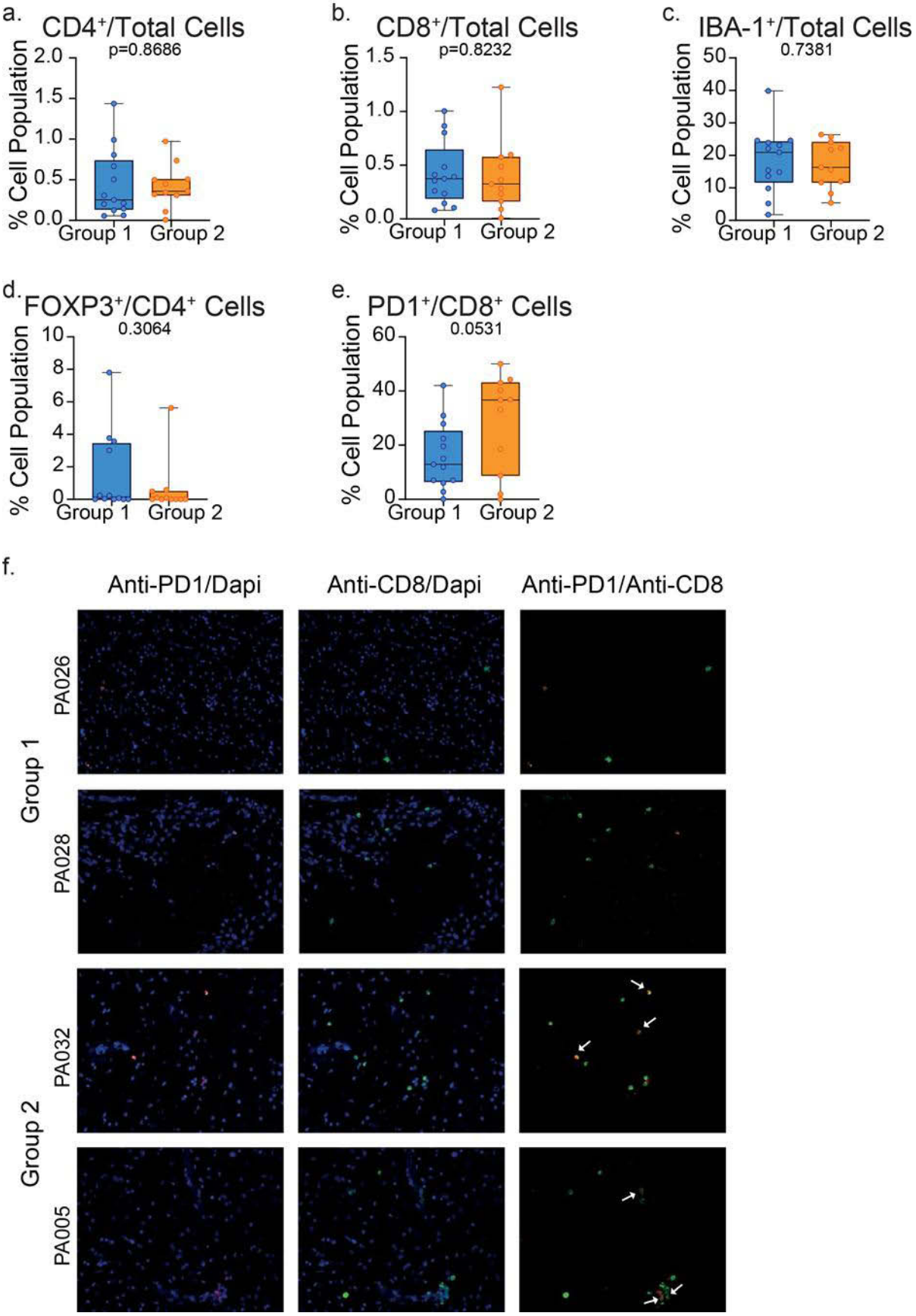
c.



Picard Supplementary Figure 9



Picard Supplementary Figure 10



Supplementary Figure Legends

Supplementary Figure S1. Distribution of PAs processed using RNA sequencing, mass spectrometry and methylation profiling. Venn diagram displaying the overlap between the different omics platforms.

Supplementary Figure S2. Similarity network and network fusion analyses of PA show two groups using multilayer proteogenomic data. a. Similarity network clustering of individual data layers of PA samples and similarity network fusion (SNF) clustering integrating all data layers are shown in the bottom row. **b-c.** Similarity network representations of transcriptome (**b.**) and proteome (**c.**) clustering show incomplete separation of PA groups. Shorter edge length and greater thickness between samples (nodes) indicates more similarity. **d.** Similarity network clustering of individual data layers of PA samples, including transcriptome, proteome and DNA methylation, and SNF clustering integrating all data layers is shown in the bottom row. **e.** SNF representation incompletely segregates samples into both PA groups. Shorter edge length and greater thickness between samples (nodes) indicates more similarity.

Supplementary Figure S3. SNF groups are recapitulated in a micro array validation cohort. 100-gene signature applied to non-overlapping transcriptomic (Kool et al.) validation cohort segregates samples into two groups. Lower panel, submap analyses show close relatedness between discovery and validation cohort.

Supplementary Figure S4. Clinical analyses of PA groups in discovery and validation cohorts. a-b. Mann-Whitney test was performed on gender (**a.**, p=0.8019) and genetic alterations (**b.**, p=0.3851) for the combined grouped samples. **c.** Combined dataset analysis of detailed location which separates optic pathway from supratentorial and infratentorial is subdivided into posterior fossa and brainstem. Location shows an

enrichment of Group 1 samples in both the optic pathway and supratentorial regions (p=0.0011, Mann-Whitney test). **d.** Multivariate analysis of PA clinical and pathological characteristics.

Supplementary Figure S5. Detailed location analyses of PA groups in discovery and validation cohorts. a. Multivariate analysis of detailed location and PA subgroups shows no benefit for survival. **b-e**. Kaplan-Meier progression-free survival curves show no significant survival difference between PA Group 1 and Group 2.

Supplementary Figure S6. Gene Set Enrichment Analysis overlap of discovery and validation RNA Sequencing cohorts. a-b. GSEA-based enrichment map representations based on ranked mRNA showing overlapping gene sets for the discovery and non-overlapping RNA sequencing (ICGC, **a.**) and microarray (Kool et al., **b.**) validation cohorts (p-value <0.001; FDR <0.05). Nodes (circles) representing enriched pathways identified in Group 1 are blue and Group 2 are orange. Edges connect pathways/nodes that share at least half of the terms defining them. Nodes grouped according to functional families are indicated on each network.

Supplementary Figure S6. Gene Set Enrichment Analysis overlap of discovery and validation mass spectrometry cohort. GSEA-based enrichment map representations based on ranked proteins showing overlapping gene sets for the discovery and non-overlapping LGG proteomic validation (CPTAC) cohort (p-value <0.001; FDR <0.05). Nodes (circles) representing enriched pathways identified in Group 1 are blue and Group 2 are orange. Edges connect pathways/nodes that share at least half of the terms defining them. Nodes grouped according to functional families are indicated on each network.

Supplementary Figure S8. Group 1 and Group 2 have different stromal/immune content based on ESTIMATE analysis. a-c. Box plots showing stromal (left panels), immune (center panels) or ESTIMATE (right panels) scores for discovery (**a.**), non-overlapping RNA sequencing (ICGC, **b.**) and microarray (Kool, **c.**) validation cohorts.

Supplementary Figure S9. CIBERSORT of scRNA based cell type signatures. Tumor, microglia, T cell and macrophage cell types were de-convoluted using an scRNA-based cell type signature from Reitman *et al*. using CiberSORT for discovery (**a.**), non-overlapping RNA sequencing (ICGC, **b.**) and microarray (Kool et al., **c.**) validation cohorts.

Supplementary Figure S10. Multiplex immune fluorescence analysis of PA samples. a-e. FFPE slides were stained using anti-CD4, anti-CD8, anti-IBA-1, anti-FOXP3, anti-PD1 or dapi. Results are expressed as percent total cells for anti-CD4, anti-CD8, anti-IBA-1 (**a-c**) and percent cell population for FOXP3+/CD4+ (**d**.) and PD1+/CD8+ (**e**.) cells. Plots were analyzed using Mann-Whitney test. **f**. representative immunofluorescence image of anti-PD1 (red) and dapi (blue) left, anti-CD8 (green) and dapi staining center, and the overlap between PD1+ and CD8+ cells right.

2.2. Manuscript 2: The long non-coding RNA *HOTAIRM1* promotes aggressiveness and radiotherapy resistance in glioblastoma

CDDpress

www.nature.com/cddis

ARTICLE OPEN

Check for updates

The long non-coding RNA *HOTAIRM1* promotes tumor aggressiveness and radiotherapy resistance in glioblastoma

Ulvi Ahmadov[1,2,3,4,15], Daniel Picard[1,2,3,4,15], Jasmin Bartl[1,2,3,4], Manuela Silginer[5], Marija Trajkovic-Arsic[6,7], Nan Qin[1,2,3,4], Lena Blümel[1,2,3,4], Marietta Wolter[4], Jonathan K. M. Lim[4], David Pauck[1,2,3,4], Alina Marie Winkelkotte[6,7], Marlen Melcher[8], Maike Langini[9,10], Viktoria Marquardt[1,2,3,4], Felix Sander[1,2,3,4], Anja Stefanski[9,10], Sascha Steltgens[4], Christina Hassiepen[11], Anna Kaufhold[1,2,3,4], Frauke-Dorothee Meyer[1,2,3,4], Annette Seibt[8], Lara Kleinesudeik[12,13], Anika Hain[14], Carsten Münk[14], Christiane Brigitte Knobbe-Thomsen[4], Alexander Schramm[11], Ute Fischer[3], Gabriel Leprivier[4], Kai Stühler[9,10], Simone Fulda[12,13], Jens T. Siveke[6,7], Felix Distelmaier[8], Arndt Borkhardt[1,2,3], Michael Weller[5], Patrick Roth[5], Guido Reifenberger[1,2,4,16] and Marc Remke[1,2,3,4,16]✉

Glioblastoma is the most common malignant primary brain tumor. To date, clinically relevant biomarkers are restricted to isocitrate dehydrogenase (IDH) gene 1 or 2 mutations and O6-methylguanine DNA methyltransferase (*MGMT*) promoter methylation. Long non-coding RNAs (lncRNAs) have been shown to contribute to glioblastoma pathogenesis and could potentially serve as novel biomarkers. The clinical significance of *HOXA* Transcript Antisense RNA, Myeloid-Specific 1 (*HOTAIRM1*) was determined by analyzing *HOTAIRM1* in multiple glioblastoma gene expression data sets for associations with prognosis, as well as, IDH mutation and *MGMT* promoter methylation status. Finally, the role of *HOTAIRM1* in glioblastoma biology and radiotherapy resistance was characterized in vitro and in vivo. We identified *HOTAIRM1* as a candidate lncRNA whose up-regulation is significantly associated with shorter survival of glioblastoma patients, independent from IDH mutation and *MGMT* promoter methylation. Glioblastoma cell line models uniformly showed reduced cell viability, decreased invasive growth and diminished colony formation capacity upon *HOTAIRM1* down-regulation. Integrated proteogenomic analyses revealed impaired mitochondrial function and determination of reactive oxygen species (ROS) levels confirmed increased ROS levels upon *HOTAIRM1* knock-down. *HOTAIRM1* knock-down decreased expression of transglutaminase 2 (TGM2), a candidate protein implicated in mitochondrial function, and knock-down of *TGM2* mimicked the phenotype of *HOTAIRM1* down-regulation in glioblastoma cells. Moreover, *HOTAIRM1* modulates radiosensitivity of glioblastoma cells both in vitro and in vivo. Our data support a role for *HOTAIRM1* as a driver of biological aggressiveness, radioresistance and poor outcome in glioblastoma. Targeting *HOTAIRM1* may be a promising new therapeutic approach.

Cell Death and Disease (2021)12:885 ; https://doi.org/10.1038/s41419-021-04146-0

INTRODUCTION

Glioblastoma is the most malignant type of astrocytic glioma and the most common malignant primary brain tumor [1]. According to the World Health Organization (WHO) classification of central nervous system tumors, glioblastomas correspond to WHO grade IV and are stratified based on their isocitrate dehydrogenase (IDH) 1 or 2 gene mutation status into two biologically and clinically distinct entities [2]. IDH-wildtype glioblastomas account for more than 90% of the tumors and preferentially manifest *de novo* with short clinical history in patients older than 50 years. In contrast, IDH-mutant glioblastomas, which have recently been re-named as astrocytoma, IDH-mutant, WHO grade 4 [3], are less common, typically occur in patients younger than 50 years and may develop from pre-

[1]Division of Pediatric Neuro-Oncogenomics, German Cancer Research Center (DKFZ), Heidelberg, Germany. [2]German Consortium for Translational Cancer Research (DKTK), partner site Essen/Düsseldorf, Düsseldorf, Germany. [3]Department of Pediatric Oncology, Hematology, and Clinical Immunology, Medical Faculty, University Hospital Düsseldorf, Düsseldorf, Germany. [4]Department of Neuropathology, Medical Faculty, Heinrich Heine University, Düsseldorf, Germany. [5]Department of Neurology, University Hospital and University of Zurich, Zurich, Switzerland. [6]Bridge Institute of Experimental Tumor Therapy, West German Cancer Center, University Medicine Essen, Essen, Germany. [7]Division of Solid Tumor Translational Oncology, German Cancer Research Center (DKFZ) and German Cancer Consortium (DKTK), partner site Essen, Heidelberg, Germany. [8]Department of General Pediatrics, Neonatology and Pediatric Cardiology, University Hospital Düsseldorf, Düsseldorf, Germany. [9]Institute for Molecular Medicine I, Medical Faculty, Heinrich Heine University, Düsseldorf, Germany. [10]Molecular Proteomics Laboratory (MPL), Biological-Medical Research Center (BMFZ), Heinrich Heine University, Düsseldorf, Germany. [11]Department of Molecular Oncology, West German Cancer Center, University Hospital Essen, Essen, Germany. [12]Institute for Experimental Cancer Research in Pediatrics, Goethe University Frankfurt, Frankfurt, Germany. [13]German Cancer Consortium (DKTK), Partner Site Frankfurt and German Cancer Research Center (DKFZ), Heidelberg, Germany. [14]Clinic for Gastroenterology, Hepatology, and Infectiology, Medical Faculty, Heinrich Heine University, Düsseldorf, Germany. [15]These authors contributed equally: Ulvi Ahmadov, Daniel Picard. [16]These authors jointly supervised this work: Guido Reifenberger, Marc Remke. ✉email: marc.remke@med.uni-duesseldorf.de

Edited by Dr Giovanni Blandino

Received: 13 July 2020 Revised: 18 June 2021 Accepted: 22 July 2021
Published online: 28 September 2021

Official journal of CDDpress

SPRINGER NATURE
CDDpress

existing IDH-mutant lower grade astrocytomas [2, 4]. Current standard therapy of glioblastoma consists of surgical resection followed by radiotherapy with concomitant and maintenance temozolomide (TMZ) chemotherapy [5]. Recently, tumor treatment fields (TTF) have been reported and approved as an additional treatment [6]. However, the outcome of IDH-wildtype glioblastoma patients remains poor, with median overall survival limited to 15-18 months and reported five-year survival rate of less than 10% [7]. The majority of glioblastomas demonstrate remarkable resistance to radiation and chemotherapy either upfront or during the course of treatment [8, 9]. Only a few alterations have been established as clinically relevant biomarkers to date, including IDH mutation as a diagnostic and prognostic biomarker, as well as aberrant promoter methylation of the *O6-methylguanine DNA methyltransferase* (*MGMT*) gene as a predictor of benefit from TMZ [10–12]. However, clinically active molecularly targeted therapy approaches are still missing [13].

To further elucidate the molecular pathogenesis of glioblastoma, we evaluated long non-coding RNA (lncRNA) expression profiles in microarray-based gene expression data sets. LncRNAs are non-coding transcripts that are longer than 200 nucleotides [14]. Recent reports have indicated important roles of aberrant lncRNA expression in tumorigenesis, progression and therapy resistance of various cancers [15–17]. We identified the lncRNA *HOXA transcript antisense RNA myeloid-specific 1* (*HOTAIRM1*) as a candidate lncRNA which maps within the *HOXA* gene cluster on the short arm of chromosome 7 and was originally implicated in myelopoiesis through modulation of gene expression in the HOXA cluster [18]. Previous studies have reported that *HOTAIRM1* expression is increased in high-grade gliomas and in recurrent compared to primary glioblastomas [19, 20]. Recently, *HOTAIRM1* has been shown to promote glioma growth and invasion through up-regulation of *HOXA1* expression [21], through long-range chromatin interactions within HOXA cluster genes [22] and regulating *SNAI2* [23]. Moreover, *HOTAIRM1* has been proposed to promote glioma growth by acting as a sponge for several tumor suppressive miRNAs [24–26]. Here, we extend these findings by providing further clinical and functional evidence implicating *HOTAIRM1* as a driver of tumor aggressiveness that contributes to radioresistance and poor outcome of glioblastoma patients.

MATERIAL AND METHODS

Cell culture

The glioblastoma cell lines T98G, U251MG and U87MG were obtained from American Type Culture Collection (ATCC, Manassas, VA, USA). LN-18 and LN-229 cell lines were provided by Dr. M. Hegi, Lausanne, while the SF126 cell line was obtained from the Japanese Collection of Research Bioresources Cell Bank (JCRB, Osaka, Japan). All lines were authenticated by short tandem repeat (STR) profiling and tested for mycoplasma contamination. Cell lines were grown in Dulbecco's modified Eagle's medium (#31966-021, DMEM, Thermo Fischer Scientific, Waltham, MA, USA) supplemented with 10% heat–inactivated fetal bovine serum (FBS, #F9665, Sigma-Aldrich, St. Louis, MO, USA) and 1% penicillin-streptomycin (#P4333, Sigma-Aldrich) in a humidified atmosphere with 5% CO_2 atmosphere at 37 °C. Stable cells were grown and/or selected with the culture media containing either blasticidin [T98G, LN-18, SF126 and U87MG] (20 µg/ml, #ant-bl-1, InvivoGen, San Diego, CA, USA) or puromycin [LN-229and U251MG] (2 µg/ml, #ant-pr-5, InvivoGen).

RNA preparation and real-time quantitative PCR

Total RNA extraction was performed by using TRIzol (#15596026, Thermo Fischer Scientific) or by using a Maxwell RSC instrument (#AS1340, RSC simplyRNA Tissue, Promega, Madison, WI, USA). Reverse transcription of total RNA (1 µg) was carried out using the MMLV-RT kit (#M3683, Promega) with random hexamers. Real-time quantitative PCR (RT-qPCR) was subsequently performed with a 1:10 dilution of reverse-transcribed cDNA using a CFX384 Touch™ Real-Time PCR Detection System (Bio-Rad, Hercules, CA, USA). RT-qPCR TagMan Universal Master Mix II (#4440038, Thermo Fischer Scientific) was employed, with TagMan probes specific for *HOTAIRM1*- (exon 1-3, #Hs.PT.58.45434173, Integrated DNA Technologies (IDT), Coralville, IA, USA), *TGM2* (#Hs.PT.58.23141755, IDT) or *PGK1* (#Hs.PT.58 v.606641, IDT). RT-qPCR results were evaluated by $2^{-\Delta\Delta Ct}$ method [27] using *PGK1* expression levels as a housekeeping gene control.

Generation of *HOTAIRM1* or *TGM2* knock-down glioblastoma cells

For transient knock-down of *HOTAIRM1* and *TGM2*,T98G, LN-229, LN-18, SF126 or U251 glioblastoma cells were seeded (100,000 cells per well) into 6-well plates the day before knock-down. Specific knock-down of *HOTAIRM1* or *TGM2* and the corresponding non-target negative controls was achieved by using *HOTAIRM1*-specific siPOOLs (#100506311, siTOOLs Technology, Planegg, Martinsried, Germany), *TGM2*-specific siPOOLs (#7052 - TGM2 (human), siTOOLs Technology) and non-target siPOOLs (Neg. control siPOOL 5 nmol, siPOOL Technology), respectively. Transfection of siPOOLs was performed using Lipofectamine RNAiMAX Transfection Reagent (#13778150, Thermo Fischer Scientific). Briefly, 7.5 µl RNAiMAX Transfection Reagent was diluted in 125 µl Opti-MEM (#31985062, Thermo Fischer Scientific) pipetted into a well with 0.5 µl (5 pM) siPOOL in 125 µl Opti-MEM.

Lentivirus-containing *HOTAIRM1* shRNA constructs (sequence is 5′-GGAGACTGGTAGCTTATTAAA-3′) and non-target negative control shRNA constructs (sequence is 5′-CCTAAGGTTAAGTCGCCCTCG-3′) were obtained from IDT and the sequences were obtain from a previous study [28]. HOTAIRM1-pLK0.1-TRC plasmid and third generation lentiviral packing plasmids (pMDL/pRRE, pRSV-Rev and pMD2.G) were transfected into HEK-293T cells using polyethylenimine (#408727, PEI, Sigma-Aldrich). Fresh culture media was added (without antibiotics) after 24 and 48 h post transfection. The virus-containing medium (48 and 72 h after transfection) was stored at −80 °C. Glioblastoma cells (LN-229, LN-18, SF126 and U87MG) were seeded in 6-well plates the day before transduction and were infected with 1.5 ml of the viral suspension containing DMEM, 10% FBS and 2 µg/ml polybrene (#107689, hexadimethrine bromide, Sigma-Aldrich) for stable cell line generation. The virus-containing medium was removed 24 h post transduction and replaced with medium containing puromycin or blasticidin for at least a week.

