

UNIVERSITÄTSKLINIKUM HAMBURG-EPPENDORF

Zentrum für Molekulare Neurobiologie (ZMNH), Hamburg

Forschungsgruppe Molekulare Neuroonkologische Pathologie

Prof. Dr. med. Julia Neumann

Multi-omic data analysis for identification of novel biomarkers and therapeutic targets in pediatric brain tumours

Cumulative Dissertation

zur Erlangung des Doktorgrades Dr. rer. biol. hum.

an der Medizinischen Fakultät der Universität Hamburg.

vorgelegt von:

Shweta Godbole

aus

Nashik, India

2024

(wird von der Medizinischen Fakultät ausgefüllt)

Angenommen von der

Medizinischen Fakultät der Universität Hamburg am:

Veröffentlicht mit Genehmigung der

Medizinischen Fakultät der Universität Hamburg.

Prüfungsausschuss, der/die Vorsitzende: Prof. Dr. med. Stefan Rutkowski

Prüfungsausschuss, zweite/r Gutachter/in: Prof. Dr. rer. nat. Harmut Schlüter

Datum der Disputation : 15 November 2024

Contents

1: Synopsis	1
1.1 Introduction	1
1.1.1 Advancements of Omics for Understanding CNS Tumours	1
1.1.1.1 DNA Methylome Data	3
1.1.1.2 Transcriptome Data	4
1.1.1.3 Proteome and N-glycan Data	5
1.1.1.4 Multi-Omic Data Analysis	5
1.1.2. Pediatric Brain Tumours	9
1.1.2.1 Medulloblastoma	9
1.1.2.2 High-Grade Glioma MYCN amplified	15
1.2: Aim of the work	16
1.2.1 Aim 1 (Project 1): Identify clinically and phenotypically relevant MB subtypes	16
1.2.2 Aim 2 (Project 2): Characterize a mouse model for the Group3 MBs with worse prognosis and survival	16
1.2.3 Aim 3 (Project 3): To understand the tumourgenesis and characterize the biology of aggressive HGG MYCN tumours	17
1.3: Material and Methods	18
1.3.1 Materials and Data	18
1.3.2 Methods	20
1.4: Summary of Results	24
1.4.1 MB project 1: Multiomic profiling of medulloblastoma reveals subtype-specific targetable alterations at the proteome and N-glycan level	24
1.4.2 MB project 2: MYC overexpression and SMARCA4 loss cooperate to drive medulloblastoma formation in mice	29
1.4.3 HGG- MYCN project 3: Mouse models of pediatric high-grade gliomas with <i>MycN</i> amplification reveal intratumoral heterogeneity and lineage signatures	31
1.5: Discussion	33
2. List of abbreviations	40
3. References	43
4. Publication List	53
4.1 MB project 1 in original format	53
4.2. MB project 2 in original format	124
4.3. HGG-MYCN amplified project in original format	137
5. Summary in German	153
6. Summary English	155

7. Statement of own contribution to the publications	157
8. Acknowledgments	159
9. Eidesstattliche Versicherung	160

1: Synopsis

1.1 Introduction

Pediatric brain tumours are a major cause of death in children among all known childhood cancers¹. These tumours are classified based on their localization, mainly supratentorial or infratentorial. They are further classified based on the age of diagnosis into three types – congenital brain tumours (CBT) (diagnosed in the first 60 days of life), infancy tumours (diagnosed in children younger than 1 year of age) and tumours occurring in the older children¹. Common treatment for most childhood brain tumours remains surgical resection or biopsy to relieve intracranial pressure caused by the tumour, followed by chemo- and/or radiotherapy²³⁻⁵. Advancements in imaging techniques, genetics, and molecular biology, have played a big role in early diagnosis of these tumours^{6,7}. Molecular analyses are typically performed using biopsy material. To increase the shelf-life of the material, without changing the underlying biological information, it is fixed in formalin and embedded in paraffin blocks (FFPE-samples). Besides histomorphological analyses, these FFPE-samples are used to study the DNA methylation patterns⁸⁻¹⁰, transcriptome signatures^{11,12} or protein abundances¹³⁻¹⁶ out of many other applications. Sub-groups and types have been identified using mainly DNA methylome and transcriptome data¹⁷. Although this has enhanced tumour classification and stratification, the route for targeted therapy for the identified subgroups and subtypes remains largely unexplored.

1.1.1 Advancements of Omics for Understanding CNS Tumours

According to the central dogma of biology, DNA gets transcribed to mRNA which undergoes translation to form proteins. Proteins are responsible for the phenotype. Epigenomic modifications such as DNA methylation regulate gene expression and are essential for normal functioning during growth, genomic imprinting and X-chromosome inactivation¹⁸. Dysregulation of DNA methylation results in diseases and plays a role in cancers¹⁹. Methylation of the promoters of certain genes will inhibit the binding of transcriptional factors which will thus prevent transcription of these genes into their respective mRNA²⁰. Since DNA methylation patterns are specific for different types of cells and tissues, they can be a great tool for the identification of subgroups and subtypes in cancers^{21,22}.

Further, the mRNA formed can undergo post-transcriptional regulation for example by microRNA, and lncRNA (long non-coding RNA) which can inhibit their translation into proteins. Since proteins are dynamic and closest to the functional element of the cell, they are essential to understand “most recent” changes in biological system. In order to fully understand the complex biological processes leading to cancer, it is essential to study the three main components of the central dogma of biology – DNA, RNA and proteins enabled with massive development in array and sequencing technologies as well as mass spectrometry-based approaches for relative quantification of proteins.

Central Nervous System (CNS) tumours remain the major cause of death in children and adolescents. Some common brain tumours which majorly affect children include medulloblastomas (MB), high-grade gliomas (HGG) especially high-grade gliomas with MYCN amplifications (HGG-MYCN), ependymomas (EPN) and embryonal tumours with multilayered rosettes (ETMRs) arranged in decreasing order of frequency. While MB and HGG are extremely rare brain tumours they make up for about 12 % of CNS tumours in children. The average age of diagnosis for MBs is 7 years while that for HGG is below 3 years of age^{23,24}. Brain tumours such as MBs and recently discovered HGG-MYCN amplified commonly occur in the cerebellum or brain stem^{23,24}. The current treatment strategy for MB patients remains complete surgical resection followed by cranial radiotherapy for patients over 3 years of age and adjuvant chemotherapy²⁵. HGG-MYCN amplified patients are treated as HGG patients which involves complete surgical resection of tumours followed by chemotherapy and radiotherapy. However, complete surgical resection of tumours in pediatric patients is often not possible and radiotherapy can have adverse effects that can have a long-term impact on the quality of life. Moreover, prognosis still remains poor with 53.8 % 5-year survival rate for MB and 20% for HGG, respectively^{26,27}.

Brain tumours are histologically heterogeneous and display large differences in clinical representation, such as the age of onset, tumour localization, extent of possible surgical resection and prognosis. Thus, new emerging technology can be used for better characterisation of these rare brain tumours. It is essential to account for all three data modality levels of the central dogma – DNA, RNA and proteins - to accurately identify the biological dysregulations and identify new treatment targets. Multi-omic

analyses can be useful for predicting alternative therapeutics for the treatment of pediatric brain tumours.

1.1.1.1 DNA Methylome Data

The field of epigenetics has greatly revolutionised the understanding of diseases²⁸. Epigenetics provides information on reversible and heritable changes in gene expression that do not involve changes to the underlying DNA sequence²⁹. DNA methylation is one of the best characterized epigenetic changes, where a cytosine (C5) base is converted to 5-methylcytosine by DNA methyl transferase enzymes (DNMTs) (MB project1: Supplementary Figure 18A, page number 123)³⁰. DNA-methylation of the MGMT promoter has been shown to be important for stratification of brain tumour patients as it predicts response to chemotherapy. The MGMT gene, a DNA repair enzyme, increases the resistance of tumours to alkylating agents, when expressed. A higher methylation of MGMT promoter is associated decreased MGMT expression and hence with better response to chemotherapy which is similar in HGGs³¹. Since the structural and functional heterogeneity of the cells is captured, DNA methylome data can be useful for the classification of tumours and the identification of new subtypes. In the brain tumour field, DNA methylome data is commonly used in diagnostics along with histology for accurate identification of the brain tumour subgroups and types and resulted in development of the brain tumour classifier (<https://www.moleculareuropathology.org/mnp/>)³². Rapid development in this field including ease of reproducible results with both Fresh Frozen (FF) and Formalin-Fixed Paraffin-Embedded (FFPE) tissues, relatively fewer amount of sample input needed and affordable costs further contributed to the inclusion of such omic data in molecular routine diagnostics³³.

1.1.1.1.1 Copy Number Variation (CNV) Analysis using DNA Methylome Data

DNA methylation data is not only useful for the identification of the epigenetic signatures but also for visualization of chromosomal copy number aberrations³⁴. As specific chromosomal alterations, such as gains and losses of chromosomal arms, or specific amplifications are associated with certain tumor classes, this information is commonly used in the diagnostic workflow^{35,36}. For example, loss of TP53, or

amplification of the MYC gene are often associated with a poor prognosis in MB and can predict tumour subtype or survival probability^{37–40}.

1.1.1.2 Transcriptome Data

Although DNA methylation data is useful for subgroup identification and diagnosis, not all changes occurring in the DNA transcribe into the mRNA. The field of transcriptomics describes the examination of whole transcriptome, accounting for approximately 80% of RNA molecules, however only 2% mRNAs that will likely be expressed and translated into proteins⁴¹. This enables the characterization of the heterogeneity of these tumours by discovery of dysregulated pathways and processes, as well as driving gene fusions which further identifies new tumour subtypes within the methylation subgroups^{42,43}. Many changes can occur at the transcriptome level, post-transcriptional modifications such as capping, splicing and polyadenylation lead to the transport pre-mRNAs from nucleus to the cytoplasm where they are used to synthesize proteins⁴⁴. Normally, pre-mRNAs undergo constitutive splicing, where in all introns are removed and exons are ligated together in order as they appear in the gene, however, under certain circumstances, there can be a deviation from this from of splicing resulting in alternatively spliced variants. In such an event, some exons are skipped that can result in formation of shorter exon lengths. This process occurs due to gene mutations such as by single nucleotide polymorphisms (SNPs) or epigenetic modifications which can affect chromatin accessibility, or due to environmental stimuli⁴⁵. This process is mediated by dynamic spliceosome mechanisms which can result in formation of proteins isoforms which can further affect the phenotype of diseases^{46,47}. Further, microRNAs are non-coding RNAs that can interact with mRNAs, inducing degradation of mRNAs and hence repression of translation to their respective protein⁴⁸ (MB project 1: Supplementary Figure 18B, page number 123). Thus, although more information can be obtained from transcriptomics compared to DNA methylome, much of this information only contributes to the biological understanding of brain tumours and treatments based on dysregulated pathways found in the transcriptome have largely been unsuccessful⁴⁹.

1.1.1.3 Proteome and N-glycan Data

To make a transition from biological process understanding to functional understanding, and to analyse patterns close to the final phenotype, global proteome analysis is essential. Global proteome analysis using mass spectrometry enabled quantitative and qualitative analysis of the proteome of samples from FF or FFPE tissues^{13,15,50,51}. Proteomics provide the final information based on all the changes incurred on different modality levels: such as at DNA and mRNA levels that contribute to the phenotype of the tumours (MB project 1: Supplementary Figure 18C, page number 123). Additionally, proteomic data enables us to identify new diagnostic biomarkers and provide a direct route to drug development since it gives us information on dysregulated pathways which can be targeted. It will also help us to identify clinically and phenotypically relevant subtypes of brain tumours and provide a better insight into the biological characteristics of these subgroups and identified subtypes^{49,52}.

Glycosylation is a common type of post-translational modification. Glycans attach to the proteins and can manipulate the function of the proteins by changing its structure, conformation or interactions with other proteins. Glycans can be mainly classified into two types: N-glycans attached through nitrogen of asparagine and O-glycans attached through oxygen of serine or threonine⁵³. Since extracellular matrix proteins are mainly bounded with N-glycans, we focus on N-glycan analysis in this thesis. N-glycans are further classified into three types – oligomannose, where only Man residues extend the core; complex, where GlcNAc-“antennae” extend the core and hybrid, which is a combination of Man and GlcNAc “antennae” which extend the core (MB project 1: Main Figure 9, page number 102)⁵⁴. These N-glycan modifications are involved in many biological processes, especially signal transduction and immune response⁵⁵. It is well-known that N-glycan modifications are involved in tumour progression and also be used as biomarkers for identification of high-grade tumours⁵⁶⁻⁵⁸.

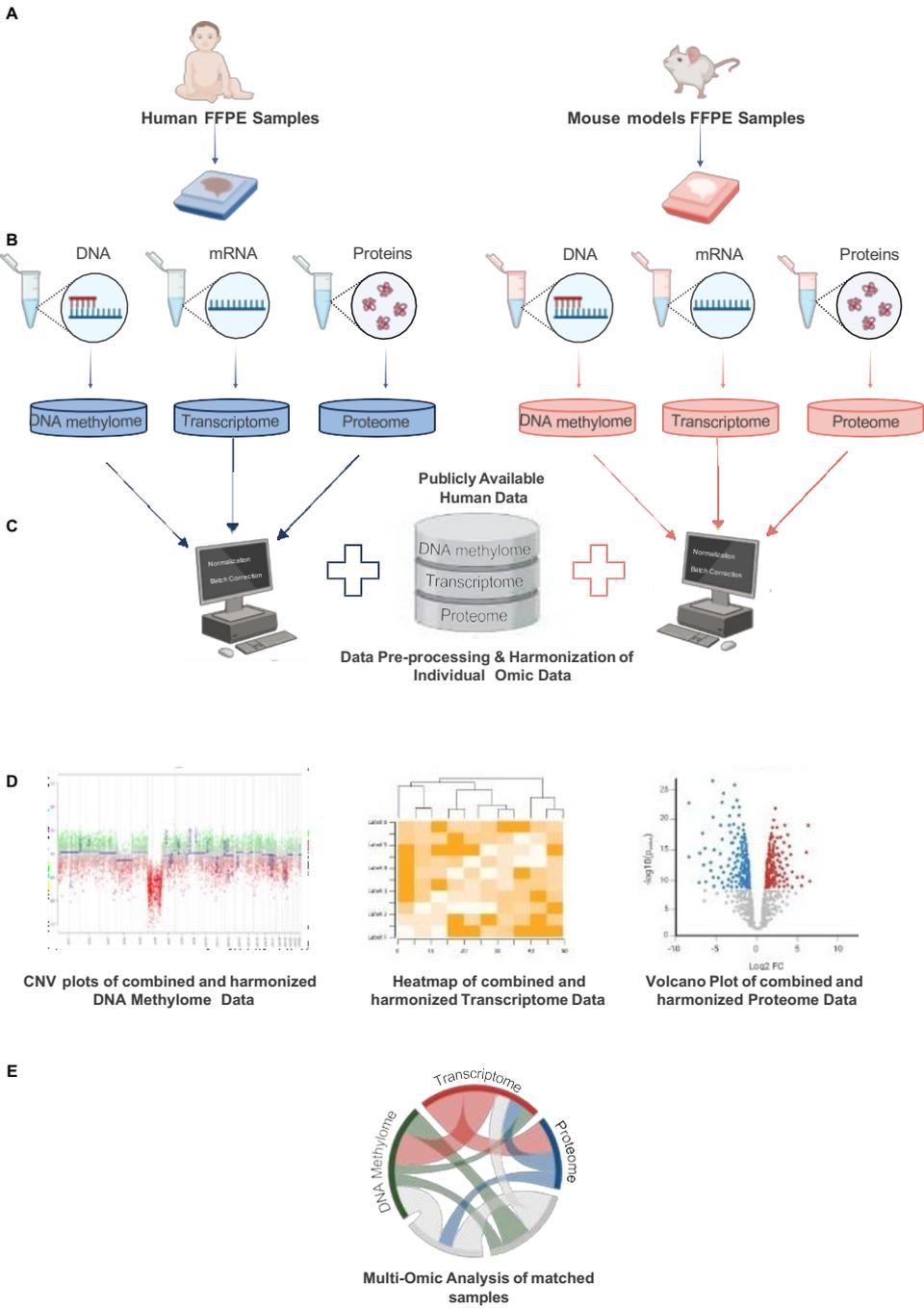
1.1.1.4 Multi-Omic Data Analysis

Combining multiple omics levels provides a better platform for the characterization of tumours. Integrative analysis with sufficient sample number provides superior results and as it combines the information of DNA, RNA and proteome level. It enables us to understand tumour biology in more detail as the relation between data types can be

explored globally and give hints concerning the potential regulatory levels (e.g. rather posttranscriptional or posttranslational regulations). The link to the proteome is finally a close view on the phenotype of the brain tumours, especially when related to histomorphology. It further enables the translation of new findings into the innovation of new therapeutics and clinically relevant biomarkers. Translational research also involves the application of this knowledge to the development of mouse models to further study tumorigenesis and provide a resource for the development and testing of identified therapeutics. It is also essential to confirm whether developed mouse-models mimic their human counterpart for which we require multiomic data integration across species.

While data integration can be extremely useful, it's highly challenging as different platforms are used for the generation of omics data. Although normalizations are often platform-specific, dimensionality reduction techniques need to be applied to reduce heterogeneity within different types of data. This will enable the identification of robust and reliable multi-omics signatures based on the phenotype of interest. We can further perform gene-specific correlations between different types of omics to identify the conservation patterns across different tumour subgroups and subtypes.

1.1.1.4.1 mixOmics DIABLO (Data Integration for Biomarker Discovery using Latent component method for Omics studies)



Created with BioRender.com

Figure 1: **A** Sample collection (human or mouse FFPE samples) **B** Sample preparation for respective omic modality **C** Data acquisition, pre-processing and normalization of individual omic data and integration with its corresponding publicly available data followed by batch effect correction **D** Downstream data analysis of combined and harmonized data for each omic type **E** Multi-omic data analysis of matched samples across omic types.

For the integration of different types of omics data, it is essential to have matched sample cases across omics. It is also essential to normalize the datasets independently before integration as normalizations are always platform-specific. Once normalized, datasets are harmonized and batch corrected to create a combined dataset (Figure 1, page number 7). Our group has previously developed algorithms for batch effect correction, such as HarmonizR, which are applied in this context⁵⁹. Data is further reduced to matched samples, and the datasets from different omics are then ready for integration (see Figure 1 page number 7). A technique called “DIABLO” can be applied, that uses dimensionality reduction wherein each dataset to be integrated will be broken down into “n” number of components^{60,61}. These components will be made up of calculated variables, called latent variables that capture the main information within the data. For example, if our dataset contains 3990 proteins, DIABLO will perform feature selection and use only a few proteins to represent the proteome data based on the phenotype of interest, e.g. the tumour subtypes. The phenotype of interest should be same for all datasets to be integrated. DIABLO aims to maximize the correlated information across the datasets and displays relevant correlations which could be useful biomarkers. Finally, DIABLO can also be used for classification using information about the phenotype of interest⁶⁰.

1.1.1.4.2 Correlation Analysis

Another important application of data integration is to identify the correlation across different omic types. This can be done on data confined to genes that are reflected in all data (as encoded DNA with promotor regions that can be differentially methylated, as mRNA transcript and finally as a protein). These analyses identify the conservation of biological features across different tumour subtypes. Pearson correlation is a technique that can determine the linear correlation between features of interest such as CpG site methylation at the DNA of a gene and its promotor in the DNA methylation data, levels of mRNA in the transcriptome data and also respective protein abundances in the proteome data. Multi-omic data integration thus gives us information about the dysregulated pathways and enables us to identify at which level the dysregulation occurred. Hypothetically, we expect there to be an inverse correlation between the CpG site DNA methylation and mRNA expression (transcriptome level) of a specific

gene. We further expect a positive correlation between the mRNA expression of a specific gene and its respective protein abundance. However, in the case of tumour biology, that is not always true⁶². Such unexpected correlations always point to the underlying biology of the tumour under question and indicate where major regulatory steps may play a role.

1.1.2. Pediatric Brain Tumours

1.1.2.1 Medulloblastoma

Traditionally, MBs were stratified based on their clinical characteristics such as age, metastasis, extent of surgical resection and histological grouping such as classic, desmoplastic or large-cell anaplastic histology to find the optimal treatment protocol and reduce the long-term effects of therapy^{63,64}. However, recent advancements in next-generation sequencing technology such as whole genome /exome sequencing, epigenomics, and transcriptomics increased understanding and stratification of MB patients into further subgroups and subtypes correlating with clinical features (see also Figure 2, page number 11).

MBs are classified into four main molecular subgroups namely, WNT-activated MB, occurring rarely but associated with a rather good prognosis and median survival of 86%, Sonic Hedgehog (SHH)-activated MB comprising approximately 30% of cases, and showing intermediate prognosis and a median overall survival of 53%, Group 4 MBs with a median overall survival of 60-70% and Group 3 MBs with a worse median survival of only 50%⁶⁵⁻⁶⁷ both occurring more frequently than WNT and SHH MBs. Further, specific mutations such as TP53 mutations in SHH MBs, MYC amplification in group 3 MBs are associated with worse prognosis in patients.

WNT MB

WNT MB patients are characterized by over-activation of the WNT pathway. This can often be due to germline mutations of the adenomatous polyposis coli gene (APC) a WNT-pathway inhibitor or somatic mutation of the beta-catenin gene (*CTNNB1*) which inturn leads to overactivation of the WNT-pathway³⁶. Monosomy of chromosome 6 is seen in 90% of WNT MB patients and is associated with a good prognosis. Currently,

monosomy of chromosome 6, nuclear beta-catenin staining, immunohistochemical staining of Dickkopf-related protein 1 (DKK1), which binds to LRP6 co-receptor and inhibits beta-catenin dependent WNT-signalling pathway, are commonly used for diagnosis of WNT MBs. While most of the WNT MB patients have classic histology, there have also been some cases which have anaplastic/ large-cell histology (LCA) and do not carry a monosomy of chromosome 6 causing them to deviate from their favourable prognosis and have a higher chance of relapse ⁶⁸. Owing to the heterogeneity within the subgroup, WNT-MBs lack a gold standard for diagnosis. Deeper epigenomic and transcriptomic studies have identified two further subtypes within the WNT MBs (Figure 2, page number 11), namely WNT α consisting of younger patients with monosomy of chromosome 6 and WNT β consisting of adult patients⁴³. Although WNT MB patients have good survival, life expectancy is reduced due to the after-effects of the treatment ⁶⁹. It is thus discussed if a proportion of patients may be over-treated using current strategies and there might be a need for therapy deacceleration ⁷⁰⁻⁷².

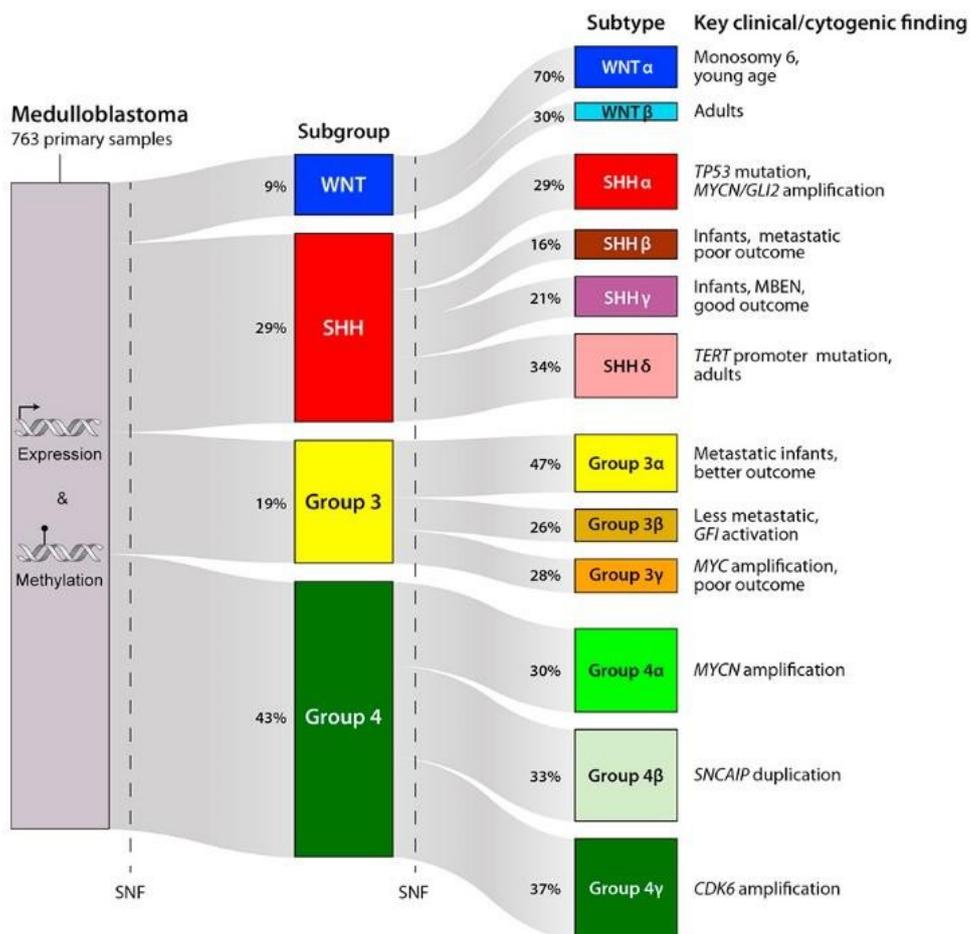


Figure 2: Figure from Cavalli et al ⁴³ (Copyright clearance from Elsevier granted, License number: 5805210453358) displaying the four main molecular subgroups identified using DNA methylation data and their corresponding transcriptomic subtypes

SHH MB

Sonic Hedgehog (SHH)- MB patients are characterized by activation of the SHH pathway due to mutations of SHH pathway genes, such as germline mutations of the SHH receptor *PTCH*, the SHH inhibitor *SUFU* and somatic mutations of *PTCH*, *SMO*, *SUFU* and/or amplifications of *GLI1*, and *MYCN* ^{73,74}. Mutations of the tumour suppressor gene *TP53* were found to be enriched in WNT and SHH MBs. SHH MB harbouring germline *TP53* mutations are often associated with poor prognosis and survival, especially in patients in the age range of 5 to 18 years³⁷. Current diagnostic criteria for SHH MBs include immunohistochemical staining for P75, GAB1, and YAP1 and loss of chromosome 9q22, as well as nodular desmoplastic histology that occurs in nearly 50% of SHH MBs^{33,68}. Furthermore, the SHH MBs also differ based on patient

age, being either predominantly present in infants or adolescents⁷⁵. All these factors led to the identification of four subtypes within this subgroup based on epigenomic and transcriptomic signatures (Figure 2, page number 11). SHH α - made up of TP53 mutated, MYCN and GLI2 amplified patients, SHH β – mainly consisted of infants and with a poor prognosis, SHH γ – also consisted of infants and MBEN histology, associated with good survival, and SHH δ – mainly adult patients with TERT mutations⁴³. SHH MBs have an intermediate prognosis, with survival rates of 81 % for TP53 wildtype cases and 41 % for TP53-mutated cases⁷⁶. However, further diagnostic criteria and especially subtype specific treatment strategies for these four molecular types are still lacking. A few studies performed proteome analysis using mass spectrometry on a small MB cohort and identified two subtypes at the proteome level within the SHH MB subgroup. SHHa – consisted of canonical activation of the SHH pathway, and PTCH mutations while SHHb was more similar to the Group 4 tumours with higher activation of glutamate, calcium and Ras signalling pathways^{77,78}. A validation of these subtypes and potential clinical implication are still lacking and are part of the analyses performed within this PhD thesis work.

Non-WNT/ Non-SHH MBs

Group 3 MBs and Group 4 MBs are very heterogenous. Group 3 tumours mostly consist of patients with LCA histology, and MYC amplifications, which can be detected using immunohistochemical staining against MYC or observed by enhanced signals in copy number analyses from DNA methylome data. Group 3 MBs have the worst survival and often a high rate of metastasis. Transcriptomic studies have identified three subtypes within this subgroup (Figure 2, page number 11). Group 3 α consisted of infants with metastasis but better outcomes, Group 3 β – less metastatic patients with GFI activation and Group 3 γ – consisted of infants and that exclusively was made up of MYC amplified patients. Out of all the three subtypes, Group 3 γ was associated with poorest survival^{43,68}.

Group 4 MB patients have classic histology and often show an isochromosome 17. While some features are common between Group 3 and Group 4 MBs, such as loss of chromosome 17p, and OTX2 amplification, they have very different prognoses. MYC amplifications are almost exclusively seen in Group 3 MBs, whereas Group 4 MBs

display *MYCN* amplifications⁷⁹. Advancements in transcriptomics identified further subtypes where Group 4 α consisted of patients with *MYCN* amplification, Group 4 β patients harboured *SNCAIP* duplications and the most common Group 4 γ subtype was made up of patients with *CDK6* amplifications (Figure 2, page number 11). Although Group 4 MBs have an occurrence rate of > 30 %, we know quite less about the tumorigenic pathways and biological dysregulation occurring in this group^{43,68,80}.

Despite the differences between Group 3 and Group 4 MBs, they are more similar to each other than to SHH or WNT MBs. With progress in epigenomics and large-scale availability of samples, there were nearly eight subtypes identified within the Group 3 and Group 4 MBs (Figure 3, page number 14). While *MYC* amplification was confined to epigenetic G3/4 subtype II, *OTX2* amplifications were seen mostly in G3/4 subtype I⁸¹. *MYCN* amplifications were seen in the G3/4 subtype V. *SMARCA4* structural variations were also seen in the G3/4 subtype II⁸¹. Further proteome analysis also identified three subtypes - G3a, G3b and G4. G3a corresponded to Group 3 MBs with *MYC* amplification, G3b without *MYC* amplification and G4 MBs which were most similar to Group 4 MB and showed enrichment of neuronal signatures and glutamate-associated pathways^{77,78}.

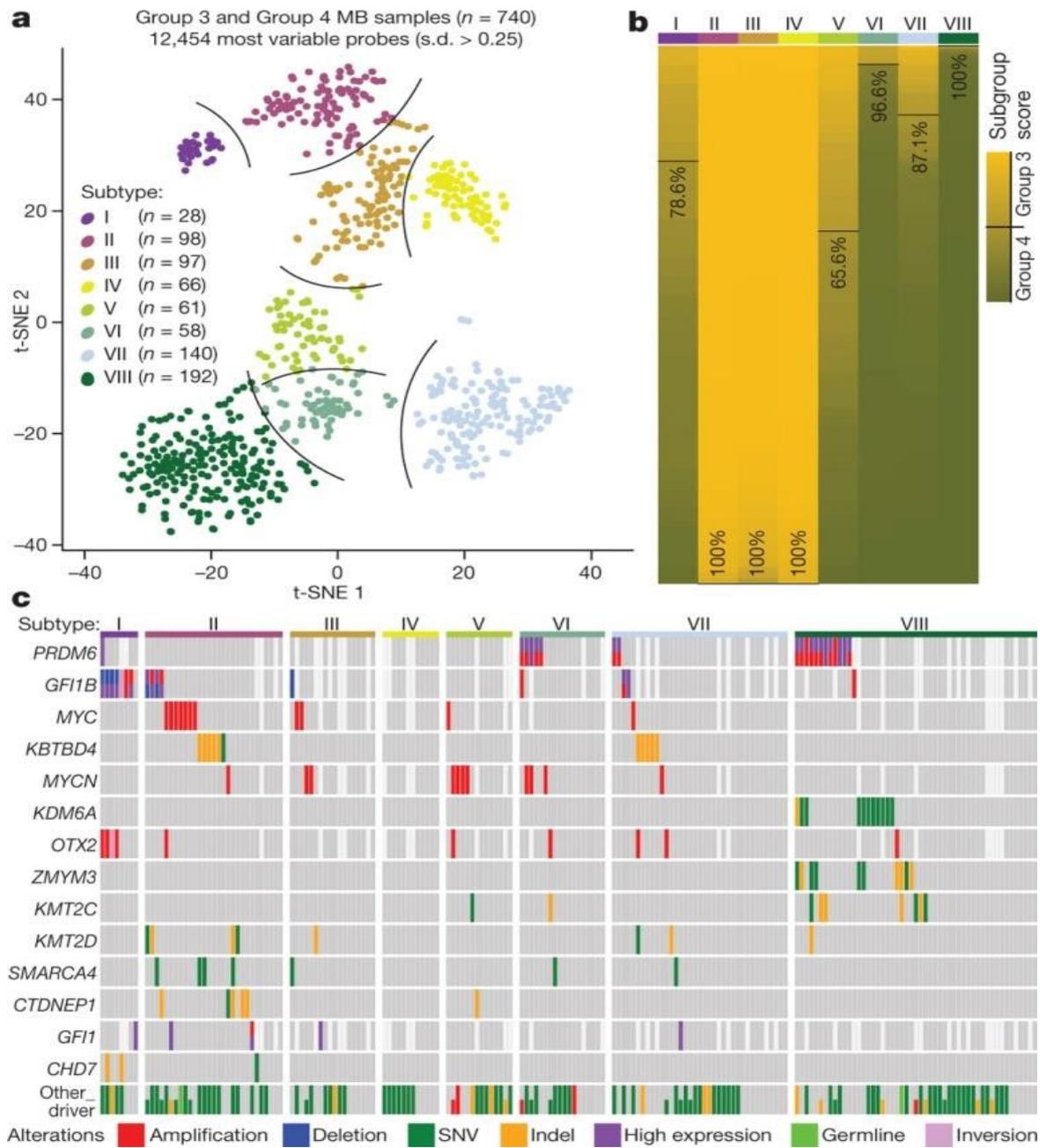


Figure 3: Figure from Northcott et al ⁸¹ (Material included under Creative Common License) displaying Group3 and Group4 DNA methylation subtypes

1.1.2.2 High-Grade Glioma MYCN amplified

Primitive neuroectodermal tumours share an undifferentiated histomorphology and have been reclassified into several different tumour subtype based on molecular parameters²². Besides the previously introduced MBs, one such entity is pediatric high-grade glioma with MYCN amplification (HGG MYCN). HGG MYCN amplified are extremely aggressive with a median survival of only 14 months^{82–84}. Histologically, they are different compared to other gliomas, characterized by undifferentiated cells, dense cell nuclei, and the absence of typical glial features. MYCN, a transcriptional factor for the MYC family is generally amplified in these tumours. Along with *MYCN* amplification, *TP53* mutations are also common. This subtype also separates from other glioma subtypes using DNA methylome profiling^{22,23}. While it can be identified as a separate entity compared to other HGG, HGG-MYCN amplified patients are still treated like HGG patients with a lack of distinct treatment regimens. Being a relatively new discovered tumour entity with aggressive growth and extremely poor survival, there is a dire need to understand the biology of these tumours and find treatment options.

1.2 : Aim of the work

In an attempt to understand the biology of the two aggressive brain tumours namely MB and HGG MYCN amplified, we performed multi-omic data integration of the DNA methylome, transcriptome and proteome data of human MB patients (shared first author paper presented in section 4.1)), and compared omic data of mouse models to human MB (co-author paper presented in section 4.2) and human HGG MYCN amplified tumors (co-author paper presented in section 4.3).

1.2.1 Aim 1 (Project 1): Identify clinically and phenotypically relevant MB subtypes

We aimed to analyse 167 cases of MBs and perform a comprehensive analysis by integrating proteome, DNA methylome, and N-glycome data. Unsupervised analyses of integrated data were performed to reveal distinct proteome subtypes that we aimed to analyse concerning pathway activation and potential biomarker representation. We further aimed to perform multiomic analysis to unveil varying degrees of conservation of proteome features across MB subtypes. Molecular features were further integrated with clinical features to identify potential risk groups⁸⁵. These findings shall contribute to our understanding of novel targetable alterations in MBs and lay the groundwork for potential immunotherapies targeting glycan structures.

1.2.2 Aim 2 (Project 2): Characterize a mouse model for the Group3 MBs with worse prognosis and survival

To further understand MB tumorigenesis and develop suitable mouse models for these rare brain tumours, we also aimed to perform multi-omic analysis by integrating mouse and human DNA methylome and transcriptome data. Group3 MBs are one of the most aggressive brain tumours with 30% of patients having genetic alterations in MYC, SMARCA4 or both genes. While previous studies have established that MYC overexpression drives MB formation in mice, the functional impact of SMARCA4 mutations remains unclear. Further, the therapeutic impact of these mutations also needs further investigation. To address this knowledge gap, we conducted experiments combining MYC overexpression with SMARCA4 loss in granule cell precursors⁸⁶. The combination of MYC overexpression and SMARCA4 loss

successfully induced tumour formation in vivo following orthotopic transplantation into recipient mice. These tumours were histologically and molecularly (RNA-seq and DNA methylome data) analysed and compared to human tumours in order to proof the suitability of SMARCA4-deficient MB mouse models as a model for Group3 MB which could be further used for targeted therapeutic strategies.

1.2.3 Aim 3 (Project 3): To understand the tumourgenesis and characterize the biology of aggressive HGG MYCN tumours

In order to delve into the process of tumourgenesis and understand the molecular biology of HGG MYCN tumours, with the ultimate aim of proposing alternative treatment strategies, we established a genetically engineered mouse model displaying MYCN overexpression and TP53 knockout in neural precursor cells ⁸⁷. In this model, all mice developed aggressive forebrain tumours early in their lifespan, and we analysed histology, DNA methylation patterns, and gene expression patterns to thoroughly compare murine tumours with HGG-MYCN tumours. The aim was to propose specific treatment strategies and develop a model system for preclinical treatment testing.

1.3 : Material and Methods

1.3.1 Materials and Data

1.3.1.1 DNA Methylome Data Generation

1.3.1.1.1 Human Samples

Human samples used for MB project 1: For the main cohort and biological validation cohort, we obtained FFPE Medulloblastoma samples dating from 1976 to 2022 from various neuropathology units in Germany, including cases collected within the HIT-MED study cohort. Additionally, single patients enrolled in the SIOP PNET5 study of the HIT-MED study centre were included. Tumour samples were obtained from patients with informed consent following the 1964 Declaration of Helsinki ethical standards. Samples were processed using standard protocols, namely, fixation of sample, dehydration of samples, embedding into formalin, and microdissection of tumour material only. DNA methylation data was generated using the Illumina Infinium HD FFPE DNA Restore Kit and EPIC BeadChip array. Data has been deposited under accession numbers GSE222478⁸⁵ and GSE243768⁸⁵.

Mouse samples used for MB project 2 and HGG-MYCN project 3: Four mouse MB Group3 tumours with MYC overexpression and SMARCA4 loss were utilized and three murine HGG-MYCN tumours were used. DNA was isolated and bisulfite conversion was performed using standard kits, followed by analysis using Illumina Mouse Methylation Bead Chip Array. Data has been deposited under accession numbers GSE235924 and GSE227413 respectively⁸⁶.

1.3.1.1.2 Publicly Available DNA Methylome Data

For MB project 1: Integration with publicly available DNA methylome data included datasets from Archer et al. (2018) EGAS0001001953 ⁷⁷ and Forget et al. (2018) GSE104728 ⁷⁸.

1.3.1.2 Transcriptome Data Generation

For human samples of MB project 1: RNA was isolated from FFPE tissue sections using the Maxwell RSC RNA FFPE Kit, followed by library preparation and sequencing. Data is deposited under accession number GSE243795⁸⁵.

For mouse samples of MB project 2 and HGG-MYCN project 3: RNA was isolated using TRIzol, followed by library preparation and sequencing. Data is deposited under accession numbers GSE235625 and GSE227413 respectively.

1.3.1.2.3 Publicly Available Transcriptome Data

Publicly available Transcriptome data MB project 1: Cavalli et al. (2017), GEO GSE85218 and GSE37382 ⁴³

Publicly available Transcriptome data MB project 2 and HGG-MYCN project 3: Sturm et al. (2016), GSE73038 ⁸⁸

1.3.1.3 Proteome Data Generation

For MB project 1: For 62 human samples, proteins were extracted from FFPE MB tissue sections and cell lines, followed by tryptic digestion and TMT labeling. LC-MS analysis was performed, and data processed using MaxQuant. Data available under PXD039319 (TMT data), and PXD048767 (validation cohorts)⁸⁵. Integration with publicly available proteome data included datasets from Archer et al. (2018), MSV000082644 ⁷⁷, Forget et al. (2018), PXD006607 ⁷⁸, Petralia et al. (2021), PXD016832 ⁵². Data was integrated and batch corrected using HarmonizR⁵⁹. Waszak et al. (2020), PXD016832 ⁸⁹ was integrated for validation.

1.3.1.4 N-Glycan Data Generation

For MB project 1: N-Glycan analysis was performed on 18 human samples, with data processed using Xcalibur Qual Browser and MaxQuant⁹⁰. Further statistical analysis was performed using Perseus software⁹¹.

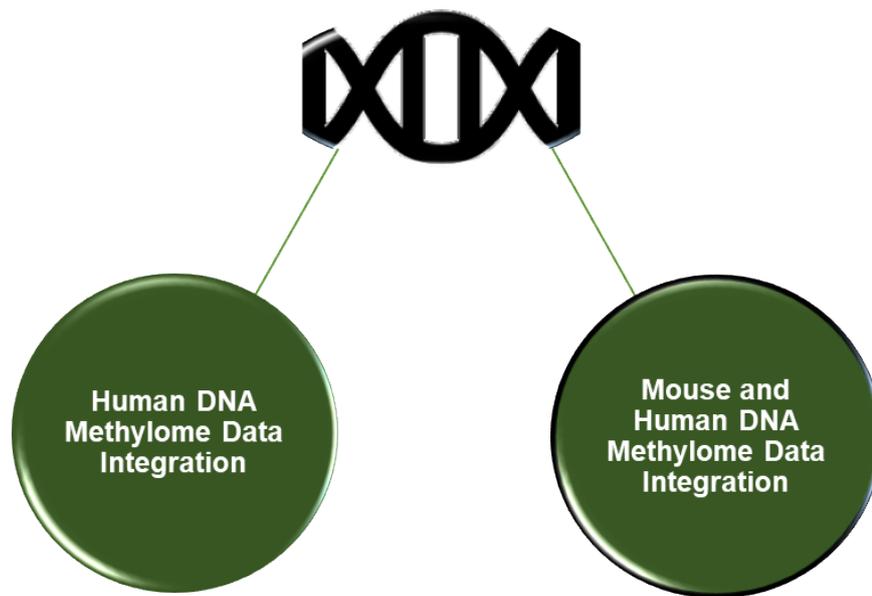
1.3.1.5 Mouse models for HGG MYCN and Group 3 MB

The experimental procedures on animals were approved by Government of Hamburg, Germany (N113/16, N050/2018, N099/2019). The mice were subjected to a 12 hour dark/light cycle, with access to food and water. Male and female mice were utilized for all experiments. hGFAP-cre⁹², Blbp-cre⁹³, Sox2-cre⁹⁴, Trp53FI/FI⁹⁵, and Isl MYCN⁹⁶ as well as Math1-cre::SmoM2FI/wt⁹⁷ were generated as previously described. Genotype was confirmed using polymerase chain reaction using DNA from ear or tail biopsies.

1.3.2 Methods

1.3.2.1 DNA Methylome Data Analysis

Data preprocessing included quality control, probe filtering, and normalization using ssNoob. Integration with mouse data involved pre-processing with minfi⁹⁸ and generation of a UMAP plot⁹⁹. For an overview see Figure 4 and for details refer to respective publications in sections 4.1, 4.2 and 4.3.



- **Human datasets from current study (MB project 1):**
GSE222478 and GSE243768
- **Publicly available human datasets for integration:**
EGAS00001001953
GSE104728
- **Human BeadChip Array:**
450K
850K
- **Package for processing raw data:**
Human:
minfi
- **Normalization technique on integrated data:**
ssNoob normalization
- **Visualizations:**
ComplexHeatmaps
UMAPs

- **Mouse dataset from current study (MB project 2 and HGG-MYCN project):**
GSE235924 and GSE227413
- **Publicly available human dataset for integration with mouse data (MB project 2 and Project 3: HGG-MYCN project):**
GSE109381
- **Mouse methylation BeadChip Array Human BeadChip Array:**
Human data : 450K array
Mouse array
- **Package for processing raw data:**
Human: minfi
Mouse: SeSAMe
- **Normalization technique on integrated data:**
Reduce data to orthologous CpG sites
Quantile normalization
- **Visualizations:**
ComplexHeatmaps
UMAPs

Figure 4: Brief outline of the DNA Methylation Data Analysis and datasets used for integration in this study **Left** in Humans **Right** in Mouse

1.3.2.2 RNA Sequencing Data Analysis

Raw fastq files were processed using FastQC and Trimmomatic for quality control in usegalaxy.eu¹⁰⁰ and alignment to reference genomes GRh38 for human and mm39.ncbiRefseq.gtf.gz for mouse respectively. Gene expression quantification and normalization were performed using featureCounts¹⁰¹ and DESeq2¹⁰², respectively.

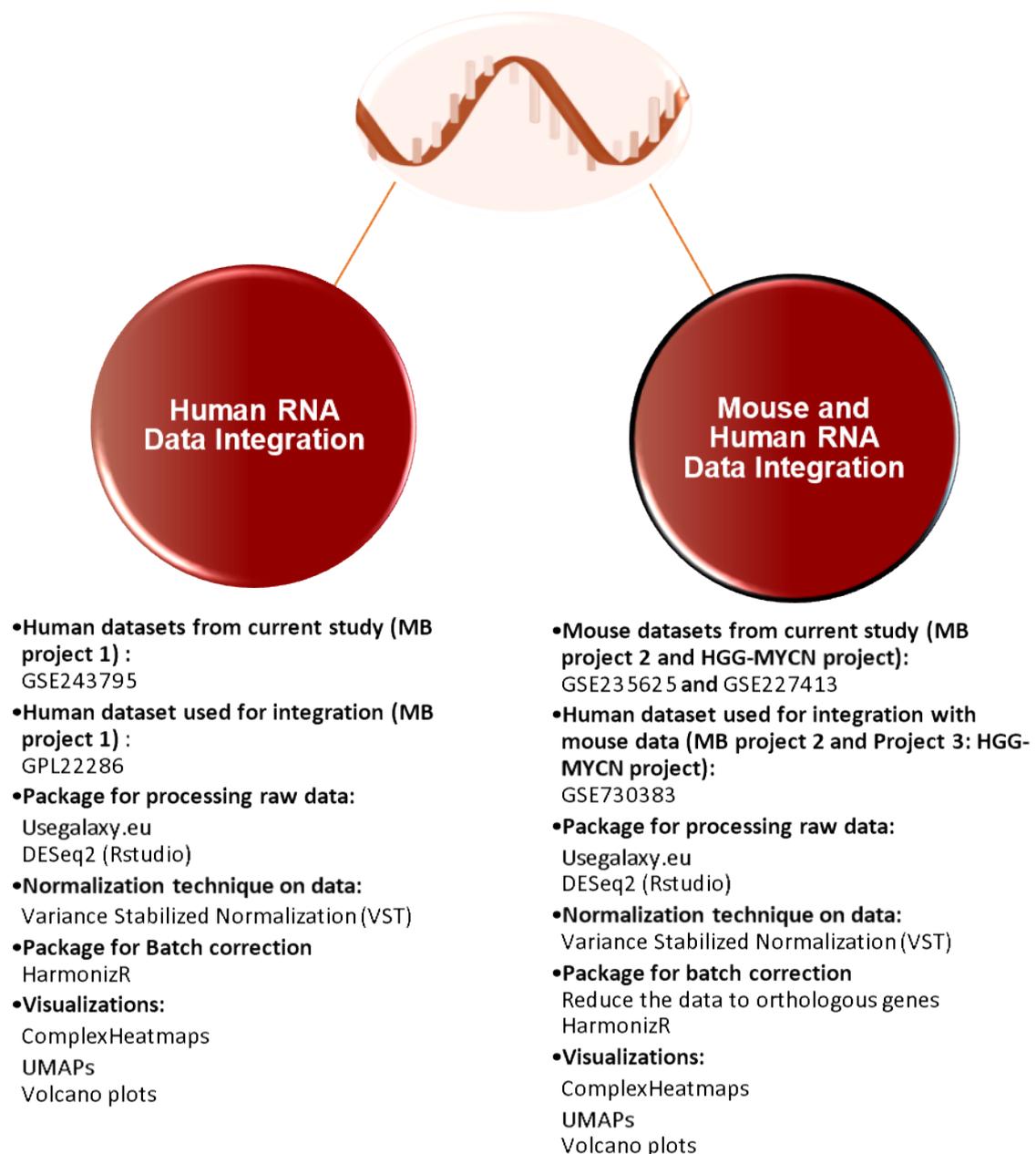


Figure 5: Brief outline of the RNA-seq Data Analysis and datasets used for integration in this study **Left** in Humans **Right** in Mouse

1.3.2.3 Proteome Data Analysis

Proteome data processing involved extraction and quantification using MaxQuant⁹⁰, followed by normalization and integration with other datasets. External validation was performed using DIA-NN and data integration with the main cohort was performed using HarmonizR⁵⁹. Further information on datasets can be found in Section: 1.3.1.1.1 and 1.3.1.1.2⁸⁵ .

1.3.2.4 N-Glycan Data Analysis

N-Glycan data ⁸⁵ were processed using Xcalibur, Glycoworkbench, and Skyline software, with further statistical analysis conducted in Perseus ⁹¹.

1.3.2.5 Data Integration and Correlation

DIABLO from mixOmics⁶¹ was used for integrating proteome and methylome data, with subsequent correlation analysis. Integration of mouse and human samples involved batch correction on orthologous genes in the RNA-seq data and orthologous CpG sites in DNA methylome data and visualization using umap and ComplexHeatmap¹⁰³.

1.3.2.6 Consensus Clustering

Consensus clustering was applied to determine the ideal number of clusters from proteome and DNA methylation data, using the ConsensusClusterPlus ¹⁰⁴ package in R.

1.3.2.7 Copy Number Variation Analysis

Copy number analysis was performed on samples with both methylation and proteomic data, involving segmentation and individual sample CNV plot generation using conumee package . Frequency plots were generated using CNAppwebtool¹⁰⁵ using the conumee segmentation data. Pearson correlation was performed between the segmenetation values in R.

1.4 : Summary of Results

1.4.1 MB project 1: Multiomic profiling of medulloblastoma reveals subtype-specific targetable alterations at the proteome and N-glycan level

In the following study, we performed multiomic characterization of MBs to identify novel biomarkers, altered pathways and clinically relevant subtypes. We aimed to improve diagnosis, classification and stratification of MB patients and to provide a basis for targeted therapies. In the following the main points of the manuscript are highlighted.

Large-Scale Proteome Analysis of MB:

- Integration of in-house proteome data generated in different measurement batches, enabled the analysis of 62 formalin-fixed paraffin-embedded (FFPE) MB tumours, comprising the four main established molecular subgroups of MB: SHH (n=57), WNT (n=19), Group 3 (n=36), and Group 4 (n=55), MB Project 1: Figure 1A, page number 87, Supplementary figures 4A, page number 109)⁸⁵.
- The age of paraffin material did not affect sample clustering or protein detection, indicating FFPE tissue's suitability for proteome analysis (MB project 1: Figure 1B,1C, page number 87, Supplementary figure 4B, page number 109)⁸⁵.
- We further expanded our cohort by integration of fresh frozen (FF)-MB proteome datasets from public repositories to increase the cohort size. After data harmonization using the HarmonizR 16,279 proteins were quantified across 167 samples, including 156 primary tumours and 11 recurrences (MB project 1: Figure 1D-1G, page number 87, Figure 2D, page number 89, Supplementary figures 1-C, page number 106 Supplementary figures 5A-B, page number 110)⁸⁵.
- We confirmed that established protein biomarkers for molecular MB subtypes such as GAB1 for SHH and WNT MBs, CTNNB1 for WNT MB and FLNA for SHH MB showed expected abundance patterns in individual studies and in the combined and harmonized data (MB project 1: Figure 1H, page number 87)⁸⁵.

Six Proteomic MB Subtypes:

- Consensus Clustering was applied to detect sample clusters within the data in an unbiased manner. Six stable proteomic clusters were identified termed pWNT, pSHHs, pSHHt, pSHHs, pG3myc, pG3, pG4⁸⁵.
- Proteome clusters were compared to established DNA methylome subtypes and transcriptomic subtypes revealing clear associations (N=117 DNA methylation data available and n= 56 RNA-seq data available, MB project 1: Figure 2A-D, page number 89 Supplementary figures 6A-C, page number 111)⁸⁵.
- Cluster Characteristics: Non-WNT/non-SHH MBs were divided into three proteome subgroups (pG3, pG3myc, pG4), while SHH MBs separated into two subgroups (pSHHs: synaptic profile, pSHHt: transcriptional profile). WNT MB formed a homogenous cluster pWNT (MB project 1: Figure 2B, page number 89). These proteome clusters were confirmed in the sole FFPE cohort, and also in matched proteome and transcriptome data (MB project 1, Supplementary figures 3A-F, page number 108)⁸⁵.
- Proteome clusters associated with clinical parameters and survival rates. pG3myc MBs showed the worst overall survival and were associated with high-risk features such as anaplastic histology and *MYC* amplification, while pWNT patients showed the best survival rate (MB project 1: Figure 2E, page number 89, Figure 6A, page number 96)⁸⁵.
- Potential protein Biomarkers for each proteome MB subtype were defined: For high-risk pG3myc MBs, Palmdelphin (PALMD) was established as a highly conserved biomarker and confirmed in case-matched samples and RNA data (MB project 1: Figure 2F, page number 89, Supplementary figure 3G, page number 108, Supplementary figures 14 A-D, page number 119)⁸⁵.
- The six proteome subtypes matched to two main proteome profiles at super-ordinate level – one profile highly enriched for a transcriptional signature (pSHHt, pWNT and pG3myc) and the other group highly enriched for a synaptic signature (pSHHs, pG3 and pG4) with cell-cycle signaling and opioid signaling pathways as potential therapy targets, respectively (MB project 1: Figure 3A-E, page number 90, Supplementary figures 7A-B, page number 112)⁸⁵.

Group-Specific Correlation of DNA Methylome and Proteome:

- Multiblock data integration and Pearson correlation analyses between DNA methylation and proteome data revealed subtype-specific correlations, indicating different conservation levels of molecular characteristics across proteomic MB subtypes. WNT MB subtype showed the highest correlation and conservation of features followed by pG3 MB whereas the other MB subtypes showed low-correlation based on CpG sites as well as differentially methylated regions (DMRs, MB project 1: Figure 4A-B, page number 92, Supplementary figures 9A-E, page number 114, Supplementary figures 10A-B, page number 115)⁸⁵.
- Identified protein biomarkers' reflection at DNA methylation level was also investigated, with generally a negative correlation for all top biomarker candidates but RPH3A protein and its respective gene's CpG sites (MB project 1: Figure 4B, 4D, page number 92)⁸⁵.

SHH MB Subtypes:

- SHH MB divided into two proteome subtypes (pSHHt and pSHHs), associated with distinct DNA methylation subtypes. (MB project 1: Figure 5A, page number 94)⁸⁵.
- Differential pathway enrichments were observed, with pSHHs showing synaptic and immunological processes upregulated, and pSHHt showing transcriptional and replicative signatures. Both subtypes showed a low correlation between DNA methylome and proteome copy number variation (CNV) profiles (MB project 1: Figure 5C, F, H, D, G, page number 94)⁸⁵.
- Nearly all *TP53*-mutated SHH MBs belonged to the pSHHt subtype (9/10), MB project 1: Figures 5A, I, page number 94)⁸⁵.
- Distinct metabolic patterns were identified, with pSHHs showing alterations in the TCA cycle and neurotransmitter metabolism (MB project 1: Supplementary figures 12A-C, page number 117)⁸⁵.

High-Risk pG3myc MBs:

- Three non-WNT/non-SHH MB subtypes were identified, with high risk pG3myc MBs characterized by enrichment of RNA processing, ubiquitinylation and sumoylation. pG3 and pG4 showed high enrichment of mRNA splicing, DNA damage response, VEGF signalling and vesicle mediated transport processes. All three subtypes showed a low correlation between the DNA methylome and proteome CNV profiles (MB project 1: Figure 2E, page number 89, 6B-J, K, Q, page number 96)⁸⁵.

pWNT MBs:

- pWNT MBs did not divide into further subtypes and showed a low abundance of proteins of the TriC/CCT multiprotein complex. Significant changes in TriC/CCT components were not reflected by other omic modalities before. TriC/CCT abundance changes potentially impact on vincristine resistance and chemotherapy response, suggesting therapy deacceleration in pWNT patients (MB project 1: Figure 7A, page number 98, Figure 8A-E, page number 100, Supplementary figures 18A-B, page number 123)⁸⁵.
- Enrichment of extracellular matrix proteins and N-glycan biogenesis, transport and presentation were observed. Comparing CNV profiles calculated from the DNA methylome and proteome data revealed a high correlation with chromosome 6 loss also being reflected at the proteome level (MB project 1: Figure 7H, I, page number 98, Figure 4A-B, page number 92, Figure 7J, page number 98)⁸⁵.
- TNC was identified as a biomarker for the pWNT subtype and was confirmed using Immunohistochemistry (MB project 1: Figure 7B-G, page number 98)⁸⁵.

N-Glycan profiles of proteome MB subtypes:

- N-Glycans represent a type of posttranslational modification. N-Glycan profiles varied across proteomic MB subtypes, with differential abundance of N-glycan species in proteome subtypes (MB project 1: Figure 9C-D, page number 102, Supplementary figures 16A-B, page number 121)⁸⁵.

- Differential N-Glycan profiles could be used as a potential immune therapy targets for MBs with high resistance to chemo and radio therapy⁸⁵.

Validation and confirmation of the results:

- The findings were validated in technical and biological validation datasets, confirming the six proteome subtypes and associated characteristics (MB project 1: Figure 10A-I, page number 104, Supplementary figures 17A-G, page number 122)⁸⁵.
- TriC/CCT complex emerged as a discriminator between pWNT and pG3myc MB subtypes, indicating potential clinical relevance which needs further *in-vitro* investigation (MB project 1: Figure 8, page number 100, Supplementary figure 15, page number 120)⁸⁵.

Overall, integrated proteome analysis provides insights into MB subtypes, their molecular characteristics, and potential therapeutic targets, paving the way for personalized treatment strategies and improved patient outcomes.

1.4.2 MB project 2: MYC overexpression and SMARCA4 loss cooperate to drive medulloblastoma formation in mice

In this study we aimed to characterize murine brain tumours with *Myc* overexpression and *SmarcA4* loss (MYC/SMARCA4 tumours) on a molecular level using RNA sequencing and investigated their similarity to human Group 3 medulloblastoma.

Characterization of MYC/SMARCA4 Tumours:

- The study utilized RNA sequencing of FFPE biopsy punches from four mouse MYC/SMARCA4 tumours. RNA from a previously established SHH MB mouse (*Math1-cre::Smo^{fl/wt}*) model and *Math1creERT2::Smarc4^{fl/fl}* cerebella at postnatal day 7 (P7) served as controls. The comparison revealed significant upregulation of *Myc* in MYC/SMARCA4 tumours, accompanied by downregulation of genes associated with neuronal development and differentiation (MB project 2: Figures 1A-J, page number 128, Figure 2A-O, page number 129, Figure 3A, page number 131)⁸⁶.
- Gene set enrichment analyses (GSEA) highlighted the upregulation of ribosome biogenesis and rRNA synthesis, a characteristic of *MYC*-driven cancers. Comparison with the SHH MB model confirmed *Myc* upregulation and downregulation of *MycN*, a target of SHH signalling. Additionally, MYC/SMARCA4 tumours exhibited upregulation of genes associated with increased malignancy in gliomas and downregulation of markers for granule cells, indicative of a less favourable prognosis (MB project 2: Figure 3B-C, page number 131)⁸⁶.

Molecular Resemblance of murine tumours to human Group 3 MB based on gene expression:

- Integration of RNA sequencing data with previously published gene expression data from human pediatric brain tumours revealed that MYC/SMARCA4

tumours exhibited molecular resemblance to human MB, particularly Group 3 MB (MB project 2: Figure 4 A, B, D, E, page number 132)⁸⁶.

- Uniform Manifold Approximation and Projection (UMAP) and Euclidean clustering analyses with human MB samples demonstrated shared similarities with both SHH MB and Group 3/4 MB. Further gene expression analysis confirmed the closest similarity to Group 3 MB, with MYC/SMARCA4 tumours clustering to this subgroup (Figure 4A, D, page number 132). These findings suggest that MYC/SMARCA4 tumours share molecular features with Group 3 MB, supporting their classification within this subgroup (Figure 2C-O, page number 129)⁸⁶.

DNA Methylation Profile Analysis:

- To further validate the classification of MYC/SMARCA4 tumours, DNA methylation analysis was performed on three mouse tumours and compared to human MB (n=228) based on orthologous CpG sites. MYC/SMARCA4 tumours clustered in close proximity to Group 3/4 MB, consistent with the gene expression analyses. Distance plot analysis confirmed the highest resemblance of MYC/SMARCA4 tumours to Group 3 MB based on DNA methylation profiles (MB project 2: Figure 5A-B, page number 133)⁸⁶.

In summary, the integrated molecular analyses indicate that MYC/SMARCA4 mouse tumours closely resemble Group 3 MB in humans based on gene expression and DNA methylation patterns. These discoveries advance our understanding of the development of Group 3 MB and their characteristics and may lead to targeted therapies tailored to specific molecular subtypes.

1.4.3 HGG- MYCN project 3: Mouse models of pediatric high-grade gliomas with *MycN* amplification reveal intratumoral heterogeneity and lineage signatures

In this study we focused on comparing tumours generated in mouse models with *MycN* overexpression and *Tp53* knockout (*MycN*/*Tp53* tumours) with human high-grade gliomas with *MYCN* amplification (HGG-MYCN), particularly regarding molecular similarities.

DNA Methylation Profiles:

- In humans, tumours were found throughout the entire brain, with the majority located in the temporal and frontal lobes (HGG-MYCN project 3: Figure 1B-G)⁸⁷.
- CNV analysis revealed imbalanced profiles with clear *MycN* amplification, confirmed by FISH and Immunohistochemistry (HGG-MYCN project 3: Figure 1H, L, page number 139)⁸⁷.
- Nuclear p53 accumulation indicative of impaired p53 function was detected by Immunohistochemistry (HGG-MYCN project 3: Figure 1I-K, page number 139).
- Bead chip arrays were used to detect the methylation status of 285,000 CpG sites in the mouse genome, with 141 orthologous CpG sites compared to human CPG sites detected via the 850k array⁸⁷.
- The methylation data of murine *MycN*/*Tp53* tumours (n=3), showed high similarity to human HGG-MYCN tumours when visualized using UMAP (HGG-MYCN project 3: Figure 2, page number 140)⁸⁷.

Transcriptomic Profiles:

- Transcriptomic profiles of murine tumours were generated and compared to gene expression data of human HGG-MYCN and other pediatric brain tumours, revealing high similarity between mouse *MycN*/*Tp53* tumours and their human HGG-MYCN counterparts using UMAP, Euclidean distance and Agreement of Differential Expression (AGDEX) analyses (HGG-MYCN project 3: Figure 3A-D, page number 141)⁸⁷.

- Gene set enrichment analysis (GSEA) for MYCN target genes showed significant enrichment of corresponding mRNAs in both human and mouse tumours (HGG-MYCN project 3: Figure 3E-F, page number 141)⁸⁷.

Comparison of murine MycN/Tp53 tumours with other tumour models:

- Comparison of *Math1-cre::SmoM2^{F/wt}* mice, a model for SHH-MB, showed expected similarities between mouse and human SHH MB tumours⁹⁷ and was used as a positive control. While our tumour model showed similarity to HGG-MYCN amplified human samples. (HGG-MYCN project: Figure 3A-D, page number 141)⁸⁷.
- Histologically too, these mouse tumours resembled their human counterparts (HGG-MYCN project: Figure 2A-T, page number 140)⁸⁷.

Overall, the study demonstrates that mouse models with *MycN* overexpression and *Tp53* knockout closely resemble their human HGG-MYCN counterparts both molecularly and histologically. These findings provide valuable insights into the biology of these rare and aggressive HGG-MYCN tumours and suggest that mouse models can serve as useful tools for studying human brain tumours and developing therapeutic strategies.

1.5 : Discussion

To address the problem of ineffective treatment strategies and improve the biological understanding of tumours we characterized the most common- MBs and highly aggressive HGG MYCN-amplified tumours. In this study we used DNA methylome and transcriptome data for both MB and HGG MYCN patients and additionally proteome and N-glycome data of MB patients to unravel phenotypic subtypes and characterize post-translational changes. We integrated different omic datasets, found clinically relevant subtypes and discovered new features which have not been described in all our projects. Finally, to understand the tumorigenesis we developed mouse models for the aggressive Group 3 MBs and HGG MYCN-amplified tumours. We integrated the human and mouse data to confirm similarity to their human counterparts. This indeed enhanced our understanding about the biological dysregulations of the two tumour types and enabled the use of these models for testing alternate therapy options.

1.5.1 Human Multi-omics analysis of MBs

1.5.1.1 Six proteome subtypes of MB belong to two main molecular profiles

Even though several subtypes have been identified for MBs at DNA methylome and transcriptome level, they proved to be unsuccessful for providing accurate therapy. One reason for this is the low mutation burden in these tumours and also that these omics are farthest from the phenotype of the disease⁵². Thus, we characterized the proteome data, integrated and harmonized our data with publicly available proteome datasets to enhance our cohort size. We performed consensus clustering analysis and identified six subtypes at proteome level which correlated with the clinical data. These six proteome subtypes belong to two main subgroups – one showing transcriptional pattern consisting of the pG3myc, pWNT and pSHHt subtypes and the other showing a synaptic pattern consisting of pG3, pG4 and pSHHs subtypes. These two main profiles that the six proteome subtypes belong to, show some common features and dysregulated pathways which can be potential therapy targets. Using Ingenuity Pathway Analysis (IPA) we found that CDK4 inhibitors can be useful for targeting the transcriptional profile subtypes (MB project 1: Figure 3E, page number 90). Various CDK4 inhibitors are already used for treatment of metastatic tumours and also used for targeting MB cells *in vivo*^{106,107} For the synaptic profile subtypes, NMDA receptor

antagonist memantine can be used (MB project 1: Figure 3D, page number 90). It is known to be neuroprotective and decreases cognitive dysfunction in patients treated with radiotherapy¹⁰⁸.

pWNT MBs did not form further subtypes. They had the highest overall survival which is in-line with literature. Staining for high expression of *CTNNB1* is commonly used for diagnosis of WNT MBs, however not all cases show a strong positive immunohistochemical staining for this marker⁶⁷. Thus, we discovered TNC as a potential biomarker for detection of pWNT MBs, wherein all cells show a strong positive staining for this protein⁸⁵. TNC, an extracellular matrix protein is associated with tumour proliferation or reduction depending on the splice variant present. WNT-pathway activation is associated with blood brain barrier disruption. This permits high response to chemotherapy¹⁰⁹. TNC could be further contributing to this feature as a high level of TNC is also associated with blood brain barrier disruption¹¹⁰. This factor also contributes to the better response of patients to chemotherapy and can be a potential marker for therapy deacceleration.

SHH MBs further divided into two subtypes namely pSHHt and pSHHs showing higher amounts of transcriptional and synaptic signalling respectively⁸⁵. pSHHt subtype mainly consisted of the *TP53* - mutated patients, higher abundance of *CHD6*, *DNAJB2* and *NNMT* associated with aberrant *TP53* – mutations, adult patients and comprised of SHH3 and SHH4 methylation subtypes (MB project 1: Figure 5A, J, page number 94). pSHHt was associated with worse prognosis compared to pSHHs MBs (MB project 1: Figure 5I, page number 94). pSHHt also resemble SHHa subtype described by Archer et al⁷⁷. This group showed high enrichment of cell cycle and DNA repair pathways. *ALDH1A3* was identified as a good diagnostic biomarker for pSHHt, confirmed using immunohistochemistry (MB project 1: Figure 5E, page number 94 Supplementary figure 11C-D, page number 116). We also confirmed higher proliferation of pSHHt compared to pSHHs MBs using Ki-67 staining (MB project 1: Supplementary figures 11E-F, page number 116). Further, recently identified *ELP1* mutations were found mutually exclusive with *TP53* mutations and were mainly a part of pSHHt subtype. The pSHHs subtype resembles SHHb subtype discovered by Archer et al⁷⁷ showing high synaptic signalling, consisting mainly of infant patients and

SHH1 and SHH2 methylation subtypes. We also found a significantly higher abundance of GABA targets, further contributing to their synaptic profile and could be good therapy targets for treatment of pSHHs MBs (MB project 1: Figure 5B, page number 94, Supplementary figure 12C, page number 117).

Non-WNT/ non-SHH MBs further divided into three subtypes. pG4 MBs, showed higher abundance of synaptic processes which is already known for Group 4 MBs^{43,77}. We also observed a higher enrichment of VEGF signalling associated with tumour angiogenesis (MB project 1: Figure 6B-C, page number 96). VEGF signalling has been proposed as a good target for Mebendazole¹¹¹. This group consisted mainly of methylation subtypes V, VI, VII and VIII. Next, we identified pG3myc tumours which consisted of not all but most of the patients with genetic *cMYC* amplification and all patients with a higher abundance of *cMYC* target proteins (MB project 1: Figure 6K, page number 96). *cMYC* amplified MBs are associated with a worse prognosis, which was also the case for pG3myc patients compared to all other proteome subtypes. Thus, proteome profiling can be useful for identification of high-risk patients which do not show genetic *MYC* amplification but still have a high abundance of *MYC* target proteins. Additionally, this group consisted of methylation subtype II which is also associated with poor prognosis. We identified PALMD as a good diagnostic biomarker for this group and confirmed with immunohistochemistry⁸⁵. Finally, the pG3 MBs were the ones with lowest cluster certainty and showed features of both pG4 and pG3myc with high enrichment of mRNA splicing, DNA damage response and SMAD signalling (MB project 1: Figure 6E-F, page number 96). They mainly consisted of methylation subtype I, IV and VIII and some *cMYC* amplified patients, SMARCA4 loss was mostly seen in pG3 MBs.

1.5.1.2 Multi-omic data integration provides insights into stream of biological information at different levels and identification of new features

In order get novel biological insights and investigate especially the pWNT MBs – with best survival and pG3myc MBs – with worse survival, we performed multi-omic analysis. We wanted to decipher features which could indeed support therapy

deacceleration in pWNT MBs and identify specific characteristics for a poor prognosis of pG3 and pG3myc MBs.

DNA methylation data is commonly used in diagnostics and has been employed for development of the brain tumour classifier³². To compare our proteome data with current gold-standard for diagnosis, we integrated our proteome data with DNA methylome data and found that the DNA methylome subtypes were reflected at proteome level. Loss of chromosome 6, a characteristic diagnostic criterion for pWNT MBs, was also reflected at the proteome level (MB project 1: Figure 7J, page number 98). Further, we investigated if the identified proteome biomarkers already reflect at the DNA methylome level. However, very few of these biomarkers showed correlation with their respective CpG sites suggesting post-transcriptional or translational changes. To investigate this further, we correlated the protein abundance of these biomarkers with their respective gene at transcriptome level. A high correlation between the transcriptome and proteome data was observed. The conservation patterns were different across the six proteome subtypes. pWNT MBs showed the highest correlation between DNA methylome, proteome and transcriptome data, followed by pG3 MBs. The other subtypes showed low correlation between DNA methylome and proteome data and high correlation between transcriptome and proteome data. Since the conservation patterns were different for different proteome subtypes, we hypothesised that post-translational changes could account for this difference in correlation. Thus, we characterised the N-glycan patterns to attribute their effect on the post-translational changes. Indeed, in-line with our previous findings, non-WNT/non-SHH proteome subtypes had highest number of N-glycans quantified. This, along with a general low correlation between omics for these subtypes supports that there could be higher amount of post-translational changes in non-WNT/non-SHH subtypes.

We discovered that, even though pWNT MBs and pG3myc MBs showed significant differences in survival, they showed highest similarity at the superordinate level, this could be attributed due to higher abundance of MYC target proteins in the pG3myc subtype and association of MYC with WNT-signaling (MB project 1: Supplementary figure 13, page number 118). In order to investigate this further, we compared the two groups across different omics. Such multi-omic data integration led to identification of

new features which have not been described for MBs before. We found a significant difference in the abundance for the TriC-/CCT complex in between pG3myc and pWNT MBs. This complex is made up of eight components TCP1-CCT8 and is associated with protein folding. Higher abundance of this complex is associated with resistance to chemotherapy^{112,113}. While it had highest abundance in pG3myc, it had lowest abundance in pWNT MBs. Further, there was no difference observed in the components of this complex at DNA methylome and transcriptome level across the six proteome subtypes (MB project 1: Figure 8C-E, page number 100, Supplementary figures 15A-B, page number 120). Thus, higher abundance of TriC/CCT complex along with identification of higher number of complex N-glycan structures in pG3myc MBs suggests that patients belonging to this group could potentially be treated with immune therapy rather than chemo or radiotherapy¹¹⁴⁻¹¹⁶. Confirming our previous proposal of therapy deacceleration for pWNT MBs, a higher abundance of TNC combined with low abundance of the TriC/CCT-complex could be a prognostic criterion which could indeed reduce therapy-based side-effect mortality in pWNT MB patients. Further validations using in-vitro experiments manipulating the TriC/CCT complex will help to strengthen these findings.

1.5.2 Cross species data integration

To understand tumorigenesis of the aggressive Group 3 MBs, mouse models with *MYC* amplification and *SMARCA4* loss were generated. *SMARCA4* loss or *MYC* amplification alone did not drive tumour formation in the mice. Genetically engineered mouse carrying both the alterations resembled human Group 3 MBs histologically. We then analysed the DNA methylome and transcriptome data of our mice and integrated this with publicly available human PNET dataset⁸⁸. While we can see that the generated mouse model clusters and lies closest to the human Group 3 MB subtype, we also observed a high similarity to the human SHH MB subtype. This could be attributed to the cellular origin of the tumours. SHH MBs are known to arise from granule cell precursors (GCPs)¹¹⁷⁻¹¹⁹ they are also among many other neural progenitors which have been targeted for modelling Group 3 MBs^{39,120,121}. Although, recent publication points that Group 3 and Group 4 MBs arise from unified rhombic lip which are not present in the mice, the current mouse model could still be useful for

exploring therapeutic potential for patients with both genetic alterations or further investigation *SMARCA4*-deficient MB mouse models.

For HGG *MYCN* amplified, *TP53* loss and forced *MYCN* expression in neural precursors were sufficient to develop large tumours in the mouse model. These tumours resemble their human counterpart histologically and based on known immunohistochemical markers. However, the age of onset differed between human and mice. While these tumours occur at a median age of eight years in humans, in mice they occurred at adolescence. This could be attributed to the species-specific differences. Further, *TP53*-loss and *MYCN*-amplification are also a characteristic mutations of other brain tumours such as MBs and EPNs, however, none of the mice had any tumours in the cerebellum or spine. While, human HGG *MYCN* amplified occur in the cerebellar regions, our mice form tumours in the olfactory bulb (OB) which could be due to differences in sensory input. It is well-known in literature that visual and hearing systems are more prone to tumour development in humans^{122,123}.

Thus, although the same genetic drivers are observed in humans, the combination of both *TP53*-loss and *MYCN*-amplification are only observed in 36% of human tumours. Therefore, even though histologically and at DNA methylome and transcriptome level, the genetically engineered mouse model is closest to its human counter-part, it is not necessarily the best method for translational evidence of this mouse model. Thus mouse-models along with human sample analysis, inclusion of cell-line experiments will better characterize the tumorigenesis and effects of different therapies. Nevertheless, currently, this mouse-model is the first ever HGG *MYCN*-amplified mouse model, that can be used for further characterizing the poorly defined brain tumour entity.

1.5.3 Limitations and future scope

In contrast to the benefits of these findings, there are also some limitations associated with our study. For all of our studies, we used FFPE tissues and micro dissected tumour dense areas. This method of micro dissecting tumour area could account for some bias in our results. While this cannot be completely avoided, we can attribute to

this bias by performing intratumoral analysis such as using spatial - transcriptomics and proteomics provided there is sufficient sample amount. The lack of correlation between DNA methylome and proteome data can also be due to the array design of the methylation chips and the CpG sites considered for correlation. Future studies focusing on CpG site-specific correlation could enable identification of CpG sites of genes which have an influence on gene expression. While we identified, TriC/CCT complex inhibitors as potential therapy targets for high-risk MBs, this needs to be confirmed using further in-vitro experiments by employing MB cell lines which have been knocked down for these complex components and then quantifying their response to chemo and radio therapeutics.

In the developed mouse models, we could identify similarities to human counterparts, however, we cannot completely overlook the species-specific differences and these cannot be completely removed by batch-effect correction. While there can be some similarities to humans, in-vitro experiments can help in confirming the dysregulated pathways. To further overcome the species-specific differences, true results can be achieved by developing tumour organoids for deeper understanding of the disease as it evolves in humans and curb the use of animal models.

In conclusion, different levels of biology identify unique features. To unravel the complexity of these diseases, multi-omic approach can be useful. It will enable us to understand the biological dysregulations and at which level the dysregulation occurs. Additionally, we can determine personalized treatment strategies and new therapeutic targets as mentioned for pG3myc and WNT MBs.

Thus, this study, provides insights and new characteristics of the six clinically relevant subtypes of MBs, which can be diagnosed using novel biomarkers. Genetically engineered mouse models of Group 3 MBs and HGG MYCN-amplified characterize these rare tumour entities further and show similarities at different levels to its human counterpart, thus being a model for further studies.

2. List of abbreviations

Congenital brain tumours (CBT)

lncRNA (long non-coding RNA)

Central Nervous system (CNS)

Medulloblastoma (MB)

High-Grade Gliomas (HGG)

High-Grade Gliomas with MYCN amplifications (HGG MYCN)

Ependymomas (EPN)

Embryonal tumours with multilayered rosettes (ETMRs)

DNA methyl transferase enzymes (DNMTs)

both fresh frozen (FF)

formalin-fixed paraffin-embedded (FFPE) tissues

DIABLO (Data Integration for Biomarker Discovery using Latent component method for Omics studies)

Wingless related Integration site (WNT)

adenomatous polyposis coli gene (APC)

beta-catenin gene (*CTNNB1*)

Dickkopf-related protein 1 (DKK1)

Large-Cell Anaplastic Histology (LCA)

Sonic Hedgehog Pathway (SHH)

Suppressor Of Fused Homolog (*SUFU*)

Patched 1 (*PTCH1*),

Smoothed, Frizzled Class Receptor (*SMO*)

GLI Family Zinc Finger 1 (*GLI1*)

MYCN Proto-Oncogene, BHLH Transcription Factor (*MYCN*)

Tumour Protein P53 (*TP53*)

GRB2 Associated Binding Protein 1 (*GAB1*),

Medulloblastoma with extensive nodularity (*MBEN*)

Telomerase Reverse Transcriptase (*TERT*)

Orthodenticle Homeobox 2 (*OTX2*)

Synuclein Alpha Interacting Protein (*SNCAIP*)

Cyclin Dependent Kinase 6 (*CDK6*)

Palmdelphin (*PALMD*)

Tenascin C (*TNC*)

SWI/SNF Related, Matrix Associated, Actin Dependent Regulator Of Chromatin, Subfamily A, Member 4 (*SMARCA4*)

Ingenuity Pathway Analysis (*IPA*)

Elongator Acetyltransferase Complex Subunit 1 (*ELP1*)

Granule Cell Precursors (*GCPs*)

Vascular Endothelial Growth Factor A (*VEGF*)

T-Complex 1 (*TCP1*)

Chaperonin Containing TCP1 Subunit 2 (*CCT2*)

Chaperonin Containing TCP1 Subunit 3 (*CCT3*)

Chaperonin Containing TCP1 Subunit 4 (*CCT4*)

Chaperonin Containing TCP1 Subunit 5 (*CCT5*)

Chaperonin Containing TCP1 Subunit 6A (*CCT6*)

Chaperonin Containing TCP1 Subunit 6B (*CCT6*)

Chaperonin Containing TCP1 Subunit 7 (*CCT7*)

Chaperonin Containing TCP1 Subunit 8 (*CCT8*)

Principle Component Analysis (PCA)

Uniform Manifold Approximation and Projection (UMAP)

Copy number variations (CNVs)

Deoxyribonucleic Acid (DNA)

Ribonucleic Acid (RNA)

Illumina BeadChip Array with 450,000 CpG sites (450K)

Illumina BeadChip Array with 850,000 CpG sites (850K/EPIC)

Alpha (α)

Beta (β)

Gamma (γ)

Delta (δ)

FISH (Fluorescence In Situ Hybridization)

Agreement of Differential Expression (AGDEX)

Gene set enrichment analysis (GSEA)

Single nucleotide polymorphism (SNP)

3. References

1. Subramanian, S. & Ahmad, T. Childhood Brain Tumors. in *Childhood Brain Tumors* (StatPearls Publishing, Treasure Island (FL), 2024).
2. DeVita Jr, V. T., Lawrence, T. & Rosenberg, S. *Cancer: Principles and Practice of Oncology (11th Edition)*. (2019).
3. Donahue, B. *et al.* Radiation therapy quality in CCG/POG intergroup 9961: implications for craniospinal irradiation and the posterior fossa boost in future medulloblastoma trials. *Front Oncol* **2**, (2012).
4. Câmara-Costa, H. *et al.* Neuropsychological Outcome of Children Treated for Standard Risk Medulloblastoma in the PNET4 European Randomized Controlled Trial of Hyperfractionated Versus Standard Radiation Therapy and Maintenance Chemotherapy. *International Journal of Radiation Oncology*Biophysics* **92**, 978–985 (2015).
5. Major, N. *et al.* The Current State of Radiotherapy for Pediatric Brain Tumors: An Overview of Post-Radiotherapy Neurocognitive Decline and Outcomes. *J Pers Med* **12**, 1050 (2022).
6. Papanicolau-Sengos, A. & Aldape, K. DNA Methylation Profiling: An Emerging Paradigm for Cancer Diagnosis. *Annual Review of Pathology: Mechanisms of Disease* **17**, 295–321 (2022).
7. Aftimos, P. G., Barthelemy, P. & Awada, A. Molecular biology in medical oncology: diagnosis, prognosis, and precision medicine. *Discov Med* **17**, 81–91 (2014).
8. Serrano, J. & Snuderl, M. Whole Genome DNA Methylation Analysis of Human Glioblastoma Using Illumina BeadArrays. in 31–51 (2018). doi:10.1007/978-1-4939-7659-1_2.
9. Kling, T., Wenger, A., Beck, S. & Carén, H. Validation of the MethylationEPIC BeadChip for fresh-frozen and formalin-fixed paraffin-embedded tumours. *Clin Epigenetics* **9**, 33 (2017).
10. Moran, S. *et al.* Validation of DNA methylation profiling in formalin-fixed paraffin-embedded samples using the Infinium HumanMethylation450 Microarray. *Epigenetics* **9**, 829–833 (2014).
11. Zong, L. *et al.* A comprehensive assessment of exome capture methods for RNA sequencing of formalin-fixed and paraffin-embedded samples. *BMC Genomics* **24**, 777 (2023).
12. Decruyenaere, P. *et al.* RNA Extraction Method Impacts Quality Metrics and Sequencing Results in Formalin-Fixed, Paraffin-Embedded Tissue Samples. *Laboratory Investigation* **103**, 100027 (2023).
13. Gustafsson, O. J. R., Arentz, G. & Hoffmann, P. Proteomic developments in the analysis of formalin-fixed tissue. *Biochim Biophys Acta Proteins Proteom* **1854**, 559–580 (2015).
14. L Avaritt, N. Decoding the Proteome in Formalin-Fixed Paraffin-Embedded (FFPE) Tissues. *J Proteomics Bioinform* **07**, (2014).
15. Magdeldin, S. & Yamamoto, T. Toward deciphering proteomes of formalin-fixed paraffin-embedded (FFPE) tissues. *Proteomics* **12**, 1045–1058 (2012).
16. Drummond, E. S., Nayak, S., Ueberheide, B. & Wisniewski, T. Proteomic analysis of neurons microdissected from formalin-fixed, paraffin-embedded Alzheimer’s disease brain tissue. *Sci Rep* **5**, 15456 (2015).
17. Ostrom, Q. T. *et al.* Alex’s Lemonade Stand Foundation Infant and Childhood Primary Brain and Central Nervous System Tumors Diagnosed in the United States in 2007–2011. *Neuro Oncol* **16**, x1–x36 (2015).

18. Jin, B., Li, Y. & Robertson, K. D. DNA Methylation: Superior or Subordinate in the Epigenetic Hierarchy? *Genes Cancer* **2**, 607–617 (2011).
19. Robertson, K. D. DNA methylation and human disease. *Nat Rev Genet* **6**, 597–610 (2005).
20. Palacios, D., Summerbell, D., Rigby, P. W. J. & Boyes, J. Interplay between DNA Methylation and Transcription Factor Availability: Implications for Developmental Activation of the Mouse *Myogenin* Gene. *Mol Cell Biol* **30**, 3805–3815 (2010).
21. Pratt, D., Sahm, F. & Aldape, K. DNA methylation profiling as a model for discovery and precision diagnostics in neuro-oncology. *Neuro Oncol* **23**, S16–S29 (2021).
22. Sturm, D. *et al.* New Brain Tumor Entities Emerge from Molecular Classification of CNS-PNETs. *Cell* **164**, 1060–1072 (2016).
23. El-Ayadi M *et al.* High-grade glioma in very young children: a rare and particular patient population. *Oncotarget* (2017).
24. Maria Yiallourou. Medulloblastoma – Brief Information. *kinderkrebsinfo* (2024).
25. Gajjar, A. *et al.* Outcomes by Clinical and Molecular Features in Children With Medulloblastoma Treated With Risk-Adapted Therapy: Results of an International Phase III Trial (SJMB03). *J Clin Oncol* **39**, 822–835 (2021).
26. Choi, J. Y. Medulloblastoma: Current Perspectives and Recent Advances. *Brain Tumor Res Treat* **11**, 28 (2023).
27. Borgenvik, A., Čančer, M., Hutter, S. & Swartling, F. J. Targeting MYCN in Molecularly Defined Malignant Brain Tumors. *Front Oncol* **10**, (2021).
28. Zoghbi, H. Y. & Beaudet, A. L. Epigenetics and Human Disease. *Cold Spring Harb Perspect Biol* **8**, a019497 (2016).
29. Wang, K. C. & Chang, H. Y. Epigenomics. *Circ Res* **122**, 1191–1199 (2018).
30. Moore, L. D., Le, T. & Fan, G. DNA Methylation and Its Basic Function. *Neuropsychopharmacology* **38**, 23–38 (2013).
31. Kurimoto, T. *et al.* Effect of O6-methylguanine-DNA methyltransferase methylation in medulloblastoma. *Mol Clin Oncol* (2017) doi:10.3892/mco.2017.1431.
32. Capper, D. *et al.* DNA methylation-based classification of central nervous system tumours. *Nature* **555**, 469–474 (2018).
33. Louis, D. N. *et al.* The 2021 WHO Classification of Tumors of the Central Nervous System : a summary. *Neuro Oncol* **23**, 1231–1251 (2021).
34. Feber, A. *et al.* Using high-density DNA methylation arrays to profile copy number alterations. *Genome Biol* **15**, R30 (2014).
35. McCabe, M. G., Bäcklund, L. M., Leong, H. S., Ichimura, K. & Collins, V. P. Chromosome 17 alterations identify good-risk and poor-risk tumors independently of clinical factors in medulloblastoma. *Neuro Oncol* **13**, 376–383 (2011).
36. Clifford, S. C. *et al.* Wnt/Wingless pathway activation and chromosome 6 loss characterize a distinct molecular sub-group of medulloblastomas associated with a favorable prognosis. *Cell Cycle* **5**, 2666–2670 (2006).
37. Zhukova, N. *et al.* Subgroup-Specific Prognostic Implications of TP53 Mutation in Medulloblastoma. *Journal of Clinical Oncology* **31**, 2927–2935 (2013).
38. Goschzik, T. *et al.* Genetic alterations of TP53 and OTX2 indicate increased risk of relapse in WNT medulloblastomas. *Acta Neuropathol* **144**, 1143–1156 (2022).
39. Kawauchi, D. *et al.* Novel MYC-driven medulloblastoma models from multiple embryonic cerebellar cells. *Oncogene* **36**, 5231–5242 (2017).
40. Staal, J. A., Pei, Y. & Rood, B. R. A proteogenomic approach to understanding myc function in metastatic medulloblastoma tumors. *Int J Mol Sci* **17**, (2016).
41. Park, E. G. *et al.* Genomic Analyses of Non-Coding RNAs Overlapping Transposable Elements and Its Implication to Human Diseases. *Int J Mol Sci* **23**, 8950 (2022).

42. Weller, M. *et al.* Molecular classification of diffuse cerebral WHO grade II/III gliomas using genome- and transcriptome-wide profiling improves stratification of prognostically distinct patient groups. *Acta Neuropathol* **129**, 679–693 (2015).
43. Cavalli, F. M. G. *et al.* Intertumoral Heterogeneity within Medulloblastoma Subgroups. *Cancer Cell* **31**, 737–754.e6 (2017).
44. Yoon, J.-H., Abdelmohsen, K. & Gorospe, M. Posttranscriptional Gene Regulation by Long Noncoding RNA. *J Mol Biol* **425**, 3723–3730 (2013).
45. Du, Y. *et al.* Differences in alternative splicing and their potential underlying factors between animals and plants. *J Adv Res* (2023) doi:10.1016/j.jare.2023.11.017.
46. Dubuc, A. M. *et al.* Subgroup-specific alternative splicing in medulloblastoma. *Acta Neuropathol* **123**, 485–499 (2012).
47. Wang, L. *et al.* The evolution of alternative splicing in glioblastoma under therapy. *Genome Biol* **22**, 48 (2021).
48. O’Brien, J., Hayder, H., Zayed, Y. & Peng, C. Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Front Endocrinol (Lausanne)* **9**, (2018).
49. Rivero-Hinojosa, S. *et al.* Proteomic analysis of Medulloblastoma reveals functional biology with translational potential. *Acta Neuropathol Commun* **6**, 48 (2018).
50. Han, X., Aslanian, A. & Yates, J. R. Mass spectrometry for proteomics. *Curr Opin Chem Biol* **12**, 483–490 (2008).
51. Wiśniewski, J. R. Proteomic sample preparation from formalin fixed and paraffin embedded tissue. *J Vis Exp* (2013) doi:10.3791/50589.
52. Petralia, F. *et al.* Integrated Proteogenomic Characterization across Major Histological Types of Pediatric Brain Cancer. *Cell* **183**, 1962–1985.e31 (2020).
53. Hu, M. *et al.* The role of N-glycosylation modification in the pathogenesis of liver cancer. *Cell Death Dis* **14**, 222 (2023).
54. Stanley, P., Moremen, K. W., Lewis, N. E., Taniguchi, N. & Aebi, M. *N-Glycans*. (2022).
55. Lei, X. *et al.* Immune cells within the tumor microenvironment: Biological functions and roles in cancer immunotherapy. *Cancer Lett* **470**, 126–133 (2020).
56. Silsirivanit, A. Glycosylation markers in cancer. in 189–213 (2019). doi:10.1016/bs.acc.2018.12.005.
57. Kim, H. *et al.* Human hepatocellular carcinomas with “Stemness”-related marker expression: keratin 19 expression and a poor prognosis. *Hepatology* **54**, 1707–1717 (2011).
58. Wu, K. *et al.* The role of CD44 in epithelial–mesenchymal transition and cancer development. *Onco Targets Ther* 3783 (2015) doi:10.2147/OTT.S95470.
59. Voß, H. *et al.* HarmonizR enables data harmonization across independent proteomic datasets with appropriate handling of missing values. *Nat Commun* **13**, 3523 (2022).
60. Singh, A. *et al.* DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics* **35**, 3055–3062 (2019).
61. Rohart, F., Gautier, B., Singh, A. & Lê Cao, K.-A. mixOmics: An R package for ‘omics feature selection and multiple data integration. *PLoS Comput Biol* **13**, e1005752 (2017).
62. Spainhour, J. C., Lim, H. S., Yi, S. V & Qiu, P. Correlation Patterns Between DNA Methylation and Gene Expression in The Cancer Genome Atlas. *Cancer Inform* **18**, 117693511982877 (2019).
63. Taylor, M. D. *et al.* Molecular subgroups of medulloblastoma: the current consensus. *Acta Neuropathol* **123**, 465–472 (2012).
64. Eberhart, C. G. *et al.* Histopathologic grading of medulloblastomas. *Cancer* **94**, 552–560 (2002).