Determination of cell viability

Cells were seeded (1000-4000 cells per well) into white-bottom 96-well plates (#136101, Thermo Fischer Scientific) and incubated for 72 h. Afterwards, cells were incubated with sterile 100 µl of 1:1 diluted (with PBS) CellTiter-Glo (#G7570, Promega) for 10 min at RT, shaken for 2 min and the absorbance was measured with a Spark 10 M microplate reader (Tecan, Maennedorf, Switzerland). All experiments were independently repeated at least three times.

Determination of cell invasion in vitro

Corning BioCoat™ Matrigel Invasion Chambers (#354480, Corning, Bedford, MA, USA) were used to determine invasive capacity of glioblastoma cells in vitro. Transwell membranes were activated by adding 500 µl serum-free medium for 2 h at room temperature. Afterwards, 750 µl medium containing 10 % FBS was added into the lower chamber. 2.5×10^5 cells in 500 µl were resuspended in serum-free medium and were added to the upper chamber. The chambers were removed after 48 h incubation (at 37 °C), the medium was removed and the membrane was washed once with PBS. Next, the cells that had migrated across the polycarbonate membrane were fixed with methanol for 2 min and stained with 1% toluidine blue for 2 min. The membrane was washed 4 times with distilled water and cells remaining in the upper chamber were removed using a cotton swab. Membranes were allowed to air dry for a minimum of one hour before being mounted with vectashield mounting medium (#H-1000, Vector laboratories, Burlingame, CA, USA) on a glass slide. Finally, six random fields were selected and imaged (20X) using an Axiovert 200 microscope (Zeiss, Oberkochen, Germany) and the AxioVision (Version 4.8) software. Invading cells were counted for each membrane. All experiments were independently repeated at least three times.

Determination of colony formation and in vitro radiosensitivity

Cells were harvested using trypsin, counted with Vi-CELL XR (Beckman Coulter, Brea, CA, USA), plated on culture-treated 100 mm dishes (500 – 1000 cells per plate) and cultured for 21 days. At the end of the incubation period, media was removed from dishes and cells were washed

once in PBS and fixed with 10% formalin for 45 min. Cells were then stained with 0.1% crystal violet for 1 h, washed in H_2O to remove excess dye and were allowed to air dry overnight. The following day, colonies that were visible to the naked eye were counted. All experiments were repeated at least three times.

For the in vitro radiation assay, cells were irradiated at 2 and 4 Gy using a Gulmay RS225 irradiation machine (Gulmay GmbH, Krefeld, Germany). Afterwards, the colony formation assay was performed as described above to evaluate the effect of the irradiation. For determing the efect of radiation on *HOTAIRM1* expression, 4 Gy irradiated cells were seeded (100,000 cells per well) into 6-well plates and samples were harvested 48 h post-seeding. HOTAIRM1 transcript levels were measured using qRT-PCR and results were validated using published data GSE153982 and GSE111247 [29].

miRNA-175-5p transient over expression

For transient over-expression of miR-17-5p mimic, 2uL of hsa-miR-17-5p mirVana mimic (50µM concentration; #4464066, Ambion) or mirVana miRNA mimic negative control #1 (50µM concentration; #4464058, Ambion) was mixed with 2uL of Lipofectamine 2000 Transfection Reagent (#11668019, Thermo Fischer Scientific) in 100µL of Opti-MEM (#31985062, Thermo Fischer Scientific), and incubated for 15 min.

Each transfection mix was placed into an individual well of a 6-well plate followed by addition of a 2 mL cell suspension of LN-229 glioblastoma cells with a final seeding density of 150,000 cells per well. Transfections were harvested 72 h post-transfection.

Luciferase reporter assay

Plasmid insert (hg38, chr20:38,137,943-38,138,030), corresponding to the predicted miR-17-5p binding site within the 3'UTR of *TGM2*, was subcloned into psiCHECK-2 vector (Promega, # C8021). The miR-17-5p binding site 5'-gtcctaagCACTTTataaa-3' was mutated to 5'- gtcctaagAAAAAAataaa-3'. The correctness of insert orientations was confirmed by sequencing. The reporter activity was measured by using Dual-Luciferase Reporter Assay System (Promega, #E1910) according to the manufacturer's instructions. The 3'UTR- Luciferase reporter gene assays were performed as described previously (Wolter et al., 2016) except for transfecting cells with 50 nM of miRNA mimic hsa-miR-17-5p (ThermoFisher Scientific, # 4464066) and control cells with 50 nM of the miRNA Negative Control #1 (ThermoFisher Scientific, # 4464058).

In vivo mouse experiments

All animal experiments were performed in accordance with the guidelines of Swiss federal law on animal protection. CD1 Foxn1 nude mice were purchased from Charles River Laboratories (Wilmington, MA, USA). Eleven 6-10-week-old mice per group were used in all experiments. Mice were anaesthetized using an intraperitoneal 3 component injection consisting of fentanyl, midazolam and medetomidin, fixed under a stereotactic device (Stoelting, Wood Dale, IL, USA) and a burr hole was drilled in the skull 2 mm lateral and 1 mm posterior to the bregma. A Hamilton syringe needle was introduced to a depth of 3 mm and LN-229 human glioma cells (75,000) in a volume of 2 µl phosphate-buffered saline (PBS) were injected into the right striatum. Local cranial radiotherapy with a single dose of 12 Gy was performed at day 15 after tumor implantation using a Gulmay 200 kV X-ray unit at 1 Gy/min at room temperature. The mice were observed daily and euthanized when neurological symptoms developed. No blinding was done for mouse experiments.

Western blot analysis

Cells were lysed and and total proteins were extracted using RIPA lysis buffer (#20-188, Merck Millipore, Burlington, MA, USA) supplemented with protease and phosphatase inhibitor cocktail from Roche (#04693132001 and #04906837001, Sigma-Aldrich). Proteins were quantified with the Bradford method using the Protein Assay Dye Reagent (#500-0006, Bio-Rad). Samples were separated by SDS-PAGE and transferred to a nitrocellulose membrane (#10600002, Sigma-Aldrich) by wet blot using the Mini Gel Tank and Blot Module (#A25977 and #B1000, Thermo Fischer Scientific). The membrane was incubated with rabbit anti-TGM2 (#3557 S, D11A6, 1:1000, Cell Signaling, Danvers, MA, USA), and mouse anti-Actin B (#MAB1501, clone 4, 1:5000, Merck Millipore) primary antibodies overnight at 4 °C. Next, the membrane incubated with species-specific, peroxidase-coupled secondary antibodies (anti-rabbit-HRP, #7074 S, 1:5000, Cell Signaling or anti-mouse-HRP, #H2014, 1:5000, Santa Cruz Biotechnology, Dallas, TX, USA) for an hour at RT. Finally, proteins were visualized using the SuperSignal West Femto Maximum Sensitivity Substrate (#34095, Thermo Fischer Scientific) and detected using the LAS-3000 Imaging System (Fujifilm, Minato, Tokyo, Japan).

RNA sequencing

Total RNA was isolated from siRNA-mediated *HOTAIRM1* knock-down and control cells of T98G, LN-229 and U251 glioblastoma cell lines 72 h post-transfection using TRIZol reagent. 500 ng total RNA was processed using the TruSeq RNA Sample Preparation v2 kit (low-throughput protocol; Illumina, San Diego, CA, USA) to prepare the barcoded libraries. Libraries were validated and quantified using either DNA 1000 or high-sensitivity chips on a Bioanalyzer (Agilent, Santa Clara, CA, USA). 7.5 pM denatured libraries were input into cBot (Illumina), followed by deep sequencing using HiSeq 2500 (Illumina) for 101 cycles, with an additional seven cycles for index reading. Fastq files were imported into Partek Flow (Partek Incorporated, St. Louis, MO, USA). Quality analysis and quality control were performed on all reads to assess read quality and to determine the amount of trimming required (both ends: 13 bases 5´ and 1 base 3´). Trimmed reads were aligned against the hg38 genome using the STAR v2.4.1d aligner. Unaligned reads were further processed using Bowtie 2 v2.2.5 aligner. Finally, aligned reads were combined before quantifying the expression against the ENSEMBL (release 84) database using the Partek Expectation-Maximization algorithm. Partek Flow default settings were used in all analyses. RNA sequencing data has been deposited in the NCBI GEO dataset repository under the identifyer GSE152147.

Mass spectrometry-based proteome analyses

For mass spectrometry (MS)-based proteome analyses, proteins were extracted from frozen cell pellets from siRNA-mediated *HOTAIRM1* knock-down and control cells of the T98G, LN-229 and U251 cell lines. In addition, proteins were extracted from shRNA-mediated *HOTAIRM1* knock-down and control cells of LN-229 cells as described. MS-based proteome analyses were performed as described before [30]. Cells were homogenized in urea buffer with a TissueLyser (Qiagen, Hilden, Germany) and subsequent sonication. After centrifugation for 15 min at 14,000 x g and 4 °C, supernatants were collected. Protein concentration was determined via Pierce 660 nm Protein Assay (Thermo Fischer Scientific) and 10 µg protein per sample were desalted through electrophoretic migration at 50 V for 10 min on a 4 −12 % Bis-Tris polyacrylamide gel (#EC62352BOX, Novex NuPAGE, Thermo Fischer Scientific). After silver staining, protein bands were cut out, reduced, alkylated and digested with trypsin before peptide extraction via sonication. Peptides were dissolved and diluted with 0.1 % TFA (v/v).

For MS-based proteome analyses, 15 µL peptide solution per sample was analyzed on a nano-high-performance liquid chromatography electrospray ionization mass spectrometer. The analytical system was composed of an RSLCnano U3000 HPLC coupled to a QExactive Plus mass spectrometer via a nano-electrospray ion source (Thermo Fischer Scientific). Injected peptides were concentrated and desalted at a flow rate of 6 µL/min on a trapping column (Acclaim PepMao C18, 2 cm×100 µm x 3 µm particle size, 100 Å pore size, Thermo Fischer Scientific) with 0.1 % TFA (v/v) for 10 min. Subsequently, peptides were separated at a constant flowrate of 300 nL/min over a 120 min gradient on an analytical column (Acclaim PepMap RSLC C18, 25 cm×75 µm x 2 µm particle size, 100 Å pore size, Thermo Fischer Scientific) at 60 °C. Separation was achieved through a gradient from 4 to 40% solvent B (solvent A: 0.1% (v/v) formic acid in water, solvent B: 0.1% (v/v) formic acid, 84% (v/v) acetonitrile in water). Afterwards, peptides were ionized at a voltage of 1,400 V and introduced into the mass spectrometer operated in positive mode. Mass spectrometry scans were recorded in profile mode in a range from 350-2000 m/z at a resolution of 70,000, while tandem mass spectra were recorded at a resolution of 17,500. Tandem mass spectra were recorded with a data-dependent Top10 method and 30% normalized collision energy. Dynamic exclusion was activated with a repeat count of 1 for 100 s.

Proteome Discoverer (version 1.4.1.14, Thermo Fisher Scientific) was applied for peptide/protein identification with Mascot (version 2.4, Matrix

Science, London, UK) as search engine employing the UniProt database (human; including isoforms; date 2016-03-01). A false discovery rate of 1% ($p \leq 0.01$) on peptide level was set as identification threshold. Proteins were quantified with Progenesis QI for Proteomics (Version 2.0, Nonlinear Dynamics, Waters Corporation, Newcastle upon Tyne, UK). The mass spectrometry proteomics data has been deposited to the ProteomeXchange Consortium via the PRIDE [31] partner repository with the data set identifier PXD020141.

Determination of superoxide levels

Cellular and mitochondrial superoxide levels were determined by labeling cells with dihydroethidium (HEt, 10 µM, 10 min, and 37 °C; #D11347, Thermo Fischer Scientific) or the mitochondria-targeted variant MitoSOX ™ Red (5 µM, 10 min, 37 °C; #M36008, Thermo Fischer Scientific) as described elsewhere [32]. The staining reactions were stopped by washing the cells three times with PBS. The red fluorescence was documented using an Axio Observer Z1 microscope (Zeiss) with the dihydroethidium filter set (F39-500, F48-515, F47-895; AHF Analysetechnik, Tuebingen, Germany). Images were analyzed and fluorescence intensity was quantified using ImageJ software (Wayne Rasband at the National Institutes of Health; http://rsbweb.nih.gov/ij/).

Bioinformatic and statistical analyses

Statistical analyses were preformed by using GraphPad Prism (version 5.0) (https://www.graphpad.com/scientific-software/prism/). Experimental data are represented as mean ± SEM based on at least three independent experiments. The two-way ANOVA test was used for statistical analyses of qRT-PCR, viability, invasion, colony formation and reactive oxygen species (ROS) staining assays. Paired T-tests or Mann-Whitney tests (non-parametric t-test) were used for comparisons between two groups for statistical analysis of in vitro radiation, western blotting quantification and N-acetyl cysteine (NAC) assays. Differences between groups were considered statistically significant at $p < 0.05$. Kaplan-Meier survival analysis were calculated using the Log Rank method.

Protein-coding genes were filtered out from the Affymetrix U133 Plus 2 array data leaving a final count of 2858 lncRNAs. Initial identification of lncRNAs in glioblastoma samples from long-term (> 36 months overall survival) versus short-term survivors (< 12 months overall survival) was carried out using GSE53733 data set based on a prospective cohort of the German Glioma Network [33], excluding the data from patients with intermediate overall survival between 12 and 36 months.

Kaplan-Meier survival curves were generated using both the GSE16011 dataset (Affymetrix U133 Plus 2 array filtered for glioblastoma samples, analyzed using R2 platform (Academic Medical Center (AMC) Amsterdam, the Netherlands)) and the TCGA dataset (Affymetrix Human Exon 1.0 ST array, https://www.cancer.gov/tcga, analyzed using IBM SPSS statistics (version 21 IBM Corporation)). The last quartile was used to define high *HOTAIRM1* expression.

Analysis of chromosome 7 gene expression relative to chromosome 7 copy number status was performed by plotting expression fold change according to disomy 7 versus trisomy 7 based on U133 Plus 2 glioblastoma datasets with copy number information (n = 29) taken from GSE7696, GSE36245 and GSE43289. In addition TCGA samples from the Affymetrix Human Exon 1.0 ST array were used in separate analyses with samples of undetermined copy number status being removed.

GeneSet Enrichment Analysis was performed using the t-value from the paired t-test for both RNA sequencing and proteomics data of the siRNA-mediated knockdown and respective controls. Gene sets were comprised of curated pathways from several databases including GO, Reactome, KEGG (March 24 2016 version; http://download.baderlab.org/EM_Genesets/current_release/Human/symbol/) and visualized using Cytoscape (www.cytoscape.org; $p \leq 0.001$, $q \leq 0.05$, similarity cutoff 0.5).

MicroRNA (miRNA) predicted to bind to *HOTAIRM1* and either of the 12 candidate proteins that were down-regulated after stable *HOTAIRM1* in LN-229 glioblastoma cells were identified using the miRanda tool (https://omictools.com/miranda-tool). Among these miRNAS, we selected those that showed inverse expression relative to *HOTAIRM1* in the investigated TCGA data set (https://www.cancer.gov/tcga).

TGM2 promoter methylation status was investigated using a TCGA data set profiled using an illumina 450 K methylation array (https://www.cancer.gov/tcga). Only a small subset of samples have *HOTAIRM1* status since sample overlap was minimal between expression and methylation data.

RESULTS

High *HOTAIRM1* expression is associated with shorter survival of glioblastoma patients

We took advantage of publically available glioma gene expression datasets to identify lncRNAs associated with overall survival of glioblastoma patients. First, we compared lncRNA expression profiles of primary glioblastoma samples from patients with long-term (overall survival > 36 months) versus short-term (overall survival < 12 months) using the German Glioma Network (GGN) cohort [33]. We found three lncRNAs that were significantly upregulated in the population of short-term survivors. Out of these lncRNAs, *HOTAIRM1* was a top candidate based on fold-change expression difference and statistical significance (Fig. 1A, Supplementary Table 1). We validated this observation across the cohort that higher *HOTAIRM1* expression also correlated with shorter overall survival (Supplementary Fig. 1A) and analyzed additional publically available data sets from two independent, non-overlapping patient cohorts published by The Cancer Genome Atlas (TCGA) consortium [34] (https://www.cancer.gov/tcga) and Gravendeel et al. (2009) [35]. Thereby, we confirmed that patients whose tumors show high *HOTAIRM1* expression levels, as defined by the upper quartile, demonstrated shorter overall survival (Fig. 1B, C). We confirmed *IDH1* mutation as a strong prognostic marker of longer survival in the GGN dataset [33] (Supplementary Fig. 1B) and found that the prognostic value of the *HOTAIRM1* expression was independent from *IDH1* mutation (Fig. 1D, E) and the *MGMT* promoter methylation status in this patient cohort (Fig. 1F, Supplementary Fig. 1). Glioblastoma patients whose tumors carried an *IDH1* mutation, *MGMT* promoter methylation and low *HOTAIRM1* expression showed the longest overall survival (Fig. 1F). Consistent with these prognostic associations, *HOTAIRM1* was recently shown by other investigators to be aberrantly expressed in glioblastoma [21] and associated with shorter survival of glioma patients [21, 24, 26].

Since the majority of glioblastomas display copy number gains of chromosome 7, often due to trisomy 7 [36], we evaluated the expression level of *HOTAIRM1* in relation to chromosome 7 copy number status in glioblastomas [37–39]. In addition to the epidermal growth factor receptor (*EGFR*) gene, i.e., the proto-oncogene most commonly amplified and overexpressed in IDH-wildtype glioblastoma [9], *HOTAIRM1* and the protein-coding gene *MEOX2* showed consistently increased expression in tumors with chromosome 7 gain when compared to tumors without this copy number increase (Fig. 1G, H). *EGFR* [40, 41] and *MEOX2* [42] overexpression have been implicated before as drivers of glioblastoma growth. *HOTAIRM1* was the only lncRNA on chromosome 7 that was significantly upregulated in gliomas with chromosome 7 gain.

HOTAIRM1 knock-down decreases glioblastoma cell viability, invasion, and clonogenicity

To determine effects of genetic knock-down of *HOTAIRM1* in glioma cells, we first performed a transient siRNA-mediated knock-down of *HOTAIRM1* in the four established glioblastoma cell lines U251MG, LN-229, LN-18, and T98G (Fig. 2A–D), which showed an intermediate expression level (Supplementary Fig. 2). Efficiency of *HOTAIRM1* knock-down was greater than 80% in each of the four cell lines (Fig. 2A). *HOTAIRM1* knock-down significantly reduced cell viability of these glioma lines by 20–30% (Fig. 2B). In addition, *HOTAIRM1* knock-down resulted in significant reduction of glioma cell invasiveness by 40–50% (Fig. 2C, Supplementary Fig. 3) and colony formation capacity by 25–40% (Fig. 2D).

To further investigate the phenotypic changes caused by *HOTAIRM1* knock-down, we generated stable knock-down glioma lines using a lentiviral shRNA approach. Twenty-four hours post transduction, glioma cell lines were selected using either clonal

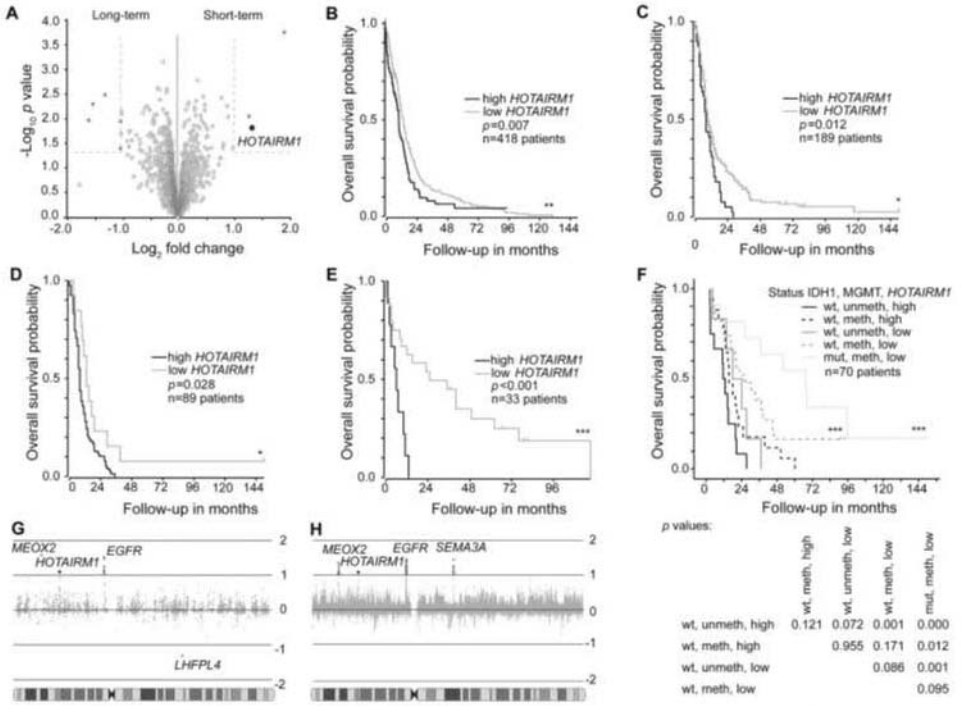


Fig. 1 Prognostic role of *HOTAIRM1* expression in glioblastoma patient datasets. A Volcano plot showing differential expression of lncRNAs in glioblastomas from patients with long-term (overall survival > 36 months) versus short-term (overall survival < 12 months) survival in the German Glioma Network (GGN) cohort[24]. The black circle highlights *HOTAIRM1* while dark gray circles represent other lncRNAs with differential expression between survival groups (±2-fold change and $p < 0.05$). Light gray circles indicate lncRNAs that are not significant. **B** Overall survival plots of glioblastoma patients from TCGA [34] (https://www.cancer.gov/tcga) and **C** Gravendeel et al. [35] stratified according to high or low *HOTAIRM1* expression levels. Cut-off for high *HOTAIRM1* was determined by upper quartile and log rank statistics were calculated. **D**, **E** Overall survival of glioblastoma patients in the Gravendeel et al. [35] cohort according to *HOTAIRM1* expression in *IDH1*-wildtype (**D**) and *IDH1*-mutant glioblastomas (**E**). **F** Overall survival of glioblastoma patients in the GGN cohort [33] stratified according to *HOTAIRM1* expression, *MGMT* promoter methylation status, and *IDH1* mutation status (wt: wild-type; mut: mutant; meth: methylated; unmeth: unmethylated). The table below the Kaplan-Meier graph lists p-values for the individual subgroups. **G**, **H** Expression of genes mapping to chromosome 7 in glioblastomas stratified according to presence or absence of chromosome 7 gain. **G** Data based on primary glioblastoma [37–39] Affymetrix U133 Plus 2 arrays or (**H**) TCGA [34] Human Exon 1.0 ST array show *HOTAIRM1* as the only lncRNA with significantly increased expression in glioblastomas with chromosome 7 gain in addition to the coding genes *EGFR, MEOX2* and *SEMA3A*. Log rank analysis for Kaplan–Meier survival plots; $***p < 0.001$, $**p < 0.01$, $*p < 0.05$.

selection with puromycin (LN-229) or selection with blasticidin for pooled populations (LN-18, SF126 and U87MG). Efficient stable knock-down was achieved for *HOTAIRM1* in all cell lines (Supplementary Fig. 4A) and the phenotypic characterization was performed as described above for transient models (Supplementary Fig. 4B–D). The stable *HOTAIRM1* knock-down cell lines corroborated the phenotype observed in transient models, i.e., cell viability was decreased by 20–40% (Supplementary Fig. 4B), while cell invasion and colony formation were reduced by 20–70% (Supplementary Fig. 4C) and 15–70% (Supplementary Fig. 4D). These findings are in line with recently published data from other groups [21, 24–26].