65. Korshunov, A. *et al.* Molecular subgroups of medulloblastoma: the current consensus. *Acta Neuropathol* **123**, 465–472 (2012).
66. Smoll, N. R. & Drummond, K. J. The incidence of medulloblastomas and primitive neuroectodermal tumours in adults and children. *Journal of Clinical Neuroscience* **19**, 1541–1544 (2012).
67. Ellison, D. W. *et al.* Medulloblastoma: clinicopathological correlates of SHH, WNT, and non-SHH/WNT molecular subgroups. *Acta Neuropathol* **121**, 381–396 (2011).
68. Taylor, M. D. *et al.* Molecular subgroups of medulloblastoma: the current consensus. *Acta Neuropathol* **123**, 465–472 (2012).
69. Nobre, L. *et al.* Pattern of Relapse and Treatment Response in WNT-Activated Medulloblastoma. *Cell Rep Med* **1**, 100038 (2020).
70. Kennedy, C. *et al.* Quality of Survival and Growth in Children and Young Adults in the PNET4 European Controlled Trial of Hyperfractionated Versus Conventional Radiation Therapy for Standard-Risk Medulloblastoma. *International Journal of Radiation Oncology*Biophysics*Physics* **88**, 292–300 (2014).
71. Ottensmeier, H. *et al.* Neuropsychological short assessment of disease- and treatment-related intelligence deficits in children with brain tumours. *Eur J Paediatr Neurol* **19**, 298–307 (2015).
72. Sabel, M. *et al.* Relapse patterns and outcome after relapse in standard risk medulloblastoma: a report from the HIT-SIOP-PNET4 study. *J Neurooncol* **129**, 515–524 (2016).
73. Yoon, J. W., Gilbertson, R., Iannaccone, S., Iannaccone, P. & Walterhouse, D. Defining a role for Sonic hedgehog pathway activation in desmoplastic medulloblastoma by identifying GLI1 target genes. *Int J Cancer* **124**, 109–119 (2009).
74. Zurawel, R. H. *et al.* Analysis of PTCH/SMO/SHH pathway genes in medulloblastoma. *Genes Chromosomes Cancer* **27**, 44–51 (2000).
75. Kool, M. *et al.* Molecular subgroups of medulloblastoma: an international meta-analysis of transcriptome, genetic aberrations, and clinical data of WNT, SHH, Group 3, and Group 4 medulloblastomas. *Acta Neuropathol* **123**, 473–484 (2012).
76. Zhukova, N. *et al.* Subgroup-Specific Prognostic Implications of TP53 Mutation in Medulloblastoma. *Journal of Clinical Oncology* **31**, 2927–2935 (2013).
77. Archer, T. C. *et al.* Proteomics, Post-translational Modifications, and Integrative Analyses Reveal Molecular Heterogeneity within Medulloblastoma Subgroups. *Cancer Cell* **34**, 396–410.e8 (2018).
78. Forget, A. *et al.* Aberrant ERBB4-SRC Signaling as a Hallmark of Group 4 Medulloblastoma Revealed by Integrative Phosphoproteomic Profiling. *Cancer Cell* **34**, 379–395.e7 (2018).
79. Menyhárt, O., Giangaspero, F. & Győrffy, B. Molecular markers and potential therapeutic targets in non-WNT/non-SHH (group 3 and group 4) medulloblastomas. *J Hematol Oncol* **12**, 29 (2019).
80. Sharma, T. *et al.* Second-generation molecular subgrouping of medulloblastoma: an international meta-analysis of Group 3 and Group 4 subtypes. *Acta Neuropathol* **138**, 309–326 (2019).
81. Northcott, P. A. *et al.* The whole-genome landscape of medulloblastoma subtypes. *Nature* **547**, 311–317 (2017).
82. Schoof, M. *et al.* Malignant gliomas with H3F3A G34R mutation or MYCN amplification in pediatric patients with Li Fraumeni syndrome. *Acta Neuropathol* **142**, 591–593 (2021).
83. Sturm, D. *et al.* Multiomic neuropathology improves diagnostic accuracy in pediatric neuro-oncology. *Nat Med* **29**, 917–926 (2023).

84. Korshunov, A. *et al.* H3-/IDH-wild type pediatric glioblastoma is comprised of molecularly and prognostically distinct subtypes with associated oncogenic drivers. *Acta Neuropathol* **134**, 507–516 (2017).
85. Voß, H. *et al.* Multiomic profiling of medulloblastoma reveals subtype-specific targetable alterations at the proteome and N-glycan level. *bioRxiv* 2023.01.09.523234 (2023) doi:10.1101/2023.01.09.523234.
86. Göbel, C. *et al.* MYC overexpression and SMARCA4 loss cooperate to drive medulloblastoma formation in mice. *Acta Neuropathol Commun* **11**, 174 (2023).
87. Schoof, M. *et al.* Mouse models of pediatric high-grade gliomas with MYCN amplification reveal intratumoral heterogeneity and lineage signatures. *Nat Commun* **14**, 7717 (2023).
88. Sturm, D. *et al.* New Brain Tumor Entities Emerge from Molecular Classification of CNS-PNETs. *Cell* **164**, 1060–1072 (2016).
89. Waszak, S. M. *et al.* Germline Elongator mutations in Sonic Hedgehog medulloblastoma. *Nature* **580**, 396–401 (2020).
90. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* **11**, 2301–2319 (2016).
91. Tyanova, S. & Cox, J. Perseus: A bioinformatics platform for integrative analysis of proteomics data in cancer research. *Methods in Molecular Biology* **1711**, 133–148 (2018).
92. Zhuo, L. *et al.* hGFAP-cre transgenic mice for manipulation of glial and neuronal function in vivo. *genesis* **31**, 85–94 (2001).
93. Hegedus, B. *et al.* Neurofibromatosis-1 Regulates Neuronal and Glial Cell Differentiation from Neuroglial Progenitors In Vivo by Both cAMP- and Ras-Dependent Mechanisms. *Cell Stem Cell* **1**, 443–457 (2007).
94. Hayashi, S., Lewis, P., Pevny, L. & McMahon, A. P. Efficient gene modulation in mouse epiblast using a Sox2Cre transgenic mouse strain. *Mech Dev* **119**, S97–S101 (2002).
95. Jonkers, J. *et al.* Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer. *Nat Genet* **29**, 418–425 (2001).
96. Althoff, K. *et al.* A Cre-conditional MYCN-driven neuroblastoma mouse model as an improved tool for preclinical studies. *Oncogene* **34**, 3357–3368 (2015).
97. Schüller, U. *et al.* Acquisition of Granule Neuron Precursor Identity Is a Critical Determinant of Progenitor Cell Competence to Form Shh-Induced Medulloblastoma. *Cancer Cell* **14**, 123–134 (2008).
98. Fortin, J.-P., Triche Jr, T. J. & Hansen, K. D. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. *Bioinformatics* **33**, 558–560 (2017).
99. Becht, E. *et al.* Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol* **37**, 38 (2019).
100. Afgan, E. *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Res* **50**, W345–W351 (2022).
101. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
102. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
103. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
104. Wilkerson, M. D. & Hayes, D. N. ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* **26**, 1572–1573 (2010).

105. Franch-Expósito, S. *et al.* CNApp, a tool for the quantification of copy number alterations and integrative analysis revealing clinical implications. *Elife* **9**, (2020).
106. Juric, V. & Murphy, B. Cyclin-dependent kinase inhibitors in brain cancer: current state and future directions. *Cancer Drug Resist* **3**, 48–62 (2020).
107. Ding, L. *et al.* The Roles of Cyclin-Dependent Kinases in Cell-Cycle Progression and Therapeutic Strategies in Human Breast Cancer. *Int J Mol Sci* **21**, (2020).
108. Brown, P. D. *et al.* Memantine for the prevention of cognitive dysfunction in patients receiving whole-brain radiotherapy: a randomized, double-blind, placebo-controlled trial. *Neuro Oncol* **15**, 1429–1437 (2013).
109. Phoenix, T. N. *et al.* Medulloblastoma Genotype Dictates Blood Brain Barrier Phenotype. *Cancer Cell* **29**, 508–522 (2016).
110. Fujimoto, M. *et al.* Deficiency of tenascin-C and attenuation of blood-brain barrier disruption following experimental subarachnoid hemorrhage in mice. *J Neurosurg* **124**, 1693–1702 (2016).
111. Bai, R.-Y., Staedtke, V., Rudin, C. M., Bunz, F. & Riggins, G. J. Effective treatment of diverse medulloblastoma models with mebendazole and its impact on tumor angiogenesis. *Neuro Oncol* **17**, 545–554 (2015).
112. Boudiaf-Benmammar, C., Cresteil, T. & Melki, R. The Cytosolic Chaperonin CCT/TRiC and Cancer Cell Proliferation. *PLoS One* **8**, (2013).
113. Zhang, Y. *et al.* Molecular chaperone CCT3 supports proper mitotic progression and cell proliferation in hepatocellular carcinoma cells. *Cancer Lett* **372**, 101–109 (2016).
114. Amon, R., Reuven, E. M., Leviatan Ben-Arye, S. & Padler-Karavani, V. Glycans in immune recognition and response. *Carbohydr Res* **389**, 115–122 (2014).
115. Kailemia, M. J., Park, D. & Lebrilla, C. B. Glycans and glycoproteins as specific biomarkers for cancer. *Anal Bioanal Chem* **409**, 395–410 (2017).
116. Thurin, M. Tumor-Associated Glycans as Targets for Immunotherapy: The Wistar Institute Experience/Legacy. *Monoclon Antib Immunodiagn Immunother* **40**, 89–100 (2021).
117. Smith, K. S. *et al.* Unified rhombic lip origins of group 3 and group 4 medulloblastoma. *Nature* **609**, 1012–1020 (2022).
118. Schüller, U. *et al.* Acquisition of Granule Neuron Precursor Identity Is a Critical Determinant of Progenitor Cell Competence to Form Shh-Induced Medulloblastoma. *Cancer Cell* **14**, 123–134 (2008).
119. Ayrault, O. *et al.* Atoh1 Inhibits Neuronal Differentiation and Collaborates with Gli1 to Generate Medulloblastoma-Initiating Cells. *Cancer Res* **70**, 5618–5627 (2010).
120. Tao, R. *et al.* MYC Drives Group 3 Medulloblastoma through Transformation of Sox2+ Astrocyte Progenitor Cells. *Cancer Res* **79**, 1967–1980 (2019).
121. Mainwaring, O. J. *et al.* ARF suppression by MYC but not MYCN confers increased malignancy of aggressive pediatric brain tumors. *Nat Commun* **14**, 1221 (2023).
122. Pan, Y. *et al.* NF1 mutation drives neuronal activity-dependent initiation of optic glioma. *Nature* **594**, 277–282 (2021).
123. Chen, P. *et al.* Olfactory sensory experience regulates gliomagenesis via neuronal IGF1. *Nature* **606**, 550–556 (2022).
124. Taylor, M. D. *et al.* Molecular subgroups of medulloblastoma: the current consensus. *Acta Neuropathol* **123**, 465–472 (2012).
125. Schwalbe, E. C. *et al.* Novel molecular subgroups for clinical classification and outcome prediction in childhood medulloblastoma: a cohort study. *Lancet Oncol* **18**, 958–971 (2017).
126. Hovestadt, V. *et al.* Medulloblastomics revisited: biological and clinical insights from thousands of patients. *Nat Rev Cancer* **20**, 42–56 (2022).

127. Northcott, P. A. *et al.* The whole-genome landscape of medulloblastoma subtypes. *Nature* **547**, 311–317 (2017).
128. Louis, D. N. *et al.* The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathologica* vol. 131 803–820 Preprint at <https://doi.org/10.1007/s00401-016-1545-1> (2016).
129. Delaidelli, A. *et al.* Clinically Tractable Outcome Prediction of Non-WNT/Non-SHH Medulloblastoma Based on TPD52 IHC in a Multicohort Study. *Clinical Cancer Research* **28**, 116–128 (2022).
130. Ramaswamy, V. *et al.* Risk stratification of childhood medulloblastoma in the molecular era: the current consensus. *Acta Neuropathol* **131**, 821–831 (2016).
131. Goschzik, T. *et al.* Molecular stratification of medulloblastoma: comparison of histological and genetic methods to detect Wnt activated tumours. *Neuropathol Appl Neurobiol* **41**, 135–144 (2015).
132. Ellison, D. W. *et al.* β -Catenin Status Predicts a Favorable Outcome in Childhood Medulloblastoma: The United Kingdom Children’s Cancer Study Group Brain Tumour Committee. *Journal of Clinical Oncology* **23**, 7951–7957 (2005).
133. Chen, Q., Tan, Z., Guan, F. & Ren, Y. The Essential Functions and Detection of Bisecting GlcNAc in Cell Biology. *Front Chem* **8**, 511 (2020).
134. Vreeker, G. C. M. *et al.* Serum N-glycan profiles differ for various breast cancer subtypes. *Glycoconj J* **38**, 387–395 (2021).
135. Rodríguez, E., Schettters, S. T. T. & van Kooyk, Y. The tumour glyco-code as a novel immune checkpoint for immunotherapy. *Nat Rev Immunol* **18**, 204–211 (2018).
136. Dotz, V. & Wuhrer, M. N-glycome signatures in human plasma: associations with physiology and major diseases. *FEBS Lett* **593**, 2966–2976 (2019).
137. Capper, D. *et al.* DNA methylation-based classification of central nervous system tumours. *Nature* **555**, 469–474 (2018).
138. Lee, H. G. *et al.* State-of-the-art housekeeping proteins for quantitative western blotting: Revisiting the first draft of the human proteome. *Proteomics* **16**, 1863–1867 (2016).
139. Voß, H. *et al.* HarmonizR enables data harmonization across independent proteomic datasets with appropriate handling of missing values. *Nat Commun* **13**, 3523 (2022).
140. Schwalbe, E. C. *et al.* Novel molecular subgroups for clinical classification and outcome prediction in childhood medulloblastoma: a cohort study. *Lancet Oncol* **18**, 958–971 (2017).
141. Menyhárt, O. & Györffy, B. Principles of tumorigenesis and emerging molecular drivers of SHH-activated medulloblastomas. *Ann Clin Transl Neurol* **6**, 990–1005 (2019).
142. Krämer, A., Green, J., Pollard, J. J. & Tugendreich, S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* **30**, 523–530 (2014).
143. Kool, M. *et al.* Genome sequencing of SHH medulloblastoma predicts genotype-related response to smoothed inhibition. *Cancer Cell* **25**, 393–405 (2014).
144. Ramaswamy, V., Nör, C. & Taylor, M. D. P53 and Medulloblastoma. *Cold Spring Harb Perspect Med* **6**, 1–9 (2016).
145. Bailey, S. *et al.* Clinical Trials in High-Risk Medulloblastoma: Evolution of the SIOP-Europe HR-MB Trial. *Cancers (Basel)* **14**, (2022).
146. Northcott, P. A. *et al.* Medulloblastoma comprises four distinct molecular variants. *Journal of Clinical Oncology* **29**, 1408–1414 (2011).
147. Menyhárt, O., Giangaspero, F. & Györffy, B. Molecular markers and potential therapeutic targets in non-WNT/non-SHH (group 3 and group 4) medulloblastomas. *J Hematol Oncol* **12**, (2019).

148. Sun, R., Kim, A. M. J. & Lim, S.-O. Glycosylation of Immune Receptors in Cancer. *Cells* **10**, (2021).
149. Klein, J. A., Meng, L. & Zaia, J. Deep Sequencing of Complex Proteoglycans: A Novel Strategy for High Coverage and Site-specific Identification of Glycosaminoglycan-linked Peptides. *Mol Cell Proteomics* **17**, 1578–1590 (2018).
150. Munkley, J. & Scott, E. Targeting Aberrant Sialylation to Treat Cancer. *Medicines (Basel)* **6**, (2019).
151. Sprung, R. W. *et al.* Equivalence of Protein Inventories Obtained from Formalin-fixed Paraffin-embedded and Frozen Tissue in Multidimensional Liquid Chromatography-Tandem Mass Spectrometry Shotgun Proteomic Analysis. *Molecular & Cellular Proteomics* **8**, 1988–1998 (2009).
152. Cook Sangar, M. L. *et al.* Inhibition of CDK4/6 by Palbociclib Significantly Extends Survival in Medulloblastoma Patient-Derived Xenograft Mouse Models. *Clinical Cancer Research* **23**, 5802–5813 (2017).
153. Vasaikar, S. *et al.* Proteogenomic Analysis of Human Colon Cancer Reveals New Therapeutic Opportunities. *Cell* **177**, 1035–1049.e19 (2019).
154. Mertins, P. *et al.* Reproducible workflow for multiplexed deep-scale proteome and phosphoproteome analysis of tumor tissues by liquid chromatography-mass spectrometry. *Nat Protoc* **13**, 1632–1661 (2018).
155. Hovestadt, V. *et al.* Decoding the regulatory landscape of medulloblastoma using DNA methylation sequencing. *Nature* **510**, 537–541 (2014).
156. Richard W Olsen and Timothy M DeLorey. No Title. in *Basic Neurochemistry: Molecular, Cellular and Medical Aspects. 6th edition.* (1999).
157. Venkatesh, H. S. *et al.* Electrical and synaptic integration of glioma into neural circuits. *Nature* **573**, 539–545 (2019).
158. Venkataramani, V. *et al.* Glioblastoma hijacks neuronal mechanisms for brain invasion. *Cell* **185**, 2899–2917.e31 (2022).
159. Akar, S., Harmankaya, İ., Uğraş, S. & Çelik, Ç. Nicotinamide N-methyltransferase expression and its association with phospho-Akt, p53 expression, and survival in high-grade endometrial cancer. *Turk J Med Sci* **49**, 1547–1554 (2019).
160. Moore, S. *et al.* The CHD6 chromatin remodeler is an oxidative DNA damage response factor. *Nat Commun* **10**, 241 (2019).
161. Alexiou, G. A. *et al.* Expression of heat shock proteins in medulloblastoma: Laboratory investigation. *Journal of Neurosurgery: Pediatrics PED* **12**, 452–457 (2013).
162. Nobre, L. *et al.* Pattern of Relapse and Treatment Response in WNT-Activated Medulloblastoma. *Cell Rep Med* **1**, (2020).
163. Meredith, D. M. & Alexandrescu, S. Embryonal and non-meningothelial mesenchymal tumors of the central nervous system - Advances in diagnosis and prognostication. *Brain Pathol* **32**, e13059 (2022).
164. D’Arcy, C. E. *et al.* Immunohistochemical and nanoString-Based Subgrouping of Clinical Medulloblastoma Samples. *J Neuropathol Exp Neurol* **79**, 437–447 (2020).
165. Northcott, P. A. *et al.* Rapid, reliable, and reproducible molecular sub-grouping of clinical medulloblastoma samples. *Acta Neuropathol* **123**, 615–626 (2012).
166. Yoshida, T., Akatsuka, T. & Imanaka-Yoshida, K. Tenascin-C and integrins in cancer. *Cell Adh Migr* **9**, 96–104 (2015).
167. Linke, F. *et al.* 3D hydrogels reveal medulloblastoma subgroup differences and identify extracellular matrix subtypes that predict patient outcome. *J Pathol* **253**, 326–338 (2021).
168. Phoenix, T. N. *et al.* Medulloblastoma Genotype Dictates Blood Brain Barrier Phenotype. *Cancer Cell* **29**, 508–522 (2016).

169. Shalabi, H., Nellan, A., Shah, N. N. & Gust, J. Immunotherapy Associated Neurotoxicity in Pediatric Oncology. *Front Oncol* **12**, (2022).
170. Schwalbe, E. C. *et al.* Rapid diagnosis of medulloblastoma molecular subgroups. *Clinical Cancer Research* **17**, 1883–1894 (2011).
171. Wang, J. *et al.* Effective inhibition of MYC-amplified group 3 medulloblastoma by FACT-targeted curaxin drug CBL0137. *Cell Death Dis* **11**, 1029 (2020).
172. Patil, S. *et al.* Combination of clotam and vincristine enhances anti-proliferative effect in medulloblastoma cells. *Gene* **705**, 67–76 (2019).
173. Sengupta, S., Pomeranz Krummel, D. & Pomeroy, S. The evolution of medulloblastoma therapy to personalized medicine. *F1000Res* **6**, 490 (2017).
174. Bassiouni, R. *et al.* Chaperonin Containing TCP-1 Protein Level in Breast Cancer Cells Predicts Therapeutic Application of a Cytotoxic Peptide. *Clin Cancer Res* **22**, 4366–4379 (2016).
175. Carr, A. C. *et al.* Targeting chaperonin containing TCP1 (CCT) as a molecular therapeutic for small cell lung cancer. *Oncotarget* **8**, 110273–110288 (2017).
176. Ohtsubo, K. & Marth, J. D. Glycosylation in Cellular Mechanisms of Health and Disease. *Cell* **126**, 855–867 (2006).
177. Greco, B. *et al.* Disrupting N-glycan expression on tumor cells boosts chimeric antigen receptor T cell efficacy against solid malignancies. *Sci Transl Med* **14**, eabg3072 (2022).
178. Marada, S. *et al.* Functional Divergence in the Role of N-Linked Glycosylation in Smoothed Signaling. *PLoS Genet* **11**, e1005473 (2015).
179. Williams, S. E. *et al.* Mammalian brain glycoproteins exhibit diminished glycan complexity compared to other tissues. *Nat Commun* **13**, 275 (2022).
180. Rodrigues, E. & Macauley, M. S. Hypersialylation in Cancer: Modulation of Inflammation and Therapeutic Opportunities. *Cancers (Basel)* **10**, (2018).
181. Northcott, P. A. *et al.* Medulloblastoma comprises four distinct molecular variants. *Journal of Clinical Oncology* **29**, 1408–1414 (2011).
182. Mcalister, G. C. *et al.* MultiNotch MS3 Enables Accurate , Sensitive , and Multiplexed Detection of Differential Expression across Cancer Cell Line Proteomes. *Anal Chem* **86**, (2014).
183. Hughes, C. S. *et al.* Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc* **14**, 68–85 (2019).
184. van Pijkeren, A. *et al.* Proteome Coverage after Simultaneous Proteo-Metabolome Liquid-Liquid Extraction. *J Proteome Res* **22**, 951–966 (2023).
185. Yurekten, O. *et al.* MetaboLights: open data repository for metabolomics. *Nucleic Acids Res* **52**, D640–D646 (2024).
186. Guan, Y., Zhang, M., Wang, J. & Schlüter, H. Comparative Analysis of Different N-glycan Preparation Approaches and Development of Optimized Solid-Phase Permethylation Using Mass Spectrometry. *J Proteome Res* **20**, 2914–2922 (2021).
187. Watanabe, Y., Aoki-Kinoshita, K. F., Ishihama, Y. & Okuda, S. GlycoPOST realizes FAIR principles for glycomics mass spectrometry data. *Nucleic Acids Res* **49**, D1523–D1528 (2021).
188. Aryee, M. J. *et al.* Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363–1369 (2014).
189. Ritchie, M. E. *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).
190. Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S. & Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods* **17**, 41–44 (2020).

191. Guan, Y. *et al.* An Integrated Strategy Reveals Complex Glycosylation of Erythropoietin Using Mass Spectrometry. *J Proteome Res* **20**, 3654–3663 (2021).
192. Community, T. G. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Res* **50**, W345–W351 (2022).
193. Triche, T. J. J., Weisenberger, D. J., Van Den Berg, D., Laird, P. W. & Siegmund, K. D. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. *Nucleic Acids Res* **41**, e90 (2013).
194. Lagnoux, A., Mercier, S. & Vallois, P. Statistical significance based on length and position of the local score in a model of i.i.d. sequences. *Bioinformatics* **33**, 654–660 (2017).
195. Schumann, Y., Neumann, J. E. & Neumann, P. Robust classification using average correlations as features (ACF). *BMC Bioinformatics* **24**, 101 (2023).
196. Tyanova, S. & Cox, J. Perseus: A bioinformatics platform for integrative analysis of proteomics data in cancer research. *Methods in Molecular Biology* **1711**, 133–148 (2018).
197. Jassal, B. *et al.* The reactome pathway knowledgebase. *Nucleic Acids Res* **48**, D498–D503 (2020).
198. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545–15550 (2005).
199. Reimand, J. *et al.* Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat Protoc* **14**, 482–517 (2019).
200. Paul Shannon, 1 *et al.* Cytoscape: A Software Environment for Integrated Models. *Genome Res* **13**, 426 (1971).
201. Reimand, J. *et al.* Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat Protoc* **14**, 482–517 (2019).
202. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676–682 (2012).

4. Publication List

4.1 MB project 1 in original format

Accepted for publication in Nature Communications

Multiomic profiling of medulloblastoma reveals subtype-specific targetable alterations at the proteome and N-glycan level

Shweta Godbole^{1, +}, Hannah Voß^{2, +}, Antonia Gocke^{1,2}, Simon Schlumbohm³, Yannis Schumann³, Bojia Peng², Martin Mynarek^{4,5}, Stefan Rutkowski⁴, Matthias Dottermusch^{1,6}, Mario M. Dorostkar^{7,8}, Andrey Korshunov^{9,10}, Thomas Mair², Stefan M. Pfister^{11,12,13}, Marcel Kwiatkowski¹⁴, Madlen Hotze¹⁴, Philipp Neumann³, Christian Hartmann¹⁵, Joachim Weis¹⁶, Friederike Liesche-Starnecker¹⁷, Yudong Guan², Manuela Moritz², Bente Siebels², Nina Struve^{5,18}, Hartmut Schlüter², Ulrich Schüller^{4,6,19}, Christoph Krisp², Julia E. Neumann^{1, 6 #}

¹ Center for Molecular Neurobiology (ZMNH), University Medical Center Hamburg-Eppendorf, Hamburg, Germany

² Section of Mass Spectrometry and Proteomics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

³ Chair for High Performance Computing, Helmut Schmidt University, Hamburg, Germany

⁴ Department of Pediatric Hematology and Oncology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

⁵ Mildred Scheel Cancer Career Center HaTriCS4, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

⁶ Institute of Neuropathology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

⁷ Center for Neuropathology, Ludwig-Maximilians-University, Munich, Germany

⁸ German Center for Neurodegenerative Diseases, Munich, Germany

⁹ Department of Neuropathology, University Hospital Heidelberg, Heidelberg, Germany

¹⁰ Clinical Cooperation Unit Neuropathology, German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany

¹¹ Hopp Children's Cancer Center at the NCT Heidelberg (KITZ), Heidelberg, Germany

¹² Division of Pediatric Neurooncology, German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany

¹³ Department of Pediatric Hematology and Oncology, Heidelberg University Hospital, Heidelberg, Germany

¹⁴ Institute of Biochemistry, University of Innsbruck, Austria

¹⁵ Department of Neuropathology, Hannover Medical School (MHH), Hannover, Germany

¹⁶ Institute of Neuropathology, RWTH Aachen University Hospital, Aachen, Germany

¹⁷ Pathology, Medical Faculty, University of Augsburg, Augsburg, Germany

¹⁸ Department of Radiotherapy & Radiation Oncology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

¹⁹ Research Institute Children's Cancer Center Hamburg, Hamburg, Germany

+ these authors contributed equally

corresponding author

Corresponding author:

Prof. Dr. med. Julia E. Neumann

Email: ju.neumann@uke.de

Postal address: Research Group Molecular Pathology in Neurooncology, Center for Molecular Neurobiology (ZMNH), University Medical Center Hamburg-Eppendorf, 20251 Hamburg, Germany

Telephone: +49 (0) 40 7410 57419

Abstract

Medulloblastomas (MBs) are malignant pediatric brain tumors that are molecularly and clinically heterogeneous. The application of omics technologies – mainly studying nucleic acids – has significantly improved MB classification and stratification, but treatment options are still unsatisfactory. The proteome and their N-glycans hold the potential to discover clinically relevant phenotypes and targetable pathways. We compile a harmonized proteome dataset of 167 MBs and integrate findings with DNA methylome, transcriptome and N-glycome data. We show six proteome MB subtypes, that can be assigned to two main molecular programs: transcription/translation (pSHHt, pWNT and pGroup3myc), and synapses/immunological processes (pSHHs, pGroup3 and pGroup4). Multiomic analysis reveals different conservation levels of proteome features across MB subtypes at the DNA methylome level. Aggressive pGroup3myc MBs and favourable pWNT MBs are most similar in cluster hierarchies concerning overall proteome patterns but show different protein abundances of the vincristine resistance-associated multiprotein complex TriC/CCT and of N-glycan turnover-associated factors. The N-glycome reflects proteome subtypes and complex-bisecting N-glycans characterize pGroup3myc tumors. Our results shed light on targetable

alterations in MB and set a foundation for potential immunotherapies targeting glycan structures.

Keywords:

Medulloblastoma subtypes, multiomics, proteome, mass spectrometry, DNA methylation, 850K, N-glycan, N-glycome, TriC/CCT, WNT, SHH, MYC, PALMD, TNC, HarmonizR

Introduction:

Medulloblastomas (MBs) are aggressive pediatric brain tumors that are histomorphologically, molecularly and clinically heterogeneous³³. Four main consensus subgroups have been described: WNT pathway activated MB (WNT MB), Sonic hedgehog pathway activated MB (SHH MB), Group 3 (G3) and Group 4 (G4) MB¹²⁴. Molecular analyses, mainly using gene expression profiling, next generation sequencing and DNA methylation analysis predict further subdivisions with distinct clinical features^{43,125–127}. Exemplary markers for poor survival comprise anaplastic histology, *MYC* amplification status, methylation subtype II/III, or *TP53* mutations in WNT and SHH MB^{25,35,38,128–130}. Conversely, methylation subtype VII, extensive nodularity (MBEN histology), a distinct whole chromosomal alteration signature in non-WNT/ non-SHH MB and WNT activation (e.g. nuclear accumulation of β -CATENIN or CTNNB1 mutations) were associated with a favourable prognosis in MB patients^{25,131,132}. The clinical association to certain methylation subtypes and chromosomal aberrations has been clearly described, however, the underlying molecular mechanisms remain to be resolved and targeted treatment options are lacking. In contrast to nucleic acids, the proteome reflects a tumor's phenotype in a more direct way and holds the potential to precisely dissect clinically relevant phenotypes and targetable alterations. Studies on small MB cohorts, using fresh-frozen (FF) tumor material, have shown that MBs display heterogeneity at the proteome level^{52,77,78}. Formalin-fixed-paraffine-embedded (FFPE) material, enables the generation of larger datasets which is essential to deal with the heterogeneity but provides challenges to proteome analysis¹⁵. In addition to protein abundance, post-translational modifications (PTM) of proteins are important to understand cell physiology and disease-related signalling^{52,77,78}. The most complex and common PTM, N-glycosylation, has not been targeted in MB yet. Changes in the N-glycome are considered potential hallmarks of cancer and N-glycan structures hold strong potential as biomarkers and immunotherapy targets^{115,133–136}.

In this work, we integrate MB proteome datasets^{52,77,78} with data of 62 FFPE MB cases and establish a joint MB proteome dataset (n = 176) that is comprehensively compared to DNA methylome data – a current gold standard for molecular brain tumor classification¹³⁷. Further, global N-glycosylation patterns of MB are assessed and correlated with identified proteome subtypes. Taken together, we present a large integrated study of the MB proteome, DNA-methylome and N-glycome, revealing new insights into MB phenotypes, potential biomarkers and therapeutic targets.

Results:

Integration of in-house proteome data and publicly available datasets enables large scale proteome analysis of MB

Proteome analysis was performed for 62 FFPE MB tumors (53 primaries, 9 recurrent cases). Additionally, 53 cases were analysed using DNA methylation profiling. Principal component analysis (PCA) and hierarchical clustering (HCL) distinguished the four main molecular subgroups of MB (SHH, WNT, G3, G4)² similarly to published FF based MB proteome datasets (Figure 1A, Supplementary Figure 1A, Supplementary Figure 3, Supplementary data file 1c, Source data are provided as a Source Data file)^{52,77,78}. Proteome data of FF and FFPE tissue from matched MB cases further showed a high correlation (Supplementary Figure 2A). The age of used paraffine material did not impact sample clustering, detected protein numbers or abundance levels of housekeeping proteins¹³⁸ (Figure 1B, Supplementary Figure 4, Supplementary Figure 17D). Proteins detected in WNT and SHH MB, showed similar tendencies in FFPE- and FF-MB datasets^{77,78} (Supplementary Figure 1B). We concluded that FFPE tissue is suitable to study proteome patterns in MB. To increase cohort size, we next integrated and harmonized FF-MB proteome datasets from public repositories^{77;78;52} (Figure 1D). Technical biases were reduced with HarmonizR¹³⁹, and harmonized samples of the joint cohort clustered according to the main MB subgroups (Figure 1 E-G, Supplementary Figure 5, Supplementary data file 1a). Established protein biomarkers for molecular MB subtypes⁶⁷, showed expected subgroup-specific abundance patterns in individual studies and in the combined and harmonized data (Figure 1H). 16,279 proteins were quantified across 167 samples (19xWNT; 57xSHH; 53xG4; 36xG3; 2xno initial main subgroup stated), including 156 primary tumors and 11 recurrences.

Six proteomic MB subtypes can be assigned to two main, potentially druggable molecular profiles

To define proteome subtypes of MB, consensus clustering was applied (Supplementary data file 1b). 6 stable clusters were identified (Figure 2 A-D). Clusters were also reflected in RNA data of matched cases (n=60, Supplementary Figure 3D-F). The assignment reliability of a sample to a respective proteome subtype was indicated as cluster certainty (Figure 2D, Supplementary data file 1c). At the proteome level, non-WNT/ non-SHH MBs divided into three groups (pG4, pG3myc and pG3, p=proteome group), while SHH MBs separated into two groups (pSHHs, pSHHt, s=synaptic profile, t=transcriptional profile). WNT MB formed a homogenous cluster (pWNT, Figure 2D). In general, a high cluster stability was given for all proteome subtypes (median 6/6), except for pG3 samples, that showed high similarity to pG4 and pG3myc respectively (median pG3 5/6, Figure 2D). Except for one case, corresponding recurrent and primary tumors were assigned to the same proteome subtype (Figure 2D). The case that switched subtype in the recurrence situation (from pSHHs to pSHHt) had a low cluster certainty in the primary sample (3/6, Figure 2D).

Proteome MB subtypes were associated with previously described DNA methylation subtypes^{43,127,140} (<https://www.moleculareuropathology.org/mnp/>)¹³⁷, Supplementary data file 1c, Supplementary Figure 6B, Figure 2D). pG3myc patients showed reduced overall survival (Figure 2E). pWNT patients showed the best overall survival rate (Figure 2E). Out of 3,996 proteins found in at least 30% of samples for each proteome subtype, 529 showed a characteristic abundance in at least one subtype. The top 5 proteins with the lowest p-value and highest mean difference were selected as

biomarker candidates (Figure 2F, Supplementary data file 2a). For high-risk non-WNT/non-SHH MBs (pG3myc) PALMD, DIEXF, MCN1, TPD52 and PYCR1 were identified. Of note, hedgehog-signalling induced proteins (MICAL1, GAB1, PDLIM3)¹⁴¹ showed a higher abundance in both, pSHHt and pSHHs. Protein biomarkers were confirmed in case-matched MB cases (FF versus FFPE tissue, n=10, Supplementary Figure 2B) and on the RNA level (Supplementary Figure 3H). Subtype assignments were confirmed in an additional published MB dataset⁸⁹ and a technical validation dataset (Supplementary data file 5j,k, Supplementary Figure 17).

The six proteome subtypes could be assigned to two superordinate clusters at the first hierarchy level in the joint (as well as all individual) datasets (Figure 3, Supplementary Figure 3). Comparing these two clusters revealed two main molecular profiles: profile 1, comprising of pG3, pG4 and pSHHs and profile 2, comprising of pWNT, pG3myc and pSHHt MBs (Figure 3A). The two clusters were confirmed in a technical validation dataset (Supplementary Figure 17). Matched RNA expression profiles also confirmed a clustering of cases according to these defined profiles (n=60, Supplementary Figure 3F). We next used gene set enrichment analysis (GSEA) to reveal potential underlying mechanisms and signalling pathways. Synaptic/immunological processes and phospholipid signalling were observed for profile 1 and a replicative/transcriptional signature was observed for profile 2 (Figure 3A,B, q-value <0.05, Supplementary data file 3a, 3b,3h, Supplementary data file 10e,f). In order to find drug targets and predict downstream effects we used the Ingenuity Pathway Analyses (IPA) tool and focused on the top two upregulated genesets based on differentially abundant proteins in profile 1 (opioid signalling and SNARE complex) and profile 2 (EIF2 signalling and cell cycle control of chromosomal replication, Figure 3 B,C, Supplementary data file 3c-g)¹⁴². Tumors of profile 1 could potentially be targeted by several drugs, including the NMDA receptor antagonist memantine. Profile 2 tumors (replicative/transcriptional signature) could - besides others - be targeted by CDK4 or DNA polymerase inhibitors (Figure 3B-E, Supplementary Figure 7, Supplementary data file 3c-g).

Group-specific correlation of the DNA methylome and the proteome reveals different conservation levels of molecular characteristics across proteomic MB subtypes.

Since DNA methylome data is routinely used in brain tumor diagnostics³³, we decided to integrate our proteome data with DNA methylome data to investigate 1) the general correlation between the two data types and 2) if protein biomarkers are reflected at DNA methylome level. To integrate the data modalities, multiblock data integration using sparse partial least squares discriminant analysis (sPLS-DA) was performed between DNA methylation data (115 samples, 10,000 differentially methylated CpG sites between the MNP v12.5 defined subtypes) and proteome data (115 samples, 3,990 quantified proteins present in 30% of samples, Supplementary Figure 8 A-C, Supplementary data file 1b,d)⁶¹. Only a fraction of features out of the 381717 probes and 3990 proteins showed correlation upon data integration using DIABLO from mixOmics, discriminating mainly the WNT subtype (Figure 4A, arrows, correlation cut-off > 0.7, Supplementary data file 4h, Supplementary Figure 9A-E). To refrain from any data bias, we next performed a MB subtype specific correlation between complete DNA methylome data (115 samples and 381,717 CpG sites) and proteome data (115 samples, 3,990 proteins, Figure 4B, C). A significantly higher number of proteins of the pWNT (38.14%, 1,552 proteins) and pG3 subtype (45.41%, 1,812 proteins) correlated with at least one CpG site of their own gene, when compared to the other subtypes

(range 1.52 – 6.49 %, Figure 4B, Supplementary data file 4b-g). Only 12.2 % - 18% of protein correlating CpG sites were located at the transcriptional start site (TSS200, TSS1500, Exon1, Figure 4B). Integrating the proteome data with DNA methylome data based on differentially methylated regions (DMR) confirmed a high correlation of features in pWNT MB (Supplementary Figure 10 A-B). Focusing on the 31 previously selected biomarker candidates (Figure 2F), we found 10 proteins correlating with CpG sites of their own gene across subtypes (Figure 4C, D, Supplementary data file 4a). In summary, DNA-methylation changes were only partly reflected at the protein level, with different feature conservation levels for different proteome subtypes.

SHH MB comprise two proteome subtypes showing a synaptic or DNA transcription/translation signature

SHH MB split into two proteome subtypes (pSHHt and pSHHs, Figure 5A). All pSHHs cases with high cluster certainty (6/6) occurred in patients below 3 years of age. The DNA methylation subtypes SHH3 (8/29) and SHH4 (9/29) were exclusively found in pSHHt MBs (Figure 5A). Methylation subtypes SHH1 and SHH2 were seen in both pSHHs and pSHHt (SHH1: $p=0.43$, SHH2: $p=0.10$, X^2 - test). We then analyzed the distribution of SHH pathway alterations, which are driver events in SHH MBs¹⁴³. *PTCH1* mutations were found exclusively but not mandatory in pSHHt tumors. *SUFU*, *SMO*, *MYCN* or *GLI2* alterations did not distribute differentially (Figure 5A). Proteome subtypes of SHH MB were not clearly separated at the transcriptome level, which is in line with previous results⁷⁷ (matched samples $n=21$, Supplementary Figure 11 A-B).

In order to analyse how copy number alterations might be reflected at the proteome level, the proteome abundance for each gene was mapped to chromosomal arms, which will be referred to as “proteome copy number variation (CNV)” henceforth. Both pSHHt and pSHHs groups showed a low overall correlation between calculated CNVs using matched DNA methylation data and proteome data ($r_{pSHHs} = 0.01$, $r_{pSHHt} = 0.20$, Figure 5D, G, Supplementary data file 5g-h).

To get insights into changed pathways in pSHH subtypes, a network clustering based on gene set enrichment using pSHHs or pSHHt specific proteins was performed (Figure 5B,C,E,F,H, Supplementary data file 5a-f, q -value < 0.05). Differential proteins in pSHHs revealed differences in synaptic, mitochondrial, and immunological processes, whereas proteins in pSHHt MB were involved in post-translational protein modification, transcription/translation, DNA repair and cell cycle. In accordance with the latter profile, pSHHt showed a significantly enhanced proliferation (assessed via ki67 staining, Supplementary Figure 11 E-F, Supplementary data file 5n). *ALDH1A31* was highly abundant in both pSHH groups (Figure 2F, Figure 5E), which could be confirmed via immunohistochemistry (Supplementary Figure 11 C,D).

Analyses of hallmark gene sets additionally revealed a distinct upregulation of proteins involved in the TCA cycle in pSHHs, indicating metabolic differences between the subtypes (Figure 5H). Subsequent analyses of metabolites and aminoacids confirmed distinct metabolic patterns in pSHHt and pSHHs (Supplementary Figure 12). Of note pSHHs showed a lower abundance of Isocitrate dehydrogenases, together with a decrease of Isocitrate, alpha-Ketoglutarate and Glutamine, indicating a higher consumption of the latter three (Supplementary Figure 12C, Supplementary data file 5l-m). Alpha-Ketoglutarate and Glutamine can be further processed to Glutamate and then GABA, which are both involved in synaptic signalling. In line with these findings,

we detected a significant increase of GABA target proteins in pSHHs (Supplementary Figure 12C).

We did not detect a significant difference in survival between pSHHs and pSHHt (Figure 5I). However, *TP53* mutations, used for stratification of high-risk SHH MB¹⁴⁴, mainly occurred in the pSHHt subtype (9/10, but differential distribution was not significant ($p=0.43$, X^2 - test)). As expected, *TP53* mutations within the pSHHt group significantly correlated with bad prognosis (Figure 5I). *TP53* mutated MBs did not form a distinct proteome cluster. However, 134 differentially abundant proteins were detected between pSHHt-*TP53* wildtype and pSHHs-*TP53* mutated MBs (Figure 5J, Supplementary data file 5i).

High-risk pG3myc MBs are characterized by a MYC profile and high abundance of Palmdelphin

We found three different non-WNT/ non-SHH MB proteome subtypes: pG3, pG3myc and pG4 (Figure 6A). pG4 exclusively included the main molecular subgroup G4, whereas pG3myc was dominated by G3 patients. pG3 included both molecular subgroups (Figure 2D). pG3myc were dominated by large cell anaplastic histology (LCA). LCA histology and *MYC* amplification are used for high-risk tumor stratification in non-WNT/ non-SHH MBs¹⁴⁵. Accordingly, *MYC* amplifications were predominantly detected in pG3myc tumors. However, not all pG3myc classified cases were *MYC* amplified. In concordance with these high-risk characteristics, a broad fraction of pG3myc cases were assigned to the methylation subtype II (16/20 cases, 80%)^{137,146} or group G3 δ^{43} (13/20 cases, 65 %, Figure 6A). Clinically, most pG3myc tumors were classified as M3 and tumors showed the worst overall survival compared to all other MB subtypes (Figure 6A, Figure 2E). Distinct protein abundance patterns and pathway enrichments were seen for pG3, pG4 and pG3myc each and all showed a low overall correlation between calculated proteome and DNA methylation CNV data (Figure 6 B-J, Supplementary data file 6a-l). Specifically, pG3myc MB showed a significant enrichment of *MYC* target proteins (FDR < 0.25; p -value < 0.0001, Figure 6K). In line with this, pG3myc MB showed a high fraction of tumor cell nuclei with accumulation of *MYC* (Supplementary Figure 13). Moreover, pG3myc MB differed from pG3 and pG4 showing enhanced signalling by ROBO receptors and an underrepresentation of proteins involved in MHCII class antigen presentation (Figure 6P, Supplementary data file 6m). To establish a diagnostic useful biomarker for histological identification of high-risk pG3myc tumors, we focused on the high differentially abundant protein Palmdelphin (PALMD, Figure 6 H). Digitally supported quantification of PALMD immunostainings confirmed a specific increase of the candidate in pG3myc tumors (Figure 6L, M, Source data are provided as a Source Data file). We additionally analysed how this biomarker is reflected at other omic levels. Indeed, a significantly higher *PALMD* mRNA expression and lower CpG site methylation was detected in pG3myc MBs compared to all other MB subtypes (Figure 6N)⁷⁷. High *PALMD* mRNA expression was also associated with poor survival in MB (Figure 6O, Supplementary Figure 14A-D). Finally, all groups displayed a low overall correlation between calculated proteome CNV and DNA methylation CNV data (Figure 6D, G, J, Supplementary data file 6j-l).

pWNT MB show low abundance of the multiprotein complex TriC/CCT

WNT MB did not divide into further subtypes based on proteome profiles (Figure 7A). Among differentially abundant proteins in comparison to other MB subtypes TNC showed the highest abundance (14.7 foldchange, Figure 7B, Supplementary data file 7a). A significantly high intensity of TNC in pWNT MB was confirmed using digitally supported immunostaining quantification (Figure 7C, D, E, Source data are provided as a Source Data file). Using a publicly available dataset⁴³, a higher mRNA expression of *TNC* in WNT MB was confirmed (Figure 7F). In contrast, CpG sites of the *TNC* gene, showed no significant difference of methylation (pWNT versus other subtypes (Figure 7G, Supplementary Figure 14A-C). GSEA revealed an enrichment of extracellular matrix proteins and N-glycan biogenesis and transport (FDR< 0.25; p-value <0.0001, Figure 7H, I, Supplementary data file 7b, c). A high overall correlation between copy number plots extracted from proteome and DNA methylome data was observed for pWNT compared to all other subtypes (Figure 7J, Supplementary data file 7d), being in line with a general increased overall correlation of proteome and DNA methylome data (Figure 4A-B).

The highest similarity of proteome profiles was observed for the pG3myc subtype, associated with high-risk features and the pWNT subtype-associated with relatively good overall survival (Figure 3A). Both subtypes showed a “transcriptional/translational” profile (Figure 3A-B) and a high abundance of MYC target proteins along with a high fraction of MYC-positive tumor cell nuclei (Figure 6K, Supplementary Figure 13). We therefore asked, what molecular changes could impact such diverse clinical behavior. Differentially abundant proteins between pG3myc and pWNT MBs included TriC/CCT proteins and the established WNT MB marker β -CATENIN¹⁴⁷ (Figure 8A, Supplementary data file 8a). Among the top discriminating gene sets was the association with TriC/CCT target proteins and asparagine linked N-glycosylation (FDR< 0.25; p-value <0.0001, Figure 8B, Supplementary data file 8b-c).

As the TriC/CCT complex has previously been reported to be associated with vincristine resistance and typical chemotherapy regimens for MB contain vincristine in the treatment combination¹⁴⁷, we further focused on this chaperonin containing multiprotein complex (Figure 8C-E, Supplementary Figure 15A). Among MB subtypes, pWNT MBs showed the lowest abundance of TriC/CCT proteins, whereas pG3myc MBs displayed the highest amount. High abundance of TriC/CCT proteins in pG3myc was confirmed at mRNA level. Matched cases, as well as publicly available transcriptome data⁴³ did not show a statistically significant downregulation of all component mRNAs in pWNT MB when compared to other subtypes (Figure 8C, Supplementary Figure 15). Further, no difference of TriC/CCT gene methylation was detected among subgroups (Figure 8C, Supplementary data file 8d). Focusing on each TriC/CCT component individually, we saw a mainly negative association between DNA methylation and RNA expression and a mainly positive one between transcriptome and proteome data - as expected (Figure 8D). However, correlation of DNA methylome and proteome data did not point in such a clear direction (Figure 8D). Consequently, only CCT2 showed a high association among all omic levels with a correlation score ≥ 0.7 (Figure 8E, Supplementary data file 8d). We therefore, identified the TriC/CCT complex as a feature discriminating pWNT and pG3myc MB.

MB subtypes show distinct N-glycan profiles

One of the major altered genesets between pWNT and pG3myc MB was N-glycosylation (Figure 8B), referring to a post translational modification which is unknown in the context of MB. As glycosylation plays a major role in immune system response and might therefore enable therapeutic options^{148,149}, we focused on this

aspect in more detail. Of note, proteins involved in all aspects of N-glycosylation (synthesis, processing, transport, and antigen presentation via MHC class II) were overrepresented in pWNT (Figure 9C). Quantitative analysis of N-glycans revealed differential N-glycosylation patterns across proteomic MB subtypes (Figure 9D-I). In total 302 N-Glycan species were identified (Figure 9 E-I; Supplementary data file 9a). For non-WNT/ non-SHH MB a higher number of N-glycans were identified in comparison to pWNT, pSHHs and pSHHt (Figure 9F, Supplementary data file 9a). At the quantitative level, proteome MB subtypes were reflected based on their N-glycan profiles (Figure 9G, Supplementary Figure 16A). 92 N-glycans were differentially abundant between the proteome MB types (Supplementary Figure 15B, Supplementary data file 9b). We identified the highest number of exclusive (complex) N-glycans in the subtypes pG3myc and pG4 ($n_{pG3myc} = 22$, $n_{pG4} = 12$, Figure 9H, I, where n represents (complex) N-glycans). Frequently described key factors in tumors are the upregulation of cancer associated sialynated N-glycans as well as aberrant fucosylation¹⁵⁰. A higher proportion of sialynated N-glycans was found in non-WNT/ non-SHH tumors (non-WNT/ non-SHH MB: 59.7-62.0% versus pWNT/pSHH: 49.5-51.9%). A significantly lower proportion of fucosylated N-glycans was detected in pSHHt, compared to all other subtypes (66.7 % (n = 74)) versus 72.1 - 80% (n = 101-174, range of the other MB subtypes, where n represents number of fucosylated N-glycans).

Taken together, integrated proteome analyses shed light on new characteristics in MB subtypes revealing potentially druggable targets. To show validity of results, we recapitulated the six proteome subtypes and two superordinate profiles found in the integrated cohort in a technical and biological validation dataset of FFPE samples ($n_{technical\ cohort}=57$, $n_{validation\ cohort}=31$, Figure 10A-G, Supplementary Figure 17, Supplementary data file 1c, Supplementary data file 10a-c,g). We further verified the differential feature conservation between DNA methylation and protein patterns in the biological validation dataset and underlined the TriC/CCT complex as a discriminator of pWNT and pG3myc MB (Figure 10H, I, Supplementary data file 10b,10h, Source data are provided as a Source Data file).

Discussion

Technical variability and missing values are a general challenge of mass spectrometry-based proteome analyses implying a need for large integrated datasets with reduction of technical biases. Using the HarmonizR integration strategy¹³⁹, we could successfully identify clinically relevant proteome subtypes of MB in a large, integrated cohort of 167 MBs. Herein, we show that FFPE material, which maintains chemical rigidity under cheap storage conditions¹³, enabled the identification and differentiation of molecular subtypes, as previously described for smaller cohorts of FF tissue^{77,78}. Respective results could moreover be confirmed in a technical and biological FFPE validation dataset. In line with previous results¹⁵¹, sample age did not impact data quality, making FFPE tissue highly suitable for large-scale analysis of rare diseases¹⁹¹⁶.

Two overriding molecular patterns were observed across MB subtypes, indicating that MB either follow a transcriptional/replicative (pWNT, pSHHt, pG3myc) or synaptic/immunological (pG4, pSHHs, pG3) profile. These profiles tempt to speculate, that MBs with a synaptic/ immunological pattern (in contrast to MBs with a transcriptional/replicative pattern) may depend more on external stimuli, such as e.g. (potential) synaptic input. Further studies are therefore needed to comprehend the

underlying functional background resulting in the observed patterns. To evaluate a therapeutic potential of these patterns, we used IPA¹⁴² and identified, besides others, CDK4 inhibitors as potential drugs for targeting the groups belonging to the transcriptional profile. Various CDK inhibitors are already FDA-approved for treatment of different types of metastatic cancers and CDK4/6 inhibition has been shown to inhibit tumor growth of medulloblastoma cells in vivo^{106,107}. In contrast, proteome subtypes belonging to the synaptic profile may be – besides others - targeted with the NMDA receptor antagonist memantine. Of note, memantine has neuroprotective properties and was shown to decrease cognitive dysfunction in patients receiving radiotherapy^{108,152}. As radiotherapy is also applied to MB patients the drug may be of specific interest, however, further studies are needed to investigate the clinical potential of the mentioned drugs for MB patients.

We found that DNA-methylation subgroups of MB - which are used for classification of brain tumors in the clinic³³ - are associated with proteome subtypes. This underlines, that the proteome harbors a great potential for identifying subtype specific therapy targets^{43,126,127,140,137}. However, only 30% of marker proteins showed a significant correlation with their respective gene's CpG sites. In general, a low correlation between proteome and methylome data was found in MB, in line with the results of previous studies on other tumor entities^{153,154}. Poor correlations might be attributed to the 850K array design since it mostly assesses promoter methylation sites whereas CpG sites correlating well with gene expression may be located further away from transcriptional start sites¹⁵⁵. Of note, correlation levels of data modalities were not evenly distributed among subtypes. Especially in pWNT tumors, proteins showed a relatively high correlation with their respective gene's CpG sites (38.9% of proteins). In addition, the commonly detected loss of chromosome 6³⁶ was also reflected in proteome data when mapping protein abundances to chromosomal arms. Molecular alterations may hence be more conserved for WNT MBs, whereas DNA-based methylation differences do not always result in an effective change in protein abundance, probably due to post transcriptional and post translational mechanisms (Supplementary Figure 18). These findings highlight the importance of proteome analysis to detect targetable alterations.

We detected two proteome SHH MB subtypes, namely pSHHs and pSHHt. While we cannot fully exclude the possibility that differences in proteome patterns could be due to variations in tissue composition, our results confirmed previously reported proteome patterns in SHH MB⁷⁷. pSHHs tumors reflect the SHHb subgroup defined by Archer et al⁷⁷ showing enrichment of synaptic pathways¹⁶. We found that pSHHs MBs are characterized by a high representation of the citric acid (TCA) cycle and respiratory electron transport, pointing at distinct metabolic profiles of SHH proteome subtypes. Metabolic analysis confirmed significant differences with isocitrate (ISO) and α -ketoglutarate (α KET) being significantly downregulated in pSHHs MBs. As pSHHs MBs also showed a high protein abundance of isocitrate dehydrogenases, this may indicate a higher consumption of these metabolites. As both α KET as well as the amino acid glutamine were significantly downregulated in pSHHs, we hypothesize that these factors might be further transformed to glutamate and further γ -Aminobutyric Acid (GABA), the latter both being linked to synaptic signalling¹⁵⁶. In line with this, pSHHs fell into the "synaptic" profile and GABA targets were significantly upregulated in these tumors. We further speculate that pSHHs tumors might be dependent on synaptic input, a principle that has been shown for other primary brain tumors, but still has to

be shown for medulloblastoma^{157,158}. pSHHt MBs showed a high abundance of proteins involved in transcription/translation, DNA repair and cell cycle. In line with this, respective MB showed an increased proliferation compared to pSHHs.

TP53-mutated SHH cases, stratified as high-risk SHH MB¹⁴⁴, did not form a distinguishable cluster. However, among others, *CHD6*, *DNAJB2* and *NNMT*, known to be associated with aberrant *TP53* expression and high tumor progression^{159–161}, showed a differential abundance comparing *TP53*-mutated to *TP53*-wildtype cases. Further, *CHD6* is suggested as a potential anti-cancer target for tumors with DNA-damage repair-associated processes¹⁶⁰. Mutations within the largest subunit of the elongator complex (*ELP1*) have lately been described in SHH MB⁸⁹. These mutations were found mutually exclusive with *TP53* mutations and *ELP1* mutated SHH MBs were characterized by translational deregulation with upregulation of factors involved in transcription and translation⁸⁹. Reanalysis of published proteome data from *ELP1* mutated SHH MB cases indeed revealed that all cases were attributed to the pSHHt MB subtype (Supplementary data file 5k)⁸⁹. As a limitation, the *ELP1* status of the SHH MB cases in our cohort was only known for n = 3 pSHHs and n = 10 pSHHt tumors (all wildtype). However, all SHH MBs with methylation subtype 3 - associated with *ELP1* mutations - fell into pSHHt^{89,137}. The clinical significance of the two proteome subtypes of SHH MB needs further validation in the future.

Current standard treatment approaches for MB (surgical removal, craniospinal irradiation and combinational chemotherapy) cause severe neuro-cognitive and neuroendocrine late effects. Due to their high responsiveness to therapy, WNT-type MBs are evaluated for therapy de-escalation¹⁶². The identification of *CTNNB1* mutations, or chromosome 6 deletion (monosomy 6) are common markers for the identification of WNT-type MB. Immunohistochemistry is used to detect nuclear β -Catenin staining in tumor cells that can be weak and found only a subset of cell nuclei^{163,164}. Here, Tenascin C (TNC) was found elevated in pWNT MBs, in line with results of previous mRNA-based analyses¹⁶⁵. TNC is a highly glycosylated extracellular matrix (ECM) protein, promoting or inhibiting proliferation and migration in cancer, depending on the present splice variant¹⁶⁶, which will be a field of further study. Besides TNC, a general enrichment of ECM proteins was detected in pWNT MBs. While the ECM has not been investigated in-depth in WNT MB, ECM components have been described to predict outcomes in MB¹⁶⁷. ECM degradation was found as a hallmark of tumor invasion, metastasis development and overall bad prognosis¹⁶⁸. WNT pathway activation dependent disruption of the blood-brain barrier (BBB)¹⁶⁸, was described to permit accumulation of high levels of intra-tumoral chemotherapy in WNT tumors, resulting in a robust therapeutic response. TNC could be another contributor to this phenotype, as high TNC levels contribute to BBB disruption^{110,168}. Furthermore, other BBB contributors, such as *EPLIN1*, *DSP* and *S100A4* were found differential in pWNT.

In line with previous results, we found three proteome subtypes of non-WNT/ non-SHH MBs⁷⁷. pG4 (predominantly comprising G4 tumors), followed the synaptic program. These findings go in line with the literature, as synaptic signatures for G4 tumors, have been described^{43,77}. In pG4 MBs, we detected a higher abundance of VEGF signalling-related proteins, previously described in the context of tumor angiogenesis. VEGF signalling can be targeted in MB using Bevacicunab or Mebendazole^{111,169} and hence might be beneficial for pG4 patients. pG3 MBs (composed of both G3 and G4 tumors)

showed the lowest cluster certainty and inherited the characteristics of both pG3myc and pG4. pG3myc tumors, showed a reduced survival rate and high-risk features, such as LCA histology and solid metastasis. Group 3 MB with *MYC* amplification are highly aggressive and exhibit a bad prognosis^{170,171}. In our cohort, more than half of the patients showed a *CMYC* amplification, while all samples showed an upregulation of *CMYC* target genes, supporting the hypothesis that besides *CMYC* amplification, changes in its phosphorylation status result in a *CMYC-driven* high-risk proteome G3 subtype⁷⁷. Therefore, proteome signatures may be additionally important for stratification of MB patients, as the current stratification scheme for high-risk MB based on (genetic) *MYC* amplification may miss these non-amplified high-risk pG3myc patients. As potential protein biomarkers for pG3myc MB, *DIEXF*, *MDN1*, *POSTN*, *TPD52* and *PALMD* were selected. *TPD52* has recently been suggested as an immunohistochemistry (IHC) marker for high-risk non-WNT/ non-SHH patients¹²⁹. *PALMD* showed the highest elevation in our cohort and was established as a suitable IHC marker for the identification of pG3myc MB. Further prospective trials need to evaluate its significance for stratification of high-risk non-WNT/ non-SHH patients. Further proposed markers for proteomic MB subtypes in this study have to be tested in prospective studies to verify their potential for classification and potential therapy prediction in the future.

High-risk pG3myc MBs showed a high resemblance to pWNT tumors with favorable outcome. Comparing both groups, revealed the components of the Tric/CCT complex to be significantly different. A high abundance of CCT complex proteins has been linked to worse prognosis in cancer and was identified as a predominate driver of Vincalcaloid resistance, including Vincristine, which is among the most frequently used drugs for MB¹⁷². The general low abundance of CCT/Tric proteins in pWNT MB could therefore be a BBB-phenotype independent explanation for the relatively high response to chemotherapy¹⁷³. The usage of CT20p, an amphipathic CCT inhibitor peptide, was described as a promising strategy for the treatment of high-risk tumors with high CCT abundance^{174,175}. Based on our data, the approach should be further investigated as a potential strategy to enhance Vincristine-mediated cytotoxicity in high-risk pG3myc MBs, which were characterized by a particularly high abundance of CCT/Tric proteins.

We further identified increased Asparagine-linked-N-glycosylation as a hallmark of WNT Medulloblastoma. Glycosylation patterns can be used as biomarkers for disease progression¹³³ and aberrant N-glycosylation patterns have been described for brain cancer¹⁷⁶. Of note, aberrant N-glycan structures in cancer could be targeted by immunotherapy and thus provide therapeutic strategies, especially for high-risk tumors that are not sensitive to classical treatment^{148,177}. As an example, chimeric-antigen-receptor (CAR)-modified T cells, that can be specifically directed against tumor-associated carbohydrate antigens (TACAs) are rapidly evolving¹¹⁶. Differential, quantitative N-glycan analysis reflected proteome MB subtypes with high similarity for pSHHt and pSHHs MBs. The latter could be related to dominant SHH activation in these groups, knowingly having an impact on N-glycosylation¹⁷⁸. 12 structures were identified only in high-risk pG3myc patients. Most of these structures are complex bisecting N-glycans, known to be associated with cell growth control and tumor progression^{133,178} and might be related to the unfavorable outcome for pG3myc patients. pG3myc-specific N-glycans do not appear in healthy brain cells, whose N-

glycome is characterized by dismissed N-glycan complexity, lack of complex N-glycans and truncated structures¹⁷⁹ and might serve as suitable immunotherapy targets for high-risk patients.

For pG4 patients, highest amounts of salivated N-Glycans were found, further supporting the immunological profile of pG4 MBs, observed at the proteome level¹⁸⁰.

Taken together, the integration of MB proteome, DNA-methylome and N-glycome data revealed (1) new insights into MB phenotypes, (2) potential biomarkers for rapid histological subtyping and for stratification, and (3) therapeutic targets for MB. Specifically, TriC/CCT-inhibitors or chimeric-antigen-receptor-modified T-cells to target tumor-specific carbohydrates may be applied for high-risk MBs. Superordinate transcription/translational or synaptic proteome profiles across subtypes further revealed targetable vulnerabilities, which may be addressed by e.g. CDK4 inhibitors or memantine.

Methods:

Subject Details:

In house patient samples for main cohort and biological validation cohort:

FFPE Medulloblastoma samples of tumors within the years 1976-2022 were obtained from tissue archives from various neuropathology units in Germany including cases that had been collected within the HIT-MED study cohort. Single patients were enrolled in the SIOP PNET5 study of the HIT-MED study center (for details see Supplementary data file 1c, Supplementary data file 11, clinical data for these patients was excluded from further analysis). All investigations were performed in accordance with local and national ethical rules of patient's material and have, therefore, been performed in accordance with the ethical standards laid down in an appropriate version of the 1964 Declaration of Helsinki. All patients gave their informed consent for scientific use of the data. All samples underwent anonymization. Tumor samples were fixed in 4 % paraformaldehyde, dehydrated, embedded in paraffin, and sectioned at 10 µm for microdissection using standard laboratory protocols. For further information on clinical details of samples, please refer to Supplementary data file 1c Supplementary data file 11.

Medulloblastoma cell lines:

The human Medulloblastoma cell lines DAOY (Ca#HTB-186) and D283med (Ca#HTB-185) were obtained from ATCC, Manassas, VA, USA. DAOY and D283med were authenticated using Eurofins using STR-profiling analysis. UW473 was kindly provided by Michael Bobola. All lines were used as Standards for TMT batches. Cells were cultivated in DMEM (Dulbecco's Modified Eagle Medium, PAN-Biotech) supplemented with 10 % FCS at 37°C, 5 % CO₂.

Publicly available datasets:

For the data integration and harmonization of in-house and publicly available DNA Methylation data the following datasets were used: Archer et al. (2018)⁷⁷: 42 FF MB samples, accessible as a subset of European Genome-phenome Archive ID: EGAS00001001953. Forget et al. (2018)⁷⁸: 38 FF MB samples, accessible via Gene Expression Omnibus (GSE104728). For the analysis of RNA Expression data, processed and normalized data from the following datasets were used: Cavalli et al. (2017)⁷⁷: 763 MB samples, accessible via Gene Expression Omnibus (GPL22286)⁴³. For the data integration and harmonization of in-house and publicly available proteome data, the following datasets were included: Archer et al. (2018)⁷⁷: 45 FF MB samples, available via the MassIVE online repository (MSV000082644, Tandem Mass Tag- (TMT) label-based protein quantification); Forget et al. (2018)⁷⁸: 39 FF MB samples, available via the PRIDE archive (PXD006607, stable isotope labeling by amino acids in cell culture- (SILAC) label-based protein quantification); Petralia et al. (2021)⁵², 23 FF MB samples, , available through the Clinical Proteomic Tumor Analysis Consortium Data Portal (<https://cptac-data-portal.georgetown.edu/cptacPublic/>) and the Proteomics Data Commons (<https://pdc.cancer.gov/pdc/>, Tandem Mass Tag- (TMT) label-based protein quantification). For validation of determined proteome subtypes, as well as the investigation of the proteome profile of ELP1 mutated SHH MB, a dataset published by Waszak et al. (2020)⁸⁹ was used (23 FF MB samples, available via the PRIDE archive (PXD016832, Data independent acquisition label free protein quantification).

Sample preparation and data acquisition:

DNA methylation profiling:

DNA methylation data was generated from FFPE tissue samples. DNA was isolated using the ReliaPrep™ FFPE gDNA Miniprep system (Promega) following the manufacturer's instructions. 100-500 ng DNA was used for bisulfite conversion using the EZ DNA Methylation Kit (Zymo Research). Then the DNA Clean & Concentrator-5 (Zymo Research) and the Infinium HD FFPE DNA Restore Kit (Illumina) were applied. Infinium BeadChip array (EPIC) using manufacturer's instructions were then used to quantify the methylation status of CpG sites on an iScan (Illumina, San Diego, USA). Data has been deposited using accession number GSE222478 and GSE243768 (linked to Series GSE243796). Additionally, previously published data measured on Infinium Human Methylation 450 BeadChip array (450K) were included from EGAS00001001953¹⁸¹, from GSE104728⁷⁸, and GSE130051⁸⁰.

Proteome profiling:

FFPE MB tissue sections were deparaffinized with N-heptane for 10 minutes and centrifuged for 10 minutes at 14,000 g. The supernatant was discarded. Proteins were extracted in 0.1 M triethyl ammonium bicarbonate buffer (TEAB) with 1% sodium deoxycholate (SDC) at 99°C for 1 hour. Sonification was performed for 10 pulses at 30% power, to degrade DNA, using a PowerPac™ HC High-Current power supply (Biorad Laboratories, Hercules, USA)) probe sonicator. For cell lines, proteins were extracted in 0.1M TEAB with 1% SDC at 99 °C for 5 minutes. Sonification was performed for 6 pulses.

The protein concentration of denatured proteins was determined by the Pierce BCA Protein assay kit (Thermo Fischer Scientific, Waltham, USA), following the

manufacturer's instructions. 60 µg of protein for each tissue lysate and 30 µg protein for each cell lysate were used for tryptic digestion. Disulfide bonds were reduced, using 10mM dithiothreitol (DTT) for 30 minutes at 60 °C. Alkylation was achieved with 20 mM iodoacetamide (IAA) for 30 minutes at 37 °C in the dark. Tryptic digestion was performed at a trypsin:protein ratio of 1:100 overnight at 37 °C and stopped by adding 1% formic acid (FA). Centrifugation was performed for 10 minutes at 14000g to pellet precipitated SDC. The supernatant was dried in a vacuum concentrator (SpeedVac SC110 Savant, (Thermo Fisher Scientific, Bremen, Germany)) and stored at -80°C until further analysis.

For the main cohort, 50 µg sample per patient and internal reference, TMT-10 plex labeling (Thermo Fischer Scientific, Waltham, USA), was performed, following the manufacturer's instruction. All 70 patients were run in 8 total TMT 10-plexes. Sample assignment to batches was performed in a semi-randomized manner, according to the four main molecular subtypes. In each batch, 1-2 internal reference samples were included, composed of equal amounts of peptide material from all 70 samples and cell lines. Isobarically labeled peptides were combined and fractionated, using high pH reversed phase chromatography (ProSwift™ RP-4H, Thermo Fischer Scientific Bremen, Germany) on an HPLC system (Agilent 12000 series, Agilent Technologies, Santa Clara, USA). Separation was performed using buffer A (10 mM ammonium hydrogen carbonate (NH₄HCO₃) in H₂O) and buffer B (10mM NH₄HCO₃ in ACN) within a 25-minute gradient, linearly increasing from 3-35% buffer B at a flow rate of 200 nl/min. In total, 13 fractions were collected for each batch, dried in a vacuum concentrator, resuspended in 0.1 % FA to a final concentration of 1mg/ml and subjected to high pH liquid chromatography coupled mass spectrometry (LC-MS). All LC-MS measurements were performed on a UPLC system (Dionex Ultimate 3000, Thermo Fisher Scientific, Bremen, Germany, trapping column: Acclaim PepMap 100 C18 trap ((100 µm x 2 cm, 100 Å pore size, 5 µm particle size); Thermo Fisher Scientific, Bremen, Germany), analytical column: Acclaim PepMap 100 C18 analytical column ((75 µm x 50 cm, 100 Å pore size, 2 µm particle size); Thermo Fisher Scientific, Bremen, Germany)), coupled to an quadrupole-orbitrap-iontrap mass spectrometer (Orbitrap Fusion, Thermo Fisher Scientific, Bremen, Germany). Separation was performed using buffer A (0.1% FA in H₂O) and buffer B (0.1% FA in H₂O) within a 60-minute gradient, linearly increasing from 2-30% buffer B at a flow rate of 300nl/min. Eluting peptides were analyzed, using a DDA based MS3 method with synchronous precursor selection (SPS), as described by McAlister et al.¹⁸². For MS – raw data please refer to the PRIDE archive ([PXD039319](https://www.ebi.ac.uk/pride/archive/projects/PXD039319)).

Proteome profiling for biological and technical validation cohort:

The deparaffinization and quantification was conducted as previously described.

20 µg of the provided samples were dissolved to a concentration of 70% ACN. 2 µL carboxylate modified magnetic beads (GE Healthcare Sera-Mag™, Chicago, USA) at 1:1 (hydrophilic/hydrophobic) in methanol were added following the SP3-protocol workflow¹⁸³. Samples were shook at 1400 rpm for 18 minutes RT and the supernatant was removed. Beads were washed two times with 100% ACN and two times with 70% EtOH. After resuspension in 50 mM ammonium bicarbonate, disulfide bonds were reduced in 10 mM DTT for 30 min, alkylated in presence of 20 mM IAA for 30 min in the dark and digested with trypsin (sequencing grade, Promega) at 1:100

(enzyme:protein) at 37 °C overnight while shaking at 1400 rpm. Peptides were bound in 95% ACN and shook at 1400 rpm for 10 min RT. The supernatant was and the beads were again two times with 100% ACN. Elution of peptides was performed with 20 µL 2% DMSO in 1% formic acid (FA). Samples were dried in a vacuum centrifuge and stored at -20°C until further use.

For the measurement samples were resuspended in 0.1 % FA to a final concentration of 1mg/ml and measured on either a Quadrupole Orbitrap hybrid mass spectrometer (QExactive, Thermo Fisher Scientific) or on a quadrupole-ion-trap-orbitrap MS (Orbitrap Fusion, Thermo Fisher) in orbitrap-orbitrap configuration. For MS – raw data please refer to the PRIDE archive [PXID048767](https://www.ebi.ac.uk/pride/archive/study/PXD048767).

Quadrupole Orbitrap hybrid mass spectrometer set-up: Chromatographic separation of peptides was achieved by nano UPLC (nanoAcquity system, Waters) with a two-buffer system (buffer A: 0.1% FA in water, buffer B: 0.1% FA in ACN). Attached to the UPLC was a peptide trap (100 µm × 20 mm, 100 Å pore size, 5 µm particle size, Acclaim PepMap 100 C18 trap, Thermo Fisher Scientific) for online desalting and purification followed by a 25-cm C18 reversed-phase column (75 µm × 200 mm, 130 Å pore size, 1.7 µm particle size, Peptide BEH C18, Waters). Peptides were separated using an 80-min gradient with linearly increasing ACN concentration from 2% to 30% ACN in 65 minutes. The eluting peptides were analyzed on a Quadrupole Orbitrap hybrid mass spectrometer (QExactive, Thermo Fisher Scientific). Here, the ions being responsible for the 15 highest signal intensities per precursor scan (1 × 10⁶ ions, 70,000 Resolution, 240ms fill time) were analyzed by MS/MS (HCD at 25 normalized collision energy, 1 × 10⁵ ions, 17,500 Resolution, 50 ms fill time) in a range of 400–1200 m/z. A dynamic precursor exclusion of 20 s was used.

Quadrupole-ion-trap-orbitrap mass spectrometer set-up: Chromatographic separation of peptides was achieved with a two-buffer system (buffer A: 0.1 % FA in water, buffer B: 0.1 % FA in ACN). Attached to the UPLC was a peptide trap (100 µm × 200 mm, 100 Å pore size, 5 µm particle size, Acclaim PepMap 100 C18 trap, Thermo Fisher Scientific) for online desalting and purification followed by a 25 cm C18 reversed-phase column (75 µm × 250 mm, 130 Å pore size, 1.7 µm particle size, Peptide BEH C18, Waters). Peptides were separated using an 80-min gradient with linearly increasing ACN concentration from 2 % to 30 % ACN in 65 minutes. Eluting peptides were ionized using a nano-electrospray ionization source (nano-ESI) with a spray voltage of 1800, transferred into the MS, and analyzed in data-dependent acquisition (DDA) mode. For each MS1 scan, ions were accumulated for a maximum of 240 milliseconds or until a charge density of 1x10⁶ ions (AGC Target) was reached. Fourier-transformation-based mass analysis of the data from the orbitrap mass analyzer was performed, covering a mass range of 400-1,200 m/z with a resolution 60,000. Peptides with charge states between 2+ - 5+ above an intensity threshold of 1x10⁵ were isolated within a 2 m/z isolation window from each precursor scan and fragmented with a normalized collision energy of 25 % using higher energy collisional dissociation (HCD). MS2 scanning was performed at a resolution of 17,500 on the quadrupole-ion-trap-orbitrap MS in orbitrap-orbitrap configuration, covering a mass range from 100 m/z and accumulated for 50 ms or to an AGC target of 1x10⁵. Already fragmented peptides were excluded for 15 seconds.

Histology and Immunohistochemistry

FFPE tissue samples were sectioned into 2 μm thick slices, according to standard laboratory protocols. Immunohistochemical stainings were performed on an automated staining machine (Ventana BenchMark XT, Roche Diagnostics, Mannheim, Germany). The following primary antibodies were used: ALDH1A3 (NBP2-15339, Novus Biologicals, 1:1000), c-myc (Z2734RL, Zeta Corporation, 1:25), TENASCIN C (SAB4200782, Sigma-Aldrich, 1:1000), PALMD (NBP2-55156, Novus Biologicals, 1:750). Further information on the antibodies and staining program can be found in Supplementary methods Table 2.

Transcriptome profiling:

Maxwell RSC RNA FFPE Kit was used to isolate RNA from $10 \times 10 \mu\text{m}$ sections of FFPE tissue (PROMEGA Maxwell RSC RNA FFPE kit). RNA 6000 Nano Chip on an Agilent 2100 Bioanalyzer (Agilent Technologies) was used to analyse RNA integrity. From 400 ng total per sample, ribosomal RNA was depleted with the help of the RiboCop rRNA Depletion Kit (Lexogen) followed by RNA sequencing library generation using the CORALL Total RNA-Seq Library Prep Kit (Lexogen), followed by the Lexogen CORALL total RNA-Seq V2 Library Prep Kit with UDIs (according to manufacture protocol, short insert size version). Illumina NextSeq2000 machine using the P3 Reagents / 100 cycle kit as paired-end sequencing 2x 57 bp (+ 2x index read 12bp). Data has been deposited under accession number GSE243795.

Metabolic and amino acid profiling:

^{13}C -Labeled Metabolite Yeast Extract (Catalog No. ISO-1, ISOTopic solutions e.U.) LOT: 20211007 and Canonical Amino Acid Mix (Catalog No. MSK-CAA-1, Cambridge Isotope Laboratories, Inc. (CIL)) were prepared according to instructions. Tissue sections of sSHH and tSHH medulloblastoma samples were deparaffinized by two 5 min washes in xylene. 20 μL of ^{13}C -Labeled Metabolite Yeast Extract and 1 μL of diluted 0.1 M Canonical Amino Acid Mix were added, samples were then homogenized in 180 μL water using the TissueLyser (Qiagen N.V., Netherlands) at 20Hz for 2 minutes. Afterwards, protein precipitation and metabolite extraction were achieved by adding ice-cold methanol twice (800 μL and 400 μL) and 80% methanol (200 μL). The supernatant was combined and dried in a vacuum concentrator centrifuge, and stored at -20°C until further use.

Polar and polar ionic metabolites were analyzed by single ion monitoring (SIM) mass spectrometry coupled to ion chromatography and IC-SIM-MS raw data processing was performed as described by van Pijkeren and Egger et al.¹⁸⁴ using a quadrupole orbitrap mass spectrometer (Exploris 480, Thermo Fisher Scientific) and an ICS-6000 (Thermo Fisher Scientific).

Amino acids were analyzed by multiple reaction monitoring (MRM) mass spectrometry using a triple quadrupole mass spectrometer coupled to ultra-high performance liquid chromatography (UPLC). Amino acids were separated using an Acquity Premier UPLC system (Waters) equipped with an Atlantis Premier BEH C18 AX column (1.7 μm , 2.1x150mm, Waters) heated to 45°C . A gradient of mobile phase A (water, 0.1% formic acid (FA)) and mobile phase B (acetonitrile, 0.1% FA) was applied as followed: 1% B at 0.350 mL/min for 1 min, to 20% B in 1 min at 0.350 mL/min, to 40%

B in 0.5 min at 0.350 mL/min, to 95% B in 1.5 min at 0.450 mL/min, hold for 0.5 min, for re-equilibration, switch to 1% B in 0.1 min at 0.450 mL/min, hold for 0.1 min at 0.450 mL/min and hold for 1.3 min at 0.350 mL/min. Samples were measured on a Xevo-TQ XS Mass spectrometer (Waters) equipped with an electrospray ionization source operated in positive ion mode. The mass spectrometer was operated in multiple reaction monitoring (MRM) mode using individual cone and collision voltages for each amino acid and its internal standard (Supplementary data file 1). Raw files were analyzed by MS Quan in waters connect (Waters, V1.7.0.7).

For MS raw data of the metabolites and amino acids please refer to MetaboLights repository¹⁸⁵ [MTBLS9830](#) and [MTBLS9836](#), respectively.

N-Glycan profiling:

100 µg of protein for 18 samples was denatured, reduced, and alkylated as described above. Samples was concentrated by 3 kDa Amicon Ultra centrifugal filters (Merck Millipore, R0NB30416) with 100 mM NH₄HCO₃ to exchange the buffer and retain globular particles above 3 kDa. Thirty units of PNGase F were added to each sample and incubated in a 37 °C Thermomixer for 24 h. After PNGase F digestion, purified N-glycans were eluted by Sep-Pak C18 cartridges (Water, WAT023590) with 5% acetic acid and dried in a speed vacuum. The purified N-glycans were then permethylated using an optimized solid-phase permethylation method and analyzed via LC-MS measurement as mentioned here ^{186 85}. Glycan data has been deposited at GlycoPOST¹⁸⁷ with the identifier [GPST000414](#).

Raw data processing:

Processing of DNA Methylation Array Data:

Idata files generated using the above protocol were processed in R (Version 4.0.5). The files were read in using the minfi package (Version 1.36.0)¹⁸⁸. Further information of quality control in supplementary methods. Differentially methylated probes/ CpG sites were found using the limma package (Version 3.46.0)¹⁸⁹, corrected for multiple testing using Benjamini Hochberg (cut-off 5% FDR). M-values of 10,000 differentially methylated CpG sites which could cluster subtypes based on biological differences were selected for further analysis. Similarly, DMR analysis was performed using DMRcate package (V4.30.0). For DMR analysis, we set a min of 10 CpGs per DMR (< 1000 nt from each other) to minimize gene overlap, which resulted in ~9000 DMRs with each DMR having 10-200 CpGs

Processing of Proteome raw data for main cohort:

Processing of Proteome raw data for the integrated cohort:

Obtained raw data from in-house generated and publicly available (Archer et al (2018)⁷⁷, TMT 10-Plex; Petralia et al. (2021)⁵², TMT 11-Plex). TMT-based LC-MS measurements were processed with the Andromeda algorithm, implemented in the MaxQuant software (Max Plank Institute for Biochemistry, Version 1.6.2.10)⁹⁰ and searched against a reviewed human database (downloaded from Uniprot February 2019, 26,659 entries).). The Carboxymethylation of cysteine residues was set as a fixed modification. Methionine oxidation, N-terminal protein acetylation and the conversion of glutamine to pyroglutamate were set as variable modifications. Peptides

with a minimum length of 6 amino acids and a maximum mass of 6,000 Da were considered. The mass tolerance was set to 10 ppm. The maximum number of allowed missed cleavages in tryptic digestion was two. A false discovery rate (FDR) value threshold <0.01, using a reverted decoy peptide databases approach, was set for peptide identification. Quantification was performed, based on TMT reporter intensities at MS3 level for LC-MS3 in-house data and at MS2 level for LC-MS2 data, acquired by Archer et al.⁷⁷ and Petralia et al.⁵² All studies were searched separately. Fractions for each TMT batch were searched jointly.

For stable isotope labeling by amino acids in cell culture (super-SILAC) data, acquired by Forget et al. (2018)⁷⁸, log₂ transformed SILAC ratios were directly obtained from the MassIVE online repository (MSV000082644).

For the external validation the dataset published by Waszak et al. (2020)⁸⁹ was used. The DIA raw data spectra were downloaded from PRIDE and processed using Data Independent Acquisition with Neural Networks (DIA-NN, version 1.8.1)¹⁹⁰. The spectra were searched against a peer reviewed human FASTA database (downloaded from UniProt April 2020, 20,365 entries). A spectral library was generated in silico by DIA-NN using the same FASTA database. Smart profiling was enabled for library generation. Methionine oxidation, carboxymethylation of cysteine residues as well as N-terminal methionine excision were set as variable modifications. The maximum number of variable modifications was set to three, the maximum number of missed cleavages was two. The peptide length range was set from 7 to 30. Mass accuracy, MS1 accuracy, and the scan window were optimized by DIA-NN. An FDR < 0.01 was applied at the precursor level - decoys were generated by mutating target precursors' amino acids adjacent to the peptide termini. Interference removal from fragment elution curves as well as normalization were disabled. Neural network classifier was set to single-pass mode and the fixed-width center of each elution peak was used for quantification.

Processing of the biological and technical validation cohorts: The spectra were searched with the Sequest algorithm integrated in the Proteome Discoverer software (v 2.4.1.15), Thermo Fisher Scientific) against a reviewed human database (downloaded from Uniprot in June 2021, Containing 20,683 entries)). Carbamidomethylation was set as fixed modification for cysteine residues and the oxidation of methionine, and pyro-glutamate formation at glutamine residues at the peptide N-terminus, as well as acetylation of the protein N-terminus were allowed as variable modifications. A maximum number of 2 missing tryptic cleavages was set. Peptides between 6 and 144 amino acids were considered. A strict cutoff (FDR<0.01) was set for peptide and protein identification. Quantification was performed using the Minora Algorithm, implemented in Proteome discoverer.

Processing of N-Glycan raw data:

N-Glycan raw data were open with Xcalibur Qual Browser (Version No 4.2.28.14). MaxQuant were used for extracting all the detected masses and *m/z* from MS raw data of permethylated reducing N-glycans. An in-house Python-script was used to extract and calculate monosaccharide compositions based on the molecular weight of each derivatized N-glycan¹⁹¹. The N-glycan structures were identified, matched to N-lycan

compositions and quantified using the Xcalibur, Glycoworkbench 2.1 and Skyline software (Version No 21.1.0.278)¹⁸⁶. Further statistical analysis was performed with the Perseus software.

Processing of raw transcriptome data:

Raw fastq files of human samples were processed in usegalaxy.eu¹⁹². Low quality reads were detected using *FastQC* (Galaxy Version 0.73+galaxy0), and *Trimmomatic* (Galaxy Version 0.38.1) was used for trimming poor quality reads (reads with average quality <20). Reads were aligned to the GRh38 human reference genome using *STAR aligner* (Galaxy Version 2.7.8a+galaxy1). Gene expression was quantified with *featureCounts* (Galaxy Version 2.0.1+galaxy2) and VST-normalized files were generated by *DEseq2* (Galaxy Version 2.11.40.7+galaxy2). Further processing of data was performed with R (v4.2.1). Transcriptome data was combined with publicly available transcriptome data⁷⁷. Batch corrected with HarmonizR¹³⁹.

Processing of DNA Methylation Array Data:

Raw signal intensities for EPIC and 450K files were read individually. Since ~ 93% of the loci of 450K array are also present on EPIC array, they can be combined using minfi's combineArrays(). After combining the two arrays they can be output as a virtual array. In this study, 450K array was the output virtual array since a greater number of samples were measured on 450K.

The detection P-value was used to identify sample quality and filter out bad quality samples (none were excluded, n=0). Further, probes having bad quality (n=49,091), probes with single nucleotide polymorphism (n = 12,868) and probes present on X and Y chromosomes (n=8,777) were filtered out. After normalization and probe filtering, the m-values $\log_2(M/U)$ where methylation intensity is denoted by M and unmethylation intensity denoted by U were used for further analysis.

Data normalization and integration:

Normalization and integration of DNA Methylation Array Data:

Single-sample noob normalization (ssNoob) was performed since we combined samples from different arrays (EPIC and 450K). The detailed method development has been mentioned^{193,194}.

Normalization and integration of Proteome data:

Prior to data integration, protein abundances were handled separately for each dataset. TMT reporter intensities were \log_2 transformed and median normalized across columns. Technical variances between TMT batches were corrected, using HarmonizR framework (Version 0.0.0.9). As described by here¹³⁹, mean subtraction across rows was applied to batch-effect corrected TMT reporter intensities to mimic SILAC ratios, prior to data integration. \log_2 transformed super SILAC ratios were median normalized across columns prior to data integration.

Processed data from individual studies was combined based on the UniProt identifier, data harmonization was performed as described above. Combined, harmonized protein abundances were mean-scaled across rows. Out of 176 analyzed cases, 9 patients were excluded from further analysis, as high blood protein yields, suppressing tumor-specific signals, were detected from LC-MS/MS measurements (Supplementary data file 1a).

For the external validation cohort protein abundances were log₂ transformed and median normalized across columns. Samples were assigned to proteome subtypes individually. Protein abundances were reduced to the 3998 proteins, considered in the main cohort. Harmonized protein abundances from the main cohort were integrated with each individual sample. Mean row normalization was performed to adjust values from validation samples to the main cohort. Pearson correlation-based hierarchical clustering, with average linkage was applied using the Perseus software (Max Plank Institute for Biochemistry, Version 1.5.8.5)⁹¹.

For biological and technical validation cohort the data was processed and harmonized as described above. For the biological validation one sample had to be excluded due to high blood protein yields as described above. The proteome subtypes for the biological validation were assigned via the ACF classifier¹⁹⁵. The proteome subtypes for the technical validation were taken from the main cohort. Protein abundances were treated as above.

Normalization of N-Glycan data:

N-Glycan intensities were log₂ transformed and median normalized across columns to compensate for injection amount variations.

Quantification and statistical analysis:

Dimensionality reduction and hierarchical clustering:

Nonlinear Iterative vertical Least Squares (NIPALS) PCA and hierarchical clustering were performed in the R software environment (version 4.1.3). For Principal component calculation and visualization, the mixOmics package (Version 6.19.4.)⁶¹ was used in Bioconductor (version 3.14). Hierarchical clustering was performed based on pheatmap package (version 1.0.12) and ComplexHeatmap (Version 2.6.2)¹⁰³. Pearson correlation was applied as a distance metric. Ward.D linkage was used. Pairwise complete correlation was used, to enable the consideration of missing values.

Consensus Clustering:

To determine the ideal number of clusters from proteome and DNA-methylation data, Consensus Clustering was applied on normalized and integrated datasets, using the ConsensusClusterPlus package (Version 1.6)¹⁰⁴, in the R software environment (version 4.1.3). In correspondence with the current maximum number of suspected MB subtypes, the number of clusters was varied from 2 to 12 and calculated with 1,000 subsamples for all combinations of two clustering methods (Hierarchical clustering (HC) and partition around medoids (PAM)) and three distance metrics (Euclidean, Spearman, Pearson). The Ward's method was applied for linkage. Missing value

tolerant pairwise complete correlation was used, to enable the consideration of missing values. For each sample, the cluster certainty was calculated by how many times under the application of different distance metrics (Euclidean, Spearman, Pearson) and clustering approaches (k-medoids, hierarchical clustering) a sample was associated with a certain cluster, while allowing a total number of six clusters.

Differential analysis and visualization:

Statistical testing was carried out, using the Perseus software¹⁹⁶. ANOVA testing was performed for the comparison across multiple subgroups/subtypes. Factors, identified with p -value < 0.05 were considered statistically significant differential abundant across groups. For the identification of subtype-specific biomarkers, Students t-testing was applied (p -value < 0.05 , Foldchange difference > 1.5). Visualization of t-test results and abundance distributions across groups was performed in PRISM (GraphPad, Version 5) and Microsoft excel (Version 16.5.).

Functional annotation of data sets:

REACTOME- based¹⁹⁷ Gene Set Enrichment Analysis was performed by using the GSEA software (version 4.1, Broad Institute, San Diego, CA, USA)¹⁹⁸. 1000 permutations were used. Permutation was performed based on gene sets. A weighted enrichment statistic was applied, using the signal-to-noise ratio as a metric for ranking genes. No additional normalization was applied within GSEA. As in default mode, gene sets smaller than 15 and bigger than 500 genes were excluded from analysis. For visualization of GSEA results, the EnrichmentMap (version 3.3)¹⁹⁹ application within the Cytoscape environment (version 3.8.2)²⁰⁰ was used. Gene sets were considered if they were identified at an FDR < 0.25 and a p -value < 0.1 . For gene-set-similarity filtering, data set edges were set automatically. A combined Jaccard and Overlap metric was used, applying a cutoff of 0.375. For gene set clustering, AutoAnnotate (version 1.3)²⁰¹ was used, using the Markov cluster algorithm (MCL). The gene-set-similarity coefficient was utilized for edge weighting.

Survival curves:

Kaplan-Meier curves were generated for the overall survival of 121 patients. All Kaplan-Meier curves and log rank test p values were generated with PRISM (GraphPad, Version 5). A conservative log-rank test (Mantel-Cox) was used for the comparison of survival curves. A significant difference between curves was assumed at a p -value < 0.05 .

Copy number frequency plots of Proteome and DNA Methylome data:

Copy number analysis was performed on samples having both methylation and proteomic data (N=115). Samples from 450K and EPIC array were read in separately as mentioned above. Data were read using `read.metharray.sheet()` and `read.metharray.exp()` using the `minfiData` package (Version 0.36.0)¹⁸⁸. For normalization, `preprocessIllumina` normalization using `MsetEx` data containing control samples for normalization of 450K array data, while for EPIC array data `minfidataEPIC`

(Version 1.16.0)¹⁸⁸ was used. IlluminaHumanMethylation450kanno.ilmn12.hg19 and IlluminaHumanMethylationEPICanno.ilm10b4.hg19 were used to generate the annotation files of 450K and EPIC array data respectively.