Proteogenomic analyses reveal evidence for mitochondrial dysfunction upon *HOTAIRM1* knock-down

To characterize molecular mechanisms underlying the observed effects of *HOTAIRM1* knock-down on glioma cells in vitro, we performed RNA sequencing and MS-based proteome analyses on three siRNA-mediated *HOTAIRM1* knock-down models (U251, LN-229, and T98G) and their respective control-transfected cell lines (Supplementary Tables 2-3). A proteogenomic approach was taken for integrative bioinformatic evaluation of the RNA and protein data sets. First, preranked GeneSet Enrichment Analysis (GSEA) was performed using the t-statistic from the T-test for both the RNA sequencing data (see Supplementary Table 4 and 5 for the lists of positively or negatively enriched genesets identified by RNA sequencing, respectively) and the mass spectrometry data (see Supplementary Table 6 and 7 for the lists of positively or negatively enriched genesets identified by proteome analyses, respectively). The GSEA output was then visualized in cytoscape. The overlapping nodes that were similarly enriched in both datasets contain genesets involved in mRNA processing and translation, as well as mitochondrial translation and mitochondrial membrane function, suggesting that *HOTAIRM1* knock-down in glioma cells interferes with mitochondrial and translational functions (Fig. 3A, Supplementary Table 8).

To verify potential mitochondrial dysfunction upon *HOTAIRM1* knock-down, we performed immunofluorescent dihydroethidium (HEt) and MitoSOX stainings to measure cytoplasmic and mitochondrial superoxide levels, which is a type of reactive oxygen species (ROS), as an indicator of mitochondrial dysfunction [43]. *HOTAIRM1* knock-down cells showed increased ROS levels

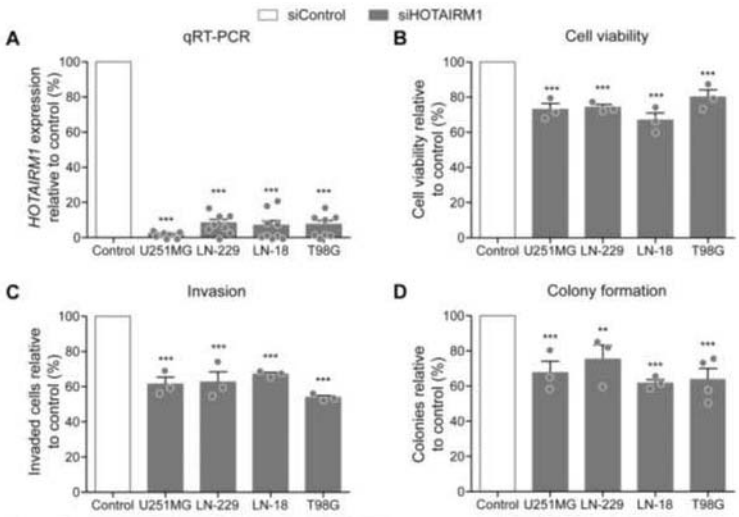


Fig. 2 *HOTAIRM1* knock-down decreases oncogenic features in glioblastoma cell lines. siRNA-mediated knock-down was achieved using siPOOLS (siTOOLs Biotech, Planegg, Germany). **A** qRT-PCR was performed using TaqMan probes against *HOTAIRM1* or *phosphoglycerate kinase 1* (*PGK1*) as a housekeeping control gene. Following *HOTAIRM1* knock-down, the four investigated glioma cell lines showed reduced cell viability as determined with the CellTiter-Glo assay (**B**), reduced invasiveness measured in Boyden chamber assays (**C**), and, finally, decreased clonogenicity as determined by colony formation assays after seeding cells at a density of 500 (U251MG and LN-18) to 1000 (LN-229 and T98G) cells per 10 cm dish (**D**). White bars indicate the results of the respective control-transfected cells set to 100%. Filled bars are results obtained with *HOTAIRM1* knock-down cells. siControl: cells transfected with non-target siPOOLS; siHOTAIRM1: cells transfected with siPOOLS against *HOTAIRM1*. Two-way ANOVA was used for statistical analyses; mean ± SEM, ***$p < 0.001$, **$p < 0.01$. $n = 4$ independent experiments for the colony formation assays for T98G cell line, $n = 3$ independent experiments per cell line and assay.

compared to control cells (Fig. 3B), a finding in line with the proteogenomic results suggesting deficient mitochondrial function in *HOTAIRM1* knock-down glioma cells. Further support for increased ROS levels as relevant driver of cellular effects caused by *HOTAIRM1* knock-down was obtained by treatment of LN-229 and LN-18 glioma cells with N-acetyl cysteine (NAC), a ROS scavenger, which showed that NAC treatment rescued the decrease in colony formation caused by *HOTAIRM1* knock-down (Fig. 3C, Supplementary Fig. 5).

HOTAIRM1 knock-down sensitizes glioblastoma cells to radiation in vitro and in vivo

Since radiation sensitivity has been associated with intracellular ROS levels [44] and since radiotherapy is an essential part of glioblastoma treatment, we investigated whether altered levels of *HOTAIRM1* affect radiosensitivity of glioblastoma cells. *HOTAIRM1* knock-down and control LN-229, SF126, and LN-18 cells were irradiated with either 2 or 4 Gy and colony formation capacity was evaluated in relation to non-irradiated cells. After 21 days, surviving colonies revealed a radiation dose-dependent significant decrease in colony formation compared to non-irradiated cells (Fig. 4A–C). *HOTAIRM1* knock-down caused reduced colony formation capacity of glioma cells already in non-irradiated cells (Fig. 2D; Supplementary Fig. 4D). However, we observed an additional, dose-dependent decrease in colony formation of *HOTAIRM1* knock-down glioma cells after irradiation compared to irradiated control-transfected glioma cells (Fig. 4A–C) which was not caused by an alteration in *HOTAIRM1* expression levels (Supplementary Fig. 6).

To validate the radiosensitizing effect of *HOTAIRM1* knock-down in vivo, we investigated an orthotopic xenograft glioma model using LN-229 control and *HOTAIRM1* knock-down cells. Mice inoculated intracerebrally with LN-229 control or *HOTAIRM1* knock-down cells were evaluated for orthotopic tumor growth and survival either with or without single irradiation with 12 Gy at day 15 post tumor cell transplantation. *HOTAIRM1* knock-down did not alter overall survival when mice were not irradiated (Fig. 4D), with control and *HOTAIRM1* knock-down LN-229-bearing mice exhibiting median overall survivals of 43 and 44.5 days, respectively. This finding contrasts with a recent study showing reduced in vivo tumor growth of U87MG glioma cells following *HOTAIRM1* knock-down [21], which might be related to the different models and experimental conditions. Importantly, however, we found that *HOTAIRM1* knock-down LN-229-bearing mice survived significantly longer following radiotherapy when compared to mice transplanted with control-transfected LN-229, as indicated by median overall survivals of 80 versus 38 days, respectively (Fig. 4E).

Transglutaminase 2 (*TGM2*) is down-regulated upon *HOTAIRM1* knock-down in glioblastoma cells

Mass spectrometry-based proteome analyses using the stable LN-229 *HOTAIRM1* knock-down model detected 16 proteins that were upregulated and 12 proteins that were down-regulated upon *HOTAIRM1* knock-down (Supplementary Tables 9–11). Transglutaminase 2 (*TGM2*) was detected as one of the 12 proteins that were down-regulated in *HOTAIRM1* knock-down LN-229 cells and also was significantly correlated with *HOTAIRM1* in the TCGA glioblastoma patient tissues (Supplementary Fig. 7A). Interestingly, *TGM2* has been shown to play a role in mitochondrial function [45] and cancer therapy resistance [46, 47]. Reduced *TGM2* mRNA and protein levels in *HOTAIRM1* stable knock-down LN-229, as well as other glioblastoma cell lines (U87MG, LN-18, and SF126) relative to control-transfected cells, were confirmed by RT-qPCR (Fig. 5A) and Western blotting (Fig. 5B–D). We corroborated reduced TGM2 expression upon siRNA-based knockdown of *HOTAIRM1*

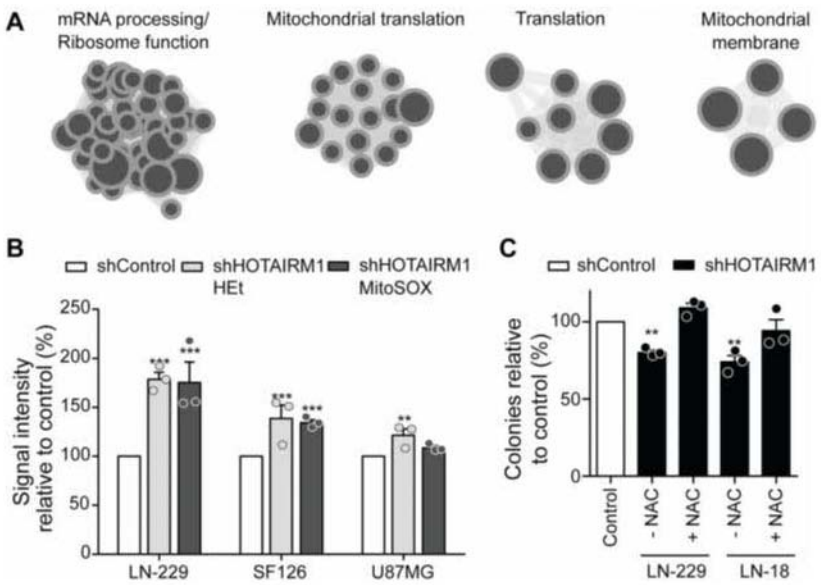


Fig. 3 Knock-down of *HOTAIRM1* results in mitochondrial dysfunction and increased reactive oxygen species (ROS). A Merged GSEA of RNA sequencing and proteomics data showing overlapping geneset clusters related to mRNA processing/ribosome function, mitochondrial translation, translation and mitochondrial membrane (see supplementary tables 2–5 for the list of the genesets). **B** Quantification of HEt (general superoxide indicator) and MitoSox (mitochondrial superoxide indicator) staining performed on stable LN-229, SF126, and U87MG *HOTAIRM1* knock-down cells and respective controls ($n = 3$). Shown are ROS levels normalized to control cells. **C** Results of colony formation assays 21 days post antioxidant NAC treatment in stable LN-229 and LN-18 *HOTAIRM1* knock-down and control cells ($n = 3$). Two-way ANOVA was used for statistical analyses; mean ± SEM, $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

in LN-229, T98G, and U87MG, while U251MG showed no detectable expression on protein level of the proposed candidate (Supplementary Fig. 7B).

siRNA-mediated *TGM2* knock-down mimics the phenotype of *HOTAIRM1* knock-down

To investigate functional roles of reduced *TGM2* expression in glioblastoma cells, we performed siRNA-mediated knock-down of *TGM2* in LN-229 and SF126 glioma cells (Fig. 6A). Similar to *HOTAIRM1* knock-down, *TGM2* knock-down resulted in significantly reduced cell viability (Fig. 6B), cell invasion (Fig. 6C), and colony formation (Fig. 6D). Although *TGM2* and *HOTAIRM1* expression is positively correlated, we did not observe the same survival benefit for *HOTAIRM1* (Supplementary Fig. 8A–D), as the strongest prognostic indicator observed was *MGMT* promoter methylation status (Supplementary Fig. 8E). The in vitro data are in line with studies reporting on tumor-promoting functions of TGM2 in other cancer models [48, 49] and suggest regulation of *TGM2* by *HOTAIRM1* as a putative mechanism driving glioma aggressiveness However, *TGM2* expression is not regulated by promoter methylation in patient samples, as analysis of glioblastoma methylation data showed that *HOTAIRM1* promoter is unmethylated in all samples, independently of high or low *HOTAIRM1* expression (Supplementary Fig. 7C).

Potential regulation of *TGM2* by *HOTAIRM1* via sponging of *hsa-miR-17-5p*

HOTAIRM1 has recently been proposed to function as a sponge for several miRNAs [24–26, 50–54] including *miR-17-5p* [50, 53], *miR-129-5p*, and *miR-495-3p* [24]. Therefore, we performed in silico analyses to determine putative miRNA binding sites that are shared between *HOTAIRM1* and transcripts of the 12 proteins with significantly reduced expression in *HOTAIRM1* knock-down LN-229 cells. The number of candidate miRNAs with putative binding sites in *HOTAIRM1* and any of the 12 candidate gene transcripts was further narrowed down by filtering for those miRNAs whose expression was inversely correlated to the expression of the 12 genes in the TCGA mRNA and miRNA glioblastoma data set (accession: phs000178.v10.p8.c1). Together, these in silico analyses revealed 15 miRNAs targeting *HOTAIRM1* and at least one of the genes of interest (Supplementary Table 12). Interestingly, 14 of the 15 miRNAs had predicted binding sites in both *TGM2* and *HOTAIRM1* (Supplementary Table 12) which includes members of the miR17-92 cluster, specifically *hsa-miR-17-5p* that has been reported to interact with *HOTAIRM1* in gastric and colorectal tumor entities [50, 53]. We found that low levels of *hsa-miR-17-5p* expression are associated with shorter survival of glioblastoma patients in the TCGA cohort (Supplementary Fig. 9A) and that expression of *hsa-miR-17-5p* is inversely correlated with *TGM2* expression in this cohort set (Supplementary Fig. 9B), however, not with *HOTAIRM1* expression (Supplementary Fig. 9C). To support these data, we performed a miR-17-5p overexpression experiment whereby miR-17-5p mimics were transfected in LN-229 cells. We were able to verify that upon increased miR-17-5p expression levels, *TGM2* expression was reduced (Supplementary Fig. 10A). To show that miR-17-5p regulates *TGM2* 3′UTR, we then performed a luciferase assay using either the predicted miR-17-5p binding site of the *TGM2* 3′UTR (17-5 wt) or a corresponding miR17-5p mutant binding site (17-5 mt) fused to Luciferase. This showed that addition of a miR-17-5p mimic led to a reduction of 17-5 wt, but not of 17-5 mt, driven Luciferase (Supplementary Fig. 10B), altogether indicating that *TGM2* transcript is a target of miR-17-5p. Since *HOTAIRM1* was shown to bind to miR-17-5p [50, 53], our data suggest that *HOTAIRM1* regulates *TGM2* expression possibly by controlling the availability of miR-17-5p.

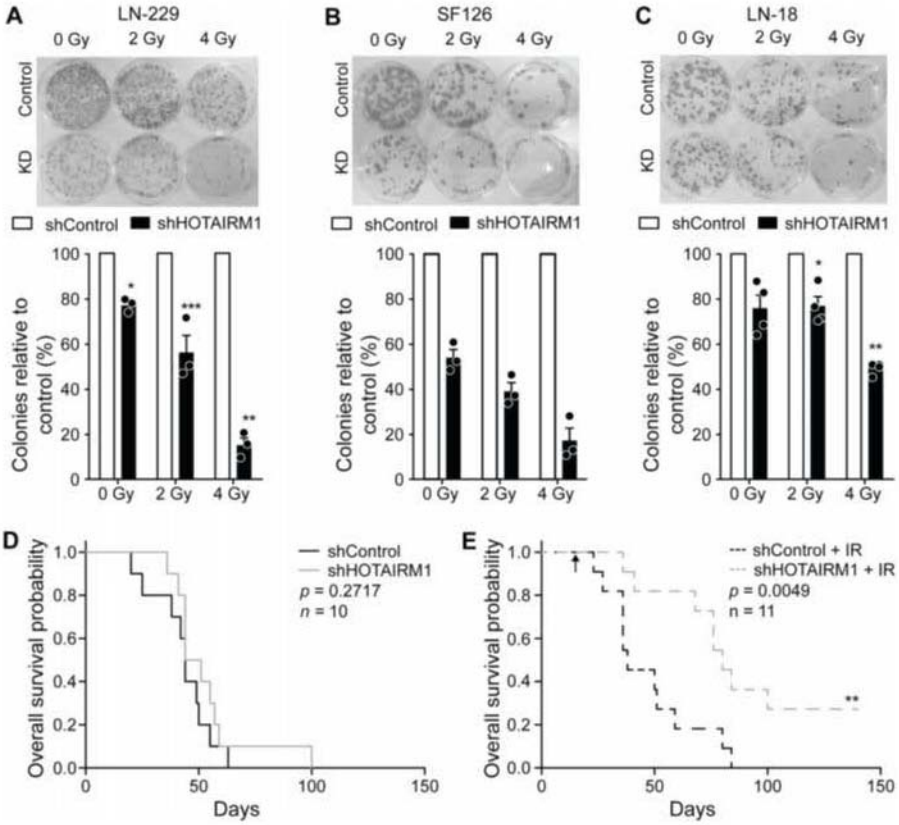


Fig. 4 *HOTAIRM1* knock-down sensitizes glioblastoma cells to radiation in vitro and in vivo. A, **B**, **C** Representative images and quantification of colony formation assays 21 days post irradiation at indicated doses for stable (**A**) LN-229, (**B**) SF126, and (**C**) LN-18 *HOTAIRM1* knock-down (KD) and control cells. Counts were normalized to the corresponding counts in isogenic controls at respective radiation dose (**D**, **E**) Kaplan–Meier survival plots of mice harboring either *HOTAIRM1* stable knock-down or control transfected LN-229 orthotopic xenografts not treated with radiation (**D**) and treated with 12 Gy radiation at day 15 (arrow) (**E**). Gray and black lines represent *HOTAIRM1* knock-down and controls, respectively. shControl: non-target shRNA; shHOTAIRM1: shRNA against *HOTAIRM1*. Two-way ANOVA was used for colony formation statistical analyses and log rank analysis for Kaplan-Meier survival plots; ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$.

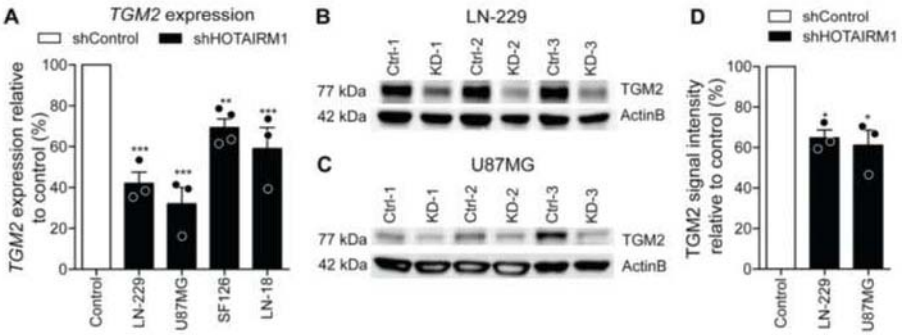


Fig. 5 *TGM2* expression is correlated with *HOTAIRM1* expression. A qRT-PCR for *TGM2* expression was performed in control and stable *HOTAIRM1* knock-down LN-229, U87MG, SF126, and LN-18 cell lines. White bar indicates the results of the respective control cells set to 100%. Note that *HOTAIRM1* knock-down significantly reduces *TGM2* mRNA levels. **B–D** Western blotting analysis of TGM2 protein expression in control versus stable *HOTAIRM1* knock-down LN-229 (**B**) and U87MG (**C**) cell lines. Shown are three independent experiments. Beta-actin (ActinB) was used as a loading control. **D** Quantification of TGM2 protein expression by western blotting analysis in control versus stable *HOTAIRM1* knock-down LN-229 and U87MG cell lines. White bar indicates the results of the respective control cells set to 100%. Two-way ANOVA was used for statistical calculation for qRT-PCR and Student's t test was used for statistical analysis. Control: non-target shRNA; KD: shRNA-mediated *HOTAIRM1* knock-down; mean ± SEM, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. $n = 3$.