Individual sample CNV plots were generated as mentioned in the Conumee package (Version 1.24.0) vignette, and the segmentation information from each sample was saved and used later for generation of cumulative CNV plot using CNAppWeb tool¹⁰⁵(cut-off $\geq |0.2|$) for gain or loss). The segmentation information for all samples belonging to one subtype were combined into a single file in subgroup specific manner and then read into CNAppWeb tool.

Combining the segmentation information from proteome data and methylome data in subgroup specific manner, pearson correlation-based distance plot was generated.

To map the protein abundancies to each of the chromosomes, protein names were converted to their respective gene names and a column containing mapping information for these genes was added. Copynumber (Version 1.30.0) package in R was used to generate segmentation information for these proteins. CNAppWeb tool using the cut-off mentioned above was used to map the protein abundancies to respective chromosomes.

Integration of Proteome and DNA Methylome data:

DIABLO from mixOmics (Version 6.19.4)⁶¹ was used for integration of proteome and methylome data to correlate the two data types. Proteome data (3990 proteins, 115 samples) and methylome data (10,000 differentially methylated CpG sites, 115 proteins) were pre-processed as mentioned above. Steps followed were same as explained in the mixOmics vignette. Briefly, datasets were integrated, an output variable containing information about which subgroup the samples belong to was also supplied. Each data set is broken down into components (5 components for this study) or latent variables which are associated with the data. Components were selected using 5-fold cross validation repeated 50 times and since the groups were imbalanced lowest overall error rate and centroid distance was used. For each dataset and for each component sparse DIABLO was applied which will select variables contributing maximally to the selected component. sPLS-DA was applied to the selected variables to generate the correlation circus plot (cut-off 0.7) which gives the variables that are either positively or negatively correlating with each other. DMRs between each methylome subtype was found in a pairwise manner, corrected for multiple testing using Benjamini Hochberg (cut-off 5% FDR) and integrated with proteome data in mixOmics.

Global correlation of Proteome and DNA Methylation data:

To check for overall correlation between the two datasets, subgroup specific (pWNT = 13, pSHHt = 29, pSHHs = 6, pG4 = 36, pG3 = 11, pG3Myc = 20) pearson correlation (cut-off 0.7) was performed between the proteome (3990 proteins and 115 samples) and methylome (381,717 probes and 115 samples) in R (Version 4.0.5). The data was subsetted for correlation value ≥ 0.7 and matches of proteins to their respective probes using python script in anaconda JupyterLab (Version 3.0.14). Non-subgroup specific pearson correlation between the proteome and methylome data was similarly performed with focus on potential biomarkers for each subgroup and their correlation

with methylation probes. Scatterplots of biomarker's protein abundance and the M-values of CpG sites of its own gene (crossing the Pearson correlation cut-off of 0.7) were plotted to confirm the correlations. For correlating DMRs and proteins, mean of all CpG sites belonging to each DMR was taken to find correlation between all DMRs and proteins

For correlation of CCT complex components, all samples for which we had all three datasets were considered (n=60) and Pearson correlation ≥ 0.7 were plotted using circlize(Version 0.4.15) and corrplot (Version 0.92) package in R (Version 4.3.0).

Quantification of Immunohistochemical stainings

Immunostained tissue sections were digitalized using a Hamamatsu NanoZoomer 2.0-HT C9600 whole slide scanner (Hamamatsu Photonics, Tokyo, Japan). Slide images were exported using NDP view v2.7.43 software. Digital image analysis was performed using ImageJ/Fiji software²⁰² after white balance correction in Adobe Photoshop 2022 (Adobe Inc., San Jose, USA). Tumor areas were labelled via manually drawn regions of interest (ROIs). Tissue areas not eligible for quantification (e.g. non-tumorous tissue, technical or digital artifacts) were excluded from the analysis. Total tumor tissue areas were measured in grayscale converted images via consistent global thresholding (0, 241) and subsequent pixel quantification within the ROIs. DAB-positive pixels (i.e. brown immunostaining) were quantified on a three-tiered intensity scale after application of the color deconvolution plugin. In detail, pixels were successively quantified within three distinct thresholds [0, 134 (strong/ 3+); 135, 182 (medium/2+); and 183, 203 (weak/ 1+)]. Based on the conventional Histo-score, pixel quantities of strong, medium and weak intensity were multiplied by three, two and one, respectively, and then summed up. The hereby generated score is referred to as a digital Histoscore (DH-score).

Data availability

Proteome data have been deposited under [PXD039319](#) (TMT data), and [PXD048767](#) (validation cohorts). DNA Methylation and RNA Seq data can be accessed via GSE243796 containing subsets GSE222478, GSE243768 and GSE243795. Metabolomics and amino acid data have been deposited to the EMBL-EBI MetaboLights database¹⁸⁵ with the identifier [MTBLS9830](#) and [MTBLS9836](#). Glycan data has been deposited at GlycoPOST¹⁸⁷ with the identifier [GPST000414](#). Source data are provided with this paper.

References

1. Subramanian, S. & Ahmad, T. Childhood Brain Tumors. in *Childhood Brain Tumors* (StatPearls Publishing, Treasure Island (FL), 2024).
2. DeVita Jr, V. T., Lawrence, T. & Rosenberg, S. *Cancer: Principles and Practice of Oncology (11th Edition)*. (2019).

3. Donahue, B. *et al.* Radiation therapy quality in CCG/POG intergroup 9961: implications for craniospinal irradiation and the posterior fossa boost in future medulloblastoma trials. *Front Oncol* **2**, (2012).
4. Câmara-Costa, H. *et al.* Neuropsychological Outcome of Children Treated for Standard Risk Medulloblastoma in the PNET4 European Randomized Controlled Trial of Hyperfractionated Versus Standard Radiation Therapy and Maintenance Chemotherapy. *International Journal of Radiation Oncology*Biophysics*Physics* **92**, 978–985 (2015).
5. Major, N. *et al.* The Current State of Radiotherapy for Pediatric Brain Tumors: An Overview of Post-Radiotherapy Neurocognitive Decline and Outcomes. *J Pers Med* **12**, 1050 (2022).
6. Papanicolau-Sengos, A. & Aldape, K. DNA Methylation Profiling: An Emerging Paradigm for Cancer Diagnosis. *Annual Review of Pathology: Mechanisms of Disease* **17**, 295–321 (2022).
7. Aftimos, P. G., Barthelemy, P. & Awada, A. Molecular biology in medical oncology: diagnosis, prognosis, and precision medicine. *Discov Med* **17**, 81–91 (2014).
8. Serrano, J. & Snuderl, M. Whole Genome DNA Methylation Analysis of Human Glioblastoma Using Illumina BeadArrays. in 31–51 (2018). doi:10.1007/978-1-4939-7659-1_2.
9. Kling, T., Wenger, A., Beck, S. & Carén, H. Validation of the MethylationEPIC BeadChip for fresh-frozen and formalin-fixed paraffin-embedded tumours. *Clin Epigenetics* **9**, 33 (2017).
10. Moran, S. *et al.* Validation of DNA methylation profiling in formalin-fixed paraffin-embedded samples using the Infinium HumanMethylation450 Microarray. *Epigenetics* **9**, 829–833 (2014).
11. Zong, L. *et al.* A comprehensive assessment of exome capture methods for RNA sequencing of formalin-fixed and paraffin-embedded samples. *BMC Genomics* **24**, 777 (2023).
12. Decruyenaere, P. *et al.* RNA Extraction Method Impacts Quality Metrics and Sequencing Results in Formalin-Fixed, Paraffin-Embedded Tissue Samples. *Laboratory Investigation* **103**, 100027 (2023).
13. Gustafsson, O. J. R., Arentz, G. & Hoffmann, P. Proteomic developments in the analysis of formalin-fixed tissue. *Biochim Biophys Acta Proteins Proteom* **1854**, 559–580 (2015).
14. L Avaritt, N. Decoding the Proteome in Formalin-Fixed Paraffin-Embedded (FFPE) Tissues. *J Proteomics Bioinform* **07**, (2014).
15. Magdeldin, S. & Yamamoto, T. Toward deciphering proteomes of formalin-fixed paraffin-embedded (FFPE) tissues. *Proteomics* **12**, 1045–1058 (2012).
16. Drummond, E. S., Nayak, S., Ueberheide, B. & Wisniewski, T. Proteomic analysis of neurons microdissected from formalin-fixed, paraffin-embedded Alzheimer’s disease brain tissue. *Sci Rep* **5**, 15456 (2015).
17. Ostrom, Q. T. *et al.* Alex’s Lemonade Stand Foundation Infant and Childhood Primary Brain and Central Nervous System Tumors Diagnosed in the United States in 2007–2011. *Neuro Oncol* **16**, x1–x36 (2015).
18. Jin, B., Li, Y. & Robertson, K. D. DNA Methylation: Superior or Subordinate in the Epigenetic Hierarchy? *Genes Cancer* **2**, 607–617 (2011).
19. Robertson, K. D. DNA methylation and human disease. *Nat Rev Genet* **6**, 597–610 (2005).
20. Palacios, D., Summerbell, D., Rigby, P. W. J. & Boyes, J. Interplay between DNA Methylation and Transcription Factor Availability: Implications for Developmental Activation of the Mouse *Myogenin* Gene. *Mol Cell Biol* **30**, 3805–3815 (2010).

21. Pratt, D., Sahm, F. & Aldape, K. DNA methylation profiling as a model for discovery and precision diagnostics in neuro-oncology. *Neuro Oncol* **23**, S16–S29 (2021).
22. Sturm, D. *et al.* New Brain Tumor Entities Emerge from Molecular Classification of CNS-PNETs. *Cell* **164**, 1060–1072 (2016).
23. El-Ayadi M *et al.* High-grade glioma in very young children: a rare and particular patient population. *Oncotarget* (2017).
24. Maria Yiallourou. Medulloblastoma – Brief Information. *kinderkrebsinfo* (2024).
25. Gajjar, A. *et al.* Outcomes by Clinical and Molecular Features in Children With Medulloblastoma Treated With Risk-Adapted Therapy: Results of an International Phase III Trial (SJMB03). *J Clin Oncol* **39**, 822–835 (2021).
26. Choi, J. Y. Medulloblastoma: Current Perspectives and Recent Advances. *Brain Tumor Res Treat* **11**, 28 (2023).
27. Borgenvik, A., Čančer, M., Hutter, S. & Swartling, F. J. Targeting MYCN in Molecularly Defined Malignant Brain Tumors. *Front Oncol* **10**, (2021).
28. Zoghbi, H. Y. & Beaudet, A. L. Epigenetics and Human Disease. *Cold Spring Harb Perspect Biol* **8**, a019497 (2016).
29. Wang, K. C. & Chang, H. Y. Epigenomics. *Circ Res* **122**, 1191–1199 (2018).
30. Moore, L. D., Le, T. & Fan, G. DNA Methylation and Its Basic Function. *Neuropsychopharmacology* **38**, 23–38 (2013).
31. Kurimoto, T. *et al.* Effect of O6-methylguanine-DNA methyltransferase methylation in medulloblastoma. *Mol Clin Oncol* (2017) doi:10.3892/mco.2017.1431.
32. Capper, D. *et al.* DNA methylation-based classification of central nervous system tumours. *Nature* **555**, 469–474 (2018).
33. Louis, D. N. *et al.* The 2021 WHO Classification of Tumors of the Central Nervous System : a summary. *Neuro Oncol* **23**, 1231–1251 (2021).
34. Feber, A. *et al.* Using high-density DNA methylation arrays to profile copy number alterations. *Genome Biol* **15**, R30 (2014).
35. McCabe, M. G., Bäcklund, L. M., Leong, H. S., Ichimura, K. & Collins, V. P. Chromosome 17 alterations identify good-risk and poor-risk tumors independently of clinical factors in medulloblastoma. *Neuro Oncol* **13**, 376–383 (2011).
36. Clifford, S. C. *et al.* Wnt/Wingless pathway activation and chromosome 6 loss characterize a distinct molecular sub-group of medulloblastomas associated with a favorable prognosis. *Cell Cycle* **5**, 2666–2670 (2006).
37. Zhukova, N. *et al.* Subgroup-Specific Prognostic Implications of TP53 Mutation in Medulloblastoma. *Journal of Clinical Oncology* **31**, 2927–2935 (2013).
38. Goschzik, T. *et al.* Genetic alterations of TP53 and OTX2 indicate increased risk of relapse in WNT medulloblastomas. *Acta Neuropathol* **144**, 1143–1156 (2022).
39. Kawachi, D. *et al.* Novel MYC-driven medulloblastoma models from multiple embryonic cerebellar cells. *Oncogene* **36**, 5231–5242 (2017).
40. Staal, J. A., Pei, Y. & Rood, B. R. A proteogenomic approach to understanding myc function in metastatic medulloblastoma tumors. *Int J Mol Sci* **17**, (2016).
41. Park, E. G. *et al.* Genomic Analyses of Non-Coding RNAs Overlapping Transposable Elements and Its Implication to Human Diseases. *Int J Mol Sci* **23**, 8950 (2022).
42. Weller, M. *et al.* Molecular classification of diffuse cerebral WHO grade II/III gliomas using genome- and transcriptome-wide profiling improves stratification of prognostically distinct patient groups. *Acta Neuropathol* **129**, 679–693 (2015).
43. Cavalli, F. M. G. *et al.* Intertumoral Heterogeneity within Medulloblastoma Subgroups. *Cancer Cell* **31**, 737–754.e6 (2017).
44. Yoon, J.-H., Abdelmohsen, K. & Gorospe, M. Posttranscriptional Gene Regulation by Long Noncoding RNA. *J Mol Biol* **425**, 3723–3730 (2013).

45. Du, Y. *et al.* Differences in alternative splicing and their potential underlying factors between animals and plants. *J Adv Res* (2023) doi:10.1016/j.jare.2023.11.017.
46. Dubuc, A. M. *et al.* Subgroup-specific alternative splicing in medulloblastoma. *Acta Neuropathol* **123**, 485–499 (2012).
47. Wang, L. *et al.* The evolution of alternative splicing in glioblastoma under therapy. *Genome Biol* **22**, 48 (2021).
48. O'Brien, J., Hayder, H., Zayed, Y. & Peng, C. Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Front Endocrinol (Lausanne)* **9**, (2018).
49. Rivero-Hinojosa, S. *et al.* Proteomic analysis of Medulloblastoma reveals functional biology with translational potential. *Acta Neuropathol Commun* **6**, 48 (2018).
50. Han, X., Aslanian, A. & Yates, J. R. Mass spectrometry for proteomics. *Curr Opin Chem Biol* **12**, 483–490 (2008).
51. Wiśniewski, J. R. Proteomic sample preparation from formalin fixed and paraffin embedded tissue. *J Vis Exp* (2013) doi:10.3791/50589.
52. Petralia, F. *et al.* Integrated Proteogenomic Characterization across Major Histological Types of Pediatric Brain Cancer. *Cell* **183**, 1962--1985.e31 (2020).
53. Hu, M. *et al.* The role of N-glycosylation modification in the pathogenesis of liver cancer. *Cell Death Dis* **14**, 222 (2023).
54. Stanley, P., Moremen, K. W., Lewis, N. E., Taniguchi, N. & Aebi, M. *N-Glycans*. (2022).
55. Lei, X. *et al.* Immune cells within the tumor microenvironment: Biological functions and roles in cancer immunotherapy. *Cancer Lett* **470**, 126–133 (2020).
56. Silsirivanit, A. Glycosylation markers in cancer. in 189–213 (2019). doi:10.1016/bs.acc.2018.12.005.
57. Kim, H. *et al.* Human hepatocellular carcinomas with “Stemness”-related marker expression: keratin 19 expression and a poor prognosis. *Hepatology* **54**, 1707–1717 (2011).
58. Wu, K. *et al.* The role of CD44 in epithelial–mesenchymal transition and cancer development. *Onco Targets Ther* 3783 (2015) doi:10.2147/OTT.S95470.
59. Voß, H. *et al.* HarmonizR enables data harmonization across independent proteomic datasets with appropriate handling of missing values. *Nat Commun* **13**, 3523 (2022).
60. Singh, A. *et al.* DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics* **35**, 3055–3062 (2019).
61. Rohart, F., Gautier, B., Singh, A. & Lê Cao, K.-A. mixOmics: An R package for ‘omics feature selection and multiple data integration. *PLoS Comput Biol* **13**, e1005752 (2017).
62. Spainhour, J. C., Lim, H. S., Yi, S. V & Qiu, P. Correlation Patterns Between DNA Methylation and Gene Expression in The Cancer Genome Atlas. *Cancer Inform* **18**, 117693511982877 (2019).
63. Taylor, M. D. *et al.* Molecular subgroups of medulloblastoma: the current consensus. *Acta Neuropathol* **123**, 465–472 (2012).
64. Eberhart, C. G. *et al.* Histopathologic grading of medulloblastomas. *Cancer* **94**, 552–560 (2002).
65. Korshunov, A. *et al.* Molecular subgroups of medulloblastoma: the current consensus. *Acta Neuropathol* **123**, 465–472 (2012).
66. Smoll, N. R. & Drummond, K. J. The incidence of medulloblastomas and primitive neuroectodermal tumours in adults and children. *Journal of Clinical Neuroscience* **19**, 1541–1544 (2012).
67. Ellison, D. W. *et al.* Medulloblastoma: clinicopathological correlates of SHH, WNT, and non-SHH/WNT molecular subgroups. *Acta Neuropathol* **121**, 381–396 (2011).

68. Taylor, M. D. *et al.* Molecular subgroups of medulloblastoma: the current consensus. *Acta Neuropathol* **123**, 465–472 (2012).
69. Nobre, L. *et al.* Pattern of Relapse and Treatment Response in WNT-Activated Medulloblastoma. *Cell Rep Med* **1**, 100038 (2020).
70. Kennedy, C. *et al.* Quality of Survival and Growth in Children and Young Adults in the PNET4 European Controlled Trial of Hyperfractionated Versus Conventional Radiation Therapy for Standard-Risk Medulloblastoma. *International Journal of Radiation Oncology*Biophysics* **88**, 292–300 (2014).
71. Ottensmeier, H. *et al.* Neuropsychological short assessment of disease- and treatment-related intelligence deficits in children with brain tumours. *Eur J Paediatr Neurol* **19**, 298–307 (2015).
72. Sabel, M. *et al.* Relapse patterns and outcome after relapse in standard risk medulloblastoma: a report from the HIT-SIOP-PNET4 study. *J Neurooncol* **129**, 515–524 (2016).
73. Yoon, J. W., Gilbertson, R., Iannaccone, S., Iannaccone, P. & Walterhouse, D. Defining a role for Sonic hedgehog pathway activation in desmoplastic medulloblastoma by identifying GLI1 target genes. *Int J Cancer* **124**, 109–119 (2009).
74. Zurawel, R. H. *et al.* Analysis of PTCH/SMO/SHH pathway genes in medulloblastoma. *Genes Chromosomes Cancer* **27**, 44–51 (2000).
75. Kool, M. *et al.* Molecular subgroups of medulloblastoma: an international meta-analysis of transcriptome, genetic aberrations, and clinical data of WNT, SHH, Group 3, and Group 4 medulloblastomas. *Acta Neuropathol* **123**, 473–484 (2012).
76. Zhukova, N. *et al.* Subgroup-Specific Prognostic Implications of TP53 Mutation in Medulloblastoma. *Journal of Clinical Oncology* **31**, 2927–2935 (2013).
77. Archer, T. C. *et al.* Proteomics, Post-translational Modifications, and Integrative Analyses Reveal Molecular Heterogeneity within Medulloblastoma Subgroups. *Cancer Cell* **34**, 396–410.e8 (2018).
78. Forget, A. *et al.* Aberrant ERBB4-SRC Signaling as a Hallmark of Group 4 Medulloblastoma Revealed by Integrative Phosphoproteomic Profiling. *Cancer Cell* **34**, 379–395.e7 (2018).
79. Menyhárt, O., Giangaspero, F. & Győrffy, B. Molecular markers and potential therapeutic targets in non-WNT/non-SHH (group 3 and group 4) medulloblastomas. *J Hematol Oncol* **12**, 29 (2019).
80. Sharma, T. *et al.* Second-generation molecular subgrouping of medulloblastoma: an international meta-analysis of Group 3 and Group 4 subtypes. *Acta Neuropathol* **138**, 309–326 (2019).
81. Northcott, P. A. *et al.* The whole-genome landscape of medulloblastoma subtypes. *Nature* **547**, 311–317 (2017).
82. Schoof, M. *et al.* Malignant gliomas with H3F3A G34R mutation or MYCN amplification in pediatric patients with Li Fraumeni syndrome. *Acta Neuropathol* **142**, 591–593 (2021).
83. Sturm, D. *et al.* Multiomic neuropathology improves diagnostic accuracy in pediatric neuro-oncology. *Nat Med* **29**, 917–926 (2023).
84. Korshunov, A. *et al.* H3-/IDH-wild type pediatric glioblastoma is comprised of molecularly and prognostically distinct subtypes with associated oncogenic drivers. *Acta Neuropathol* **134**, 507–516 (2017).
85. Voß, H. *et al.* Multiomic profiling of medulloblastoma reveals subtype-specific targetable alterations at the proteome and N-glycan level. *bioRxiv* 2023.01.09.523234 (2023) doi:10.1101/2023.01.09.523234.
86. Göbel, C. *et al.* MYC overexpression and SMARCA4 loss cooperate to drive medulloblastoma formation in mice. *Acta Neuropathol Commun* **11**, 174 (2023).

87. Schoof, M. *et al.* Mouse models of pediatric high-grade gliomas with MYCN amplification reveal intratumoral heterogeneity and lineage signatures. *Nat Commun* **14**, 7717 (2023).
88. Sturm, D. *et al.* New Brain Tumor Entities Emerge from Molecular Classification of CNS-PNETs. *Cell* **164**, 1060–1072 (2016).
89. Waszak, S. M. *et al.* Germline Elongator mutations in Sonic Hedgehog medulloblastoma. *Nature* **580**, 396–401 (2020).
90. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* **11**, 2301–2319 (2016).
91. Tyanova, S. & Cox, J. Perseus: A bioinformatics platform for integrative analysis of proteomics data in cancer research. *Methods in Molecular Biology* **1711**, 133–148 (2018).
92. Zhuo, L. *et al.* hGFAP-cre transgenic mice for manipulation of glial and neuronal function in vivo. *genesis* **31**, 85–94 (2001).
93. Hegedus, B. *et al.* Neurofibromatosis-1 Regulates Neuronal and Glial Cell Differentiation from Neuroglial Progenitors In Vivo by Both cAMP- and Ras-Dependent Mechanisms. *Cell Stem Cell* **1**, 443–457 (2007).
94. Hayashi, S., Lewis, P., Pevny, L. & McMahon, A. P. Efficient gene modulation in mouse epiblast using a Sox2Cre transgenic mouse strain. *Mech Dev* **119**, S97–S101 (2002).
95. Jonkers, J. *et al.* Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer. *Nat Genet* **29**, 418–425 (2001).
96. Althoff, K. *et al.* A Cre-conditional MYCN-driven neuroblastoma mouse model as an improved tool for preclinical studies. *Oncogene* **34**, 3357–3368 (2015).
97. Schüller, U. *et al.* Acquisition of Granule Neuron Precursor Identity Is a Critical Determinant of Progenitor Cell Competence to Form Shh-Induced Medulloblastoma. *Cancer Cell* **14**, 123–134 (2008).
98. Fortin, J.-P., Triche Jr, T. J. & Hansen, K. D. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. *Bioinformatics* **33**, 558–560 (2017).
99. Becht, E. *et al.* Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol* **37**, 38 (2019).
100. Afgan, E. *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Res* **50**, W345–W351 (2022).
101. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
102. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
103. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
104. Wilkerson, M. D. & Hayes, D. N. ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* **26**, 1572–1573 (2010).
105. Franch-Expósito, S. *et al.* CNApp, a tool for the quantification of copy number alterations and integrative analysis revealing clinical implications. *Elife* **9**, (2020).
106. Juric, V. & Murphy, B. Cyclin-dependent kinase inhibitors in brain cancer: current state and future directions. *Cancer Drug Resist* **3**, 48–62 (2020).
107. Ding, L. *et al.* The Roles of Cyclin-Dependent Kinases in Cell-Cycle Progression and Therapeutic Strategies in Human Breast Cancer. *Int J Mol Sci* **21**, (2020).
108. Brown, P. D. *et al.* Memantine for the prevention of cognitive dysfunction in patients receiving whole-brain radiotherapy: a randomized, double-blind, placebo-controlled trial. *Neuro Oncol* **15**, 1429–1437 (2013).

109. Phoenix, T. N. *et al.* Medulloblastoma Genotype Dictates Blood Brain Barrier Phenotype. *Cancer Cell* **29**, 508–522 (2016).
110. Fujimoto, M. *et al.* Deficiency of tenascin-C and attenuation of blood-brain barrier disruption following experimental subarachnoid hemorrhage in mice. *J Neurosurg* **124**, 1693–1702 (2016).
111. Bai, R.-Y., Staedtke, V., Rudin, C. M., Bunz, F. & Riggins, G. J. Effective treatment of diverse medulloblastoma models with mebendazole and its impact on tumor angiogenesis. *Neuro Oncol* **17**, 545–554 (2015).
112. Boudiaf-Benmammar, C., Cresteil, T. & Melki, R. The Cytosolic Chaperonin CCT/TRiC and Cancer Cell Proliferation. *PLoS One* **8**, (2013).
113. Zhang, Y. *et al.* Molecular chaperone CCT3 supports proper mitotic progression and cell proliferation in hepatocellular carcinoma cells. *Cancer Lett* **372**, 101–109 (2016).
114. Amon, R., Reuven, E. M., Leviatan Ben-Arye, S. & Padler-Karavani, V. Glycans in immune recognition and response. *Carbohydr Res* **389**, 115–122 (2014).
115. Kailemia, M. J., Park, D. & Lebrilla, C. B. Glycans and glycoproteins as specific biomarkers for cancer. *Anal Bioanal Chem* **409**, 395–410 (2017).
116. Thurin, M. Tumor-Associated Glycans as Targets for Immunotherapy: The Wistar Institute Experience/Legacy. *Monoclon Antib Immunodiagn Immunother* **40**, 89–100 (2021).
117. Smith, K. S. *et al.* Unified rhombic lip origins of group 3 and group 4 medulloblastoma. *Nature* **609**, 1012–1020 (2022).
118. Schüller, U. *et al.* Acquisition of Granule Neuron Precursor Identity Is a Critical Determinant of Progenitor Cell Competence to Form Shh-Induced Medulloblastoma. *Cancer Cell* **14**, 123–134 (2008).
119. Ayrault, O. *et al.* Atoh1 Inhibits Neuronal Differentiation and Collaborates with Gli1 to Generate Medulloblastoma-Initiating Cells. *Cancer Res* **70**, 5618–5627 (2010).
120. Tao, R. *et al.* MYC Drives Group 3 Medulloblastoma through Transformation of Sox2+ Astrocyte Progenitor Cells. *Cancer Res* **79**, 1967–1980 (2019).
121. Mainwaring, O. J. *et al.* ARF suppression by MYC but not MYCN confers increased malignancy of aggressive pediatric brain tumors. *Nat Commun* **14**, 1221 (2023).
122. Pan, Y. *et al.* NF1 mutation drives neuronal activity-dependent initiation of optic glioma. *Nature* **594**, 277–282 (2021).
123. Chen, P. *et al.* Olfactory sensory experience regulates gliomagenesis via neuronal IGF1. *Nature* **606**, 550–556 (2022).
124. Taylor, M. D. *et al.* Molecular subgroups of medulloblastoma: the current consensus. *Acta Neuropathol* **123**, 465–472 (2012).
125. Schwalbe, E. C. *et al.* Novel molecular subgroups for clinical classification and outcome prediction in childhood medulloblastoma: a cohort study. *Lancet Oncol* **18**, 958–971 (2017).
126. Hovestadt, V. *et al.* Medulloblastomics revisited: biological and clinical insights from thousands of patients. *Nat Rev Cancer* **20**, 42–56 (2022).
127. Northcott, P. A. *et al.* The whole-genome landscape of medulloblastoma subtypes. *Nature* **547**, 311–317 (2017).
128. Louis, D. N. *et al.* The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathologica* vol. 131 803–820 Preprint at <https://doi.org/10.1007/s00401-016-1545-1> (2016).
129. Delaidelli, A. *et al.* Clinically Tractable Outcome Prediction of Non-WNT/Non-SHH Medulloblastoma Based on TPD52 IHC in a Multicohort Study. *Clinical Cancer Research* **28**, 116–128 (2022).
130. Ramaswamy, V. *et al.* Risk stratification of childhood medulloblastoma in the molecular era: the current consensus. *Acta Neuropathol* **131**, 821–831 (2016).

131. Goschzik, T. *et al.* Molecular stratification of medulloblastoma: comparison of histological and genetic methods to detect Wnt activated tumours. *Neuropathol Appl Neurobiol* **41**, 135–144 (2015).
132. Ellison, D. W. *et al.* β -Catenin Status Predicts a Favorable Outcome in Childhood Medulloblastoma: The United Kingdom Children’s Cancer Study Group Brain Tumour Committee. *Journal of Clinical Oncology* **23**, 7951–7957 (2005).
133. Chen, Q., Tan, Z., Guan, F. & Ren, Y. The Essential Functions and Detection of Bisecting GlcNAc in Cell Biology. *Front Chem* **8**, 511 (2020).
134. Vreeker, G. C. M. *et al.* Serum N-glycan profiles differ for various breast cancer subtypes. *Glycoconj J* **38**, 387–395 (2021).
135. Rodríguez, E., Schettters, S. T. T. & van Kooyk, Y. The tumour glyco-code as a novel immune checkpoint for immunotherapy. *Nat Rev Immunol* **18**, 204–211 (2018).
136. Dotz, V. & Wuhrer, M. N-glycome signatures in human plasma: associations with physiology and major diseases. *FEBS Lett* **593**, 2966–2976 (2019).
137. Capper, D. *et al.* DNA methylation-based classification of central nervous system tumours. *Nature* **555**, 469–474 (2018).
138. Lee, H. G. *et al.* State-of-the-art housekeeping proteins for quantitative western blotting: Revisiting the first draft of the human proteome. *Proteomics* **16**, 1863–1867 (2016).
139. Voß, H. *et al.* HarmonizR enables data harmonization across independent proteomic datasets with appropriate handling of missing values. *Nat Commun* **13**, 3523 (2022).
140. Schwalbe, E. C. *et al.* Novel molecular subgroups for clinical classification and outcome prediction in childhood medulloblastoma: a cohort study. *Lancet Oncol* **18**, 958–971 (2017).
141. Menyhárt, O. & Györffy, B. Principles of tumorigenesis and emerging molecular drivers of SHH-activated medulloblastomas. *Ann Clin Transl Neurol* **6**, 990–1005 (2019).
142. Krämer, A., Green, J., Pollard, J. J. & Tugendreich, S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* **30**, 523–530 (2014).
143. Kool, M. *et al.* Genome sequencing of SHH medulloblastoma predicts genotype-related response to smoothened inhibition. *Cancer Cell* **25**, 393–405 (2014).
144. Ramaswamy, V., Nör, C. & Taylor, M. D. P53 and Medulloblastoma. *Cold Spring Harb Perspect Med* **6**, 1–9 (2016).
145. Bailey, S. *et al.* Clinical Trials in High-Risk Medulloblastoma: Evolution of the SIOP-Europe HR-MB Trial. *Cancers (Basel)* **14**, (2022).
146. Northcott, P. A. *et al.* Medulloblastoma comprises four distinct molecular variants. *Journal of Clinical Oncology* **29**, 1408–1414 (2011).
147. Menyhárt, O., Giangaspero, F. & Györffy, B. Molecular markers and potential therapeutic targets in non-WNT/non-SHH (group 3 and group 4) medulloblastomas. *J Hematol Oncol* **12**, (2019).
148. Sun, R., Kim, A. M. J. & Lim, S.-O. Glycosylation of Immune Receptors in Cancer. *Cells* **10**, (2021).
149. Klein, J. A., Meng, L. & Zaia, J. Deep Sequencing of Complex Proteoglycans: A Novel Strategy for High Coverage and Site-specific Identification of Glycosaminoglycan-linked Peptides. *Mol Cell Proteomics* **17**, 1578–1590 (2018).
150. Munkley, J. & Scott, E. Targeting Aberrant Sialylation to Treat Cancer. *Medicines (Basel)* **6**, (2019).
151. Sprung, R. W. *et al.* Equivalence of Protein Inventories Obtained from Formalin-fixed Paraffin-embedded and Frozen Tissue in Multidimensional Liquid Chromatography-Tandem Mass Spectrometry Shotgun Proteomic Analysis. *Molecular & Cellular Proteomics* **8**, 1988–1998 (2009).

152. Cook Sangar, M. L. *et al.* Inhibition of CDK4/6 by Palbociclib Significantly Extends Survival in Medulloblastoma Patient-Derived Xenograft Mouse Models. *Clinical Cancer Research* **23**, 5802–5813 (2017).
153. Vasaikar, S. *et al.* Proteogenomic Analysis of Human Colon Cancer Reveals New Therapeutic Opportunities. *Cell* **177**, 1035–1049.e19 (2019).
154. Mertins, P. *et al.* Reproducible workflow for multiplexed deep-scale proteome and phosphoproteome analysis of tumor tissues by liquid chromatography-mass spectrometry. *Nat Protoc* **13**, 1632–1661 (2018).
155. Hovestadt, V. *et al.* Decoding the regulatory landscape of medulloblastoma using DNA methylation sequencing. *Nature* **510**, 537–541 (2014).
156. Richard W Olsen and Timothy M DeLorey. No Title. in *Basic Neurochemistry: Molecular, Cellular and Medical Aspects. 6th edition.* (1999).
157. Venkatesh, H. S. *et al.* Electrical and synaptic integration of glioma into neural circuits. *Nature* **573**, 539–545 (2019).
158. Venkataramani, V. *et al.* Glioblastoma hijacks neuronal mechanisms for brain invasion. *Cell* **185**, 2899–2917.e31 (2022).
159. Akar, S., Harmankaya, İ., Uğraş, S. & Çelik, Ç. Nicotinamide N-methyltransferase expression and its association with phospho-Akt, p53 expression, and survival in high-grade endometrial cancer. *Turk J Med Sci* **49**, 1547–1554 (2019).
160. Moore, S. *et al.* The CHD6 chromatin remodeler is an oxidative DNA damage response factor. *Nat Commun* **10**, 241 (2019).
161. Alexiou, G. A. *et al.* Expression of heat shock proteins in medulloblastoma: Laboratory investigation. *Journal of Neurosurgery: Pediatrics PED* **12**, 452–457 (2013).
162. Nobre, L. *et al.* Pattern of Relapse and Treatment Response in WNT-Activated Medulloblastoma. *Cell Rep Med* **1**, (2020).
163. Meredith, D. M. & Alexandrescu, S. Embryonal and non-meningothelial mesenchymal tumors of the central nervous system - Advances in diagnosis and prognostication. *Brain Pathol* **32**, e13059 (2022).
164. D’Arcy, C. E. *et al.* Immunohistochemical and nanoString-Based Subgrouping of Clinical Medulloblastoma Samples. *J Neuropathol Exp Neurol* **79**, 437–447 (2020).
165. Northcott, P. A. *et al.* Rapid, reliable, and reproducible molecular sub-grouping of clinical medulloblastoma samples. *Acta Neuropathol* **123**, 615–626 (2012).
166. Yoshida, T., Akatsuka, T. & Imanaka-Yoshida, K. Tenascin-C and integrins in cancer. *Cell Adh Migr* **9**, 96–104 (2015).
167. Linke, F. *et al.* 3D hydrogels reveal medulloblastoma subgroup differences and identify extracellular matrix subtypes that predict patient outcome. *J Pathol* **253**, 326–338 (2021).
168. Phoenix, T. N. *et al.* Medulloblastoma Genotype Dictates Blood Brain Barrier Phenotype. *Cancer Cell* **29**, 508–522 (2016).
169. Shalabi, H., Nellan, A., Shah, N. N. & Gust, J. Immunotherapy Associated Neurotoxicity in Pediatric Oncology. *Front Oncol* **12**, (2022).
170. Schwalbe, E. C. *et al.* Rapid diagnosis of medulloblastoma molecular subgroups. *Clinical Cancer Research* **17**, 1883–1894 (2011).
171. Wang, J. *et al.* Effective inhibition of MYC-amplified group 3 medulloblastoma by FACT-targeted curaxin drug CBL0137. *Cell Death Dis* **11**, 1029 (2020).
172. Patil, S. *et al.* Combination of clotam and vincristine enhances anti-proliferative effect in medulloblastoma cells. *Gene* **705**, 67–76 (2019).
173. Sengupta, S., Pomeranz Krummel, D. & Pomeroy, S. The evolution of medulloblastoma therapy to personalized medicine. *F1000Res* **6**, 490 (2017).

174. Bassiouni, R. *et al.* Chaperonin Containing TCP-1 Protein Level in Breast Cancer Cells Predicts Therapeutic Application of a Cytotoxic Peptide. *Clin Cancer Res* **22**, 4366–4379 (2016).
175. Carr, A. C. *et al.* Targeting chaperonin containing TCP1 (CCT) as a molecular therapeutic for small cell lung cancer. *Oncotarget* **8**, 110273–110288 (2017).
176. Ohtsubo, K. & Marth, J. D. Glycosylation in Cellular Mechanisms of Health and Disease. *Cell* **126**, 855–867 (2006).
177. Greco, B. *et al.* Disrupting N-glycan expression on tumor cells boosts chimeric antigen receptor T cell efficacy against solid malignancies. *Sci Transl Med* **14**, eabg3072 (2022).
178. Marada, S. *et al.* Functional Divergence in the Role of N-Linked Glycosylation in Smoothed Signaling. *PLoS Genet* **11**, e1005473 (2015).
179. Williams, S. E. *et al.* Mammalian brain glycoproteins exhibit diminished glycan complexity compared to other tissues. *Nat Commun* **13**, 275 (2022).
180. Rodrigues, E. & Macauley, M. S. Hypersialylation in Cancer: Modulation of Inflammation and Therapeutic Opportunities. *Cancers (Basel)* **10**, (2018).
181. Northcott, P. A. *et al.* Medulloblastoma comprises four distinct molecular variants. *Journal of Clinical Oncology* **29**, 1408–1414 (2011).
182. Mcalister, G. C. *et al.* MultiNotch MS3 Enables Accurate, Sensitive, and Multiplexed Detection of Differential Expression across Cancer Cell Line Proteomes. *Anal Chem* **86**, (2014).
183. Hughes, C. S. *et al.* Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc* **14**, 68–85 (2019).
184. van Pijkeren, A. *et al.* Proteome Coverage after Simultaneous Proteo-Metabolome Liquid-Liquid Extraction. *J Proteome Res* **22**, 951–966 (2023).
185. Yurekten, O. *et al.* MetaboLights: open data repository for metabolomics. *Nucleic Acids Res* **52**, D640–D646 (2024).
186. Guan, Y., Zhang, M., Wang, J. & Schlüter, H. Comparative Analysis of Different N-glycan Preparation Approaches and Development of Optimized Solid-Phase Permethylation Using Mass Spectrometry. *J Proteome Res* **20**, 2914–2922 (2021).
187. Watanabe, Y., Aoki-Kinoshita, K. F., Ishihama, Y. & Okuda, S. GlycoPOST realizes FAIR principles for glycomics mass spectrometry data. *Nucleic Acids Res* **49**, D1523–D1528 (2021).
188. Aryee, M. J. *et al.* Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363–1369 (2014).
189. Ritchie, M. E. *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).
190. Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S. & Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods* **17**, 41–44 (2020).
191. Guan, Y. *et al.* An Integrated Strategy Reveals Complex Glycosylation of Erythropoietin Using Mass Spectrometry. *J Proteome Res* **20**, 3654–3663 (2021).
192. Community, T. G. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Res* **50**, W345–W351 (2022).
193. Triche, T. J. J., Weisenberger, D. J., Van Den Berg, D., Laird, P. W. & Siegmund, K. D. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. *Nucleic Acids Res* **41**, e90 (2013).
194. Lagnoux, A., Mercier, S. & Vallois, P. Statistical significance based on length and position of the local score in a model of i.i.d. sequences. *Bioinformatics* **33**, 654–660 (2017).

195. Schumann, Y., Neumann, J. E. & Neumann, P. Robust classification using average correlations as features (ACF). *BMC Bioinformatics* **24**, 101 (2023).
196. Tyanova, S. & Cox, J. Perseus: A bioinformatics platform for integrative analysis of proteomics data in cancer research. *Methods in Molecular Biology* **1711**, 133–148 (2018).
197. Jassal, B. *et al.* The reactome pathway knowledgebase. *Nucleic Acids Res* **48**, D498–D503 (2020).
198. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545–15550 (2005).
199. Reimand, J. *et al.* Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat Protoc* **14**, 482–517 (2019).
200. Paul Shannon, 1 *et al.* Cytoscape: A Software Environment for Integrated Models. *Genome Res* **13**, 426 (1971).
201. Reimand, J. *et al.* Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat Protoc* **14**, 482–517 (2019).
202. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676–682 (2012).

Acknowledgements

We thank Tasja Lempertz, Carolina Janko, Ulrike Rumpf, Karin Gehlken, Celina Soltwedel, Ann-Kathleen Leptien, Dr. Patrick Bluemke, and Paula Nissen for skillful and kind support. JN is funded by the Deutsche Forschungsgemeinschaft (DFG, Emmy Noether program), the Hamburger Krebsgesellschaft e.V. and the Erich und Gertrud Roggenbuck-foundation.

Author contributions

S.G, H.V., A.G. and J.N. wrote and reviewed the manuscript. J.N. planned and designed the study. S.G. and H.V. conducted experiments. H.V, S.S., B.P., T.M., H.S., C.K., N.S., B.S. and Y.G. analyzed proteome and glycosylation data. A.G and S.G. generated and analyzed biological and technical validation data. U.S. and S.G analyzed methylation data. S.G. and Y.S. integrated proteome and methylome data. M.D. performed digitally supported quantification of IHC. S.P., S.R., M.M., MM. D., A.K., C.H., J.W., and F.L-S. analyzed and interpreted histological, molecular and clinical data. A.G., M.M, M.K. and M.H. generated and analyzed the metabolomic and amino acid data. All authors reviewed the manuscript and approved its final version.

Declaration of interests

The authors declare no competing interests.

Figures

Figure 1

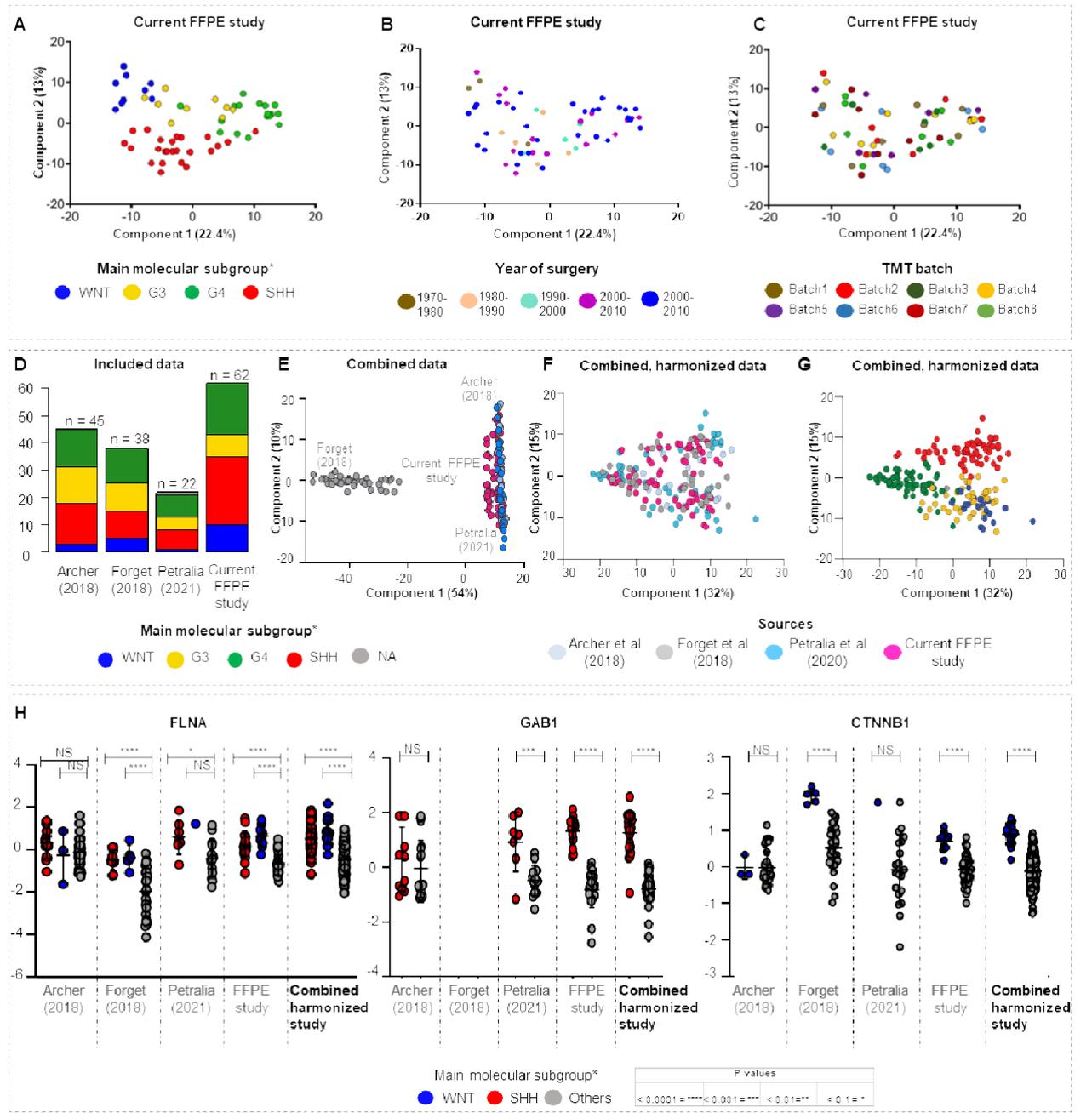


Figure 1: Harmonization and integration of proteome Medulloblastoma (MB) datasets

NIPALS principal component analyses (PCA) of measured FFPE samples (n=62) with assignment to **(A)** the four main molecular MB subgroups¹²⁴, **(B)** age of measured samples, **(C)** measured TMT batch. **(D)** Overview of analysed datasets. PCA of data before **(E)** and after **(F,G)** data harmonization using ComBat in the HarmonizR framework annotated for the source of the samples **(F)** and for main molecular MB

subgroups (G, n=167). **(H) Protein abundance of the the WNT and SHH MB marker FILAMIN A** (nSHH_Archer = 15, nWNT_Archer = 3, nOthers_Archer = 27, nSHH_Forget = 10, nWNT_Forget = 5, nOthers_Forget = 23, nSHH_Petralia = 7, nWNT_Petralia = 1, nOthers_Petralia = 14, nSHH_FFPE = 25, nWNT_FFPE = 10, nOthers_FFPE = 27, nSHH_combined = 57, nWNT_combined = 19, nOthers_combined = 91, two-tailed, unpaired t-test, pshhArchervsOtherArcher = n.s., pWNTArchervsOtherArcher = n.s., pshhForgetvsOtherForget < 0.0001, pWNTForgetvsOtherForget < 0.0001, pshhPetraliavsOtherPetralia = 0.02, pwntPetraliavsOtherPetralia = n.s., pshhFFPEvsOtherFFPE < 0.0001, pwntFFPEvsOtherFFPE < 0.0001, pshhcombinedvsOthercombined < 0.0001, pshhcombinedvsOthercombined < 0.0001) **SHH MB marker GAB1** (nSHH_Archer = 15, nOthers_Archer = 30, nSHH_Forget = 10, nOthers_Forget = 28, nSHH_Petralia = 7, nOthers_Petralia = 15, nSHH_FFPE = 25, nOthers_FFPE = 37, nSHH_combined = 57, nOthers_combined = 110, two-tailed, unpaired t-test, pshhArchervsOtherArcher = n.s., pshhPetraliavsOtherPetralia = 0.008, pshhFFPEOtherFFPE < 0.0001, pshhcombinedOthercombined < 0.0001), and the **WNT MB marker CTNNB1** (nWNT_Archer = 3, nOthers_Archer = 42, nWNT_Forget = 5, nOthers_Forget = 33, nWNT_Petralia = 1, nOthers_Petralia = 21, nWNT_FFPE = 10, nOthers_FFPE = 52, nWNT_combined = 19, nOthers_combined = 148, two-tailed, unpaired t-test, pwntArchervsOtherArcher = n.s., pwntForgetvsOtherForget < 0.0001, pwntPetraliavsOtherPetralia = n.s., pwntFFPEOtherFFPE < 0.0001, pwntcombinedOthercombined < 0.0001). Data are presented as mean values +/- SD in each dataset individually and in the joint dataset after harmonization PCAs are based on ≥70% valid values, *: p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, n.d.= not detected, n.s.= not significant, n represents biologically independent human samples

Figure 2

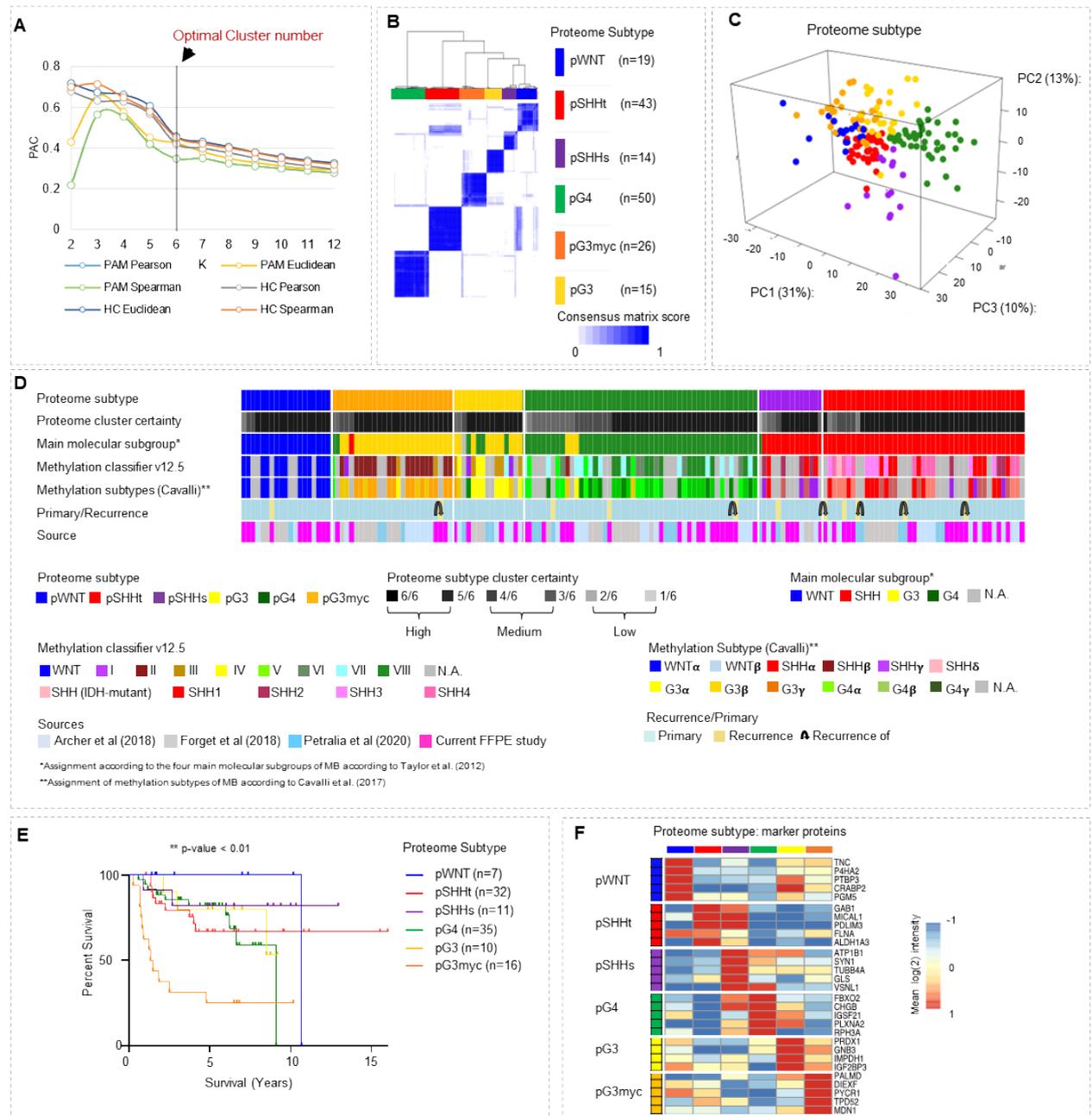


Figure 2: MB segregate into six proteome subtypes

(A) Proportion of ambiguous clustering (PAC) scores for k=2-12 in consensus clustering, using different cluster algorithms ($n_{MB} = 167$, based on $\geq 30\%$ valid values). **(B)** Optimal clustering of proteome data. Consensus scores shown in color scale from white (samples never cluster together) to blue (samples always cluster together). Six proteome subtypes, pWNT, pSHH-t, pSHH-s, pGroup3myc, pGroup3 and pGroup4, were defined. **(C)** Visualization of the first three principal components. **(D)** Clinical sample information. **(E)** Log-rank (Mantel-Cox) test comparing the survival curves of proteome subtypes (p value < 0.001, overall χ^2 -square test). **(F)** Group specific mean log₂ protein intensity of protein subtype marker candidate proteins. n represents biologically independent human samples.

Figure 3

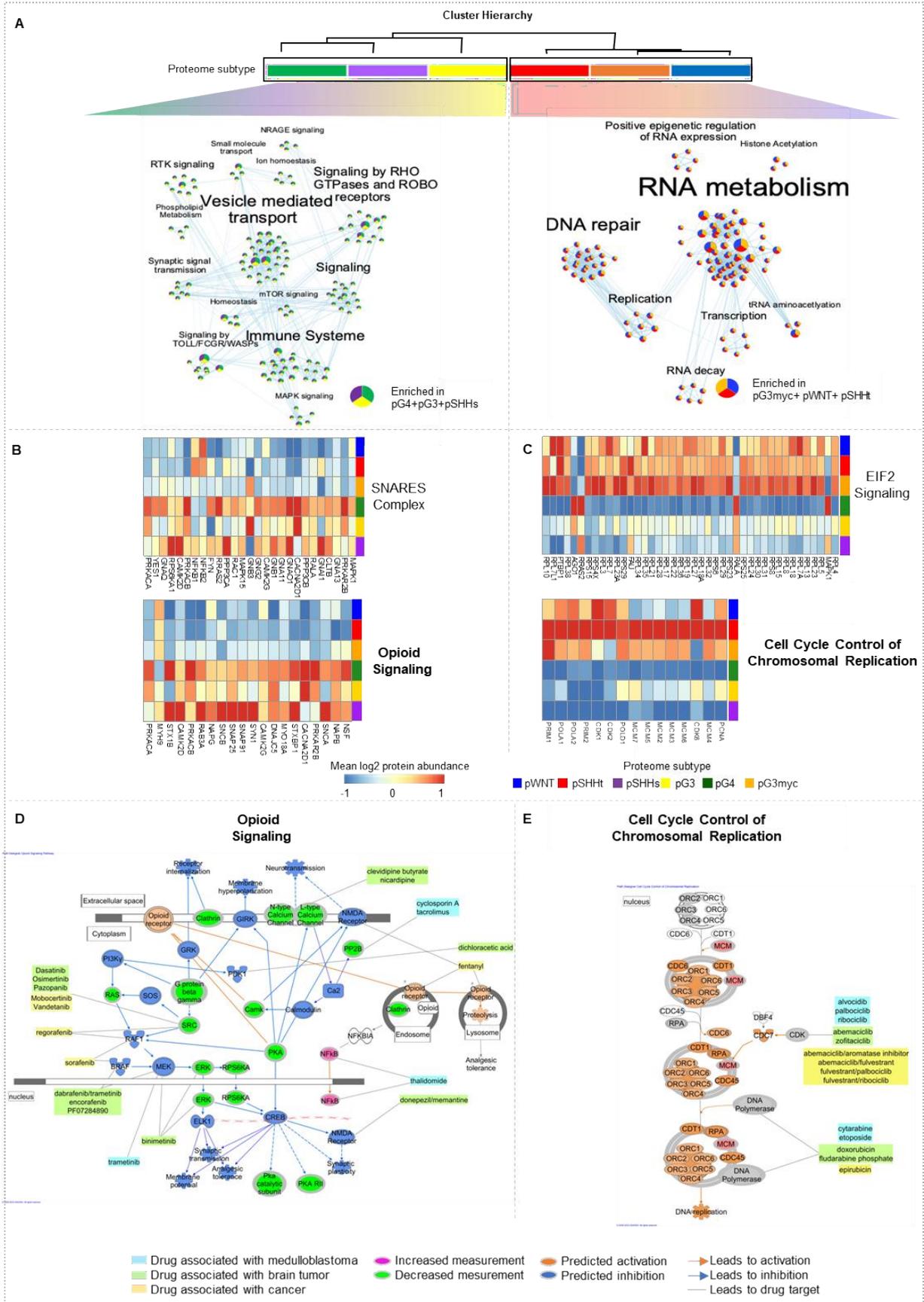


Figure 3: Proteome subtypes of MB can be assigned to two main profiles

(A) Proteome cluster similarity hierarchy based on stepwise increasing k-means execution from k=2-6 with network analyses showing gene set overlap dependent MCL clustering of enriched gene sets, comparing pG3, pG4 and pSHHs (n=79, profile 1), to pWNT, pG3myc, pSHHt (n=88, profile 2). Gene set enrichment analysis (GSEA) was based on REACTOME pathways for all analysis. **(B,C)** Top two upregulated genesets based on differentially abundant proteins using Ingenuity Pathway Analyses (IPA) in profile 1 (opioid signaling and SNARE complex (two-tailed, unpaired t-test, log₂ FC > 1.5 and p-value < 0.05)) **(B)** and profile 2 (EIF2 signaling and cell cycle control of chromosomal replication (log₂ FC > 1.5 and p-value < 0.05) **(C)**),. (D,E) IPA based pathway analyses of opioid signalling (D) and cell cycle control of chromosomal replication E) indicating therapeutic targets with respective drugs. n represents biologically independent human samples.

Figure 4

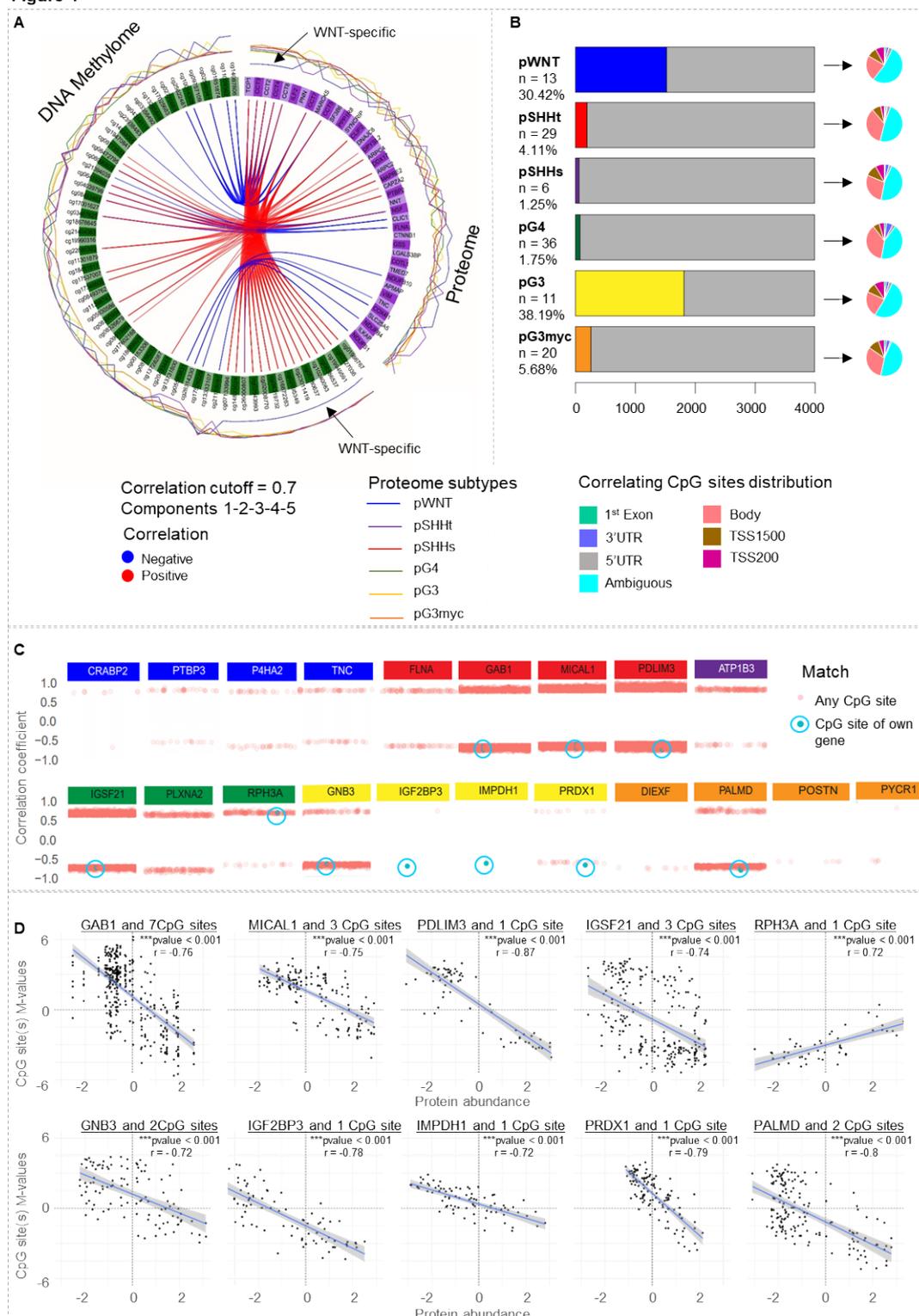


Figure 4: Correlation between DNA methylome and proteome features

(A) Circular plot from mixOmics analyses based on selected features of the first five components from proteome and methylome data. The plot illustrates features with correlation $r > 0.7$ represented on side quadrants. Proteome group specific feature levels are shown in the outer circle. **(B)** Proteome subtype specific Pearson correlation calculated between matched proteins and CpG methylation sites. The number of

proteins correlating with CpG site methylation of their own gene ($r > 0.7$) is shown in colour. The pie chart shows the distribution of correlating CpG sites concerning the position in a gene. **(C)** Subtype independent Pearson correlation between 3,990 proteins and 381,717 methylation probes focusing on subtype specific biomarkers. Pearson Correlations > 0.7 are shown, CpG sites correlating with the corresponding gene are highlighted in blue. Some biomarkers correlated with more than one CpG site of their own gene (GAB1: 7, GNB3: 2, IGSF21: 3, MICAL1: 3, and PALMD: 2). **(D)** Scatterplot of the 10 biomarker proteins correlating with the CpG site(s) of their own gene (pearson correlation > 0.7 , $p < 0.001$). The linear regression line was aligned for all correlating CpG site(s), SE = 0.95.

Figure 5

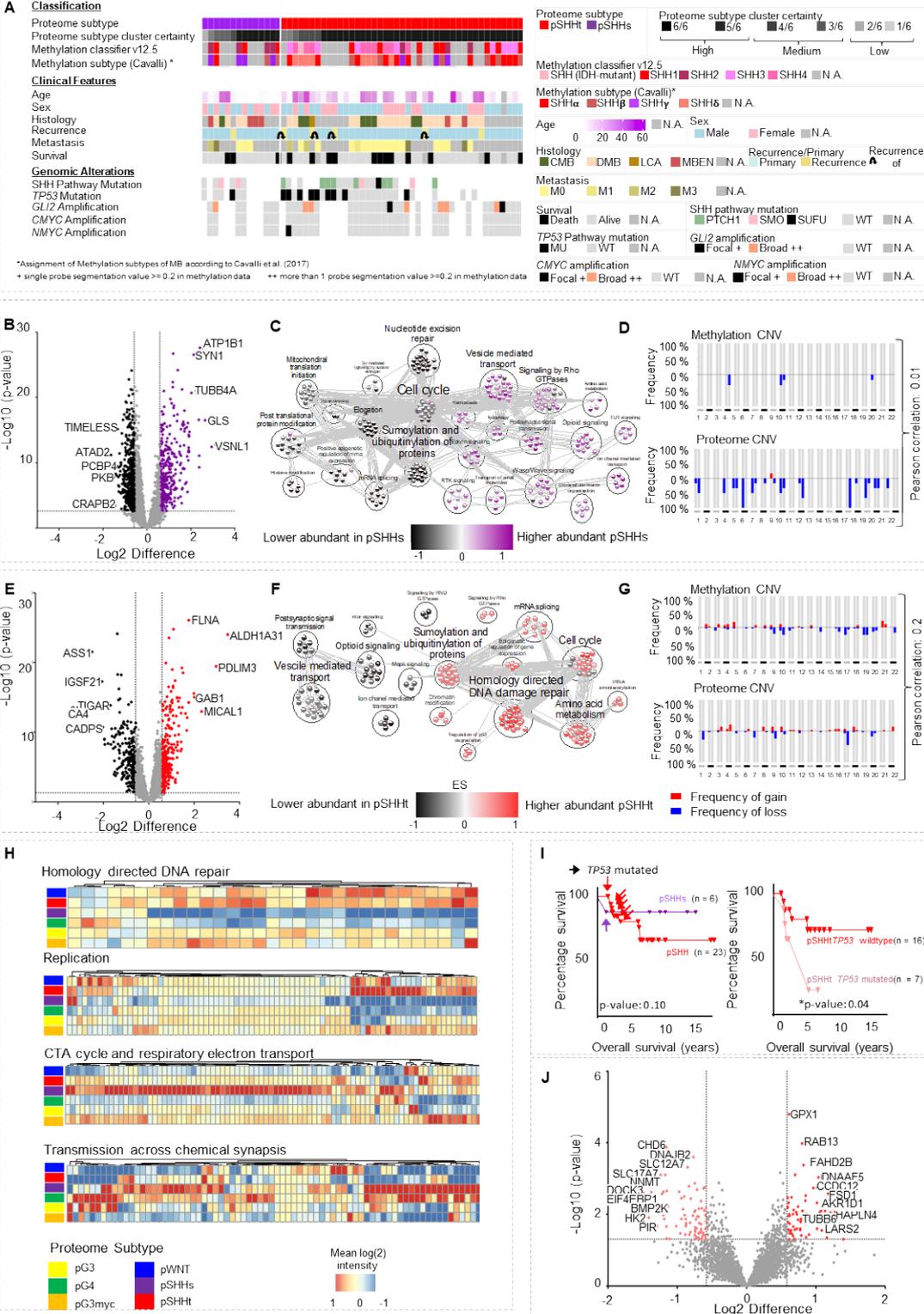


Figure 5: SHH MB comprise two proteome MB subtypes

(A) Histological, molecular, and clinical characteristics of the MB subtypes pSHHt (n=43) and pSHHs (n=14). **(B)** Volcano plot showing differentially abundant proteins comparing pSHHs tumors to all other proteome subtypes (two-tailed, unpaired t-test, p-value < 0.05; log2FC > 1.5). **(C)** MCL clustering of enriched gene sets in pSHHs MBs. **(D)** Copy number variations (CNV) plots of matched pSHHs MB (n=6) calculated from

either DNA methylation or proteome data with pearson correlation between both omic types ($r=0.01$). **(E)** Differentially abundant proteins when comparing pSHHt tumors to all other proteome subtypes (two-tailed, unpaired t-test, $p\text{-value}<0.05$; $\log_2\text{FC} > 1.5$). **(F)** MCL clustering of enriched gene sets in pSHHt. **(G)** CNV plots for matched pSHHt MBs ($n= 29$) calculated from either DNA methylation or proteome data with pearson correlation between both omic types ($r=0.2$) **(H)** Heatmaps showing mean MB subtype protein abundance hallmark genesets homology directed repair (GSEA differential expression analysis normalized enrichment score (NES), $\text{NES}_{\text{pSHHt}} = 2.2$, $p = <0.0001$, $\text{FDR} < 0.25$), replication ($\text{NES}_{\text{pSHHt}} = 2.2$, $p = 0.01$), TCA cycle and respiratory electron transport ($\text{NES}_{\text{pSHHs}} = 3.9$, $p = <0.0001$, $\text{FDR} < 0.25$) and transmission across chemical synapses ($\text{NES}_{\text{pSHHs}} = 3.2$, $p = <0.0001$, $\text{FDR} < 0.25$) based on differentially abundant proteins. **(I)** Overall survival of pSHHt MB ($n= 23$) and pSHHs MB ($n=5$) and overall survival of pSHHt MB depended on TP53 mutation status. TP53 mutated cases displayed a significantly worse survival (Mantel cox test $p\text{-value} = 0.04$). **(J)** Volcano plot, showing differentially abundant proteins when comparing TP53 mutated cases to wildtype cases in pSHHt tumors (two-tailed, unpaired t-test, $p\text{-value}<0.05$; $\log_2\text{FC} > 1.5$). n represents biologically independent human samples.

Figure 6

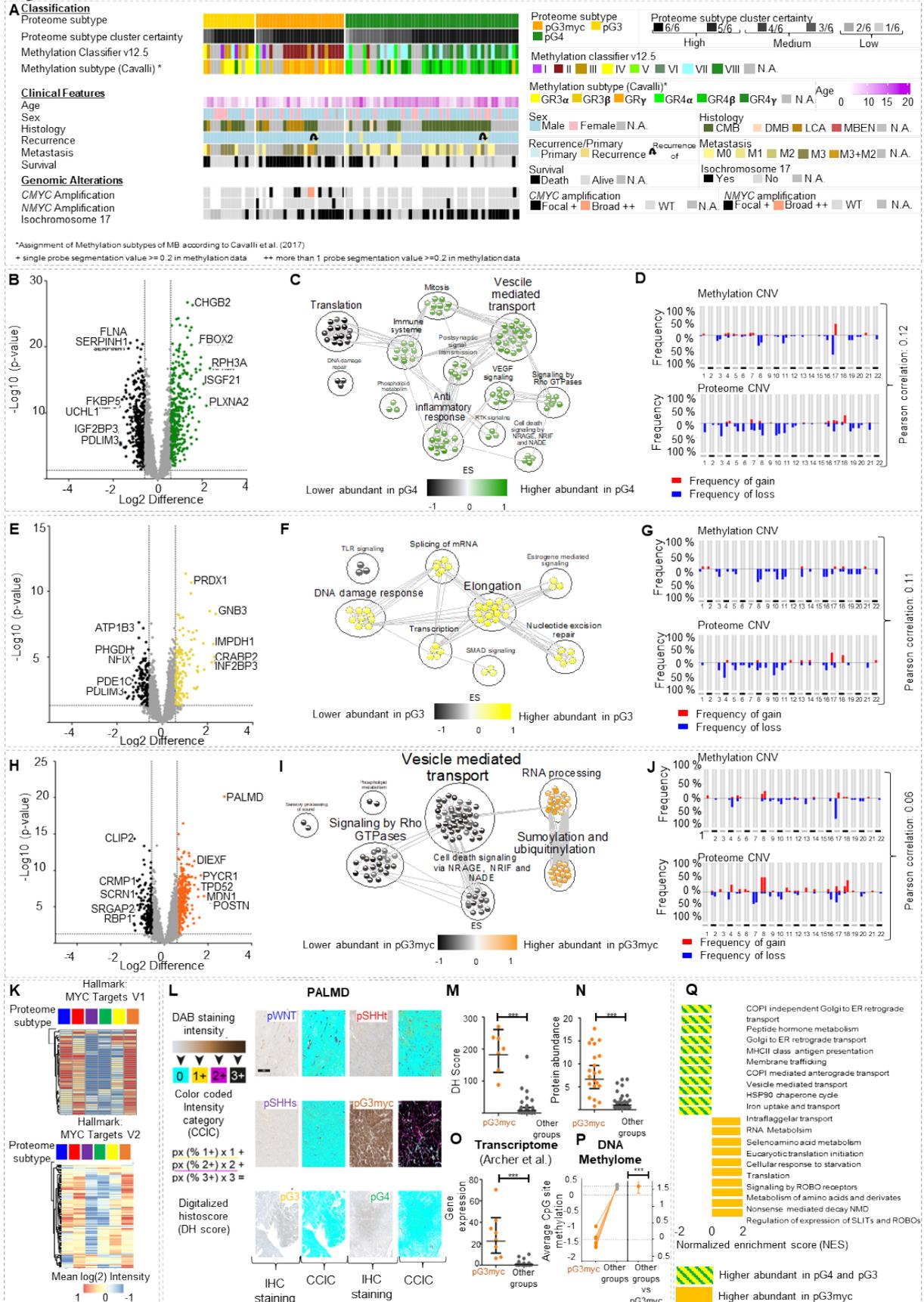


Figure 6: pGroup3-Myc tumors display an enhanced MYC target protein profile and can be identified by Palmdelphin (PALMD) staining

(A) Histological, molecular, and clinical characteristics of the MB subtypes pG3myc (n=26), pG3 (n=15) and pG4 (n=40). **(B)** Volcano plot showing differentially abundant proteins when comparing pG4 tumors to all other proteome subtypes (two-tailed, unpaired t-test, p-value<0.05; log2FC > 1.5). **(C)** MCL clustering of enriched gene sets in pG4 MB. **(D)** CNV plots of pG4 MBs (n=40) calculated from either DNA methylation or proteome data with pearson correlation between both omic types (r=0.12). **(E)** Differentially abundant proteins when comparing pG3 tumors to all other proteome subtypes (two-tailed, unpaired t-test, p-value<0.05; log2FC > 1.5). **(F)** MCL clustering of enriched gene sets in pG3 MB. **(G)** DNA methylation or proteome CNV plots of pG3 MB (n=11) with pearson correlation between both omic types (r=0.11). **(H)** Differentially abundant proteins in pG3myc MB. Palmdelphin (PALMD) was highly abundant in pG3myc tumors (two-tailed, unpaired t-test, p-value<0.05; log2FC > 1.5). **(I)** MCL clustering of enriched gene sets, in pG3myc MB. **(J)** DNA methylation or proteome CNV plots of pG3myc MB (n=20) with pearson correlation between both omic types (r=0.06). **(K)** Mean protein abundance in MB subtypes for hallmark gene sets MYC Targets V1 and MYC Targets V2. **(L)** Scheme and representative images of digitally supported immunostaining intensity quantification of PALMD immunostainings in MB. Quantified pixels of different staining intensities were used to calculate a digital Histo-score (DHS), **(M)** Significantly enhanced digital histoscore for PALMD in pG3myc MB ($n_{pG3myc} = 7$) compared to all other MB subtypes ($n_{Others} = 22$, p<0.0001, data are presented as mean values +/- SD). **(N)** Protein abundance for PALMD in pG3myc MB ($n_{pG3myc} = 21$) compared to all other MB subtypes ($n_{Others} = 84$, unpaired t-test, p<0.0001, data are presented as mean values +/- SD). **(O)** PALMD gene expression in pGroup3myc MBs ($n_{pG3myc} = 6$) compared to all other MB subtypes ($n_{Others} = 30$, two-tailed, unpaired t-test, p<0.0001, data extracted from Archer et al. 2018, data are presented as mean values +/- SD). **(P)** Average DNA methylation at CpG sites of the PALMD gene (Mean M-values of $n_{pG3myc} = 6$ CpG sites shown, two-tailed, unpaired t-test, pvalue < 0.001, data are presented as mean values +/- SD). pGroup3myc MBs show significant lower levels of methylation (two-tailed, unpaired t-test, p<0.0001). **(Q)** GSEA showing the top 10 up or downregulated pathways comparing pG3myc MB to pG3/4 MB (GSEA differential expression analysis normalized enrichment score (NES), p<0.01, FDR<0.25). n represents biologically independent human samples. For immunostaining, each sample was stained once.

Figure 7

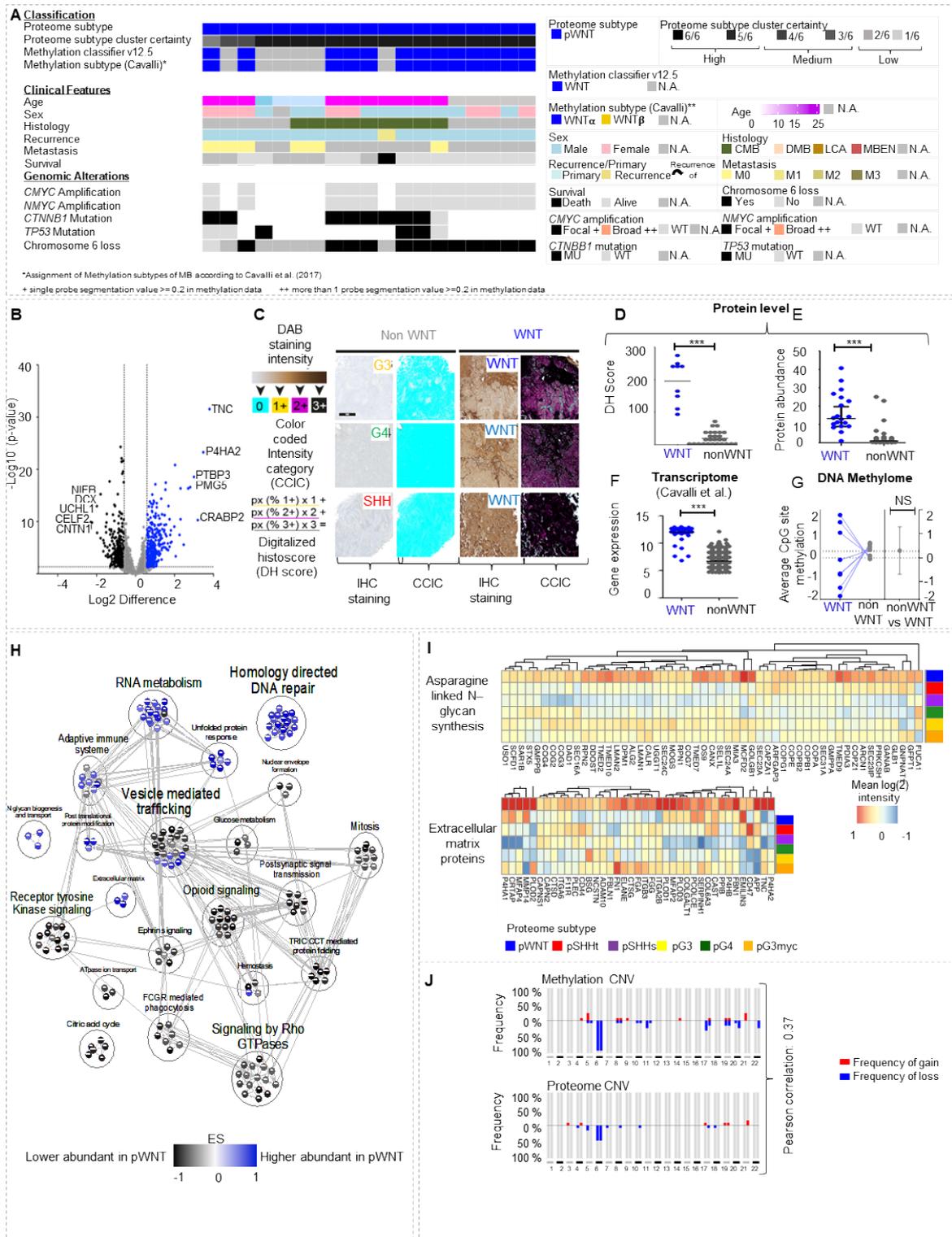


Figure 7: pWNT MB show high feature conservation and can be identified by Tenascin C (TNC) staining.

(A) Histological, molecular, and clinical characteristics of the pWNT MB subtype (n=19). (B) Differentially abundant proteins when comparing pWNT tumors to all other proteome subtypes (two-tailed, unpaired t-test, p-value<0.05; log2FC > 1.5). (C) Scheme and representative images of digital quantification of TNC immunostainings in MB. (D) Significantly enhanced DHS for TNC in pWNT MB (n_{pWNT}=9) compared to

all other MB subtypes ($n_{\text{others}}=28$, two-tailed, unpaired t-test, $p<0.0001$, data are presented as mean values \pm SD). **(E)** Protein abundance for TNC in pWNT MBs ($n_{\text{pWNT}}=19$) compared to all other MB subtypes ($n_{\text{others}}=148$, two-tailed, unpaired t-test, $p < 0.0001$, data are presented as mean values \pm SD). **(F)** *TNC* gene expression in WNT MBs and other MB subtypes in a published dataset of MB ($n_{\text{WNT}} = 70$, $n_{\text{nonWNT}} = 693$, two-tailed, unpaired t-test, $p < 0.001$, data are presented as mean values \pm SD)⁴³. **(G)** Average DNA methylation at CpG sites of the *TNC* gene (mean value for $n_{\text{pWNT}}=8$ CpG sites shown, two-tailed, unpaired t-test, $p = \text{n.s.}$, data presented as mean values \pm SD). **(H)** MCL clustering of eEnriched gene sets, comparing pWNT to all other subtypes in GSEA. **(I)** Heatmaps showing mean protein abundance in MB subtypes for hallmark genesets specifically enriched in pWNT MB (GSEA differential expression analysis normalized enrichment score (NES), $\text{NES}_{\text{Glycan}}=2.2$, $p_{\text{Glycan}}=<0.001$; $\text{NES}_{\text{EMP}}=1.7$, $p_{\text{EMP}}=0.02$). **(J)** CNV plots of pWNT MBs ($n=8$) calculated from either DNA methylation or proteome data with pearson correlation between both omic types ($r=0.37$). n represents biologically independent human samples. For immunostaining, each sample was stained once. $n.s.=$ not significant.

Figure 8

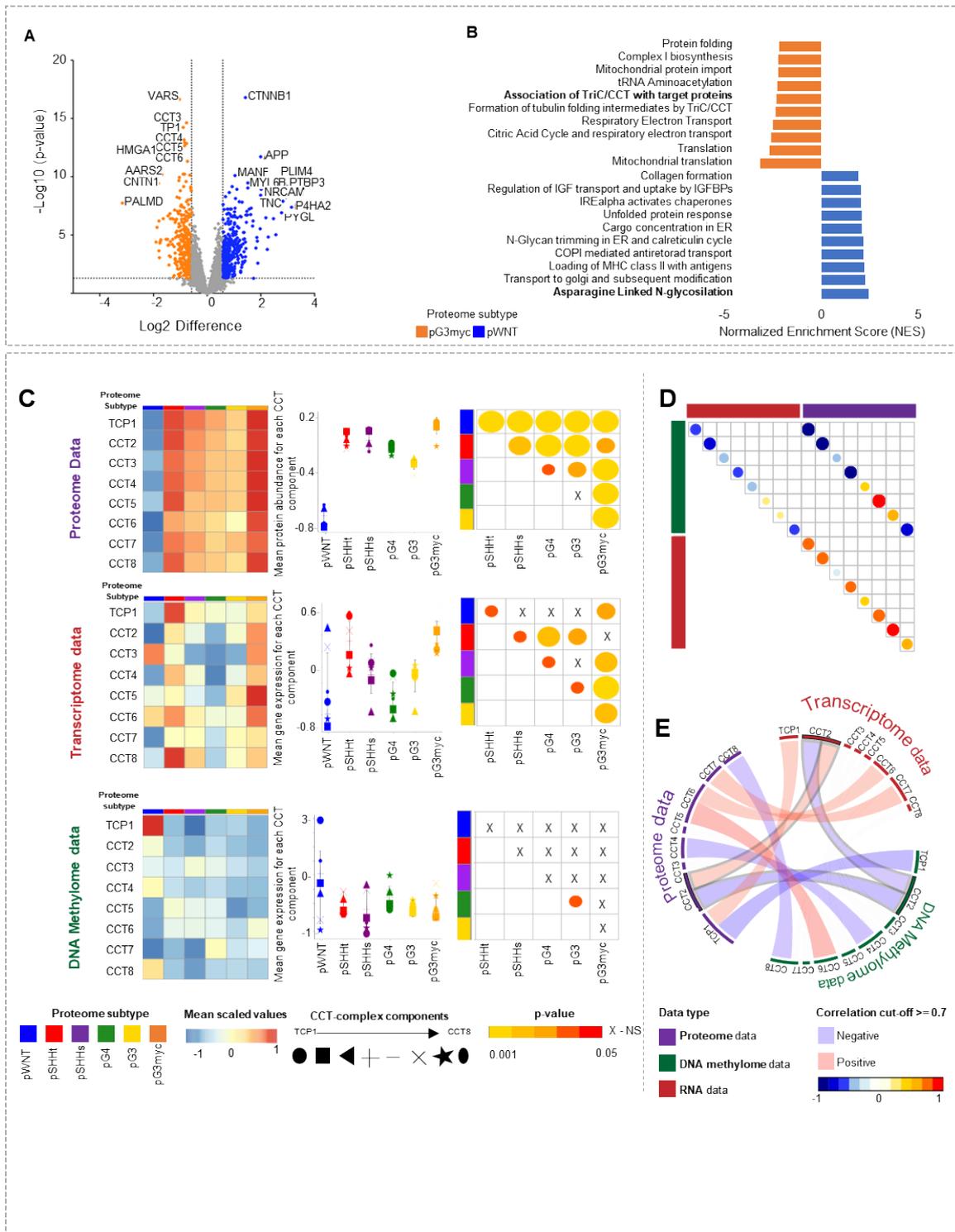


Figure 8: Differential proteomics reveal low abundance of all multiprotein complex TriC/CCT components as a hallmark of pWNT MB

(A) Differentially abundant proteins when comparing pWNT (n=19) to pG3myc (n=26) MB (two-tailed, unpaired t-test, p-value<0.05; log2FC > 1.5). (B) GSEA showing the top 10 up or downregulated pathways comparing pG3myc MB to pWNT (GSEA differential expression analysis normalized enrichment score (NES), p<0.05, FDR<0.25). (C) Mean protein abundances, gene expression values and methylation at CpG sites for all components of the tailless complex polypeptide 1 ring complex/

Chaperonin containing tailless complex polypeptide 1 (TriC/CCT) per proteome subtype in matched cases ($n_{pWNT}=4$, $n_{pSHHt}=14$, $n_{pSHHs}=4$, $n_{pG3}=6$, $n_{pG4}=17$, $n_{p3Myc}=11$, data are presented as mean values +/- SD. Left: Heatmaps. Middle: Quantification (two-tailed, unpaired t-test). Right: p-values when comparing subtypes ($p_{pWNTvspSHHt} < 0.0001$, $p_{pWNTvspSHHs} < 0.0001$, $p_{pWNTvspG3} < 0.0001$, $p_{pWNTvspG3myc} < 0.0001$, $p_{pWNTvspG4} < 0.0001$, $p_{pSHHtvspSHHs} < 0.001$, $p_{pSHHtvspG3} < 0.0001$, $p_{pSHHtvspG3myc} < 0.0001$, $p_{pSHHtvspG4} < 0.01$, $p_{pSHHsvspG3} < 0.01$, $p_{pSHHsvspG3myc} < 0.0001$, $p_{pSHHsvspG4} < 0.05$, $p_{pG3vsG4} = n.s.$, $p_{pG3vspG3myc} < 0.0001$, $p_{pG4vspG3myc} < 0.0001$). **(D)** Correlation plot displaying mean correlation for each component in all three omic types. **(E)** Circus plot displaying correlations ≥ 0.7 for each component's protein, gene and CpG site. Only CCT2 significantly correlated on all three levels. n represents biologically independent human samples. n.s= not significant.

Figure 9

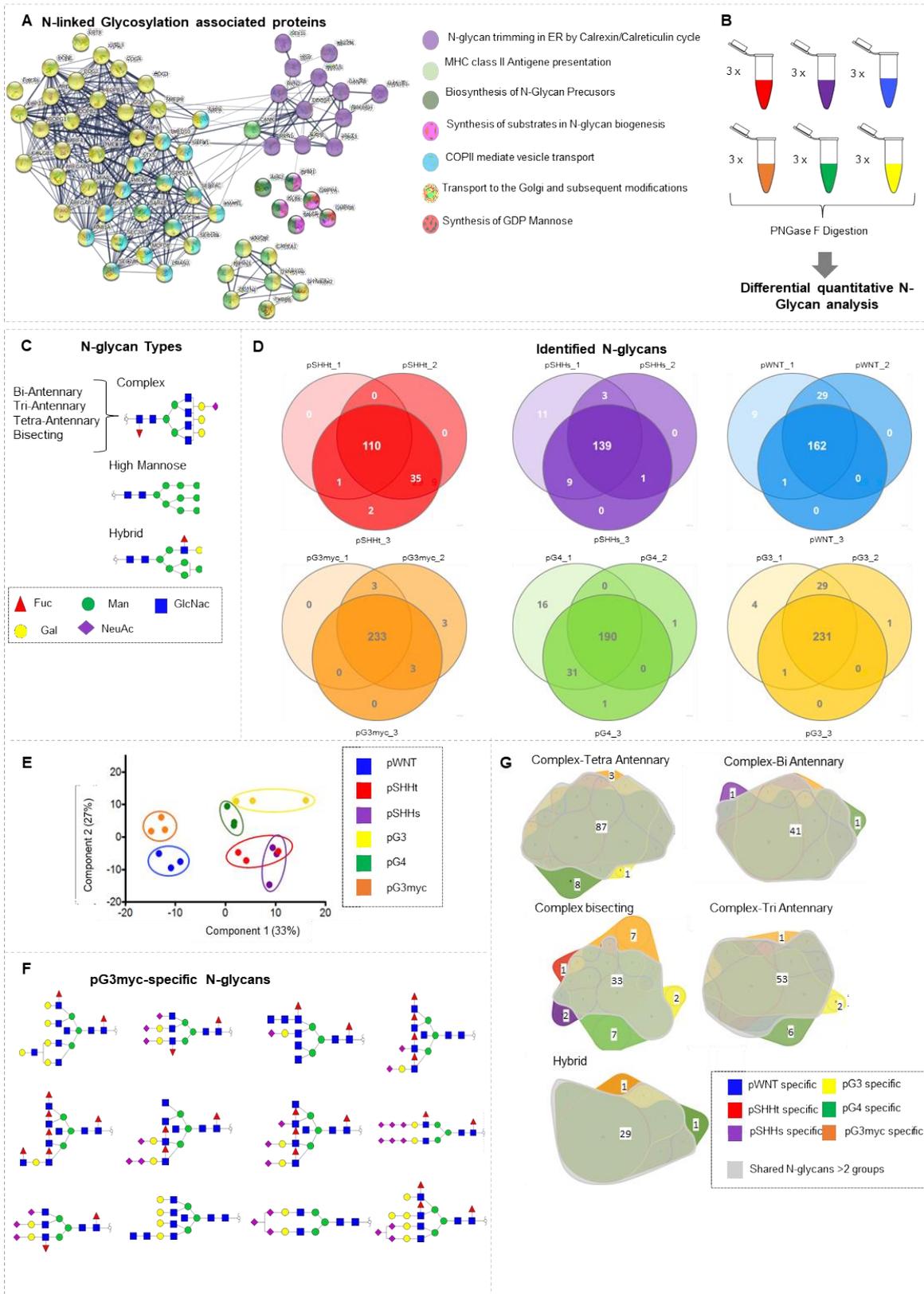


Figure 9: N-glycan analysis reveals significant differences across N-glycan profiles of proteomic MB subtypes

(A) STRING network analyses of differentially abundant proteins involved in N-linked glycosylation. **(B)** Scheme of N-glycan analyses. **(C)** Schematic visualization of N-glycan types. **(D)** Venn diagrams showing overlap of identified glycans per MB proteome subtype ($n_{pWNT}=3$, $n_{pSHHt} = 3$, $n_{pSHHs}=3$, $n_{pG3} = 3$, $n_{pG3myc} = 3$, $n_{pG4}=3$). **(E)** PCA, based on N-glycan abundances, illustrating the separation of proteome MB subtypes at the N-glycan level. **(F)** 2D Structure visualization for pG3myc-specific N-glycans. GlcNAc=N-Acetylglucosamine; Gal=Galactose; Fuc=Fucose; ManNAc=N-Acetylmannosamine; Neu5Ac=N-Acetylneuraminic acid. **(G)** Venn Diagram, comparing the identified hybrid-Type and complex N-glycans between proteome subtypes. n represents biologically independent human samples.