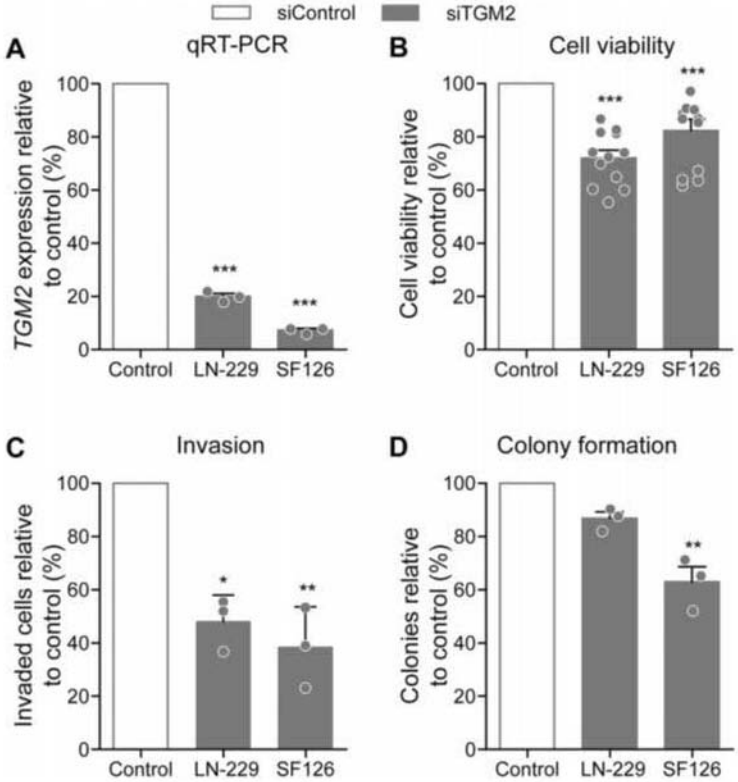


Fig. 6 *TGM2* knock-down decreases oncogenic features in glioblastoma cell lines. A qRT-PCR analysis confirms siRNA-mediated knock-down of TGM2 in LN-229 and SF126 glioma cells. Shown are *TGM2* mRNA levels normalized to *PGK1* mRNA levels relative to control transfected cells set to 100%. Results of determination of cell viability using CellTiter-Glo assays (**B**), cell invasion using Boyden chamber assays (**C**), and colony formation propensity using colony formation assays (**D**) in control versus stable *HOTAIRM1* knock-down LN-229 and SF126 cell lines. White bars represent control cells normalized to 100%, filled bars represent *TGM2* knock-down cells. Two-way ANOVA was used for statistical analyses; mean ± SEM, $***p < 0.001$, $**p < 0.01$. $n = 3$.

DISCUSSION

We have collected compelling evidence from several independent data sets that high *HOTAIRM1* expression is linked to clinical aggressiveness and shorter survival of glioblastoma patients and that gene copy number gain is a likely cause of increased *HOTAIRM1* expression levels in glioblastoma. After targeting *HOTAIRM1* expression in glioblastoma cell lines, the oncogenic potential of these cells was diminished and RNA sequencing and mass spectrometry data suggested impaired mitochondrial function. As it has been shown that cells are more sensitive after temozolomide treatment [24], we decided to focus on the novel finding of mitochondiral disfunction induced upon *HOTAIRM1* deficiency. Not only did we validate this finding in vitro, but our data indicate that high expression of *HOTAIRM1* supports radio-resistance of glioblastoma cells, which in turn may contribute to shorter patient survival as seen in our in vivo model. Since radiation sensitivity has been associated with intracellular ROS levels [44], the modulation of ROS levels by *HOTAIRM1* shown in our study suggests *HOTAIRM1*-mediated reduction of intracellular ROS as a potential mechanism contributing to glioma radioresistance.

To determine additional factors implicated in the *HOTAIRM1* mode of action, stable control, and *HOTAIRM1* knockdown cell lines have been profiled by proteomic analysis, pointing to *TGM2* as a potential mediator of radioresistance. TGM2 is localized in mitochondria as well as in the cytoplasm, endoplasmic reticulum, and plasma membranes [55]. The function of TGM2 in mitochondria is an emerging field [45] and data indicate that TGM2 plays a role in metabolism and mitochondrial respiration [56]. Our data shows that *HOTAIRM1* promotes *TGM2* expression in glioblastoma cells, which is related to miRNA-mediated mechanisms implicating *hsa-miR-17-5p*. Our analysis predicted binding sites for *hsa-miR-17-5p* in both *HOTAIRM1* and *TGM2* mRNAs. This microRNA, which has been reported to be upregulated by irradiation in glioblastoma [57], has been linked to glioma recurrence [58]. In addition, *TGM2* confers radio-resistance in different types of cancer cells [47, 59]. Collectively, these data suggest that *HOTAIRM1* may promote glioma growth and therapy resistance by sponging *hsa-miR-17-5p*, and thereby increasing *TGM2* transcript and protein levels. Thus, in addition to the epigenetic modulation of *HOXA1* and the sponging *hsa-miR-129-5p* and *hsa-miR-495-3p*, sponging of *hsa-miR-17-5p* (and potentially other *TGM2*-binding miRNAs) by *HOTAIRM1* may cause increased *TGM2* mRNA and protein expression in glioblastoma. However, *HOTAIRM1* may also affect other down-stream targets as suggested by the absense of TGM2 expresion in U251MG, as the siRNA-based *HOTAIRM1* depletion had similar antitumoral effects in this model. Furthermore, we showed *HOTAIRM1*-mediated TGM2 depletion, which might occur due to hsa-miR-17-5p modulation at the translational level as *TGM2* mRNA levels were not affected in T98G, while a reduction in protein levels was detected.

In summary, we confirm and extend recent data implicating *HOTAIRM1* as an oncogenic lncRNA driving tumor growth, therapy resistance, and poor prognosis of glioblastoma. Moreover, our data suggest a novel role for *HOTAIRM1* in regulating mitochondrial function and ROS levels in glioblastoma cells by modulating expression of *TGM2*, potentially by functioning as a miRNA sponge.

REFERENCES

1. Ostrom QT, Cioffi G, Gittleman H, Patil N, Waite K, Kruchko C, et al. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2012–2016. Neuro Oncol. 2019;21:v1–v100.
2. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization classification of tumors of the central nervous system: a summary. Acta Neuropathol. 2016;131:803–20.
3. Brat DJ, Aldape K, Colman H, Figrarella-Branger D, Fuller GN, Giannini C, et al. cIMPACT-NOW update 5: recommended grading criteria and terminologies for IDH-mutant astrocytomas. Acta Neuropathol. 2020;139:603–8.
4. Reifenberger G, Wirsching HG, Knobbe-Thomsen CB, Weller M. Advances in the molecular genetics of gliomas - implications for classification and therapy. Nat Rev Clin Oncol. 2017;14:434–52.
5. Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N. Engl J Med. 2005;352:987–96.
6. Stupp R, Taillibert S, Kanner A, Read W, Steinberg D, Lhermitte B, et al. Effect of tumor-treating fields plus maintenance temozolomide vs maintenance temozolomide alone on survival in patients with glioblastoma: a randomized clinical trial. JAMA 2017;318:2306–16.
7. Weller M, van den Bent M, Tonn JC, Stupp R, Preusser M, Cohen-Jonathan-Moyal E, et al. European Association for Neuro-Oncology (EANO) guideline on the diagnosis and treatment of adult astrocytic and oligodendroglial gliomas. Lancet Oncol. 2017;18:e315–e29.
8. Shah AH, Graham R, Bregy A, Thambuswamy M, Komotar RJ. Recognizing and correcting failures in glioblastoma treatment. Cancer Invest. 2014;32:299–302.
9. Aldape K, Zadeh G, Mansouri S, Reifenberger G, von Deimling A. Glioblastoma: pathology, molecular mechanisms and markers. Acta Neuropathol. 2015; 129:829–48.
10. Hegi ME, Diserens AC, Gorlia T, Hamou MF, de Tribolet N, Weller M, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. N. Engl J Med. 2005;352:997–1003.
11. Weller M, Tabatabai G, Kastner B, Felsberg J, Steinbach JP, Wick A, et al. MGMT promoter methylation is a strong prognostic biomarker for benefit from dose-intensified temozolomide rechallenge in progressive glioblastoma: the DIRECTOR trial. Clin Cancer Res. 2015;21:2057–64.
12. Wick W, Weller M, van den Bent M, Sanson M, Weiler M, von Deimling A, et al. MGMT testing—the challenges for biomarker-based glioma treatment. Nat Rev Neurol. 2014;10:372–85.
13. Le Rhun E, Preusser M, Roth P, Reardon DA, van den Bent M, Wen P, et al. Molecular targeted therapy of glioblastoma. Cancer Treat Rev. 2019;80:101896.
14. Sigova AA, Mullen AC, Molinie B, Gupta S, Orlando DA, Guenther MG, et al. Divergent transcription of long noncoding RNA/mRNA gene pairs in embryonic stem cells. Proc Natl Acad Sci USA. 2013;110:2876–81.
15. Ayers D, Vandesompele J Influence of microRNAs and Long Non-Coding RNAs in Cancer Chemoresistance. Genes (Basel). 2017;8.
16. Sahu A, Singhal U, Chinnaiyan AM. Long noncoding RNAs in cancer: from function to translation. Trends Cancer. 2015;1:93–109.
17. Heery R, Finn SP, Cuffe S, Gray SG. Long Non-Coding RNAs: Key Regulators of Epithelial-Mesenchymal Transition, Tumour Drug Resistance and Cancer Stem Cells. Cancers (Basel). 2017;9.
18. Zhang X, Lian Z, Padden C, Gerstein MB, Rozowsky J, Snyder M, et al. A myelopoiesis-associated regulatory intergenic noncoding RNA transcript within the human HOXA cluster. Blood 2009;113:2526–34.
19. Chen Y, Wu JJ, Lin XB, Bao Y, Chen ZH, Zhang CR, et al. Differential lncRNA expression profiles in recurrent gliomas compared with primary gliomas identified by microarray analysis. Int J Clin Exp Med 2015;8:5033–43.
20. Zhang X, Sun S, Pu JK, Tsang AC, Lee D, Man VO, et al. Long non-coding RNA expression profiles predict clinical phenotypes in glioma. Neurobiol Dis. 2012;48:1–8.
21. Li Q, Dong C, Cui J, Wang Y, Hong X. Over-expressed lncRNA HOTAIRM1 promotes tumor growth and invasion through up-regulating HOXA1 and sequestering G9a/EZH2/Dnmts away from the HOXA1 gene in glioblastoma multiforme. J Exp Clin Cancer Res. 2018;37:265.
22. Shi T, Guo D, Xu H, Su G, Chen J, Zhao Z, et al. HOTAIRM1, an enhancer lncRNA, promotes glioma proliferation by regulating long-range chromatin interactions within HOXA cluster genes. Mol Biol Rep. 2020;47:2723–33.
23. Xie P, Li X, Chen R, Liu Y, Liu D, Liu W, et al. Upregulation of HOTAIRM1 increases migration and invasion by glioblastoma cells. Aging (Albany NY) 2020;13:2348–64.
24. Liang Q, Li X, Guan G, Xu X, Chen C, Cheng P, et al. Long non-coding RNA, HOTAIRM1, promotes glioma malignancy by forming a ceRNA network. Aging (Albany NY). 2019;11.
25. Hao Y, Li X, Chen H, Huo H, Liu Z, Chai E. Over-expression of long noncoding RNA HOTAIRM1 promotes cell proliferation and invasion in human glioblastoma by up-regulating SP1 via sponging miR-137. Neuroreport 2020;31:109–17.
26. Lin YH, Guo L, Yan F, Dou ZQ, Yu Q, Chen G. Long non-coding RNA HOTAIRM1 promotes proliferation and inhibits apoptosis of glioma cells by regulating the miR-873-5p/ZEB2 axis. Chin Med J (Engl) 2020;133:174–82.
27. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2 (-Delta Delta C (T)) Method. Methods 2001;25:402–8.
28. Zhang X, Weissman SM, Newburger PE. Long intergenic non-coding RNA HOTAIRM1 regulates cell cycle progression during myeloid maturation in NB4 human promyelocytic leukemia cells. RNA Biol. 2014;11:777–87.
29. Rajaraman S, Canjuga D, Ghosh M, Codrea MC, Sieger R, Wedekink F, et al. Measles virus-based treatments trigger a pro-inflammatory cascade and a distinctive immunopeptidome in glioblastoma. Mol Ther Oncolytics 2019;12:147–61.
30. Poschmann G, Seyfarth K, Besong Agbo D, Klafki HW, Rozman J, Wurst W, et al. High-fat diet induced isoform changes of the Parkinson's disease protein DJ-1. J Proteome Res. 2014;13:2339–51.
31. Vizcaino JA, Csordas A, Del-Toro N, Dianes JA, Griss J, Lavidas I, et al. 2016 update of the PRIDE database and its related tools. Nucleic Acids Res. 2016;44:11033.
32. Distelmaier F, Valsecchi F, Liemburg-Apers DC, Lebiedzinska M, Rodenburg RJ, Heil S, et al. Mitochondrial dysfunction in primary human fibroblasts triggers an adaptive cell survival program that requires AMPK-alpha. Biochim Biophys Acta 2015;1852:529–40.
33. Reifenberger G, Weber RG, Riehmer V, Kaulich K, Willscher E, Wirth H, et al. Molecular characterization of long-term survivors of glioblastoma using genome- and transcriptome-wide profiling. Int J Cancer 2014;135:1822–31.
34. Cancer Genome Atlas Research N. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature 2008;455:1061–8.
35. Gravendeel LA, Kouwenhoven MC, Gevaert O, de Rooi JJ, Stubbs AP, Duijm JE, et al. Intrinsic gene expression profiles of gliomas are a better predictor of survival than histology. Cancer Res. 2009;69:9065–72.
36. Bigner SH, Mark J, Burger PC, Mahaley MS Jr., Bullard DE, Muhlbaier LH, et al. Specific chromosomal abnormalities in malignant human gliomas. Cancer Res. 1988;48:405–11.
37. Murat A, Migliavacca E, Gorlia T, Lambiv WL, Shay T, Hamou MF, et al. Stem cell-related "self-renewal" signature and high epidermal growth factor receptor expression associated with resistance to concomitant chemoradiotherapy in glioblastoma. J Clin Oncol. 2008;26:3015–24.
38. Sturm D, Witt H, Hovestadt V, Khuong-Quang DA, Jones DT, Konermann C, et al. Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. Cancer Cell 2012;22:425–37.
39. Vital AL, Tabernero MD, Castrillo A, Rebelo O, Tao H, Gomes F, et al. Gene expression profiles of human glioblastomas are associated with both tumor cytogenetics and histopathology. Neuro Oncol. 2010;12:991–1003.
40. Chakravarti A, Chakladar A, Delaney MA, Latham DE, Loeffler JS. The epidermal growth factor receptor pathway mediates resistance to sequential administration of radiation and chemotherapy in primary human glioblastoma cells in a RAS-dependent manner. Cancer Res. 2002;62:4307–15.
41. Mazzoleni S, Politi LS, Pala M, Cominelli M, Franzin A, Sergi Sergi L, et al. Epidermal growth factor receptor expression identifies functionally and molecularly distinct tumor-initiating cells in human glioblastoma multiforme and is required for gliomagenesis. Cancer Res. 2010;70:7500–13.
42. Tachon G, Masliantsev K, Rivet P, Petropoulos C, Godet J, Milin S, et al. Prognostic significance of MEOX2 in gliomas. Mod Pathol. 2019;32:774–86.
43. Bhat AH, Dar KB, Anees S, Zargar MA, Masood A, Sofi MA, et al. Oxidative stress, mitochondrial dysfunction and neurodegenerative diseases; a mechanistic insight. Biomed Pharmacother. 2015;74:101–10.
44. An Z, Yu JR, Park WY. Rosiglitazone enhances radiosensitivity by inhibiting repair of DNA damage in cervical cancer cells. Radiat Environ Biophys. 2017;56:89–98.
45. Lai TS, Lin CJ, Wu YT, Wu CJ. Tissue transglutaminase (TG2) and mitochondrial function and dysfunction. Front Biosci (Landmark Ed) 2017;22:1114–37.
46. Li C, Cai J, Ge F, Wang G. TGM2 knockdown reverses cisplatin chemoresistance in osteosarcoma. Int J Mol Med 2018;42:1799–808.
47. Yin J, Oh YT, Kim JY, Kim SS, Choi E, Kim TH, et al. Transglutaminase 2 inhibition reverses mesenchymal transdifferentiation of glioma stem cells by regulating C/EBPbeta signaling. Cancer Res. 2017;77:4973–84.

48. Li Z, Xu X, Bai L, Chen W, Lin Y. Epidermal growth factor receptor-mediated tissue transglutaminase overexpression couples acquired tumor necrosis factor-related apoptosis-inducing ligand resistance and migration through c-FLIP and MMP-9 proteins in lung cancer cells. J Biol Chem. 2011;286:21164–72.
49. Hidaka H, Seki N, Yoshino H, Yamasaki T, Yamada Y, Nohata N, et al. Tumor suppressive microRNA-1285 regulates novel molecular targets: aberrant expression and functional significance in renal cell carcinoma. Oncotarget. 2012;3:44–57.
50. Lu R, Zhao G, Yang Y, Jiang Z, Cai J, Zhang Z, et al. Long noncoding RNA HOTAIRM1 inhibits cell progression by regulating miR-17-5p/ PTEN axis in gastric cancer. J Cell Biochem. 2019;120:4952–65.
51. Zheng M, Liu X, Zhou Q, Liu G. HOTAIRM1 competed endogenously with miR-148a to regulate DLGAP1 in head and neck tumor cells. Cancer Med. 2018.
52. Xiao Y, Yan X, Yang Y, Ma X. Downregulation of long noncoding RNA HOTAIRM1 variant 1 contributes to osteoarthritis via regulating miR-125b/BMPR2 axis and activating JNK/MAPK/ERK pathway. Biomed Pharmacother. 2019;109:1569–77.
53. Ren T, Hou J, Liu C, Shan F, Xiong X, Qin A, et al. The long non-coding RNA HOTAIRM1 suppresses cell progression via sponging endogenous miR-17-5p/ B-cell translocation gene 3 (BTG3) axis in 5-fluorouracil resistant colorectal cancer cells. Biomed Pharmacother. 2019;117:109171.
54. Chen ZH, Wang WT, Huang W, Fang K, Sun YM, Liu SR, et al. The lncRNA HOTAIRM1 regulates the degradation of PML-RARA oncoprotein and myeloid cell differentiation by enhancing the autophagy pathway. Cell Death Differ. 2017;24:212–24.
55. Bianchi N, Beninati S, Bergamini CM. Spotlight on the transglutaminase 2 gene: a focus on genomic and transcriptional aspects. Biochem J. 2018;475:1643–67.
56. Rodolfo C, Mormone E, Matarrese P, Ciccosanti F, Farrace MG, Garofano E, et al. Tissue transglutaminase is a multifunctional BH3-only protein. J Biol Chem. 2004;279:54783–92.
57. Chaudhry MA, Sachdeva H, Omaruddin RA. Radiation-induced micro-RNA modulation in glioblastoma cells differing in DNA-repair pathways. DNA Cell Biol. 2010;29:553–61.
58. Malzkorn B, Wolter M, Liesenberg F, Grzendowski M, Stuhler K, Meyer HE, et al. Identification and functional characterization of microRNAs involved in the malignant progression of gliomas. Brain Pathol. 2010;20:539–50.
59. Arienti C, Tesei A, Carloni S, Ulivi P, Romeo A, Ghigi G, et al. SLUG silencing increases radiosensitivity of melanoma cells in vitro. Cell Oncol (Dordr) 2013;36:131–9.

ACKNOWLEDGEMENTS

We thank Sarah Göbbels for her technical help and Thomas Hielscher for his support with statistical analyzes. This project was supported by a joint grant from the Deutsche Forschungsgemeinschaft (German Research Foundation) and the Swiss National Science Foundation to M.R., G.R., P.R., and M.W. (RE2857/2-1, RE938/4-1; SNF 310030E_170717). This work was also partly supported by the Deutsche Forschungsgemeinschaft (German Research Foundation) [RE2857/4-1, SCHE 656/2-1 (KFO 337)]. J.T.S. is supported by the German Cancer Consortium (DKTK), the Deutsche Forschungsgemeinschaft (DFG) grant SI1549/3-1 (DFG/GRC-CRU337) and SI1549/4-1, and the German Cancer Aid (Grant no. 70112505; PIPAC consortium).

AUTHOR CONTRIBUTIONS

U.A., J.B., M.S., M.M., M.L., A.S., M.W. J.K.M.L., D.P., F.S., A.F. and F.D.M. performed the experiments. U.A. and D.P. analyzed the data and wrote the paper. D.P. performed all data analyses. J.B., G.R. and G.L. edited the manuscript. M.T.A., N.Q., A.M.W., L.B., V.M., S.S., C.H., L.K., A.H., C.M., C.B.K.T., A.S., U.F., K.S., S.F., J.T.S., F.D. and A.B. revised the article critically for important intellectual content. G.R., P.R., M.W. and M.R. provided guidance on experimental technology. M.R. is the corresponding author and designed the research. All authors read and approved the final manuscript.

FUNDING

Open Access funding enabled and organized by Projekt DEAL.

ETHICS STATEMENT

All animal experiments were performed in accordance with the guidelines of Swiss federal law on animal protection. No human subjects were part of this study.