Figure 10

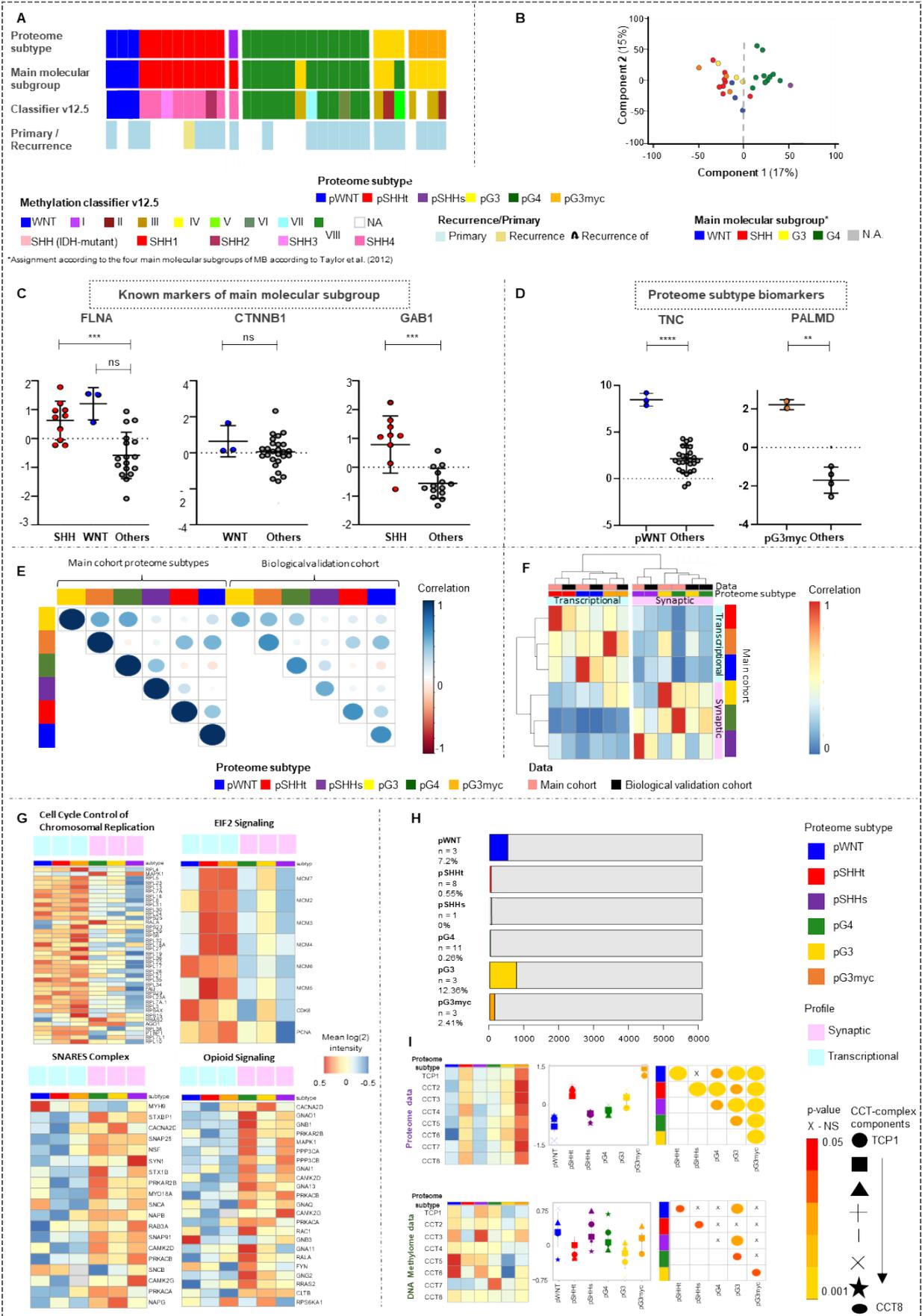
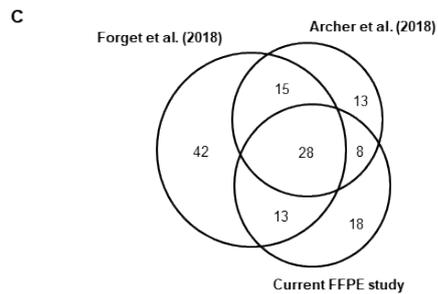
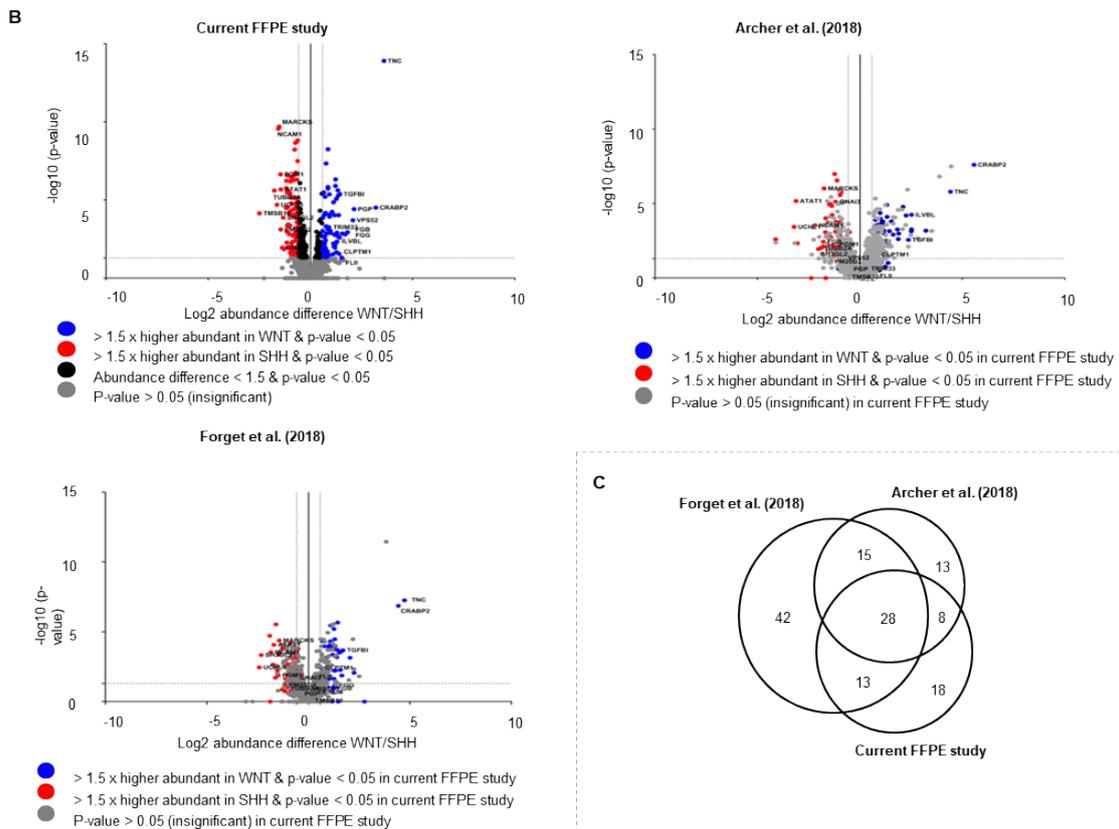
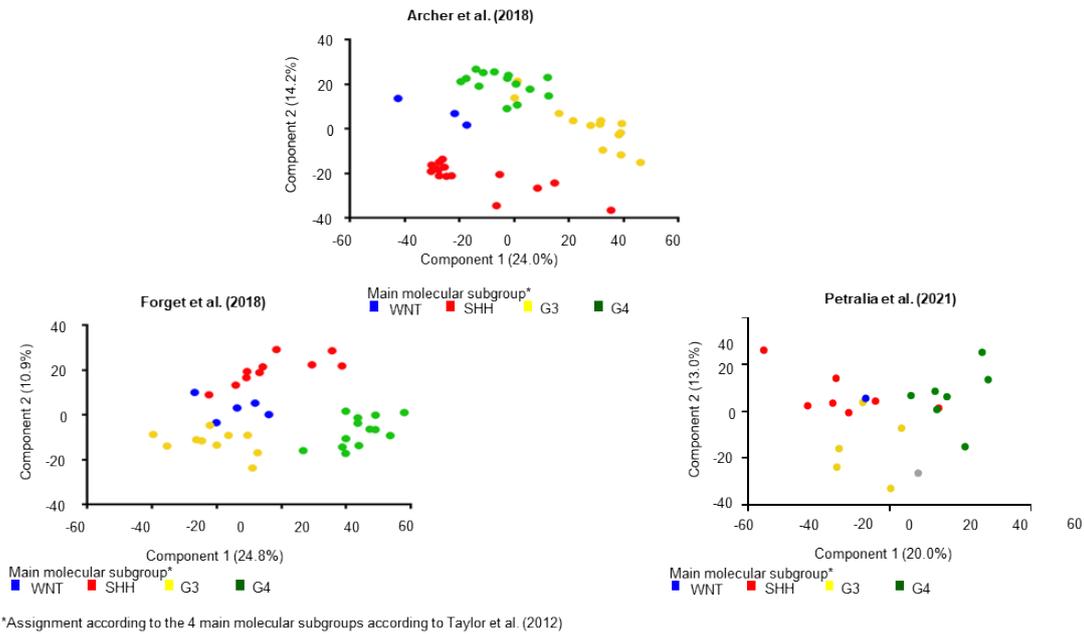


Figure 10: Confirmation of proteome subtypes and differential feature conservation in an independent biological FFPE dataset

(A) Clinical sample information with proteome subtype assignments using ACF based classification¹⁹⁵ **(B)** PCA, based on proteins found in $\geq 70\%$ samples, illustrating the separation of proteome MB subtypes **(C)** Protein abundances of established biomarkers WNT and SHH biomarker FLNA ($n_{WNT} = 3$, $n_{SHH} = 9$, $n_{Others} = 18$, two-tailed, unpaired t-test, $p_{pWNTvsOthers} = NS$, $p_{pSHHvsOthers} < 0.001$), WNT biomarker CTNNB1 ($n_{WNT} = 3$, $n_{Others} = 27$, two-tailed, unpaired t-test, $p_{pWNTvsOthers} = n.s.$) and SHH biomarker GAB1 ($n_{SHH} = 9$, $n_{Others} = 21$, two-tailed, unpaired t-test, $p_{pSHHvsOthers} < 0.001$, data are represented as mean values \pm SD). **(D)** Significant higher abundance of TNC ($n_{pWNT} = 3$, $n_{Others} = 27$, two-tailed, unpaired t-test, $p_{pWNTvsOthers} < 0.0001$) and PALMD ($n_{pG3myc} = 3$, $n_{pOthers} = 27$, two-tailed, unpaired t-test, $p_{pG3mycvsOthers} < 0.01$) in pWNT and the pG3myc subtype, respectively. Data are represented as mean values \pm SD. **(E)** Correlation plot displaying mean Pearson correlation per subtype between the integrated cohort and the biological validation cohort. **(F)** Hierarchical clustering of biological validation samples with samples from the main cohort (Pearson correlation and ward.D2 linkage) **(G)** Heatmaps showing mean protein abundance for the top hit gene sets enriched in the transcriptional (top) and synaptic profile (bottom). **(H)** Bar plot displaying proteome subtype specific pearson correlation calculated for matched samples between proteins and CpG sites ($r > 0.7$, $n=29$, total number of samples having both DNA methylome and proteome data, 5880 proteins and 549,089 CpG sites). The number of proteins correlating with CpG site of their own gene are shown in colour. **(I)** Left : Heatmaps for Mean protein abundancies, gene expression values and methylation at CpG sites for all components of the tailless complex polypeptide 1 ring complex/ Chaperonin containing tailless complex polypeptide 1 (TriC/CCT) per proteome subtype in matched cases ($n=29$, $n_{pWNT} = 3$, $n_{pSHHt} = 8$, $n_{pSHHs} = 2$, $n_{pG3} = 3$, $n_{pG3myc} = 3$, $n_{pG4}=11$) samples having both DNA methylome and proteome data). Middle : Quantification (two-tailed, unpaired t-test, data are presented as mean values \pm SD) Right : p-values when comparing subtypes ($p_{pWNTvspSHHt} < 0.0001$, $p_{pWNTvspSHHs} = NS$, $p_{pWNTvspG3} < 0.0001$, $p_{pWNTvspG3myc} < 0.0001$, $p_{pWNTvspG4} < 0.001$, $p_{pSHHtvspSHHs} < 0.001$, $p_{pSHHtvspG3} < 0.001$, $p_{pSHHtvspG3myc} < 0.0001$, $p_{pSHHtvspG4} < 0.01$, $p_{pSHHsvspG3} < 0.001$, $p_{pSHHsvspG3myc} < 0.0001$, $p_{pSHHsvspG4} < 0.05$, $p_{pG3vsG4} < 0.01$, $p_{pG3vspG3myc} < 0.0001$, $p_{pG4vspG3myc} < 0.0001$). n represents biologically independent human samples. n.s.= not significant.

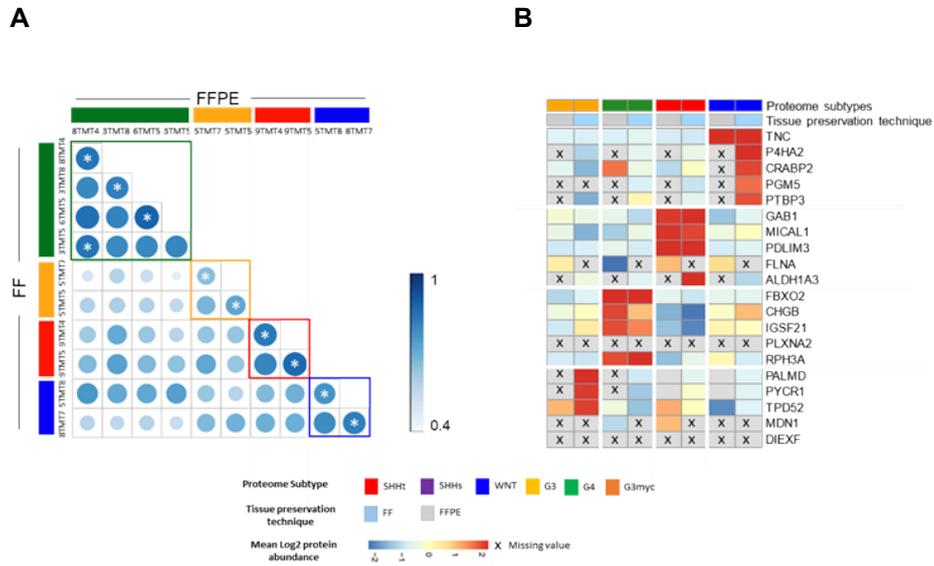
Supplementary Figures

A PCA for individual studies and differentially abundant proteins identified in each studies individually

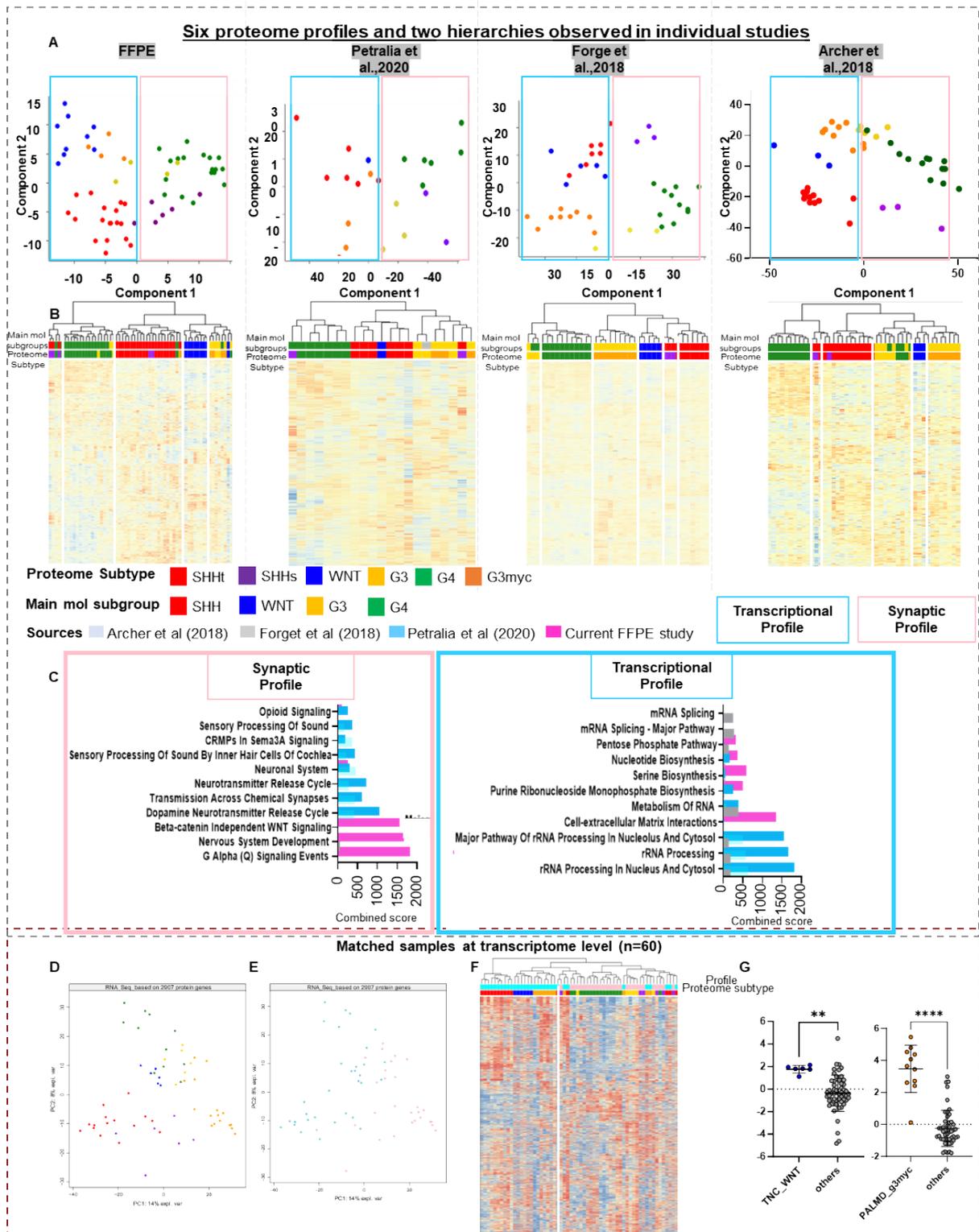


Supplementary Figure 1: A Scatter Plot visualization of the first 2 principal components in NIPALS PCA, based all proteins, quantified in individual cohorts (Archer et al (2018): 14522 proteins, n = 45, Forget et al (2018): 3890 proteins, n = 38, Petralia et al (2021): 5680 proteins, n = 22). Samples were colored according to the main molecular subtype. **B** Volcano Plot visualization of two sided, two-sided unpaired, t-testing results, showing the abundance distribution of differential abundant proteins between SHH and WNT type MBs in the current FFPE study and the fresh frozen tissue datasets, provided by Archer et al. (2018) and Forget et al. (2018). For T-testing proteins were reduced to 2968 proteins, identified in the current FFPE study (n=62). Proteins with a p-value < 0.05, that showed a foldchange > 1.5 between tested groups were considered statistically significant and differential abundant. **C** Venn diagram, comparing statistically significant differential abundant proteins between SHH and WNT type MBs between three different proteome studies on human MB tissue (Forget et al. (2018); Archer et al. (2018); current FFPE study). To enable comparability across studies proteins were reduced to 1594 proteins, found at least 2 times in all studies. . n represents biologically independent human samples.

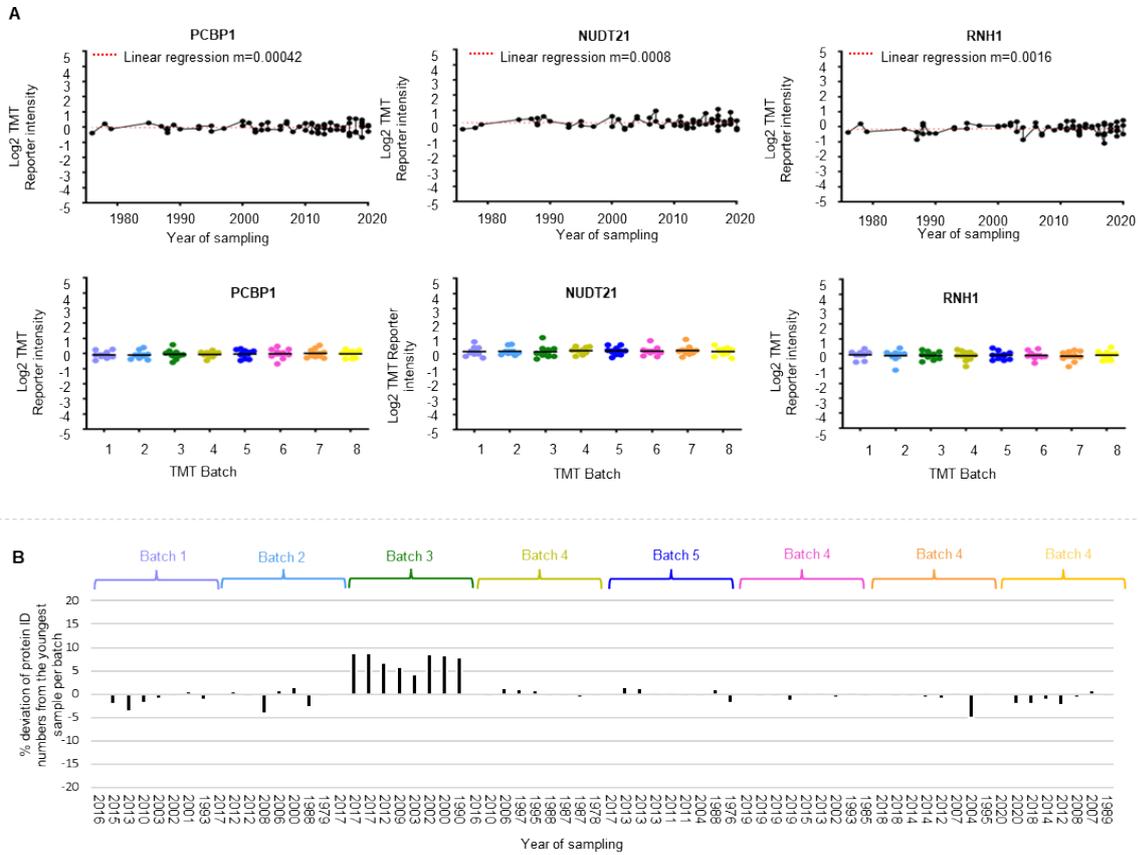
Correlation between matched FF and FFPE samples



Supplementary Figure 2: A Corrplot showing pearson correlation between FF sample (n = 10) on Y-axis and their respective FFPE samples (n = 10) on the X-axis annotated for proteome subtype. Asterisks mark highest Pearson score. **B** Mean of biomarkers for each proteome subtype (defined in the study). n represents biologically independent human samples.

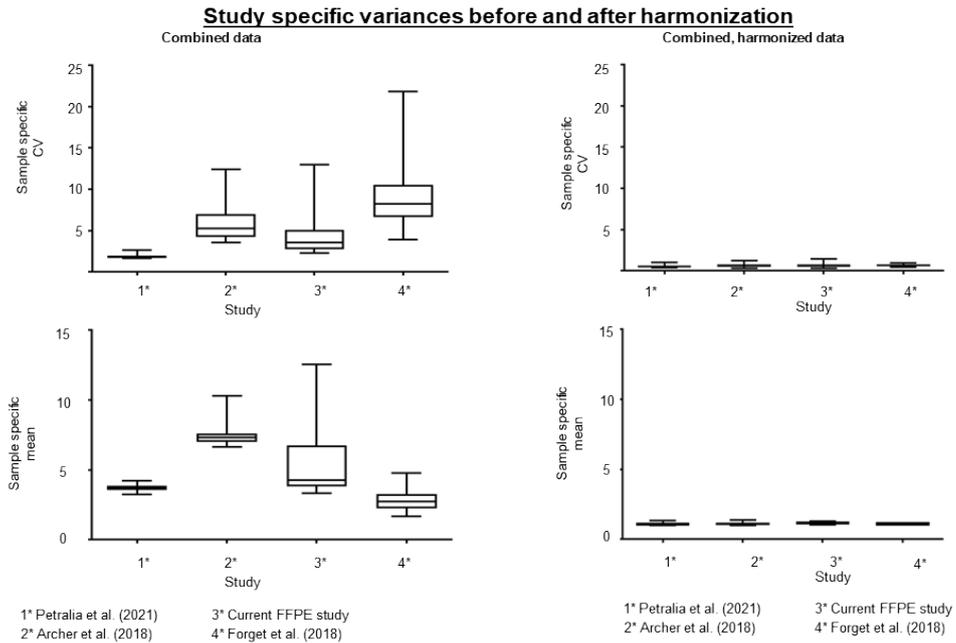


Protein abundances for housekeeping proteins and number of proteins quantified based on year

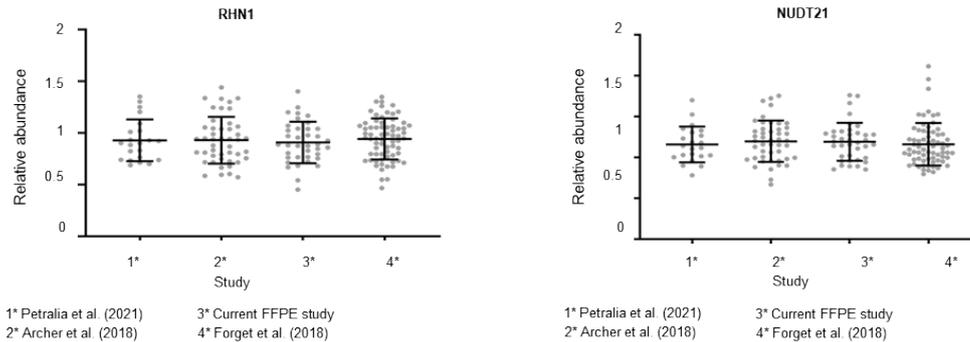


Supplementary Figure 4: A Log₂ TMT Reporter intensity for selected housekeeping proteins per measured TMT Batch and across sampling years. Normalized, TMT Batch-Effect corrected protein abundances were used. Each dot represents one independent biological sample. **B** Percentage deviation of the total number of identified proteins from the identified proteins of the youngest sample per batch.

A



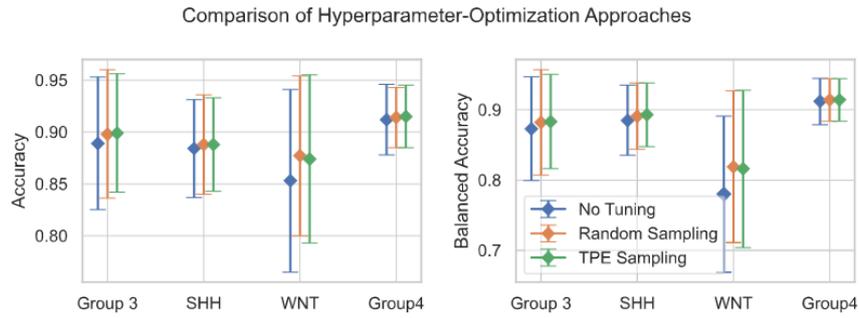
B



Supplementary Figure 5: A Sample specific CV and mean for each batch prior to and after the usage of ComBat (L/S scaling, parametric Bayesian framework) in the HarmonizR framework prior to and after data harmonization across studies. In boxplots, 50 % of the data points are inside the box (Q1 (Quartile 1) being the lower bound of the box (25 %), Q3 being the upper bound of the box (75 %)). Whiskers show all values beyond the box without outliers. Outliners were defined as $Q3 + 1.5 * IQR$ (Interquartile range) (upper outlier) and $Q1 - 1.5 * IQR$ (lower outlier). IQR being $Q1 - Q3$. **B** Protein abundance distribution of housekeeping proteins NUDT21 and RHN1 across different studies after data harmonization, data presented as \pm SD.

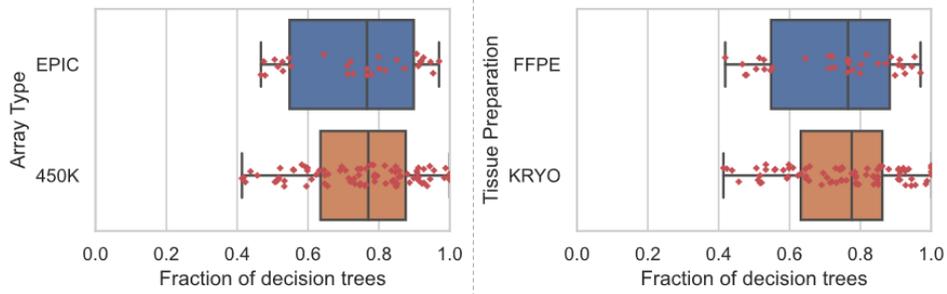
Classification of samples based on Random forest classifier

A

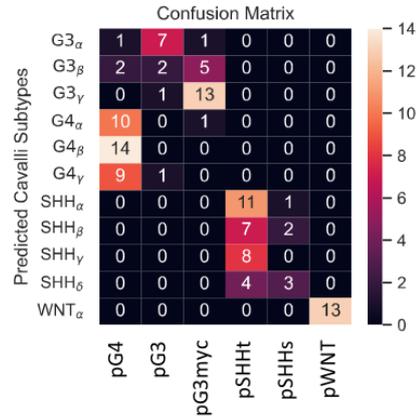


B

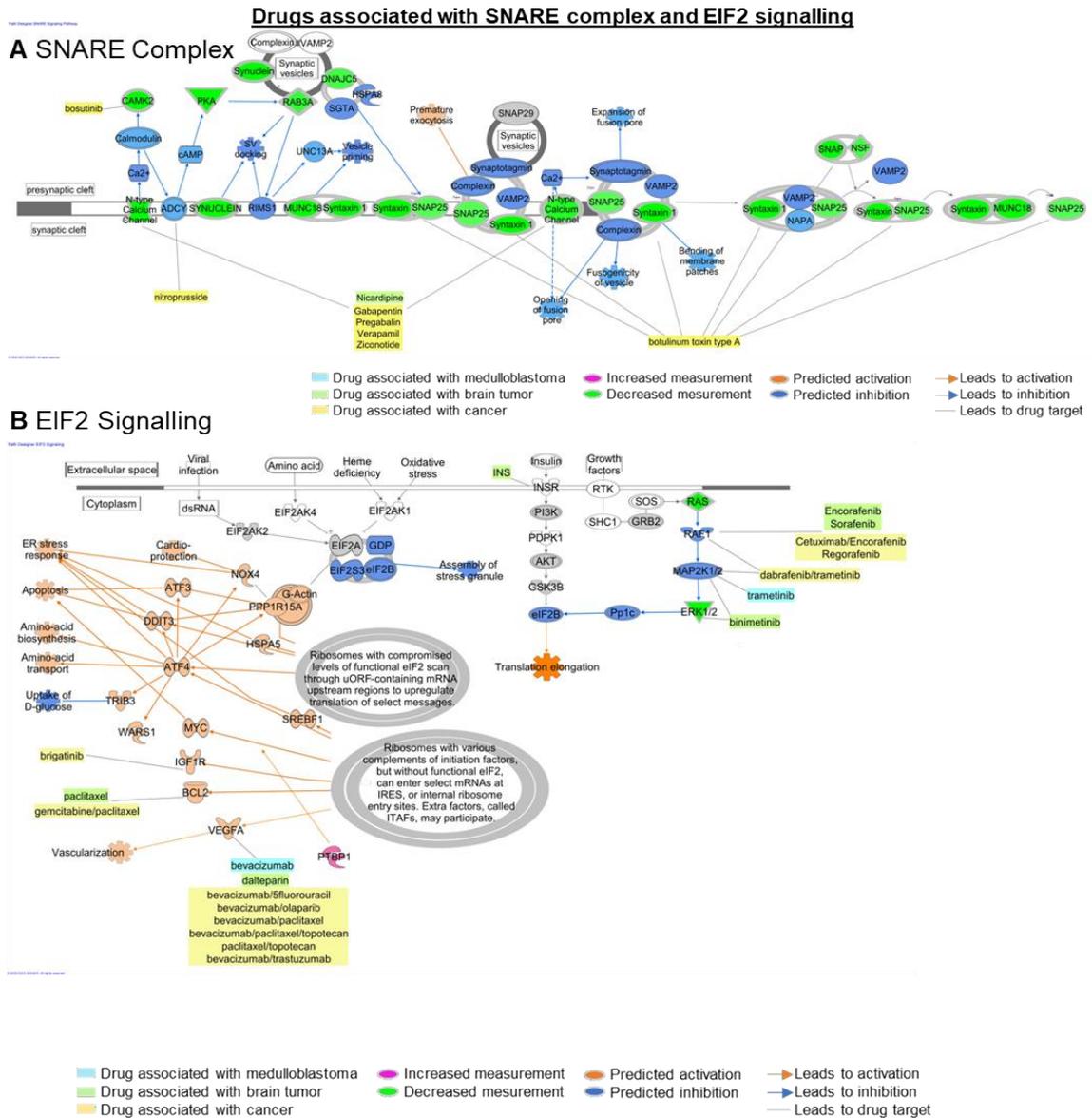
Estimate of Batch Effects



C

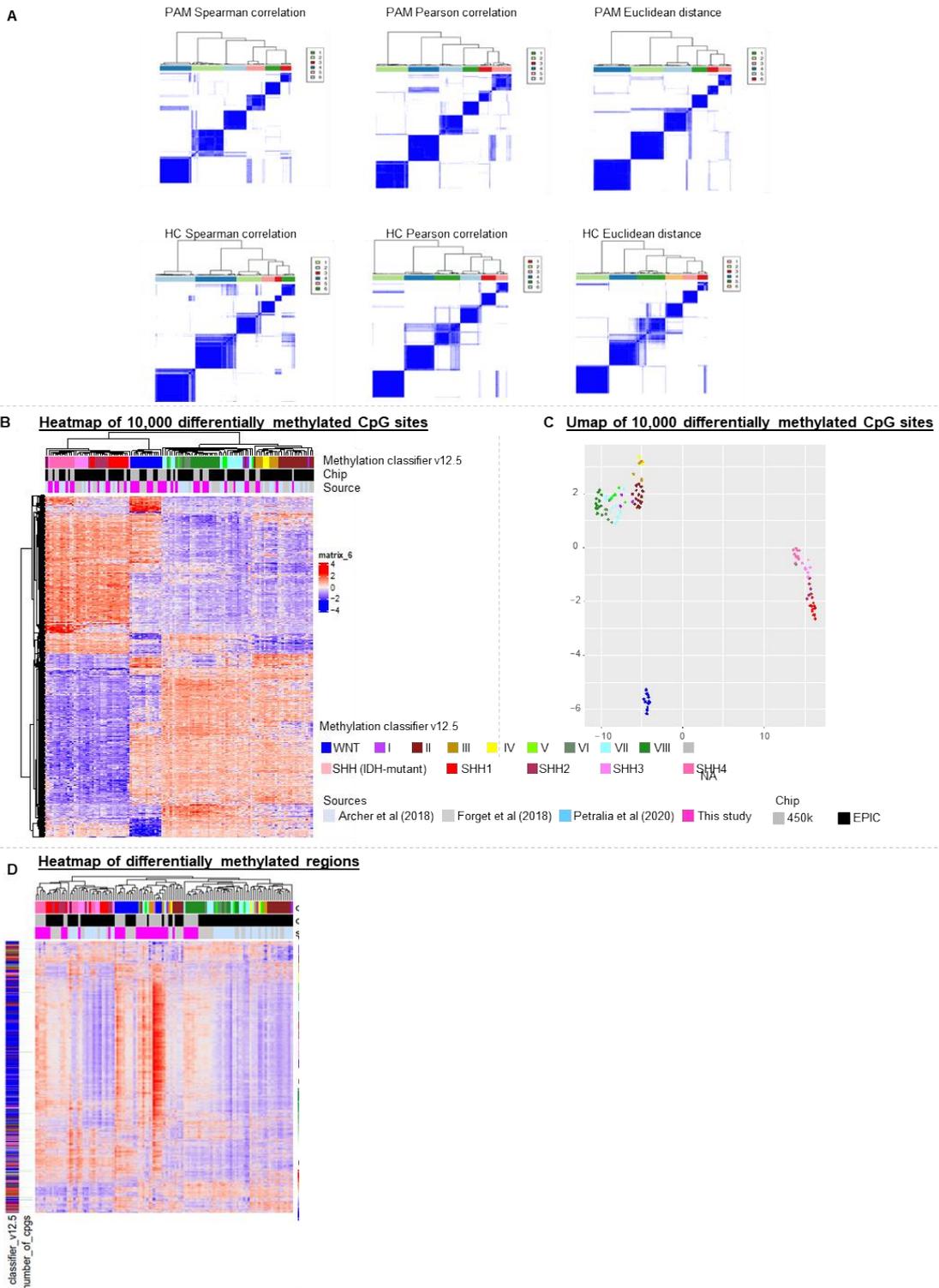


Supplementary Figure 6: **A** Comparison of different approaches to hyperparameter-optimization for the RandomForest-algorithm in order to classify samples into subgroups based on DNA-methylation as described by Cavalli et al (2017), data presented as mean \pm SD. In general, TPE-Sampling yields the best accuracy (left) and balanced accuracy (right) on randomly drawn validation sets. **B** Estimation of the impact of array type and type of tissue preparation on the classification. The number of DecisionTrees voting for the predicted class shows no significant relationship with both processing parameters (Kruskal-Wallis-H-test p-value 0.511 for the array type and 0.540 for the type of tissue preparation), dots represent single value, box represents Median with Q1 to Q3, error bars show Q1-1.5*IQR and Q3+1.5*IQR. IQR=inter-quartile range. **C** Confusion matrix for the predicted Cavalli-subtype and the proteomic subtypes from this study. n represents biologically independent human samples.



Supplementary Figure 7: A Drugs associated with SNARE complex and **B** EIF2 signalling found using Ingenuity pathway analysis.

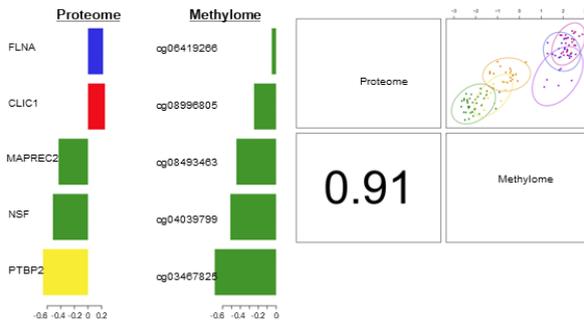
Consensus clustering at DNA methylome level



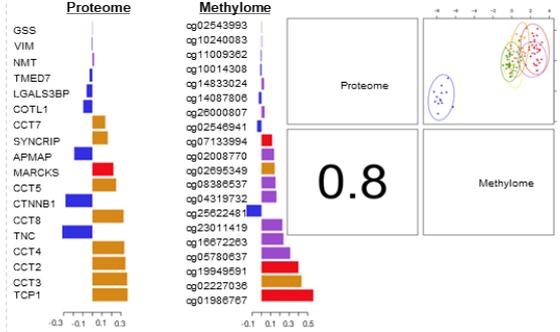
Supplementary Figure 8: **A** The optimal clustering methylome data, as determined using Pearson correlation, Spearman correlation and Euclidean distance as distance metric for hierarchical clustering (HC) and k-meloids clustering (PAM) respectively. k , number of clusters. Consensus scores are indicated using a color scale from white (samples never cluster together) to blue (samples always cluster together). **B** Heatmap of 10,000 most differentially methylated CpG sites for $n = 117$ samples (M -values) hierarchically clustered using Pearson correlation and ward.D linkage (moderated t -statistic corrected for multiple testing using Bonferroni-Hochberg correction and p -value < 0.05). **C** Umap based on these differentially methylated CpG sites. **D** Hierarchical clustering of 10,000 differentially methylated DMRs, clustering based on the 14-subtypes using MNP v12.5 RandomForest Classifier. n represents biologically independent human samples.

Component-wise loading obtained from mixOmics

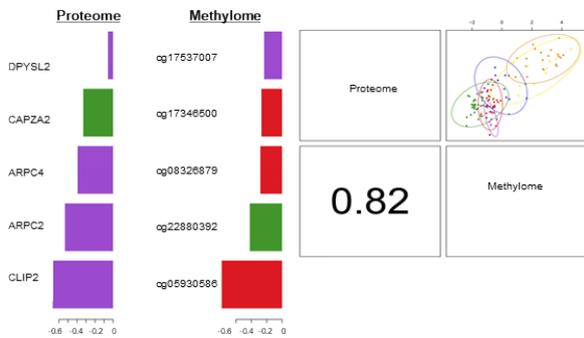
A Component 1 loadings and PCA



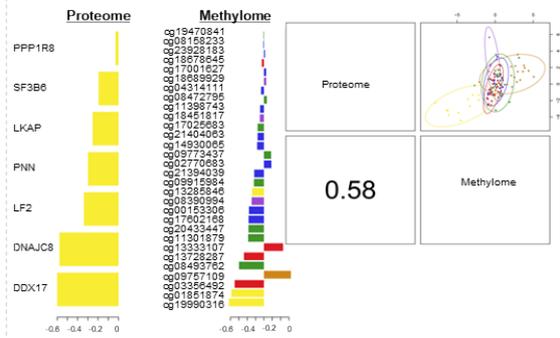
B Component 2 loadings and PCA



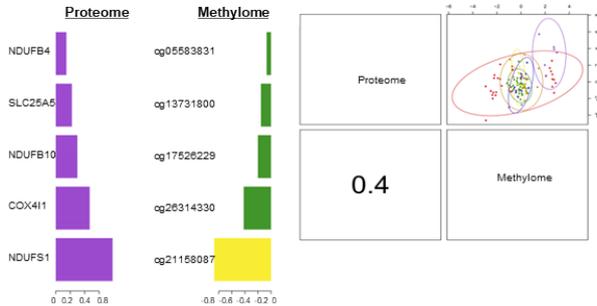
C Component 3 loadings and PCA



D Component 4 loadings and PCA



E Component 5 loadings and PCA

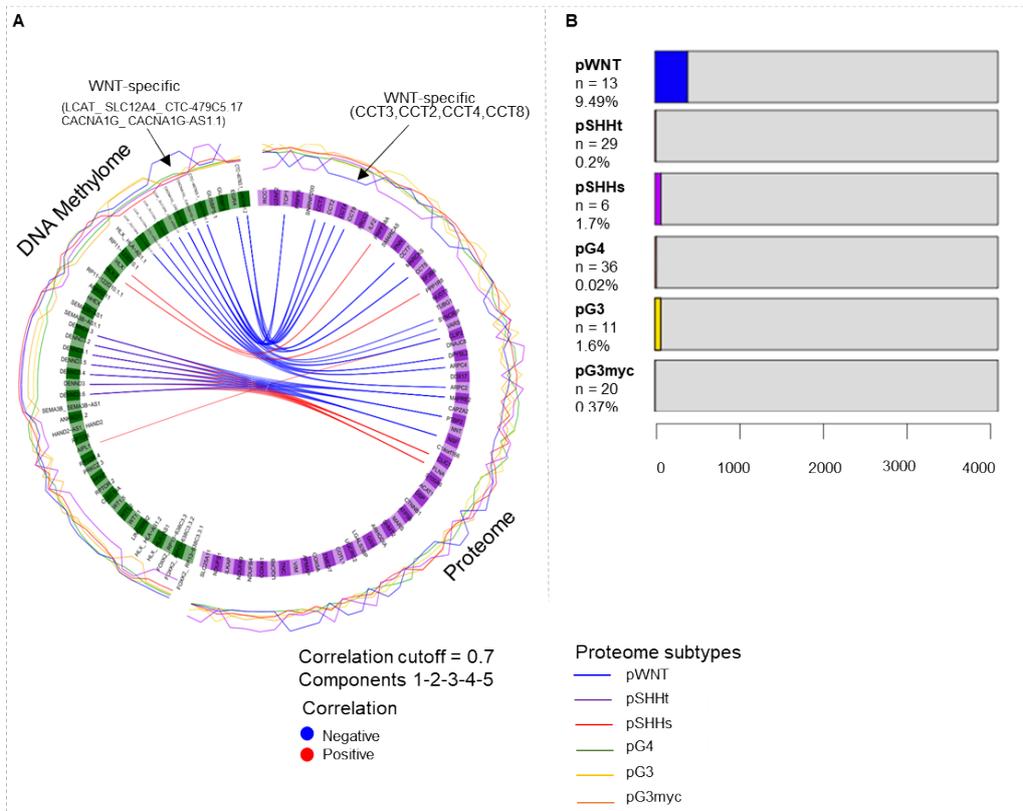


Proteome Subtype

- pWNT
- pG3
- pSHHT
- pSHHS
- pG4
- pG3myc

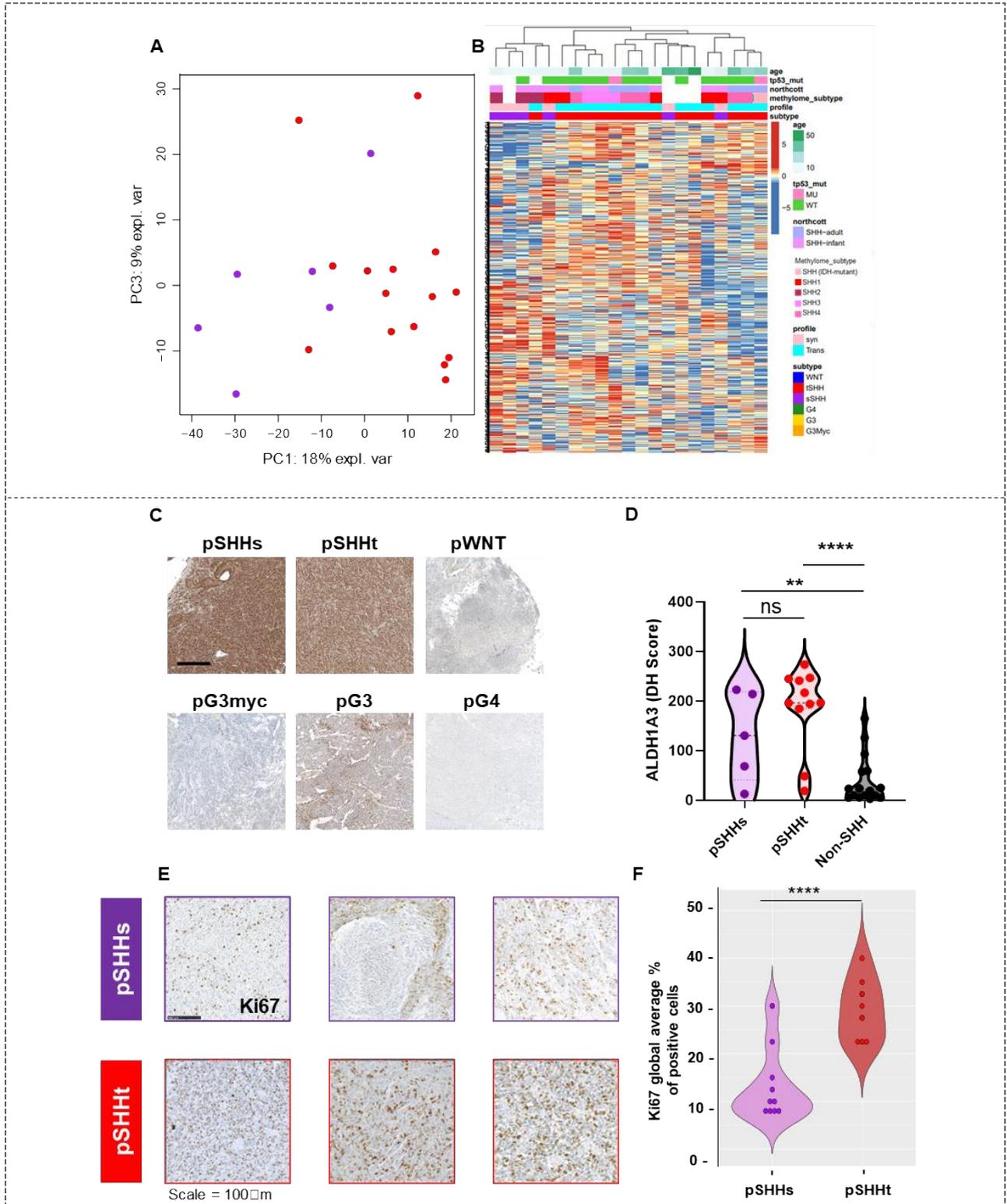
Supplementary Figure 9: A-E Integration of DNA methylome (n = 115 samples, 10,000 differentially methylated CpG sites) and proteome data (115 samples, 3990 proteins) was achieved using sparse partial least squares discriminant analysis (sPLS-DA). This integration reduced the data into six components (artificial variables made up of weights of original variables contributing to the respective component). A-E The barplots represent the loading value (defined as the weight assigned to each feature to determine their contribution to the selected component) and the colors represent the proteome subtype to which the feature contributes. The length of the bar plot represents the absolute loading value and the direction depicts whether the loading value is positive or negative. The corresponding PCA shows how well the selected features for each component separate the six proteome subtypes. . n represents biologically independent human samples.