COMPETING INTERESTS

G.R. has received honoraria for advisory boards from Abbvie. M.W. has received research grants from Abbvie, Adastra, Bristol Meyer Squibb (BMS), Dracen, Merck, Sharp & Dohme (MSD), Merck (EMD), Novocure, Piqur and Roche, and honoraria for lectures or advisory board participation or consulting from Abbvie, Basilea, Bristol Meyer Squibb (BMS), Celgene, Merck, Sharp & Dohme (MSD), Merck (EMD), Novocure, Orbus, Roche and Tocagen. P.R. has received honoraria for advisory board participation and lectures from Bristol-Myers Squibb, Covagen, Debiopharm, Medac, Merck, MSD, Novocure, QED, Roche and Virometix. C.M. is supported by the Heinz Ansmann foundation. JTS reports the following disclosures: Bristol-Myers Sqibb, Celgene, Roche (Research Funding); AstraZeneca, Bristol-Myers Squibb, Celgene, Immunocore, Novartis, Roche, Shire (Consulting or advisory role); AstraZeneca, Aurikamed, Baxalta, Bristol Myers Squibb, Celgene, Falk Foundation, iomedico, Immunocore, Novartis, Roche, Shire (honoraria); minor equity in iTheranostics and Pharma15 (<3%) and member of the Board of Directors for Pharma15, all outside the submitted work. The other authors report no conflict of interest.

ADDITIONAL INFORMATION

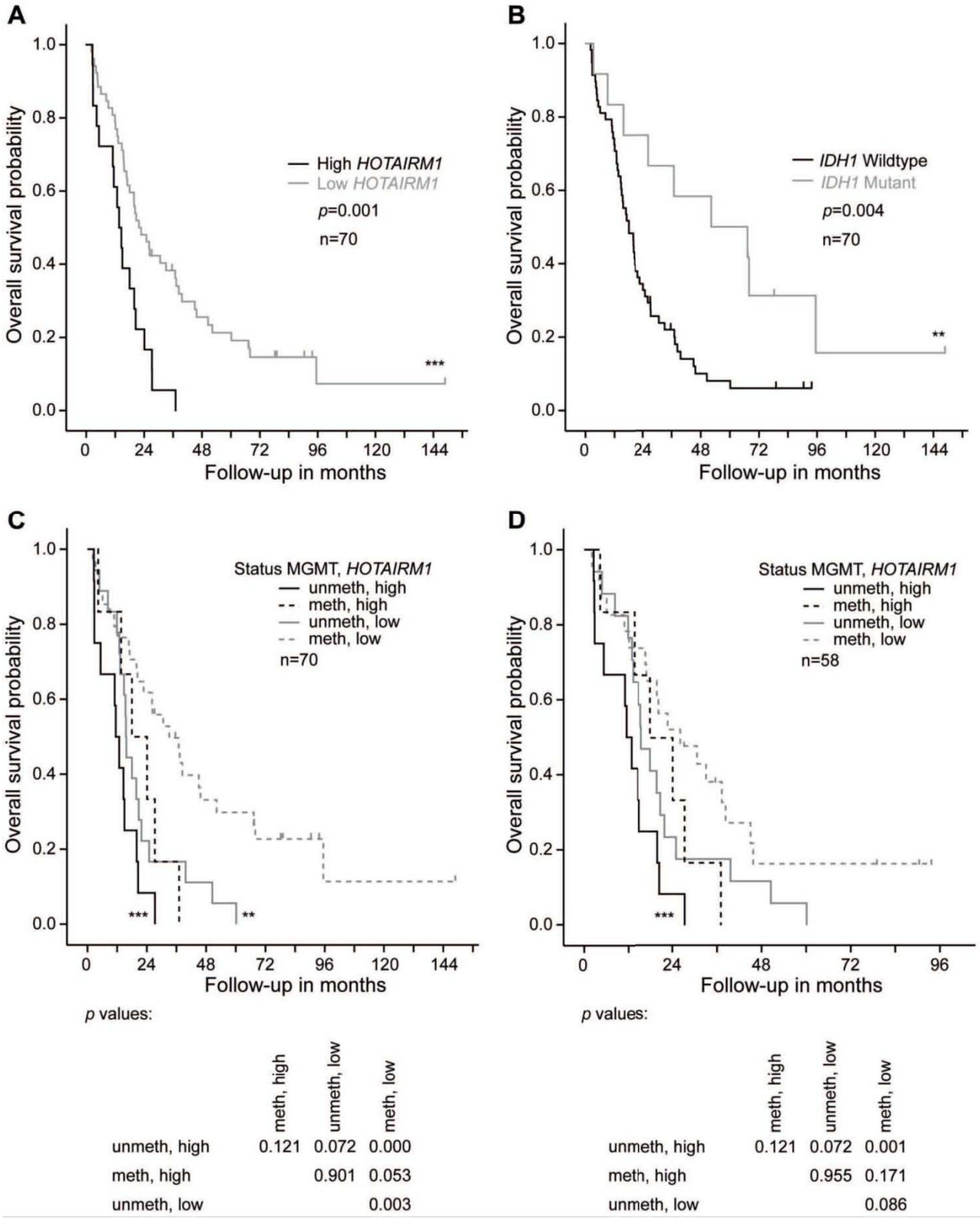
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41419-021-04146-0.

Correspondence and requests for materials should be addressed to Marc Remke.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Figure 1



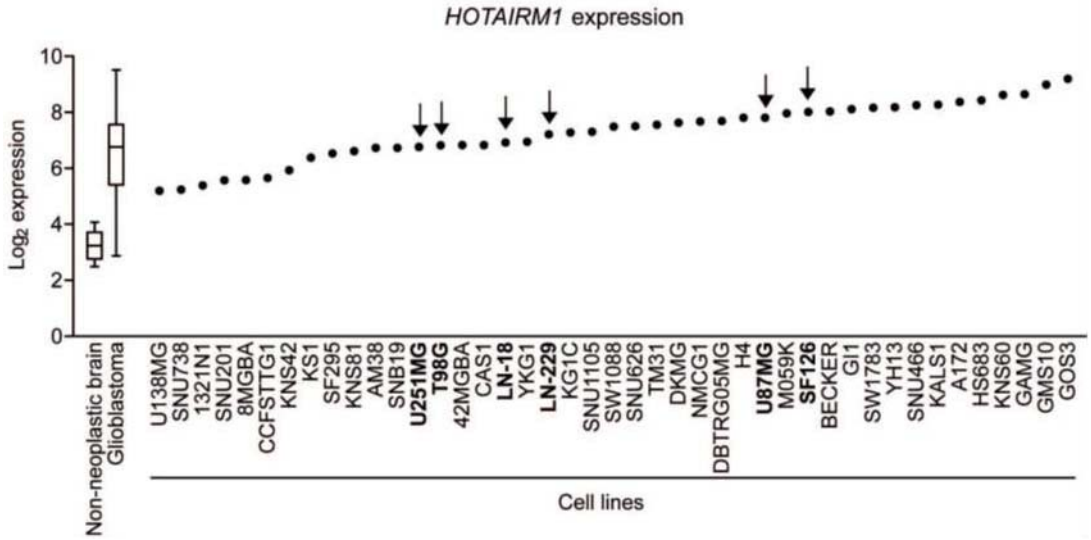
p values:

	meth, high	unmeth, low	meth, low
unmeth, high	0.121	0.072	0.000
meth, high		0.901	0.053
unmeth, low			0.003

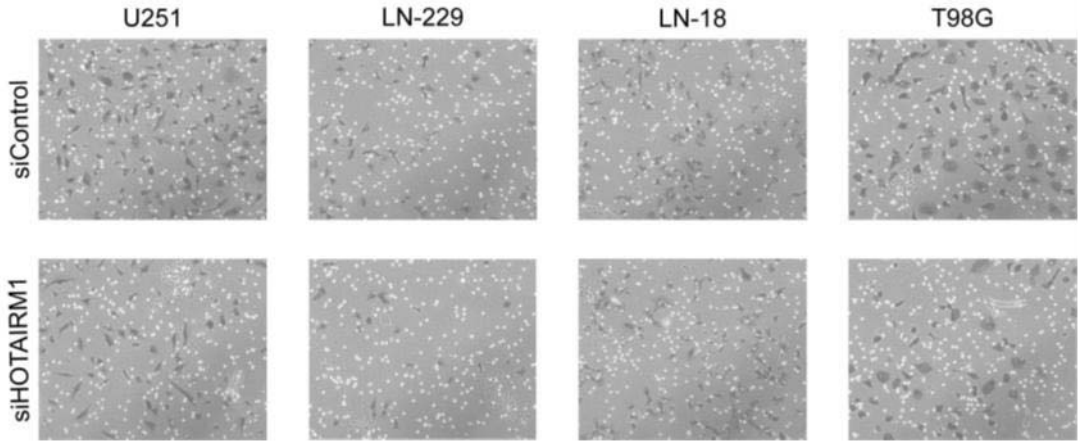
p values:

	meth, high	unmeth, low	meth, low
unmeth, high	0.121	0.072	0.001
meth, high		0.955	0.171
unmeth, low			0.086

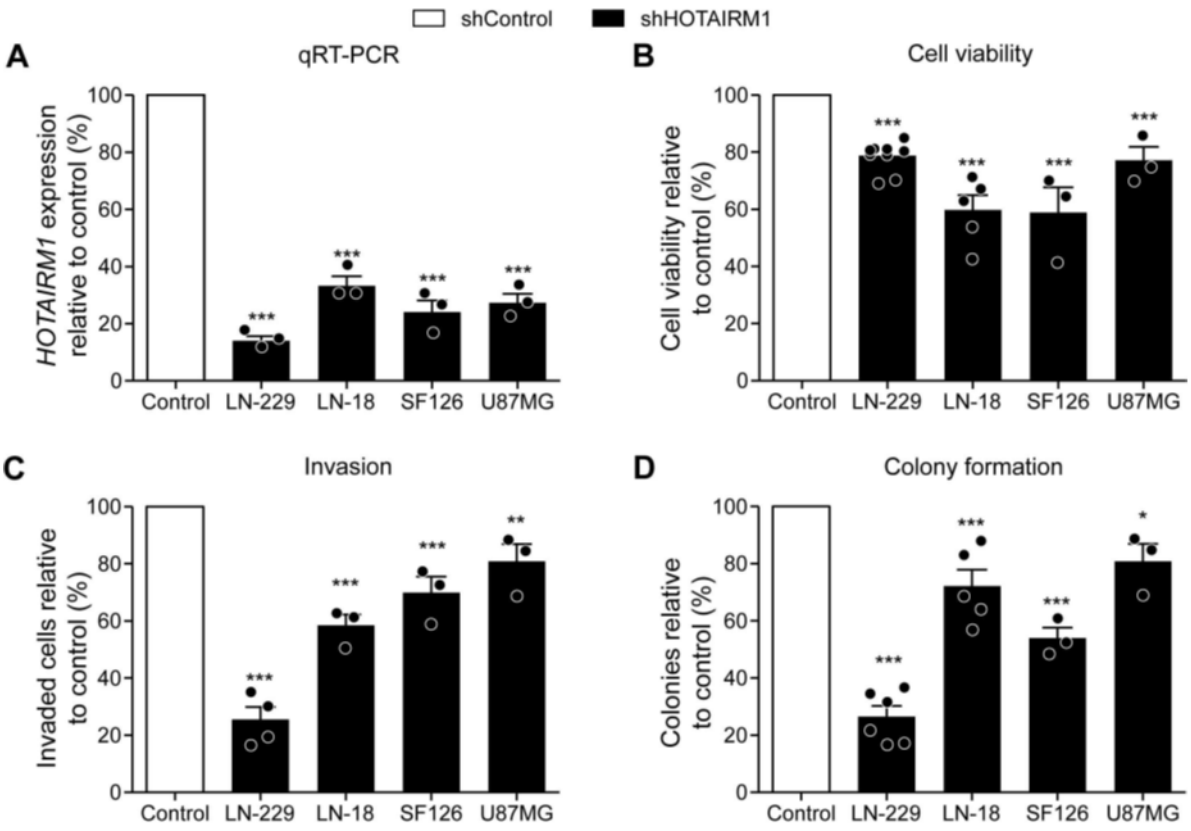
Supplementary Figure 2



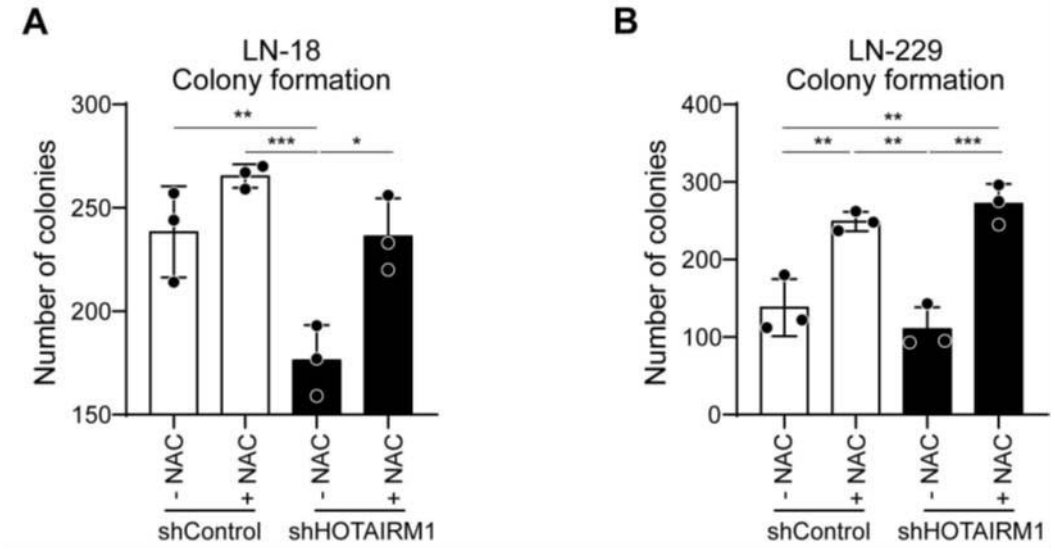
Supplementary Figure 3



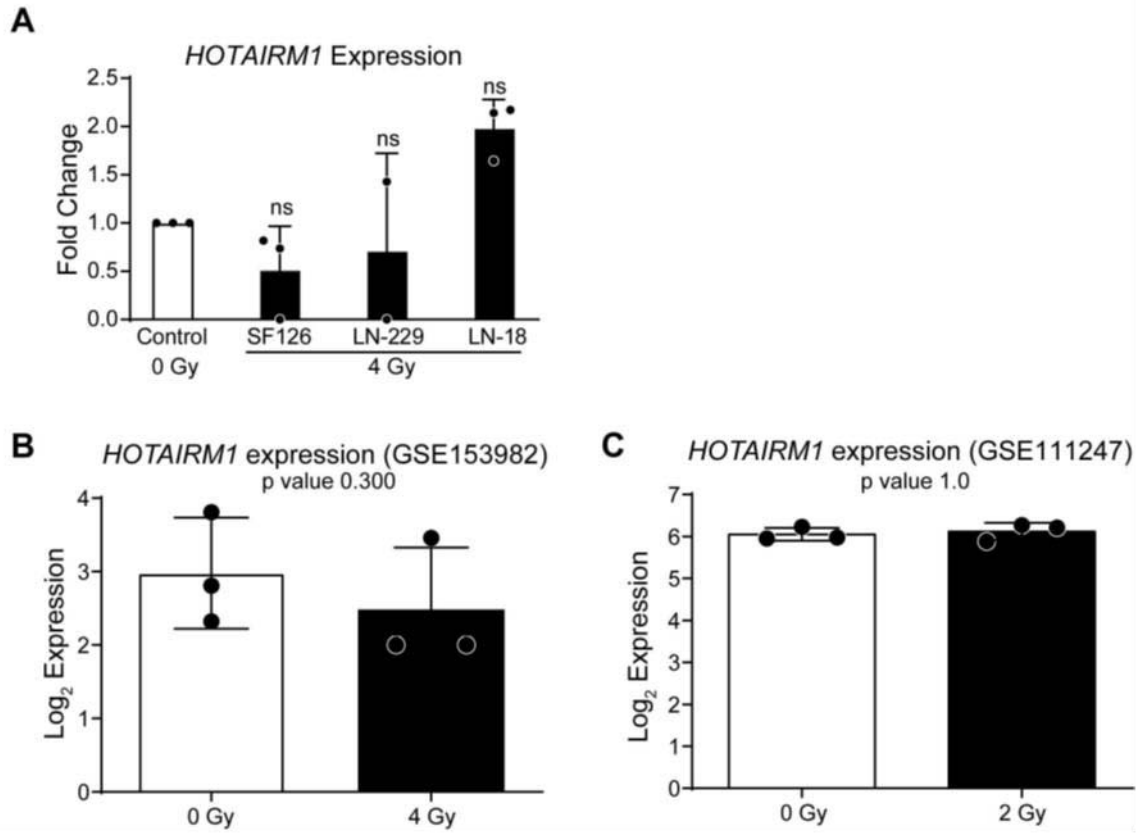
Supplementary Figure 4



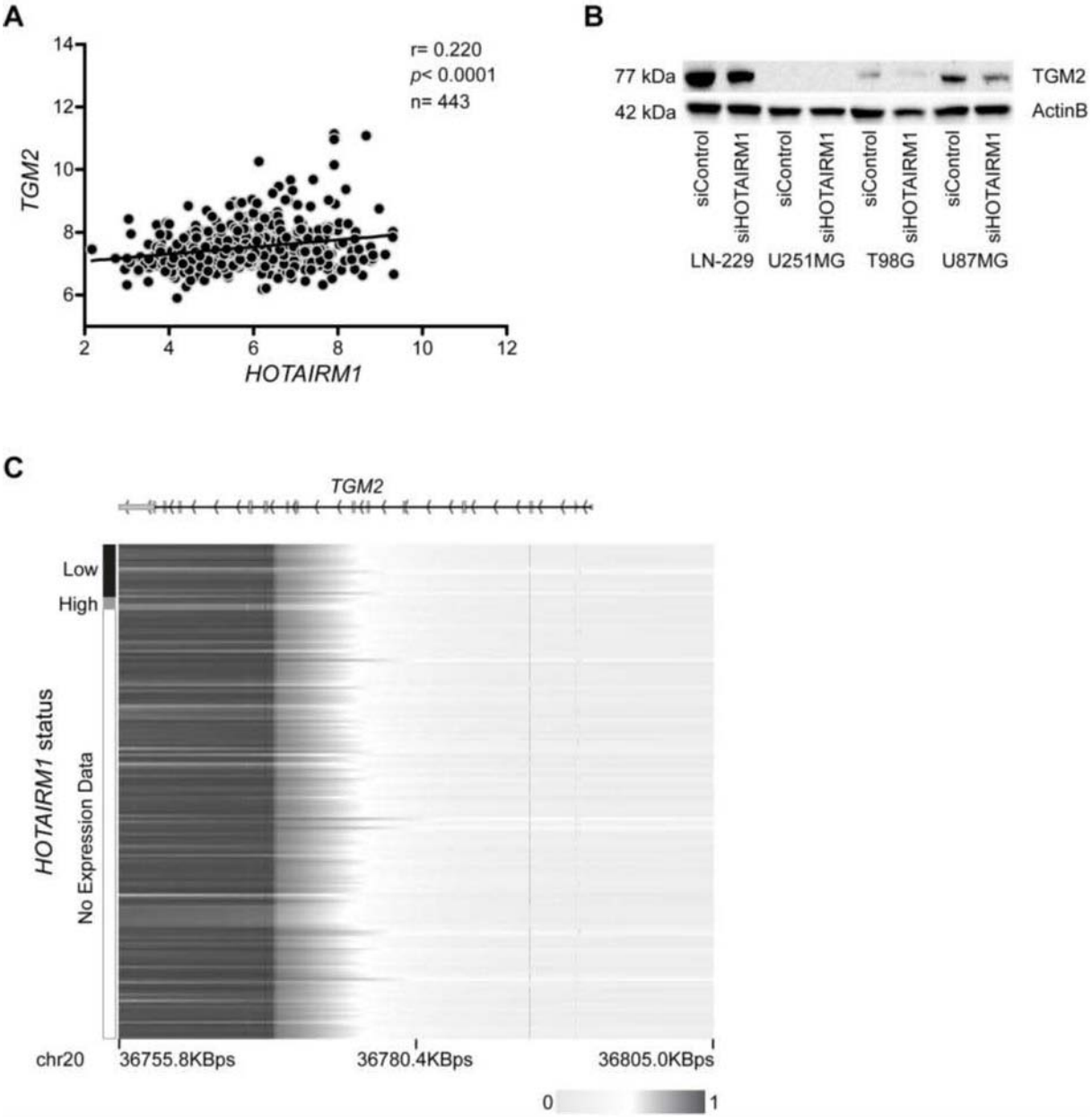
Supplementary Figure 5



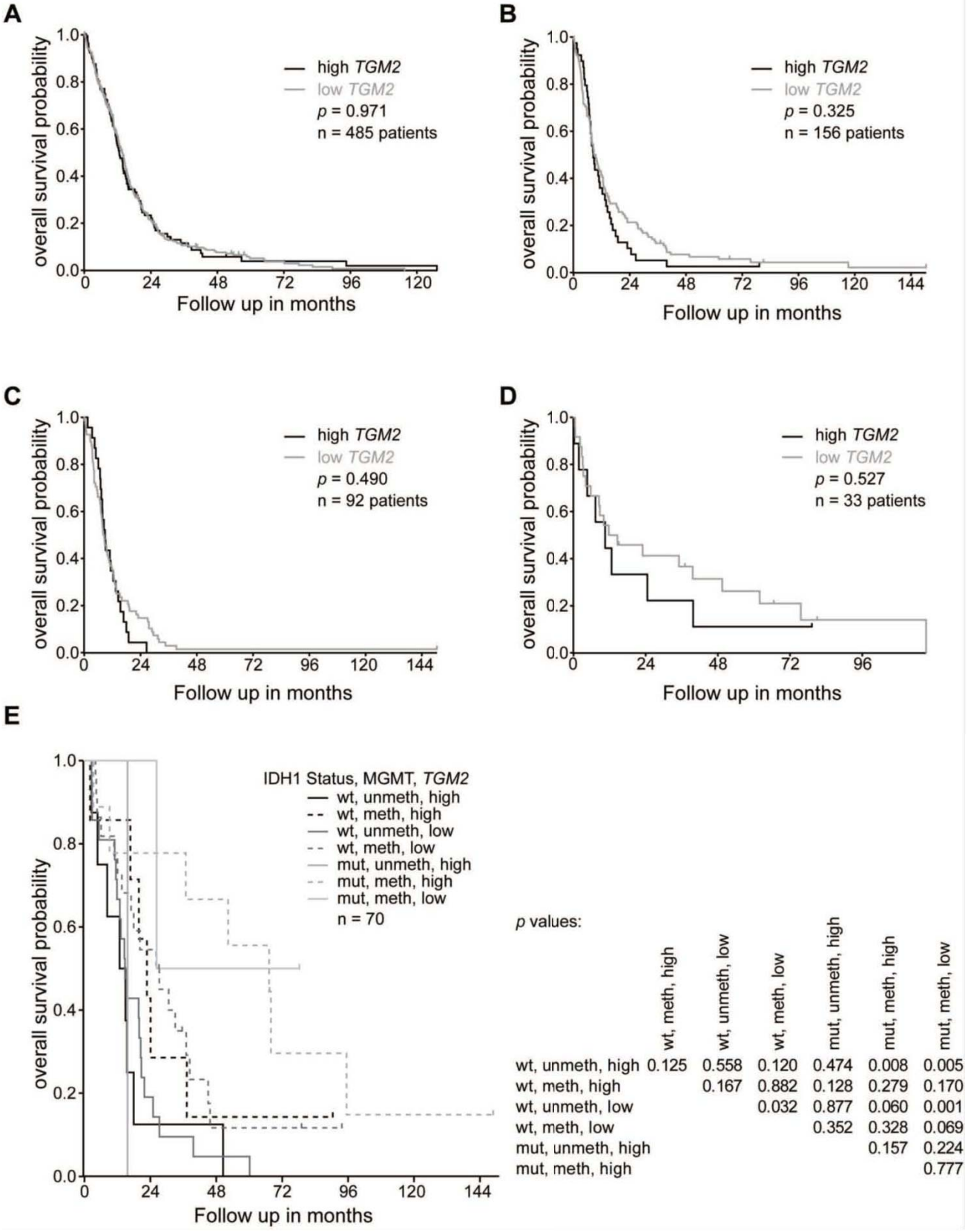
Supplementary Figure 6



Supplementary Figure 7



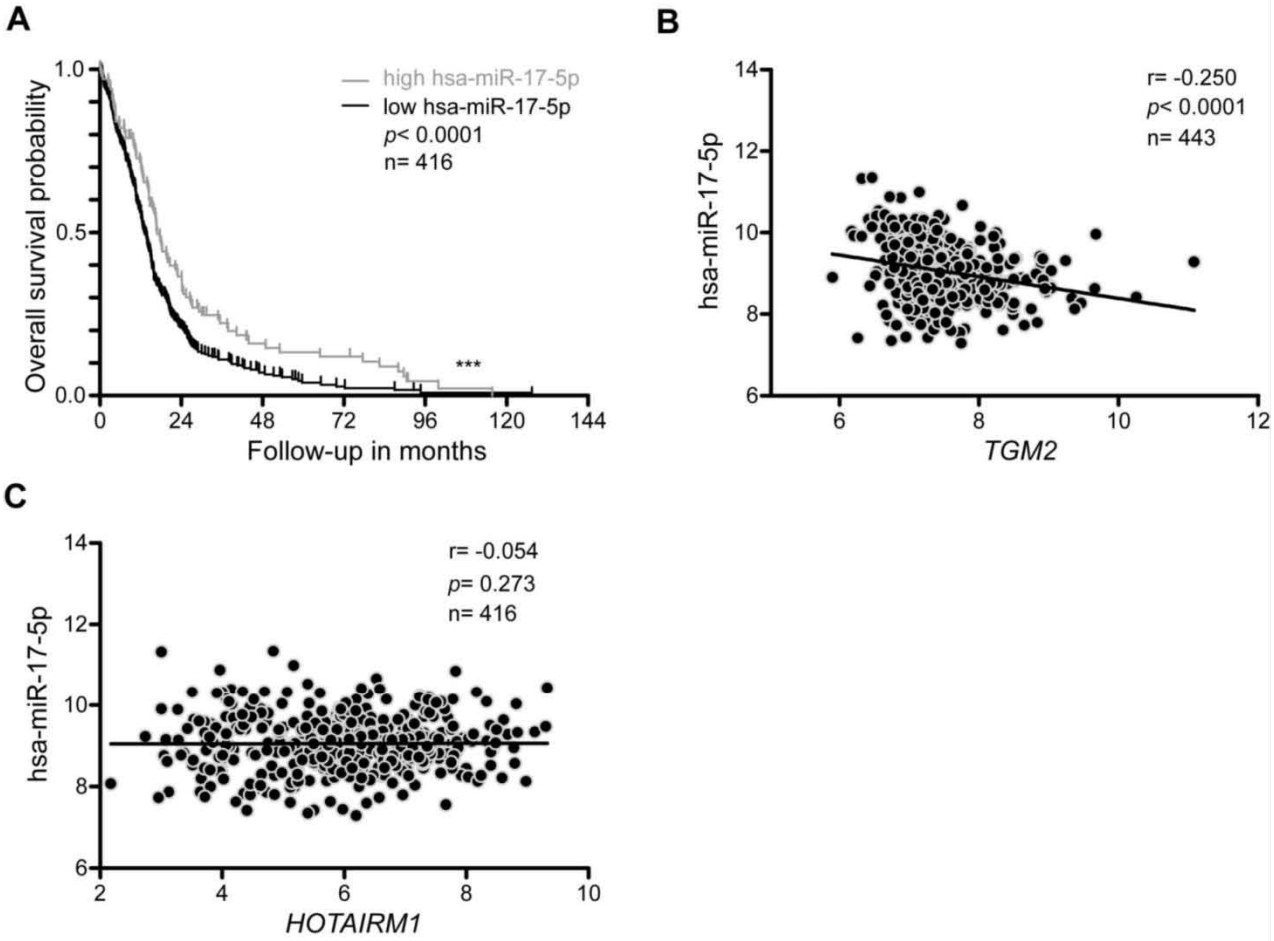
Supplementary Figure 8



p values:

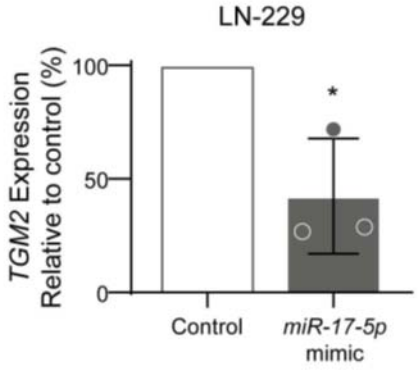
	wt, meth, high	wt, unmeth, low	wt, meth, low	mut, unmeth, high	mut, meth, high	mut, meth, low
wt, unmeth, high	0.125	0.558	0.120	0.474	0.008	0.005
wt, meth, high		0.167	0.882	0.128	0.279	0.170
wt, unmeth, low			0.032	0.877	0.060	0.001
wt, meth, low				0.352	0.328	0.069
mut, unmeth, high					0.157	0.224
mut, meth, high						0.777

Supplementary Figure 9

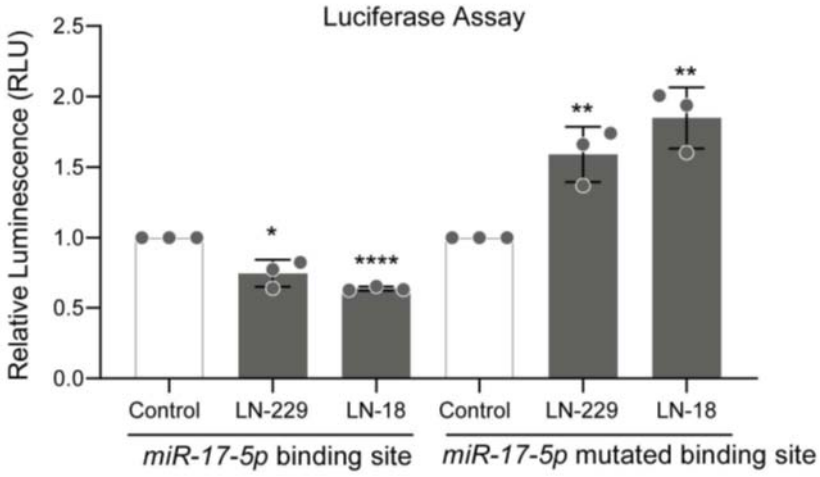


Supplementary Figure 10

A



B



Supplementary Figure Legends

Supplementary Figure 1

High *HOTAIRM1* expression is associated with shorter survival of glioblastoma patients. (**A**) Kaplan-Meier survival curves of 70 glioblastoma patients from the GGN cohort (1) stratified according to *HOTAIRM1* expression. Black and grey lines indicate patients whose tumors had high or low *HOTAIRM1* expression, respectively. Cut-off was determined by final quartile and log rank statistics were calculated. (**B**) Kaplan-Meier survival plot of the same patient cohort stratified according to *IDH* mutation status; black and grey lines indicate patients with IDH-wildtype or IDH-mutant tumors, respectively. (**C**) Kaplan-Meier survival plot of the patient cohort stratified according to *HOTAIRM1* expression (black and grey lines, high and low *HOTAIRM1*, repsectively) and *MGMT* protomer metylation (full and dashed lines, methylated vs unmethylated MGMT, respectively). (**D**) Kaplan-Meier survival plot of (**C**) after excluding the patients with IDH-mutant tumors.

Supplementary Figure 2

***HOTAIRM1* expression levels in various glioma cell lines are similar to those seen in glioblastoma tissue samples.** *HOTAIRM1* expression is shown using the Gravendeel *et al*. dataset (2) for non-neoplastic brain samples (n=8) and glioblastoma tissue samples (n=159), as well as, the Cancer Cell Line Encyclopedia (CCLE, GSE36133) glioma cell lines (n=43) profiled using the Affymetrix U133 Plus 2 array. Arrows and bold names indicate the glioblastoma cell lines used in this study.

Supplementary Figure 3

***HOTAIRM1* knock-down decreases invasiveness in established glioblastoma cell lines.** Invasiveness was measured using Boyden chamber assays. Representative images (10X) from one of three replicates of the Boyden chamber assays performed for each cell line are shown. siControl: cells transfected with non-target siRNA pool; siHOTAIRM1: cells transfected with siRNA pool against *HOTAIRM1.*

Supplementary Figure 4

Stable *HOTAIRM1* knock-down decreases oncogenic features in glioblastoma cell lines. Stable knock-down was achieved using shRNA against *HOTAIRM1*. (**A**) qRT-PCR was performed using TaqMan probes against *HOTAIRM1* or *phosphoglycerate kinase 1* (*PGK1*) (housekeeping gene). (**B**) Cell viability was measured using CellTiter-Glo assays. (**C**) Invasiveness was measured using Boyden chamber assays. (**D**) Colony formation assays were done by seeding cells at a density of 500 (U251MG and LN-18) to 1000 (LN-229, U87MG and SF126) cells into 10 cm dishes. Histogram bars are as follows: outline bars represent results obtained for control-transfected cells set to 100% and filled bars are results obtained for the respective glioma cells after *HOTAIRM1* knock-down relative to the control cells. shControl: cells transduced with non-target shRNA; shHOTAIRM1: cells transduced with shRNA against *HOTAIRM1*. Two-way ANOVA was used for statistical analyses; mean +/-SEM, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. n = 3.

Supplementary Figure 5

Increased reactive oxygen species after antioxidant treatment of *HOTAIRM1* knockdown compared to control. (**A-B**) Results of colony formation assays 21 days post antioxidant NAC treatment in stable LN-229 (**A**) and LN-18 (**B**) *HOTAIRM1* knock-down and control cells (n = 3) (non-normalized data for Figure 3). Two-way ANOVA was used for statistical analyses; mean +/-SEM, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Supplementary Figure 6

Radiation of established glioma cell lines does not alter *HOTAIRM1* expression levels. Expression of *HOTAIRM1* was measured in gliomablastoma cells lines by qRT-PCR (SF126, LN229 or LN18 (**A)**) or by Affymetrix U133P2 arrays in two separate datasets (HK-374 (**B)** and LNZ308 **(C)**) 48 hours post 2 or 4Gy radiation. Two-way ANOVA was used for statistical analyses; mean +/-SEM (**A**) or Mann-Whitney test, mean +/-SD (**B-C**).

Supplementary Figure 7

HOTAIRM1 and TGM2 expression levels are correlated. (**A**) Dot plot showing Pearson correlation of *TGM2* and *HOTAIRM1* expression in glioblastomas of the TCGA cohort (3). (**B**) Western blotting analysis of TGM2 protein expression in control versus transient *HOTAIRM1* knock-down of LN-229, U251MG, T98G and U87MG cell lines. Beta-actin (ActinB) was used as a loading control. (**C**) Heatmap showing the methylation status of the *TGM2* gene and probes up to 5kb upstream, and the relationship to *HOTAIRM1* expression, where available. Two-way ANOVA was used for statistical analyses; mean +/-SEM, * $p < 0.05$

Supplementary Figure 8

Prognostic role of *TGM2* expression in glioblastoma patient datasets. (**A** Overall survival plots of glioblastoma patients from TCGA (https://www.cancer.gov/tcga) and **(B)** Gravendeel *et al*. stratified according to high or low *TGM2* expression levels. Cut-off for high *TGM2* was determined by upper quartile and log rank statistics were calculated. (**C,D**) Overall survival of glioblastoma patients in the Gravendeel *et al*. cohort according to *TGM2* expression in *IDH1*-wildtype (**C**) and *IDH1*-mutant glioblastomas (**D**). (**E**) Overall survival of glioblastoma patients in the GGN cohort (1) stratified according to *TGM2* expression, *MGMT* promoter methylation status, and *IDH1* mutation status (wt: wild-type; mut: mutant; meth: methylated; unmeth: unmethylated). The table right of the Kaplan-Meier graph lists p-values for the individual subgroups. Log rank analysis for Kaplan-Meier survival plots**.**

Supplementary Figure 9

Decreased *hsa-miR-17-5p* expression associates with shorter survival of glioblastoma patients. (**A**) Kaplan-Meier plot illustrating overall survival of IDH-wildtype glioblastoma patients from the TCGA cohort (3) stratified according to *hsa-miR-17-5p* expression. Black and grey lines indicate patients whose tumors showed low and high *hsa-miR-17-5p* expression, respectively. Cut-off was determined by final quartile and log rank statistics were calculated. (**B**-**C**) Dot plots showing Pearson correlations of *hsa-miR-17-5p* and *TGM2* expression (**B**), and *hsa-miR-17-5p* and *HOTAIRM1* expression (**C**) in glioblastomas of the TCGA cohort[3].

Supplementary Figure 10

TGM2 expression is dependent on *hsa-miR-17-5p*. (**A**) qRT-PCR of *TGM2* expression levels of LN-229 cells transfected with either a *hsa-miR-17-5p* mimic or a non-target negative control. (**B**) Dual-Glo luciferase assay where LN229 cells were transfected with a *hsa-miR-17-5p* mimic or a non-target

negative control, as well as, a construct containing either the *TGM2 hsa-miR-17-5p* binding site or a *TGM2 hsa-miR-17-5p* mutated binding. Data is normlized to firefly luciferase for transfection efficiency. Student T-test were used for statistical analyses; mean +/-SEM, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

References

1. Reifenberger G, Weber RG, Riehmer V, Kaulich K, Willscher E, Wirth H, et al. Molecular characterization of long-term survivors of glioblastoma using genome- and transcriptome-wide profiling. Int J Cancer. 2014;135(8):1822-31.
2. Gravendeel LA, Kouwenhoven MC, Gevaert O, de Rooi JJ, Stubbs AP, Duijm JE, et al. Intrinsic gene expression profiles of gliomas are a better predictor of survival than histology. Cancer Res. 2009;69(23):9065-72.
3. Cancer Genome Atlas Research N. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature. 2008;455(7216):1061-8.

3. Discussion

The work summarized in the first paper of this thesis was aimed at elucidating biological and clinical differences in PAs (2). The approach undertaken was to utilize tumors from all age groups, instead of restricting the cohort to either pediatric or adult tumors (21, 24, 31-33, 100-104). In addition and the first of its kind for PA, SNF, a multi-layer approach, was used to analyze these tumors as single layer analyses did not provide consistent clustering and allowed for the study of additional cohorts (21, 66, 100-104).

Indeed, two distinct groups were observed using SNF of the RNA sequencing transcriptome and the proteome, and, using a small number of genes or proteins (100 in total for each layer of data) based on the SNF groups, these groups were not only extended to samples with just a single layer of data, but to three additional datasets: 1- proteomic data (63), 2- RNA sequencing data (14) and 3- microarray transcriptome data (13), showing the robustness of this technique. Although epigenetic information was available for this cohort, addition of this layer to the SNF destroyed the grouping observed with just transcriptome and proteome layers. This is in keeping with previous studies that could not show biological heterogeneity using only epigenetic information (21, 100, 101, 104).

When analyzing the clinical data, trends were consistently observed between datasets. However, since the original dataset comprised only 62 samples, a combined analytic approach of all four datasets was undertaken to reveal associations with the clinical data. In fact, when combining the data we observed that supratentorial tumors were predominantly from Group 1. When supratentorial and infratentorial were further divided into optic pathway, supratentorial, posterior fossa and brainstem, supratentorial tumors were still predominantly from Group 1, even after removing samples in the optic pathway.

Next, age of diagnosis was assessed between the two groups. Interestingly, Group 1 had a significantly younger mean age (approximately 7 years of age) but was not exclusively composed of pediatric patients as there were 3 patients above age of 40. In Group 2, patients tended to be older with the majority of adult patients found in this

group, but the average age was still only approximately 12 years of age, showing that this group contained many pediatric patients as well.

Interestingly, neither gender nor type of alterations differed significantly between PA groups. As PA is a MAPK driven tumor (8, 14, 17, 21), all samples with mutation or fusion information showed an alteration in the MAPK pathway. However, except for a non-significant trend in *BRAF* fusions, the alterations were evenly distributed between Group 1 and 2, pointing to a different mechanism for differentiating these groups and confirming that neither of these groups would benefit more from MAPK inhibiters compared to the other group.

Finally, as PAs have a low mortality rate and these tumors tend to recur (2, 5, 8), progression-free survival was analyzed. Group 1 patients, younger and with tumors predominantly found in the supratentorial, had a worse progression-free survival. This is interesting since it has been previously reported that adults have a worse survival compared to pediatric patients (9-11). Notably, when performing a hazard ratio analysis, only the PA group factor was statistically significant, suggesting that these groups are indeed an important segregation for PA.

To elucidate the biological heterogeneity between the two PA groups, Protein/RNA ratios, Ingenuity Pathway Analysis (IPA) and GeneSet Enrichment Analysis (GSEA) were performed. First, the Protein/RNA ratios generated for each group showed that Group1 was driven by the transcriptome, whereas Group 2 was driven by the proteome, which was important to understanding the pathway analyses. Specifically, when looking at the IPA using a stringent cutoff (fold change ±2, q value <0.05), the top terms in the transcriptome analysis were focused on the immune system for Group 1. Conversely, only a few terms were present for Group 2 and these were not connected to each other. The opposite was true for the proteome analysis where Group1 had few terms and Group 2 had terms which were themed, in this instance around mitochondrion biosynthetic pathways.

These data were corroborated in the GSEA and focused on the immune response in the initial analysis, since two independent datasets verified these results. In addition, deconvolution algorithms ESTIMATE and CIBERSOT showed a significant immune related or T-cell signaling in the RNA sequencing transcriptome data (105, 106). This phenotype was not observed in the microarray data, which could be due to the limited

resolution of microarray data in comparison to RNA sequencing. Finally, using multiplex immune fluorescence analysis of PA samples, a strong trend of anti-PD1^{+}/anti-CD8^{+} cells was present in Group 2, suggesting an exhausted immune population that was not present in Group 1. This would also suggest that Group 1 has an active immune population, but a larger cohort must be tested to prove this.

On the other hand, Group 2 results were not highlighted in the publication since a proteome validation dataset was not included initially. The gene sets which were enriched in Group 2, which included terms for “Cellular Respiration” and “TCA metabolism”, are in line with the IPA analysis and suggested that Group 2 PAs have an increase in energy compared to Group 1. It would be interesting to test mitochondria inhibitors in Group 2 PAs as there has been some data suggesting that these might be useful in a cancer setting (107).

As shown in the appendix (section 4), high-throughput drug screening on non-overlapping primary cell cultures was performed. Pediatric PAs were considered as a single entity since these samples were not transcriptionally profiled and groups could not be attributed. Therefore, a further study of adult and pediatric samples which are grouped would be an asset to the field of PAs. To date, very few cell lines have been developed for PAs, in part due to oncogene-induced senescence which prevents cells from propagating (8, 108). Selt et al have overcome this limitation by forcing cells to grow using a SV40-TAg lentivirus (108). It would be of great interest to profile these cell lines, group them and process them using the high-throughput drug screening available at the Heinrich Heine University core facility to determine if there are pharmacological differences which could be exploited between the two groups.

In summary, although this work is not the first to stratify PAs (21, 66, 100-104), the analysis performed in this thesis describes subdividing PAs into two groups that have been consistently applied to three validation datasets, which is novel for the field of PAs. It will be important to explore the immune profile of Group 1 samples in a larger cohort and potentially the use of mitochondria inhibitors in Group 2. Although PAs are not an aggressive tumor, these data may help physicians provide stratification for this disease and, in the future, the best possible treatment.

The work summarized in the second manuscript was aimed at identifying a biomarker which would distinguish short- *versus* long-term survivors of GB. The initial analysis

used to identify *HOTAIRM1* used microarray data generated by the German Glioma Network included several hundred lncRNAs, however, if this analysis had been performed with RNA sequencing data, the number of lncRNAs could have been closer to 100,000 lncRNAs (109). Nevertheless, *HOTAIRM1* clearly showed an increase in expression in the more aggressive short-term samples. Although this study is a bit older and the glioblastoma samples consisted of both IDH-mutant and -wildtype samples, there was distinct survival benefit observed in the IDH-wildtype GB patients. Additionally, the high *HOTAIRM1* expression correlated with decreased survival in two additional datasets (110, 111).

Although this publication was not the first to show that *HOTAIRM1* was associated with an aggressive behavior, other publications did not focus on the function of *HOTAIRM1* (112-114). A number of cell lines were used to phenotype *HOTAIRM1* knockdowns, specifically U251MG, LN-229, LN-18 and T98G were used in transient knockdown experiments and LN-229, LN-18, SF126 and U87MG were used for stable knockdown experiments. LN-229 and LN-18 were the only cell lines used in both the transient and stable experiments. *HOTAIRM1* could not be overexpressed in U251MG and T98G and, therefore, additional cell lines were added to the overall study. Since there are numerous GB cell lines, the focus when adding new cell lines to the study was the ability to grow *in vivo*. However, neither SF126 nor U87MG would grow *in vivo* in when implanted in mice and, therefore, only one mouse model, LN-229, was included in the publication.

After observing a decrease in cell viability, invasion and colony formation, an attempt to describe the mechanism of *HOTAIRM1* was undertaken. Transient knockdowns of U251MG, LN-229 and T98G were profiled using RNA sequencing and proteomics. As seen in manuscript 1, using both RNA sequencing and proteomics is imperative to get a complete overview when using pathway analyses. However, in this case, the focus was on RNA sequencing and proteomic overlapping pathways in an effort to reduce the potential, available pathways. Indeed, this approach left us with four clusters in the GSEA, two of which involved mitochondria and was validated in experiments using superoxide indicators, anti-oxidant treatments and irradiation.

To determine additional factors implicated in the *HOTAIRM1* mode of action, the proteomics data were re-analyzed with a focus on potential *HOTAIRM1* partners. The top candidate was transmembrane protein 87A (TMEM87A) but this gene did not

show any significance when looking at survival. In an attempt to link the list of proteins to *HOTAIRM1*, potential microRNA binding sites were determined for *HOTAIRM1* and the 12 proteins which were differentially regulated when *HOTAIRM1* was knockdown. Using only microRNAs overlapping with *HOTAIRM1* and microRNA/protein pairs which were negatively correlated in the TCGA dataset (110), Transglutaminase 2 (TGM2) was the top candidate (8 out of the top 10 pairs), though it was only the 5th on the list of differentially expressed proteins. Not only did TGM2 and the top microRNA, *has-miR-17-5p*, show a similar survival as *HOTAIRM1*, TGM2 can be located in the mitochondria as well and is known to be involved in mitochondria respiration (115, 116).

In summary, utilizing long non-coding RNA to determine biomarkers of short survival in glioblastoma has led to the discovery that the *HOTAIRM1/hsa-miR-17-5p/TGM2* axis is an important factor for shortened survival. In fact, this process can be repeated for any number of studies, especially if RNA sequencing was used to generate expression data.

As a broader perspective on the key studies conducted during my thesis, the integration of proteomics and genomics, known as proteogenomics, represents a promising approach in the diagnosis and treatment of brain tumor patients (117). In this last part of the discussion, I will focus on how proteogenomics may revolutionize the care of brain tumor patients by offering enhanced diagnostic accuracy and personalized treatment strategies (63, 118).

One of the key advantages of proteogenomics lies in its ability to provide a comprehensive view of the molecular landscape of brain tumors. By analyzing genomic alterations in combination with the corresponding protein expression profiles, clinician-scientists gain a more nuanced understanding of the underlying biology driving tumor development and progression. This holistic approach enables the identification of novel biomarkers that can aid in early detection, classification, and prognostication of brain tumors (119).

In the realm of diagnostics, proteogenomic analysis holds the potential to overcome some of the limitations of traditional histopathological methods (61, 120). Brain tumors are notorious for their heterogeneity, both within individual tumors (intratumoral heterogeneity) and across different patients (intertumoral heterogeneity) (2). Proteogenomics offers a more precise classification of tumors by capturing the

dynamic interplay between genetic mutations and protein expression patterns, allowing for tailored therapeutic interventions. In addition, proteogenomic profiling can uncover targetable molecular aberrations that may otherwise not be detected, thereby expanding the repertoire of therapeutic options available to patients (119).

In the era of precision medicine, proteogenomics is paving the way for personalized treatment strategies in children and adults with brain tumors. By stratifying patients based on their molecular profiles, clinician-scientists can match individuals with the most effective therapies while minimizing the risk of side effects. For example, identifying specific protein targets or signaling pathways dysregulated in a patient's tumor can guide the selection of immunotherapies or targeted therapies tailored to their molecular profile (121). Moreover, proteogenomic insights can inform the design of clinical trials, facilitating the development of novel therapeutics and accelerating the translation of research findings into clinical practice.

Monitoring of treatment response and disease progression constitutes another promising application of proteogenomics in brain tumor management. By tracking changes in the proteomic landscape over time, clinicians can determine treatment efficacy and detect early signs of recurrence or resistance. This real-time monitoring enables timely and individualized adjustments to treatment regimens, which will ideally maximize the therapeutic benefit, minimize harmful side effects and improve patient outcomes (122).

However, several challenges must be addressed to fully realize the potential of proteogenomics in the clinical setting. These challenges include standardizing protocols for sample collection and analysis, integrating multi-omics data from diverse platforms, and ensuring the scalability and affordability of proteogenomic technologies. Furthermore, robust bioinformatics tools and computational algorithms are required to interpret complex proteogenomic datasets and extract clinically relevant insights (122).

In conclusion, proteogenomics holds tremendous promise for transforming the diagnosis and treatment of brain tumor patients, which is urgently required for high-grade glioma patients in particular. By integrating genomic and proteomic information, this approach offers a more comprehensive understanding of tumorigenesis and opens novel avenues for personalized medicine. As technology continues to advance and our understanding of the molecular basis of brain tumors

grows more profound, proteogenomics will play an increasingly pivotal role in improving patient care and outcomes.

4. Appendix

Until now, PA is known as a low-grade, single pathway disease with low heterogeneity (7, 8). Using similarity network fusion to perform a proteogenomic analysis of PA, we have shown that PA can be segregated into two groups and these groups can be identified in validation cohorts using a 100 gene expression signature. Although not tested in this thesis, the top expressing genes based on p-value might represent promising biomarkers for segregating PAs. Elastin microfibril interfacer 2 (EMILIN2) is located in the extracellular matrix and is closely associated with immune infiltration (123). Inositol 1,4,5-triphosphate receptor associated 1 (IRAG1, also known as MRVI1) is a substrate of cGMP-dependent kinase-1 (PKG1) and is involved in calcium release (124). These genes are in line with the IPA and GSEA findings seen in our manuscript as these results showed a clear immune response for Group 1 and action potential/ion transmembrane activity for Group 2.

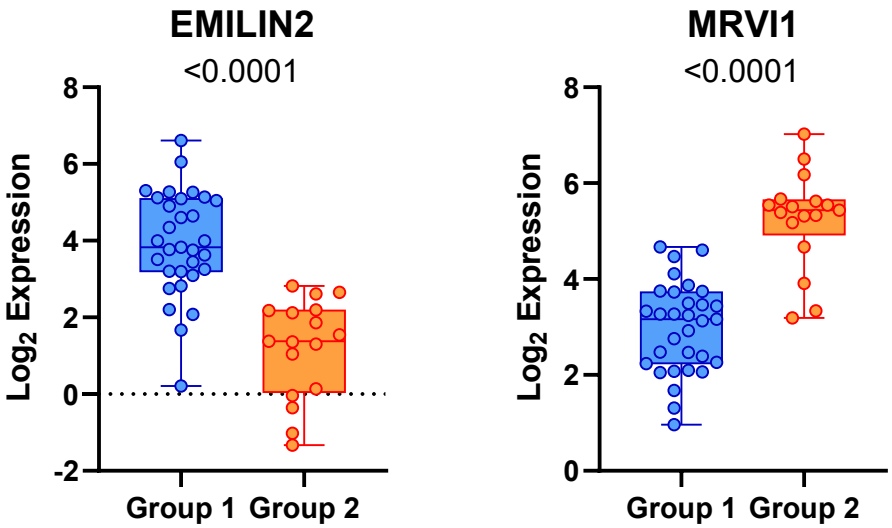


Figure 3.1: Potential biomarkers for Group 1 and 2 PAs.
Box plots showing top genes for each group based on p-value, y-axis shows log_2 expression and x-axis shows PA groups.

In addition to this, we also performed high-throughput drug screening on non-overlapping primary cell cultures. Although these were all pediatric samples, we were unable to group these PAs since there was no expression data. Instead, we

compared the results of the drug screening with those of glioblastoma and medulloblastoma samples. We identified R788 disodium hexahydrate as a compound, which was more active in the benign PAs.

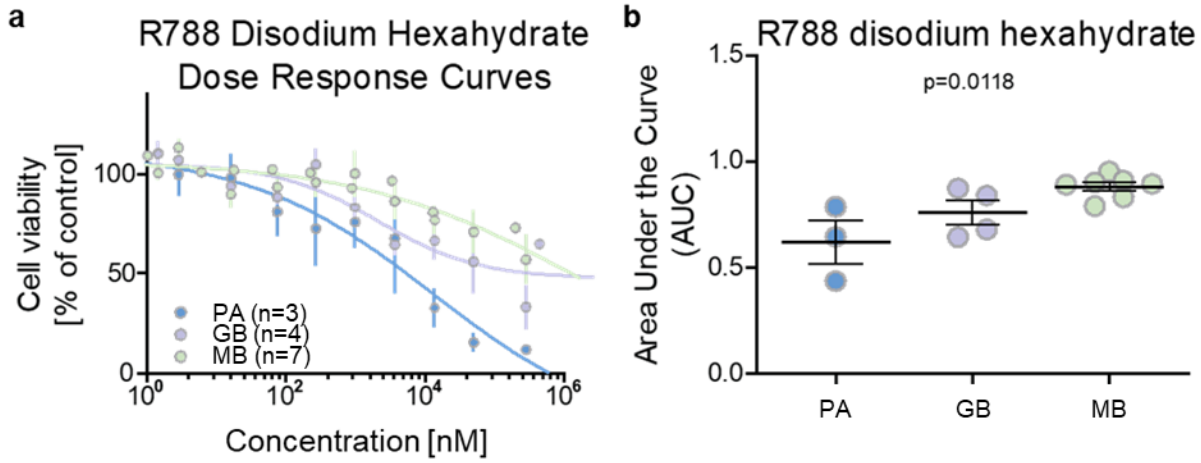


Figure 3.2: Drug screening of patient derived primary cultures
a. Dose response curves of R788 disodium hexahydrate for pilocytic astrocytoma (PA), glioblastoma (GB) and medulloblastoma (MB). Y-axis shows cell viability as a percentage of the untreated cells and x-axis shows $\log_{10}$ concentration in nanomole. **b**. Dot plot of quantified area under the curve values for each tumor type. Fisher's exact test was used to compare the drug response of PA versus GB and MB cultures. Y-axis shows area under the curve measurement (0=sensitive and 1=resistant) and the x-axis shows the different tumor entities (courtesy of the High-thoughput Drug Screening core facility).

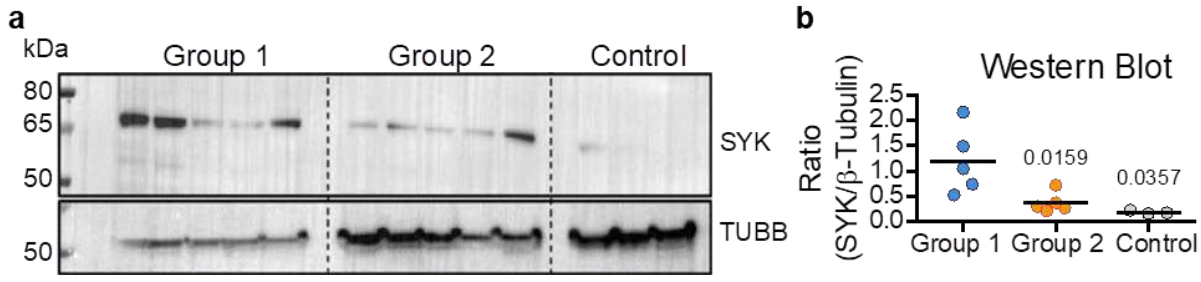


Figure 3.3: Western blot analysis of SYK.
a. Western blot analysis of 5 Group 1, 5 Group 2 and 3 Control samples immunoblotted with anti-SYK antibody (above) and the loading control anti-beta-tubulin (TUBB, below). Size in kilo Daltons is shown on the left. **b.** SYK expression is displayed as ratio of SYK over beta-tubulin (y-axis). Group 1 has significantly more SYK compared to Group 2 or Control (x-axis) (courtesy of the Molecular Proteomics Laboratory core facility).

We were then able to show that SYK is increased in Group 1 PAs compared to Group 2 and to control fetal brain tissue. SYK expression is increased upon JAK1 and TLR3 signaling, which leads to an increase in type 1 interferon genes (125, 126).

These results are in line with our bioinformatics pathway analysis. Unfortunately, with the addition of new samples, there was no longer a significant difference between PAs and the more aggressive glioblastoma and medulloblastoma. There is still the possibility that Group 1 would be more sensitive to SYK inhibitors compared to Group 2. Unfortunately, we did not have the means to distinguish the primary cultures into groups up to now.

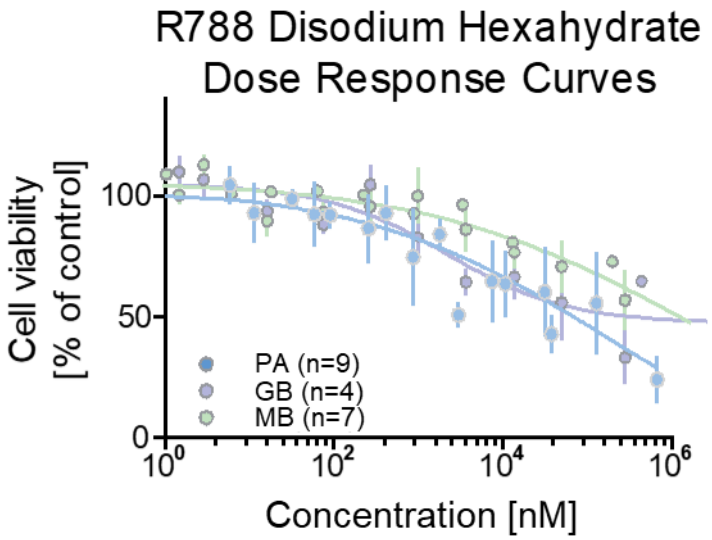


Figure 3.4: Updated drug screening of patient-derived primary cultures
Dose response curves of R788 disodium hexahydrate for PA, glioblastoma (GB) and medulloblastoma (MB). Y-axis shows cell viability as a percentage of the untreated cells and x-axis shows $\log_{10}$ concentration in nanomole (courtesy of the High-thoughput Drug Screening core facility).

5. Author contribution to manuscripts

Title: Integrative multi-omics reveals two biologically distinct groups of pilocytic astrocytoma

Published in: Acta Neuropathologica

Contribution: Daniel Picard substantially conceived and designed the study. He collected samples from the University Hospital Düsseldorf and the Centro Hospitalar Lisboa Norte. He performed all bioinformatics analyses. In addition, he created all the figures and performed the statistical analyses. Furthermore, he wrote substantial parts of the manuscript.

Title: The long non-coding RNA *HOTAIRM1* promotes tumor aggressiveness and radiotherapy resistance in glioblastoma

Published in: Cell Death and Disease

Contribution: Daniel Picard substantially conceived and designed the study. He acquired the bioinformatic data and performed all bioinformatics analyses. In addition, he created all the figures and performed the statistical analyses. Furthermore, he wrote substantial parts of the manuscript.

6. References

1. Weller M, van den Bent M, Tonn JC, Stupp R, Preusser M, Cohen-Jonathan-Moyal E, et al. European Association for Neuro-Oncology (EANO) guideline on the diagnosis and treatment of adult astrocytic and oligodendroglial gliomas. Lancet Oncol. 2017;18(6):e315-e29.

2. Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, et al. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. Neuro Oncol. 2021;23(8):1231-51.

3. Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. The New England journal of medicine. 2005;352(10):987-96.

4. Ostrom QT, Price M, Neff C, Cioffi G, Waite KA, Kruchko C, et al. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2015-2019. Neuro Oncol. 2022;24(Suppl 5):v1-v95.

5. Ostrom QT, Price M, Ryan K, Edelson J, Neff C, Cioffi G, et al. CBTRUS Statistical Report: Pediatric Brain Tumor Foundation Childhood and Adolescent Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2014-2018. Neuro Oncol. 2022;24(Suppl 3):iii1-iii38.

6. Knight J, De Jesus O. Pilocytic Astrocytoma. StatPearls. Treasure Island (FL)2022.

7. Louis DN, Ohgaki H, Wiestler OD, Cavenee WK, Burger PC, Jouvet A, et al. The 2007 WHO classification of tumours of the central nervous system. Acta Neuropathol. 2007;114(2):97-109.

8. Milde T, Rodriguez FJ, Barnholtz-Sloan JS, Patil N, Eberhart CG, Gutmann DH. Reimagining pilocytic astrocytomas in the context of pediatric low-grade gliomas. Neuro Oncol. 2021;23(10):1634-46.

9. Johnson DR, Brown PD, Galanis E, Hammack JE. Pilocytic astrocytoma survival in adults: analysis of the Surveillance, Epidemiology, and End Results Program of the National Cancer Institute. J Neurooncol. 2012;108(1):187-93.

10. Tomita Y, Hibler EA, Suruga Y, Ishida J, Fujii K, Satomi K, et al. Age is a major determinant for poor prognosis in patients with pilocytic astrocytoma: a SEER population study. Clin Exp Med. 2022.

11. Yang W, Porras JL, Khalafallah AM, Sun Y, Bettegowda A, Mukherjee D. Comparison of adult and pediatric pilocytic astrocytomas using competing risk analysis: A population-based study. Clin Neurol Neurosurg. 2022;212:107084.

12. Mair MJ, Wohrer A, Furtner J, Simonovska A, Kiesel B, Oberndorfer S, et al. Clinical characteristics and prognostic factors of adult patients with pilocytic astrocytoma. J Neurooncol. 2020;148(1):187-98.

13. Buhl JL, Selt F, Hielscher T, Guiho R, Ecker J, Sahm F, et al. The Senescence-associated Secretory Phenotype Mediates Oncogene-induced Senescence in Pediatric Pilocytic Astrocytoma. Clin Cancer Res. 2019;25(6):1851-66.

14. Jones DT, Hutter B, Jager N, Korshunov A, Kool M, Warnatz HJ, et al. Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. Nat Genet. 2013;45(8):927-32.

15. Packer RJ, Pfister S, Bouffet E, Avery R, Bandopadhayay P, Bornhorst M, et al. Pediatric low-grade gliomas: implications of the biologic era. Neuro Oncol. 2017;19(6):750-61.

16. Collins VP, Jones DT, Giannini C. Pilocytic astrocytoma: pathology, molecular mechanisms and markers. Acta Neuropathol. 2015;129(6):775-88.

17. Bornhorst M, Frappaz D, Packer RJ. Pilocytic astrocytomas. Handb Clin Neurol. 2016;134:329-44.

18. Cler SJ, Skidmore A, Yahanda AT, Mackey K, Rubin JB, Cluster A, et al. Genetic and histopathological associations with outcome in pediatric pilocytic astrocytoma. J Neurosurg Pediatr. 2022:1-9.

19. Younes ST, Herrington B. In silico analysis identifies a putative cell-of-origin for BRAF fusion-positive cerebellar pilocytic astrocytoma. PLoS One. 2020;15(11):e0242521.

20. Ruda R, Capper D, Waldman AD, Pallud J, Minniti G, Kaley TJ, et al. EANO - EURACAN - SNO Guidelines on circumscribed astrocytic gliomas, glioneuronal, and neuronal tumors. Neuro Oncol. 2022;24(12):2015-34.

21. Reitman ZJ, Paolella BR, Bergthold G, Pelton K, Becker S, Jones R, et al. Mitogenic and progenitor gene programmes in single pilocytic astrocytoma cells. Nat Commun. 2019;10(1):3731.

22. Okonechnikov K, Joshi P, Sepp M, Leiss K, Sarropoulos I, Murat F, et al. Mapping pediatric brain tumors to their origins in the developing cerebellum. Neuro Oncol. 2023.

23. Capper D, Jones DTW, Sill M, Hovestadt V, Schrimpf D, Sturm D, et al. DNA methylation-based classification of central nervous system tumours. Nature. 2018;555(7697):469-74.

24. Metais A, Bouchoucha Y, Kergrohen T, Dangouloff-Ros V, Maynadier X, Ajlil Y, et al. Pediatric spinal pilocytic astrocytomas form a distinct epigenetic subclass from pilocytic astrocytomas of other locations and diffuse leptomeningeal glioneuronal tumours. Acta Neuropathol. 2023;145(1):83-95.

25. Collins KL, Pollack IF. Pediatric Low-Grade Gliomas. Cancers (Basel). 2020;12(5).

26. Nelson AJ, Zakaria R, Jenkinson MD, Brodbelt AR. Extent of resection predicts risk of progression in adult pilocytic astrocytoma. Br J Neurosurg. 2019;33(3):343-7.

27. Santos AN, Dieckmann C, Rauschenbach L, Oppong MD, Dinger TF, Deuschl C, et al. Long-term outcome after management of pilocytic astrocytoma in the posterior fossa in a pediatric population. IBRO Neurosci Rep. 2022;13:388-92.

28. Fangusaro J, Onar-Thomas A, Young Poussaint T, Wu S, Ligon AH, Lindeman N, et al. Selumetinib in paediatric patients with BRAF-aberrant or neurofibromatosis type 1-associated recurrent, refractory, or progressive low-grade glioma: a multicentre, phase 2 trial. Lancet Oncol. 2019;20(7):1011-22.

29. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature. 2013;500(7463):415-21.

30. Jones DT, Gronych J, Lichter P, Witt O, Pfister SM. MAPK pathway activation in pilocytic astrocytoma. Cell Mol Life Sci. 2012;69(11):1799-811.

31. Fontebasso AM, Shirinian M, Khuong-Quang DA, Bechet D, Gayden T, Kool M, et al. Non-random aneuploidy specifies subgroups of pilocytic astrocytoma and correlates with older age. Oncotarget. 2015;6(31):31844-56.

32. Jacob K, Albrecht S, Sollier C, Faury D, Sader E, Montpetit A, et al. Duplication of 7q34 is specific to juvenile pilocytic astrocytomas and a hallmark of cerebellar and optic pathway tumours. Br J Cancer. 2009;101(4):722-33.

33. Pfister S, Janzarik WG, Remke M, Ernst A, Werft W, Becker N, et al. BRAF gene duplication constitutes a mechanism of MAPK pathway activation in low-grade astrocytomas. J Clin Invest. 2008;118(5):1739-49.

34. Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, et al. Mutations of the BRAF gene in human cancer. Nature. 2002;417(6892):949-54.

35. Sadighi Z, Slopis J. Pilocytic astrocytoma: a disease with evolving molecular heterogeneity. J Child Neurol. 2013;28(5):625-32.

36. Ryall S, Zapotocky M, Fukuoka K, Nobre L, Guerreiro Stucklin A, Bennett J, et al. Integrated Molecular and Clinical Analysis of 1,000 Pediatric Low-Grade Gliomas. Cancer Cell. 2020;37(4):569-83 e5.

37. Zwaig M, Baguette A, Hu B, Johnston M, Lakkis H, Nakada EM, et al. Detection and genomic analysis of BRAF fusions in Juvenile Pilocytic Astrocytoma through the combination and integration of multi-omic data. BMC Cancer. 2022;22(1):1297.

38. Subbiah V, Westin SN, Wang K, Araujo D, Wang WL, Miller VA, et al. Targeted therapy by combined inhibition of the RAF and mTOR kinases in malignant spindle cell neoplasm harboring the KIAA1549-BRAF fusion protein. Journal of Hematol Oncol. 2014;7(8).

39. Cancer Genome Atlas N. Genomic Classification of Cutaneous Melanoma. Cell. 2015;161(7):1681-96.

40. Shoushtari AN, Chatila WK, Arora A, Sanchez-Vega F, Kantheti HS, Rojas Zamalloa JA, et al. Therapeutic Implications of Detecting MAPK-Activating Alterations in Cutaneous and Unknown Primary Melanomas. Clin Cancer Res. 2021;27(8):2226-35.

41. Gronych J, Korshunov A, Bageritz J, Milde T, Jugold M, Hambardzumyan D, et al. An activated mutant BRAF kinase domain is sufficient to induce pilocytic astrocytoma in mice. J Clin Invest. 2011;121(4):1344-8.

42. Bergoug M, Doudeau M, Godin F, Mosrin C, Vallee B, Benedetti H. Neurofibromin Structure, Functions and Regulation. Cells. 2020;9(11).

43. Tamura R. Current Understanding of Neurofibromatosis Type 1, 2, and Schwannomatosis. Int J Mol Sci. 2021;22(11).

44. Costa AA, Gutmann DH. Brain tumors in Neurofibromatosis type 1. Neurooncol Adv. 2019;1(1):vdz040.

45. Cowzer D, Zameer M, Conroy M, Kolch W, Duffy AG. Targeting KRAS in Pancreatic Cancer. J Pers Med. 2022;12(11).

46. Li Y, Wei C, Wang W, Li Q, Wang ZC. Tropomyosin receptor kinase B (TrkB) signalling: targeted therapy in neurogenic tumours. J Pathol Clin Res. 2022.

47. Sorokin M, Rabushko E, Rozenberg JM, Mohammad T, Seryakov A, Sekacheva M, et al. Clinically relevant fusion oncogenes: detection and practical implications. Ther Adv Med Oncol. 2022;14:17588359221144108.

48. Bale TA. FGFR- gene family alterations in low-grade neuroepithelial tumors. Acta Neuropathol Commun. 2020;8(1):21.

49. Sievert AJ, Lang SS, Boucher KL, Madsen PJ, Slaunwhite E, Choudhari N, et al. Paradoxical activation and RAF inhibitor resistance of BRAF protein kinase fusions characterizing pediatric astrocytomas. Proc Natl Acad Sci U S A. 2013;110(15):5957-62.

50. Nobre L, Zapotocky M, Ramaswamy V, Ryall S, Bennett J, Alderete D, et al. Outcomes of BRAF V600E Pediatric Gliomas Treated With Targeted BRAF Inhibition. JCO Precis Oncol. 2020;4.

51. Hargrave DR, Bouffet E, Tabori U, Broniscer A, Cohen KJ, Hansford JR, et al. 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. 2019;25(24):7303-11.

52. Bouffet E, Geoerger B, Moertel C, Whitlock JA, Aerts I, Hargrave D, et al. Efficacy and Safety of Trametinib Monotherapy or in Combination With Dabrafenib in Pediatric BRAF V600-Mutant Low-Grade Glioma. J Clin Oncol. 2023;41(3):664-74.

53. Horin P. [Biological principles of chemotherapy: cellular response as a result of its evolutionary memory]. Cas Lek Cesk. 2001;140(19):579-82.

54. Huang Y, Franklin J, Gifford K, Roberts BL, Nicolette CA. A high-throughput proteo-genomics method to identify antibody targets associated with malignant disease. Clin Immunol. 2004;111(2):202-9.

55. Ahn JM, Kim MS, Kim YI, Jeong SK, Lee HJ, Lee SH, et al. Proteogenomic analysis of human chromosome 9-encoded genes from human samples and lung cancer tissues. J Proteome Res. 2014;13(1):137-46.

56. Anurag M, Jaehnig EJ, Krug K, Lei JT, Bergstrom EJ, Kim BJ, et al. Proteogenomic Markers of Chemotherapy Resistance and Response in Triple-Negative Breast Cancer. Cancer Discov. 2022;12(11):2586-605.

57. Archer TC, Ehrenberger T, Mundt F, Gold MP, Krug K, Mah CK, et al. Proteomics, Post-translational Modifications, and Integrative Analyses Reveal Molecular Heterogeneity within Medulloblastoma Subgroups. Cancer Cell. 2018;34(3):396-410 e8.

58. Cao L, Huang C, Cui Zhou D, Hu Y, Lih TM, Savage SR, et al. Proteogenomic characterization of pancreatic ductal adenocarcinoma. Cell. 2021;184(19):5031-52 e26.

59. Krug K, Jaehnig EJ, Satpathy S, Blumenberg L, Karpova A, Anurag M, et al. Proteogenomic Landscape of Breast Cancer Tumorigenesis and Targeted Therapy. Cell. 2020;183(5):1436-56 e31.

60. Wang LB, Karpova A, Gritsenko MA, Kyle JE, Cao S, Li Y, et al. Proteogenomic and metabolomic characterization of human glioblastoma. Cancer Cell. 2021;39(4):509-28 e20.

61. Li Y, Lih TM, Dhanasekaran SM, Mannan R, Chen L, Cieslik M, et al. Histopathologic and proteogenomic heterogeneity reveals features of clear cell renal cell carcinoma aggressiveness. Cancer Cell. 2023;41(1):139-63 e17.

62. Forget A, Martignetti L, Puget S, Calzone L, Brabetz S, Picard D, et al. Aberrant ERBB4-SRC Signaling as a Hallmark of Group 4 Medulloblastoma Revealed by Integrative Phosphoproteomic Profiling. Cancer Cell. 2018;34(3):379-95 e7.

63. Petralia F, Tignor N, Reva B, Koptyra M, Chowdhury S, Rykunov D, et al. Integrated Proteogenomic Characterization across Major Histological Types of Pediatric Brain Cancer. Cell. 2020 183(7):1962-85.

64. Mertins P, Mani DR, Ruggles KV, Gillette MA, Clauser KR, Wang P, et al. Proteogenomics connects somatic mutations to signalling in breast cancer. Nature. 2016;534(7605):55-62.

65. Zhang H, Liu T, Zhang Z, Payne SH, Zhang B, McDermott JE, et al. Integrated Proteogenomic Characterization of Human High-Grade Serous Ovarian Cancer. Cell. 2016;166(3):755-65.

66. Wang B, Mezlini AM, Demir F, Fiume M, Tu Z, Brudno M, et al. Similarity network fusion for aggregating data types on a genomic scale. Nat Methods. 2014;11(3):333-7.

67. Monti S, Tamayo P, Mesirov J, Golub T. Consensus clustering: a resampling-based method for class discovery and visualization. Mach Learn. 2003;51:91-118.

68. MacQueen JB. Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability. University of California Press. 1967;5(1):281-97.

69. Zhong Y, Peng Y, Lin Y, Chen D, Zhang H, Zheng W, et al. MODILM: towards better complex diseases classification using a novel multi-omics data integration learning model. BMC medical informatics and decision making. 2023;23(1):82.

70. Hobbs BD, Morrow JD, Wang XW, Liu YY, DeMeo DL, Hersh CP, et al. Identifying chronic obstructive pulmonary disease from integrative omics and clustering in lung tissue. BMC pulmonary medicine. 2023;23(1):115.

71. Roussel MF, Stripay JL. Modeling pediatric medulloblastoma. Brain pathology. 2020;30(3):703-12.

72. Tamimi AF, Juweid M. Epidemiology and Outcome of Glioblastoma. In: De Vleeschouwer S, editor. Glioblastoma. Brisbane (AU)2017.

73. Brown NF, Ottaviani D, Tazare J, Gregson J, Kitchen N, Brandner S, et al. Survival Outcomes and Prognostic Factors in Glioblastoma. Cancers (Basel). 2022;14(13).

74. Oh JE, Ohta T, Nonoguchi N, Satomi K, Capper D, Pierscianek D, et al. Genetic Alterations in Gliosarcoma and Giant Cell Glioblastoma. Brain pathology. 2016;26(4):517-22.

75. Ohgaki H, Kleihues P. The definition of primary and secondary glioblastoma. Clin Cancer Res. 2013;19(4):764-72.

76. Herrlinger U, Tzaridis T, Mack F, Steinbach JP, Schlegel U, Sabel M, et al. Lomustine-temozolomide combination therapy versus standard temozolomide therapy in patients with newly diagnosed glioblastoma with methylated MGMT promoter (CeTeG/NOA-09): a randomised, open-label, phase 3 trial. Lancet. 2019;393(10172):678-88.

77. Devaux Y, Zangrando J, Schroen B, Creemers EE, Pedrazzini T, Chang CP, et al. Long noncoding RNAs in cardiac development and ageing. Nature reviews Cardiology. 2015;12(7):415-25.

78. Liu D. Tumors and Cancers. Tumors and Cancers: Central and Peripheral Nervous Systems 2017(CRC Press).

79. Kung JT, Colognori D, Lee JT. Long noncoding RNAs: past, present, and future. Genetics. 2013;193(3):651-69.

80. Wu H, Yang L, Chen L. The Diversity of Long Noncoding RNAs and Their Generation. Trends Genet. 2017;3:540-52.

81. Liao Q, Liu C, Yuan X, Kang S, Miao R, Xiao H, et al. Large-scale prediction of long non-coding RNA functions in a coding-non-coding gene co-expression network. Nucleic acids research. 2011;39(9):3864-78.

82. Michelhaugh SK, Lipovich L, Blythe J, Jia H, Kapatos G, Bannon MJ. Mining Affymetrix microarray data for long non-coding RNAs: altered expression in the nucleus accumbens of heroin abusers. Journal of neurochemistry. 2011;116(3):459-66.

83. Lee C, Kikyo N. Strategies to identify long noncoding RNAs involved in gene regulation. Cell Biosci. 2012;2(37).

84. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, et al. Landscape of transcription in human cells. Nature. 2012;489(7414):101-8.

85. Huarte M. The emerging role of lncRNAs in cancer. Nature medicine. 2015;21(11):1253-61.

86. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell. 2007;129(7):1311-23.

87. Ji F, Wuerkenbieke D, He Y, Ding Y, Du R. Long Noncoding RNA HOTAIR: An Oncogene in Human Cervical Cancer Interacting With MicroRNA-17-5p. Oncology research. 2018;26(3):353-61.

88. Camacho CV, Choudhari R, Gadad SS. Long noncoding RNAs and cancer, an overview. Steroids. 2018;133:93-5.

89. Gutmann DH, Rasmussen SA, Wolkenstein P, MacCollin MM, Guha A, Inskip PD, et al. Gliomas presenting after age 10 in individuals with neurofibromatosis type 1 (NF1). Neurology. 2002;59(5):759-61.

90. Hutchinson JN, Ensminger AW, Clemson CM, Lynch CR, Lawrence JB, Chess A. A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. BMC genomics. 2007;8:39.

91. Fang J, Qiao F, Tu J, Xu J, Ding F, Liu Y, et al. High expression of long non-coding RNA NEAT1 indicates poor prognosis of human cancer. Oncotarget. 2017;8(28):45918-27.

92. Wang Z, Zou Q, Song M, Chen J. NEAT1 promotes cell proliferation and invasion in hepatocellular carcinoma by negative regulating miR-613 expression. Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie. 2017;94:612-8.

93. Jen J, Tang YA, Lu YH, Lin CC, Lai WW, Wang YC. Oct4 transcriptionally regulates the expression of long non-coding RNAs NEAT1 and MALAT1 to promote lung cancer progression. Molecular cancer. 2017;16(1):104.

94. Zhang C, Li JY, Tian FZ, Zhao G, Hu H, Ma YF, et al. Long Noncoding RNA NEAT1 Promotes Growth and Metastasis of Cholangiocarcinoma Cells. Oncology research. 2018;26(6):879-88.

95. Li W, Zhang Z, Liu X, Cheng X, Zhang Y, Han X, et al. The FOXN3-NEAT1-SIN3A repressor complex promotes progression of hormonally responsive breast cancer. J Clin Invest. 2017;127(9):3421-40.

96. Li X, Wang S, Li Z, Long X, Guo Z, Zhang G, et al. The lncRNA NEAT1 facilitates cell growth and invasion via the miR-211/HMGA2 axis in breast cancer. International journal of biological macromolecules. 2017;105(Pt 1):346-53.

97. Benetatos L, Vartholomatos G, Hatzimichael E. MEG3 imprinted gene contribution in tumorigenesis. International journal of cancer. 2011;129(4):773-9.

98. Wang Y, Wang Y, Li J, Zhang Y, Yin H, Han B. CRNDE, a long-noncoding RNA, promotes glioma cell growth and invasion through mTOR signaling. Cancer letters. 2015;367(2):122-8.

99. Xu CH, Xiao LM, Liu Y, Chen LK, Zheng SY, Zeng EM, et al. The lncRNA HOXA11-AS promotes glioma cell growth and metastasis by targeting miR-130a-5p/HMGB2. European review for medical and pharmacological sciences. 2019;23(1):241-52.

100. Lambert SR, Witt H, Hovestadt V, Zucknick M, Kool M, Pearson DM, et al. Differential expression and methylation of brain developmental genes define location-specific subsets of pilocytic astrocytoma. Acta Neuropathol. 2013;126(2):291-301.

101. Rodriguez FJ, Giannini C, Asmann YW, Sharma MK, Perry A, Tibbetts KM, et al. Gene expression profiling of NF-1-associated and sporadic pilocytic astrocytoma identifies aldehyde dehydrogenase 1 family member L1 (ALDH1L1) as an underexpressed candidate biomarker in aggressive subtypes. J Neuropathol Exp Neurol. 2008;67(12):1194-204.

102. Sharma MK, Mansur DB, Reifenberger G, Perry A, Leonard JR, Aldape KD, et al. Distinct genetic signatures among pilocytic astrocytomas relate to their brain region origin. Cancer Res. 2007;67(3):890-900.

103. Wong KK, Chang YM, Tsang YT, Perlaky L, Su J, Adesina A, et al. Expression analysis of juvenile pilocytic astrocytomas by oligonucleotide microarray reveals two potential subgroups. Cancer Res. 2005;65(1):76-84.

104. Zakrzewski K, Jarzab M, Pfeifer A, Oczko-Wojciechowska M, Jarzab B, Liberski PP, et al. Transcriptional profiles of pilocytic astrocytoma are related to their three different locations, but not to radiological tumor features. BMC Cancer. 2015;15:778.

105. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nature biotechnology. 2019;37(7):773-82.

106. Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4:2612.

107. Bueno MJ, Ruiz-Sepulveda JL, Quintela-Fandino M. Mitochondrial Inhibition: a Treatment Strategy in Cancer? Current oncology reports. 2021;23(4):49.

108. Selt F, El Damaty A, Schuhmann MU, Sigaud R, Ecker J, Sievers P, et al. Generation of patient-derived pediatric pilocytic astrocytoma in-vitro models using SV40 large T: evaluation of a modeling workflow. J Neurooncol. 2023;165(3):467-78.

109. Uszczynska-Ratajczak B, Lagarde J, Frankish A, Guigo R, Johnson R. Towards a complete map of the human long non-coding RNA transcriptome. Nature reviews Genetics. 2018;19(9):535-48.

110. Cancer Genome Atlas Research N. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature. 2008;455(7216):1061-8.

111. Gravendeel LA, Kouwenhoven MC, Gevaert O, de Rooi JJ, Stubbs AP, Duijm JE, et al. Intrinsic gene expression profiles of gliomas are a better predictor of survival than histology. Cancer Res. 2009;69(23):9065-72.

112. Li Q, Dong C, Cui J, Wang Y, Hong X. Over-expressed lncRNA HOTAIRM1 promotes tumor growth and invasion through up-regulating HOXA1 and sequestering G9a/EZH2/Dnmts away from the HOXA1 gene in glioblastoma multiforme. Journal of experimental & clinical cancer research : CR. 2018;37(1):265.

113. Liang Q, Li X, Guan G, Xu X, Chen C, Cheng P, et al. Long non-coding RNA, HOTAIRM1, promotes glioma malignancy by forming a ceRNA network. Aging. 2019;11(17):6805-38.

114. Lin YH, Guo L, Yan F, Dou ZQ, Yu Q, Chen G. Long non-coding RNA HOTAIRM1 promotes proliferation and inhibits apoptosis of glioma cells by regulating the miR-873-5p/ZEB2 axis. Chinese medical journal. 2020;133(2):174-82.

115. Bianchi N, Beninati S, Bergamini CM. Spotlight on the transglutaminase 2 gene: a focus on genomic and transcriptional aspects. The Biochemical journal. 2018;475(9):1643-67.

116. Rodolfo C, Mormone E, Matarrese P, Ciccosanti F, Farrace MG, Garofano E, et al. Tissue transglutaminase is a multifunctional BH3-only protein. The Journal of biological chemistry. 2004;279(52):54783-92.

117. Nesvizhskii AI. Proteogenomics: concepts, applications and computational strategies. Nat Methods. 2014;11(11):1114-25.

118. Wang Y, Luo R, Zhang X, Xiang H, Yang B, Feng J, et al. Proteogenomics of diffuse gliomas reveal molecular subtypes associated with specific therapeutic targets and immune-evasion mechanisms. Nat Commun. 2023;14(1):505.

119. Joshi SK, Piehowski P, Liu T, Gosline SJC, McDermott JE, Druker BJ, et al. Mass Spectrometry-Based Proteogenomics: New Therapeutic Opportunities for Precision Medicine. Annual review of pharmacology and toxicology. 2024;64:455-79.

120. Wang JM, Hong R, Demicco EG, Tan J, Lazcano R, Moreira AL, et al. Deep learning integrates histopathology and proteogenomics at a pan-cancer level. Cell reports Medicine. 2023;4(9):101173.

121. Rivero-Hinojosa S, Grant M, Panigrahi A, Zhang H, Caisova V, Bollard CM, et al. Proteogenomic discovery of neoantigens facilitates personalized multi-antigen targeted T cell immunotherapy for brain tumors. Nat Commun. 2021;12(1):6689.

122. Mani DR, Krug K, Zhang B, Satpathy S, Clauser KR, Ding L, et al. Cancer proteogenomics: current impact and future prospects. Nature reviews Cancer. 2022;22(5):298-313.

123. Zhao G, Zheng J, Tang K, Chen Q. EMILIN2 is associated with prognosis and immunotherapy in clear cell renal cell carcinoma. Front Genet. 2022;13:1058207.

124. Schlossmann J, Ammendola A, Ashman K, Zong X, Huber A, Neubauer G, et al. Regulation of intracellular calcium by a signalling complex of IRAG, IP3 receptor and cGMP kinase Ibeta. Nature. 2000;404(6774):197-201.

125. Mielcarska MB, Bossowska-Nowicka M, Gregorczyk-Zboroch KP, Wyzewski Z, Szulc-Dabrowska L, Gierynska M, et al. Syk and Hrs Regulate TLR3-Mediated Antiviral Response in Murine Astrocytes. Oxid Med Cell Longev. 2019;2019:6927380.

126. Tassiulas I, Hu X, Ho H, Kashyap Y, Paik P, Hu Y, et al. Amplification of IFN-alpha-induced STAT1 activation and inflammatory function by Syk and ITAM-containing adaptors. Nat Immunol. 2004;5(11):1181-9.