MixOmcs using DMRs instead of CpGs



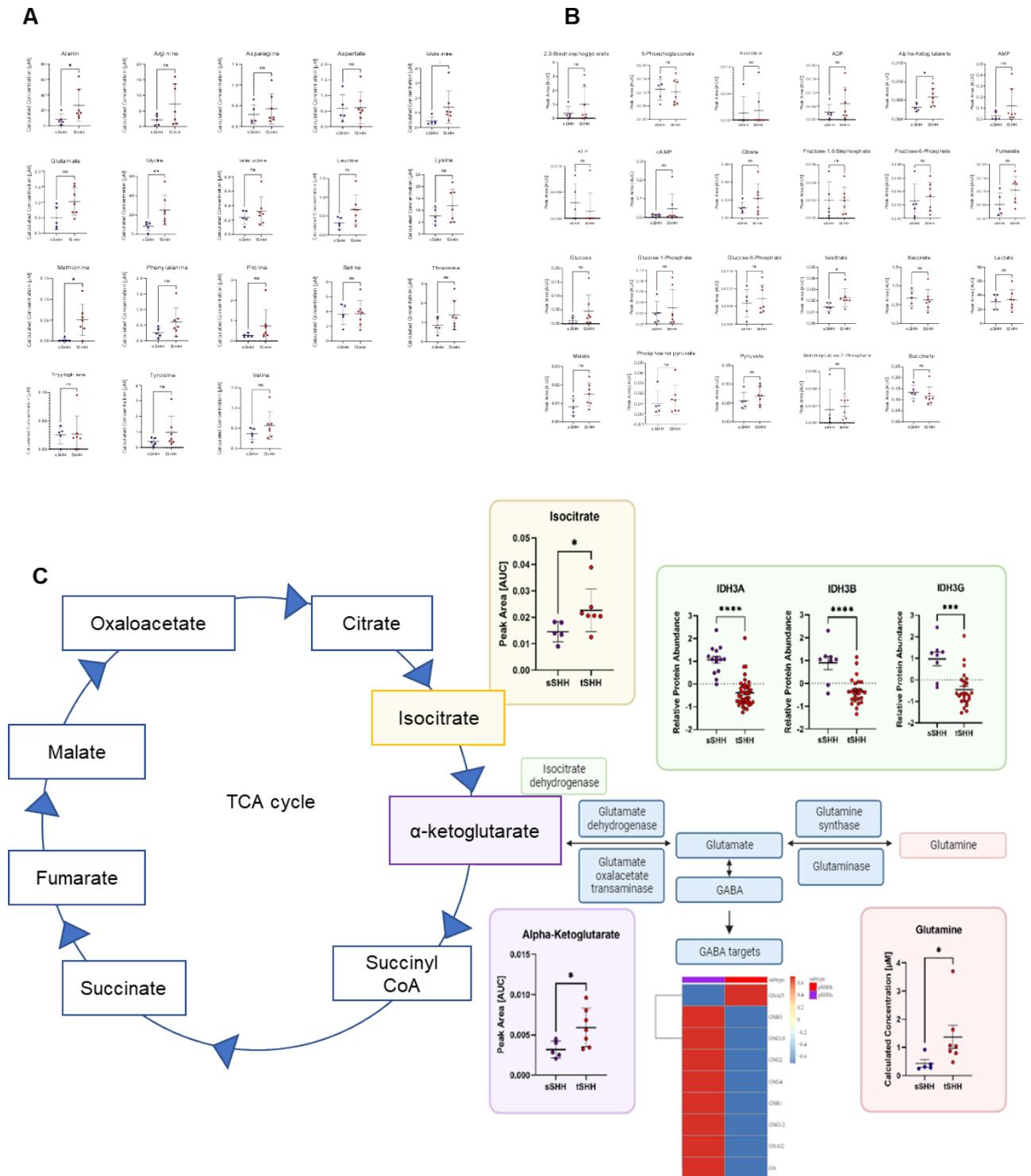
Supplementary Figure 10: A Circos plot using 10,000 differentially methylated regions integrated with 3990 proteins from the main cohort (n = 117 samples). Plotted here are features selected on 5 components showing a correlation of > 0.7
B Barplot representing the number of proteins correlating with the differentially methylated region belonging to their own gene. (pearson correlation > 0.7 are plotted). n represents biologically independent human samples.

Differences in pSHHt and pSHHs at transcriptome level and using Ki67 stainings

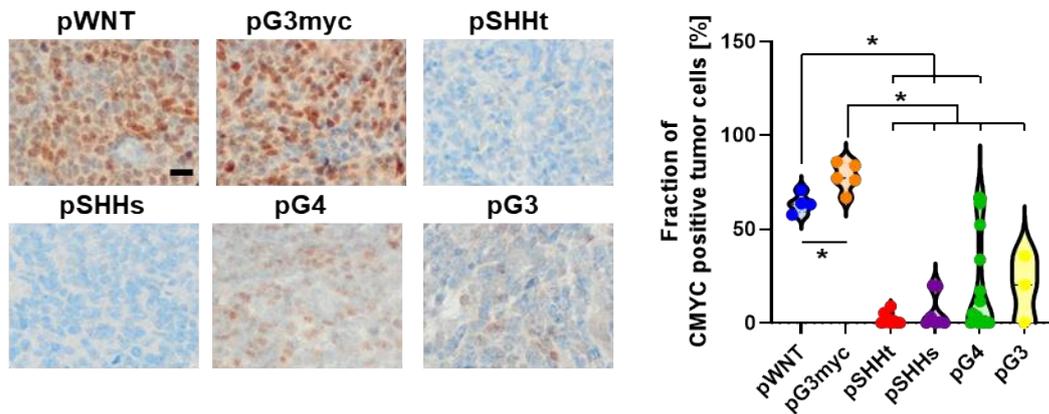


Supplementary Figure 11: **A** PCA and **B** Hierarchical clustering using 100% valid values, Pearson correlation and ward.D2 linkage only in SHH samples (n=21) **C** Representative images of immunohistochemical stainings against ALDH1A3. A strong cytoplasmic signal was detected in pSHHt and pSHHs MBs. **D** Digital quantification of ALDH1A3 DAB signals showed a significantly enhanced digital histoscore (DH Score) for pSHHt and pSHHs compared to all other MB subtypes ($n_{pSHHs}=5$, $n_{pSHHt}=11$, $n_{non-SHH}=16$, $p_{pSHHt_vs_pSHHs}=0.0224$, $p_{pSHHs_vs_non-SHH}=0.009$, $p_{pSHHt_vs_non-SHH}<0.0001$, unpaired t-test). Scale bar is 500 μ m. **E** Representative images for Ki67 stainings. Higher number of cells stained positive for Ki67 in pSHHt compared to pSHHs. **F** pSHHt subtypes show significantly higher percentage of Ki67 positive cells compared to pSHHs group ($n_{pSHHs}=6$, $n_{pSHHt}=6$, $p_{pSHHt_vs_pSHHs}<0.0001$, unpaired t-test. .n represents biologically independent human samples.

Amino acid and metabolite distribution



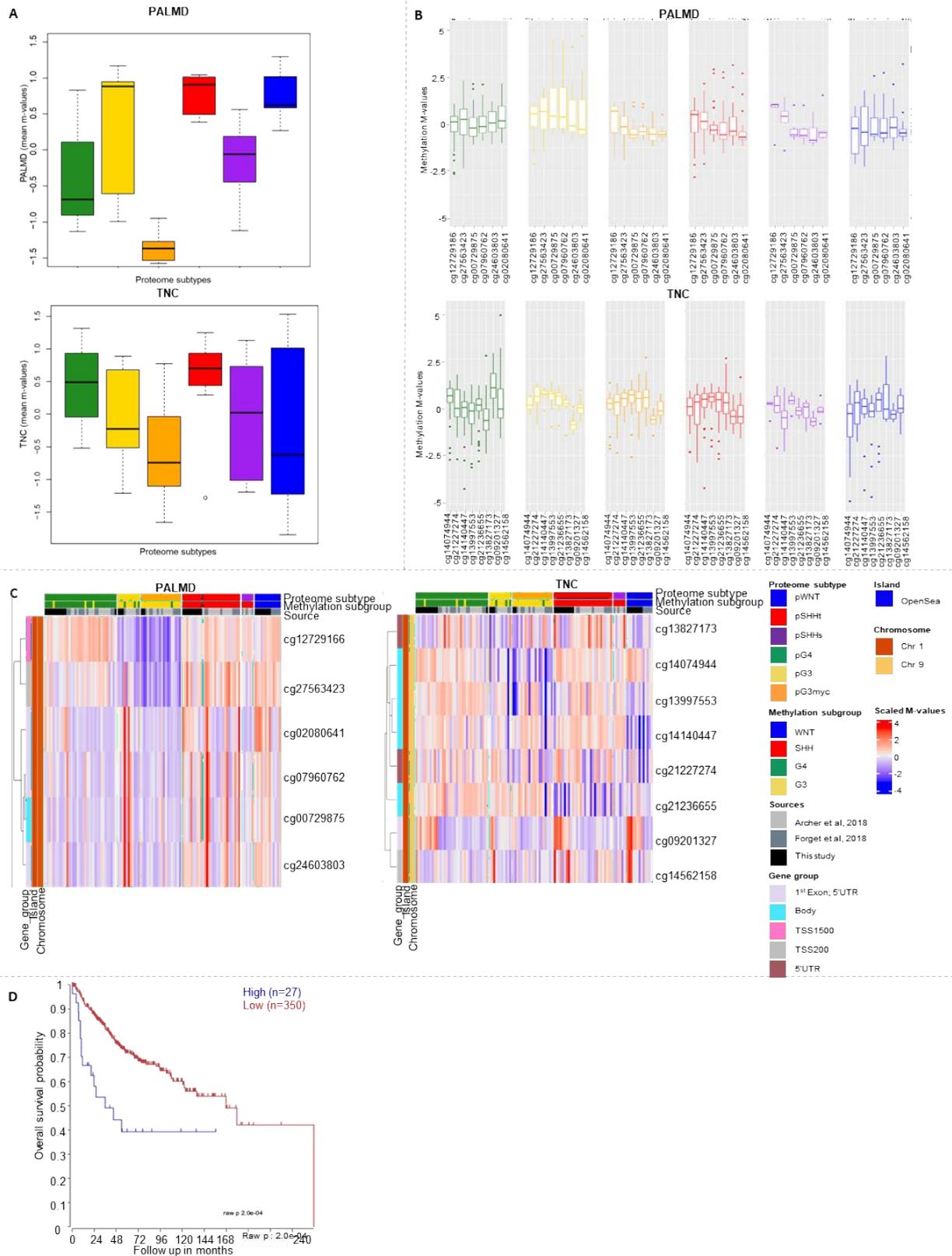
Immunohistochemical stainings for cMYC



Supplementary Figure 13A:

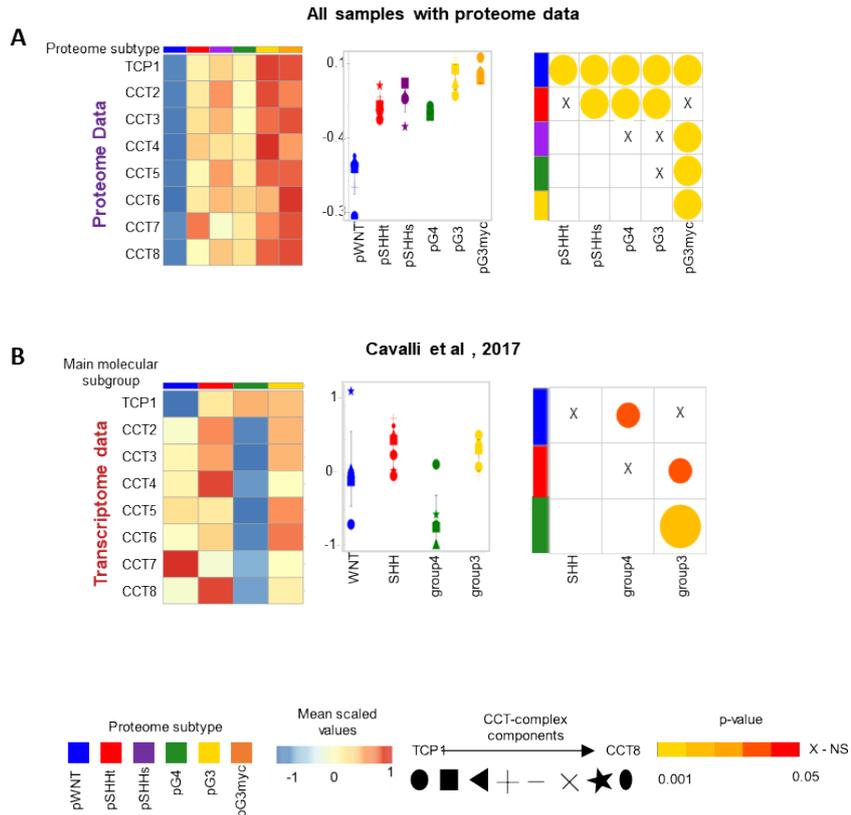
Representative images of Immunohistochemical stainings against CMYC with quantification of the fraction of CMYC positive tumor cells. A strong nuclear signal was detected in a majority of tumor cells in pWNT (n=4) and G3myc (n=5) MBs (Mean_{pWNT}: 63.9 %, Mean_{pG3myc} = 78.0 %), whereas all other MB subtypes showed lower fractions (mean_{pSHHs} =6.2 %, mean_{pSHHt} = 2.0 %, mean_{pG3} =18.7 %, mean_{pG4} = 18.9 %). *: P_{pG3myc_vs_pWNT}=0.03, P_{pG3_vs_pWNT}=NS, P_{pG4_vs_pWNT}=0.02, P_{pWNT_vs_pSHHt}=0.002, P_{pWNT_vs_pSHHs}=0.006, P_{pG3myc_vs_pSHHt}=0.001, P_{pG3myc_vs_pSHHs}=0.002, P_{pG3myc_vs_pG3}=0.03, P_{pG3myc_vs_pG4} <0.0001,, unpaired Mann-Whitney-test). Scale bar is 10 μ m . . n represents biologically independent human samples.

PALMD and TNC methylome data



Supplementary Figure 14: A Subtype specific boxplot showing mean methylation mvalues of all CpG sites for PALMD and TNC genes respectively across proteome subtypes (boxplots represents Median with Q1 to Q3, error bars show Q1-1.5*IQR and Q3+1.5*IQR. IQR=inter-quartile range), for each subtype. B Boxplots displaying methylation mvalue of all individual CpG sites of PALMD and TNC in six proteome MB subtypes (n=117) (boxplots represents Median with Q1 to Q3, error bars show Q1-1.5*IQR and Q3+1.5*IQR. IQR=inter-quartile range), . C Heatmap displaying the methylation of individual CpG sites of PALMD and TNC along with CpG site information: gene group, island and chromosome. D Survival plot showing differences in survival based on PALMD gene expression from the Cavalli dataset (Cavalli et al,2018, $n_{high} = 27$, $n_{low} = 350$). The plot was generated with the 'R2: Genomics Analysis and Visualization Platform (<http://r2.amc.nl>). n represents biologically independent human samples.

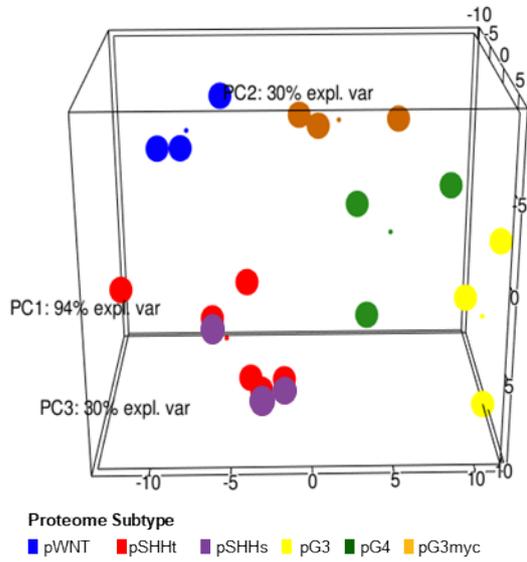
CCT- old cohort and cavalli



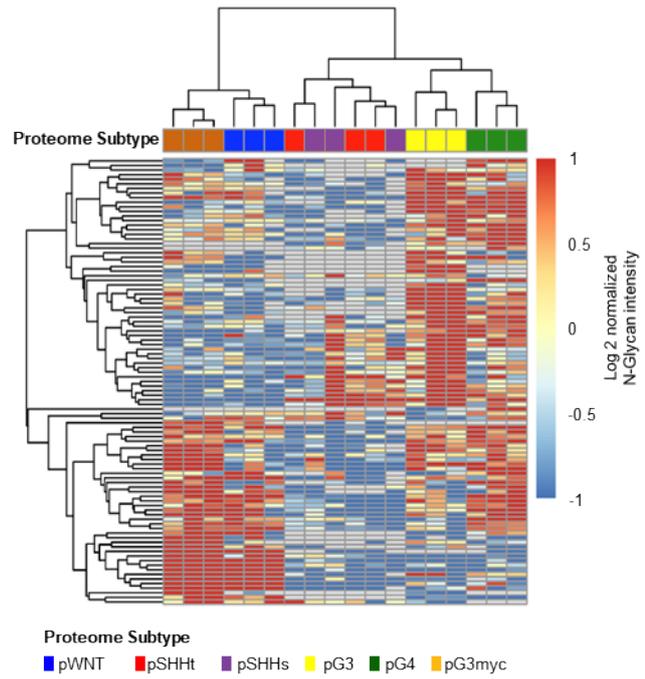
Supplementary Figure 15: **A** Left : Heatmap displaying CCT-complex components abundance across the six- proteome subtypes (mean per subtype, n=167) Middle : quantification as dotplots for mean per CCT- component per, data are presented as +/- SD and Right : corplot displaying the pvalues of the t-test results for each comparison in whole proteome data subtype ($p_{pWNTvsSHH} < 0.0001$, $p_{pWNTvsSHHs} < 0.0001$, $p_{pWNTvspG3} < 0.0001$, $p_{pWNTvspG3myc} < 0.0001$, $p_{pWNTvspG4} < 0.0001$, $p_{pSHHtvspSHHs} < 0.001$, $p_{pSHHtvspG3} < 0.0001$, $p_{pSHHtvspG3myc} = NS$, $p_{pSHHtvspG4} < 0.01$, $p_{pSHHsvspG3} < 0.01$, $p_{pSHHsvspG3myc} < 0.0001$, $p_{pSHHsvspG4} < 0.0001$, $p_{pG3vsG4} = NS$, $p_{pG3vspG3myc} < 0.0001$, $p_{pG4vspG3myc} < 0.0001$) and **B** Left: Heatmap CCT-complex components abundance across the four MB sugroups Cavalli data (n= 763, WNT=70, SHH=223 ,G4=326,G3=144)[1] Middle : quantification as dotplots for mean per CCT- component per, data are presented as +/- SD and Right: corplot displaying the pvalues of the t-test results for each comparison based on MB subgroups ($p_{WNTvsSHH} = NS$, $p_{WNTvsG4} < 0.01$, $p_{WNTvsG3} = 0.03$, $p_{SHHvsG4} = NS$, $p_{SHHvsG3} = 0.04$, $p_{G4vsG3} < 0.0001$). . n represents biologically independent human samples.

PCA and heatmap using Nglycan data

A

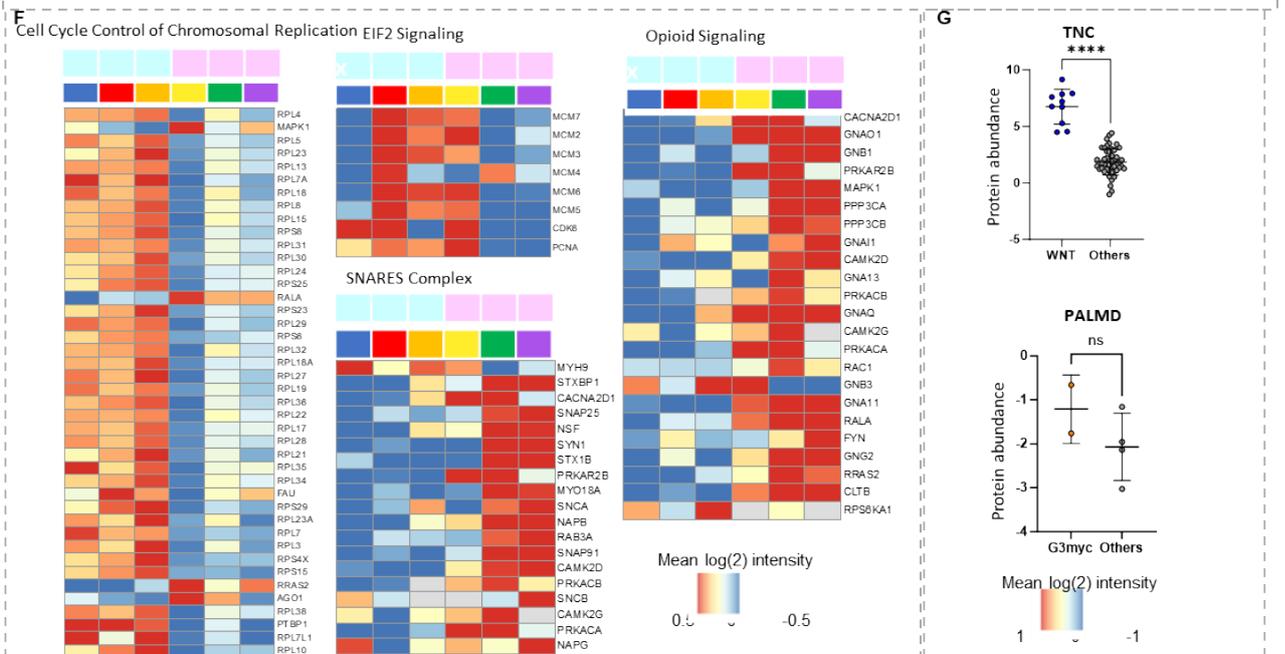
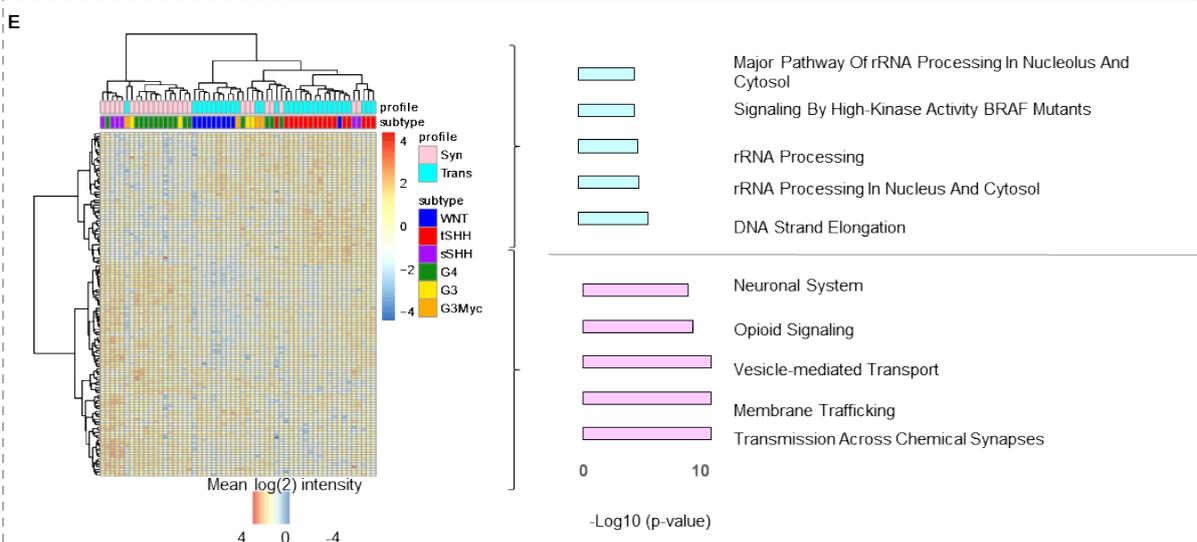
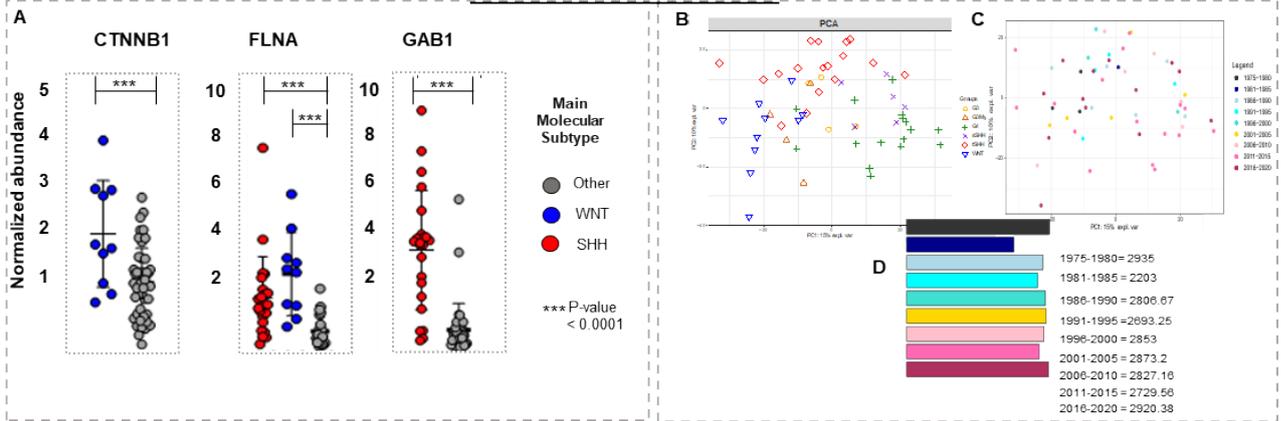


B



Supplementary Figure 16: A 3D Scatter Plot visualization of the first 3 principal components in NIPALS PCA, based on 302 N-glycans, found in the current FFPE cohort ($n_{\text{pWNT}}=3$, $n_{\text{pSHHt}}=3$, $n_{\text{pSHHs}}=3$, $n_{\text{pG3}}=3$, $n_{\text{pG3myc}}=3$, $n_{\text{pG4}}=3$). Samples were colored according to the proteome subtype. Heatmap visualization of Pearson correlation-based hierarchical clustering with Ward.D linkage, based on 92 ANOVA significant N-glycans between different proteome subtypes of MB. N-glycan abundances have been mean normalized across rows prior to visualization. .n represents biological independent human samples.

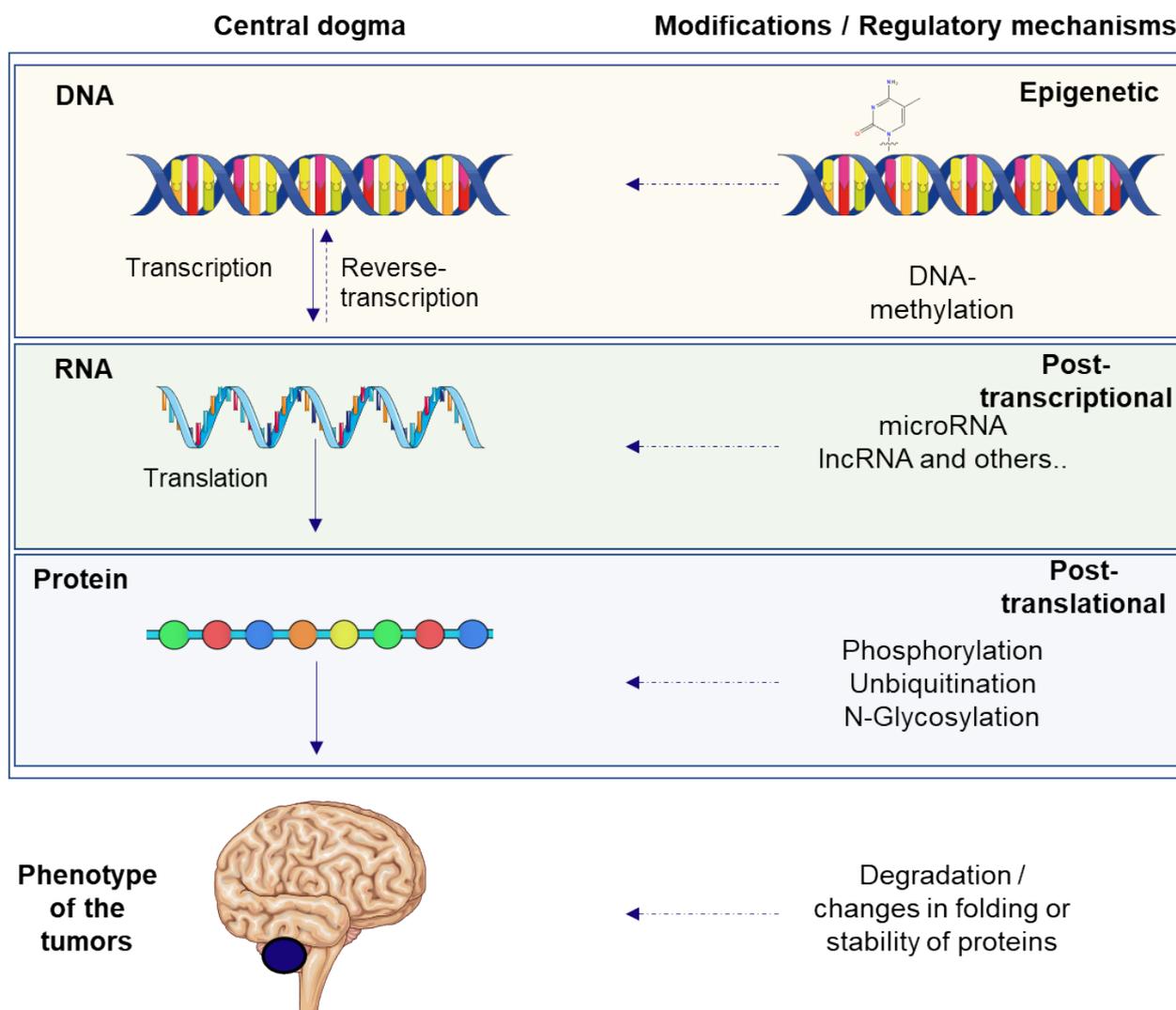
Technical validation cohort



Supplementary Figure 17 : A Housekeeping genes distribution in technical validation cohort (n=57, pWNT=10, pSHH=17, pSHS=6, pG4 = 16, pG3 = 4, G3myc = 4), two-tailed, unpaired-test (CTNNB1: $p_{pWNTvsOthers} < 0.001$, FLNA : $p_{pWNTvsOthers} < 0.001$, : $p_{pSHHvsOthers} < 0.001$, GAB1:

$p_{pSHHvsOthers} < 0.001$ B PCA reflecting the six proteome subtypes C PCA based on sample age D number of proteins quantified across samples from different age (mean of samples quantified in the given age range were considered) E Hierarchical clustering of samples based on proteome subtypes F Enriched pathways in samples belonging to synaptic and transcriptional profiles G Biomarkers for pWNT (TNC, $n_{pWNT} = 10$, $n_{Others} = 47$, : $p_{pWNTvsOthers} < 0.001$) and pG3myc (PALMD, $n_{pG3myc} = 2$, $n_{Others} = 4$, : $p_{pG3mycvsOthers} = NS$ (PALMD was not detected in other samples)) identified in the main cohort. ns represents biologically independent human samples.

Theoretical Central dogma of brain tumour biology



Supplementary Figure 18 : Central dogma of biology-

The first level is the DNA (in yellow) :in terms of cancers the most common modification is epigenetic modification. This modification can result in silencing of gene (mostly tumor suppressors) by inhibiting it's transcription to mRNA if methylation occurs at promoter region or can result the gene to be constantly active (mostly oncogenes) if the modification occurs on gene body. The DNA is then transcribed to RNA and hence modifications of the DNA have higher probability of being reflected at the RNA level.

The second level is RNA (in green) : Once the DNA is transcribed to RNA, there can be several post-transcriptional modifications such as inhibition or initiation of translation of mRNA to proteins by several microRNAs, lncRNAs or siRNAs. Thus changes at mRNA level have a higher probability of being detected at the proteome level.

The third level is Protein (in light blue) : Here the translated amino acids form disulphide bridges to in turn form proteins which will be responsible for the phenotype of healthy / disease condition. However, the protein structures can also undergo further post-translational modifications such as N-glycosylation or phosphorylation etc, which can indeed affect the stability and structure of proteins. This can result in certain deformed structures which will result in a diseased phenotype.

Hence, not all changes occurring at DNA level will be transmitted to RNA level, and definitely not to the proteome level as there are several steps of modifications or regulations that occur in between.

Similarly, not all changes at RNA level will be reflected at the proteome level due to possible post transcriptional changes. But the proteome level is the final level which is then responsible for the disease or healthy phenotype, and also has the possibility to undergo post translational modifications.

These modifications are known to be more active in case of a disease condition. But it is important to note that some modifications are also essential for normal function and protection from the disease

Parts of art were created with <https://bioicons.com/>. CC 0 · CC by SA

4.2. MB project 2 in original format

Göbel et al. *Acta Neuropathologica Communications* (2023) 11:174 <https://doi.org/10.1186/s40478-023-01654-2>

RESEARCH

Open Access



MYC overexpression and SMARCA4 loss cooperate to drive medulloblastoma formation in mice

Carolin Göbel^{1,2}, Shweta Godbole³, Melanie Schoof^{1,2}, Dörthe Holdhof^{1,2}, Catena Kresbach^{1,2,4},
Carolin Loose^{1,2}, Julia Neumann^{3,4} and Ulrich Schüller^{1,2,4}

Abstract

Group 3 medulloblastoma is one of the most aggressive types of childhood brain tumors. Roughly 30% of cases carry genetic alterations in *MYC*, *SMARCA4*, or both genes combined. While overexpression of *MYC* has previously been shown to drive medulloblastoma formation in mice, the functional significance of *SMARCA4* mutations and their suitability as a therapeutic target remain largely unclear. To address this issue, we combined overexpression of *MYC* with a loss of *SMARCA4* in granule cell precursors. Both alterations did not increase proliferation of granule cell precursors in vitro. However, combined *MYC* overexpression and *SMARCA4* loss successfully induced tumor formation in vivo after orthotopic transplantation in recipient mice. Resulting tumors displayed anaplastic histology and exclusively consisted of *SMARCA4*-negative cells although a mixture of recombined and non-recombined cells was injected. These observations provide first evidence for a tumor-promoting role of a *SMARCA4* deficiency in the development of medulloblastoma. In comparing the transcriptome of tumors to the cells of origin and an established Sonic Hedgehog medulloblastoma model, we gathered first hints on deregulated gene expression that could be specifically involved in *SMARCA4*/*MYC* driven tumorigenesis. Finally, an integration of RNA sequencing and DNA methylation data of murine tumors with human samples revealed a high resemblance to human Group 3 medulloblastoma on the molecular level. Altogether, the development of *SMARCA4*-deficient medulloblastomas in mice paves the way to deciphering the role of frequently occurring *SMARCA4* alterations in Group 3 medulloblastoma with the perspective to explore targeted therapeutic options.

Keywords Group 3 medulloblastoma, *MYC*, BRG1, BAF complex, Chromatin remodeling

*Correspondence:

Ulrich Schüller
u.schueller@uke.de

¹Department of Pediatric Hematology and Oncology, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, Hamburg 20251, Germany

²Research Institute Children's Cancer Center Hamburg, Martinistrasse 52, Building N63 (LIV), Hamburg D-20251, Germany

³Center for Molecular Neurobiology, Falkenried 94, Hamburg 20251, Germany

⁴Institute of Neuropathology, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, Hamburg 20251, Germany



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

The BAF (BRG1/BRM-associated factor) chromatin remodeling complex greatly influences gene expression in mammals by regulating accessibility of DNA regions for the binding of transcription factors [2, 12]. Its catalytic activity depends on the presence of one of the mutually exclusive ATPase subunits SMARCA2 or SMARCA4 (SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily A, member 4; also known as BRG1 [BRAHMA related gene 1]) [31, 39]. In contrast to SMARCA2, SMARCA4 has proven indispensable in embryonic development as *Smarca4*-deficient mouse blastocysts die during the peri-implantation stage [8, 53]. Moreover, several mouse models have confirmed that functional SMARCA4 is essential for cerebellar development, neurogenesis, and gliogenesis [24, 25, 40, 43]. Consequently, alterations in the *SMARCA4* gene have been associated with a variety of intellectual disorders such as the Coffin-Siris syndrome and autism spectrum disorders [6, 68]. Moreover, deleterious *SMARCA4* mutations can be found throughout various cancer entities, including small cell carcinoma of the ovary, hypercalcemic type (SCCOHT), non-small cell lung cancer (NSCLC), pancreatic cancer, hepatocellular carcinoma, head and neck cancer, and atypical/teratoid rhabdoid tumors (ATRT) [15, 23, 27, 52]. In NSCLC and ATRT, *SMARCA4* alterations are associated with a significantly worse prognosis than *SMARCA4* wild-type cases [15, 20]. Medulloblastomas (MB), the most common malignant pediatric brain tumors, mainly show somatic heterozygous missense mutations of *SMARCA4*, which are suggested to have a dominant-negative effect resulting in a loss of function [15, 29, 44, 55]. MB can be divided into four main molecular subgroups according to their transcriptome and global DNA methylation: Sonic Hedgehog (SHH), Wingless/Int-1 (WNT), Group 3, and Group 4 [9, 65]. Alterations of *SMARCA4* mostly affect WNT and Group 3 MB, occurring in around 20% and 9-15% of cases, respectively, which places it among the most frequently mutated genes in both subgroups [17, 44]. However, the functional significance of these *SMARCA4* alterations in tumor development remains unknown. In this study, we focused on Group 3 MB, which mostly affect younger children and infants and show the worst prognosis of all subgroups with a median 5-year survival below 60% [10, 17, 30, 65]. Therefore, effective treatment regimens including targeted therapies are urgently needed. Besides *SMARCA4* mutations, recurrent alterations in Group 3 MB include *MYC* amplifications in 15-20% of cases, which correlate with poor survival [10, 17, 34]. *MYC* and *SMARCA4* can also be concurrently altered as detected in around 1-6% of Group 3 MB [17, 29, 44, 55]. Several mouse models have convincingly demonstrated a tumor-driving role of *MYC*

in the development of Group 3 MB [5, 32, 33, 47, 64]. However, none of the previously developed Group 3 MB mouse models include alterations of *Smarca4*. In this study, we present a new MB mouse model with combined *MYC* overexpression and *SMARCA4* loss in granule cell precursors (GCPs) and provide evidence for a tumor-promoting role of a *SMARCA4* deficiency in MB.

Materials and methods

Transgenic animals

All experimental procedures on animals were approved by the Government of Hamburg, Germany (N113/16, N050/2018, N099/2019) and were performed according to national regulations. Mice were kept on a 12 h dark/light cycle, and water and food were available *ad libitum*. Animals of both sexes were used for experiments. The strain *Smarca4^{fl/fl}* (also known as *Brg1^{fl/fl}*) has been previously generated and described [28, 62], *Math1-creER^{T2}* mice were obtained from Jackson Laboratories, ME, USA (#7684) [37], and *CD1^{nu/nu}* mice were obtained from Charles River Laboratories, MA, USA (#086) [45]. *Math1creER^{T2}* and *Smarca4^{fl/fl}* mice were maintained on a *C57Bl6/J* background. Genotyping was performed by PCR using genomic DNA from ear or tail biopsies with the following primer pairs (5'-3'): *cre* (fw): TCCGGGCTGCCACGACCAA, *cre* (rv): GGCGCGGCAACACCATTTT, *Smarca4* floxed (fw): GTCATACTTATGTCATAGCC, *Smarca4* floxed (rv): GCCTTGCTCAAAGTATAAG, *Smarca4* recombined (fw): GATCAGCTCATGCCCTAAGG, *Smarca4* recombined (rv): GCCTTGCTCAAAGTATAAG. To induce *Smarca4* recombination in *Math1-creER^{T2}::Smarca4^{fl/fl}* mice, pups received a single dose of 0.4 mg tamoxifen dissolved in corn oil by intraperitoneal injection at postnatal day 3 (P3).

Lentivirus production

A lentiviral plasmid driving overexpression of *MYC* was generated by cloning the murine *Myc* gene from a previously described MSCV-MYC-IRES-RFP construct [32] into a self-designed lentiviral expression vector backbone (pLV-CMV-IRES-GFP) ordered from VectorBuilder, IL, USA. Production and titration of second generation lentiviral particles was performed by transfection of HEK293T cells as previously described [57]. Viral particles were concentrated by ultracentrifugation and stored at -80 °C before transduction.

Culture of granule cell precursors (GCPs)

Primary murine GCPs were isolated from *Math1-creER^{T2}::Smarca4^{fl/fl}* or *Smarca4^{fl/fl}* pups at P7 or P8 as previously described [42]. Lentiviral transduction of GCPs with *MYC* (pLV-CMV-MYC-IRES-GFP) or Mock (pLV-CMV-IRES-GFP) constructs was

performed 4 h after isolation with addition of protamine sulfate (8 µg/ml) and centrifugation at 2,000 rpm for 1 h. Medium was changed the next morning with concurrent exchange of FCS-supplemented medium to serum-free medium containing 3 µg/mL SHH protein. Bromodeoxy- uridine (BrdU) was added to the cells at a concentration of 25 µg/mL for 2 h before fixation of cells. For orthotopic transplantation, transduced GCPs were dissociated with Accutase 24 h after isolation and were washed and resuspended in a solution of 3:1 medium and Matrigel on ice.

Stereotactic transplantations

During stereotactic transplantations, recipient mice (6-week-old *CD1^{nu/nu}*) were anesthetized by isoflurane inhalation. They additionally received analgesia by subcutaneous injections of carprofen (6 mg/kg) before transplantation and on the day after. For the procedure, mice were placed in a stereotactic frame (David Kopf Instruments, CA, USA) on a heating pad, and eye ointment was applied to avoid dehydration. Local anesthesia (2% lidocaine) was applied before performing a skin incision and puncturing the skull for injection. A total of 1.5×10^6 cells were injected using a Hamilton syringe (World Precision Instruments, FL, USA) at coordinates x: +1 mm, y: -1 mm, and z: -2 mm from the lambda suture at 30° from the skull surface. Mice were monitored daily for any sign of tumor development within the following six months.

Immunohistochemistry (IHC)

For histological examination of brains, tissue was fixed in 4% formaldehyde for at least 12 h. The tissue was dehydrated, embedded in paraffin, and sectioned at 2 µm according to standard protocols. Hematoxylin and eosin (HE) stainings were applied according to standard protocols. 3,3'-Diaminobenzidine (DAB) stainings were performed on a Ventana Benchmark system using the ultraView or OptiView DAB detection kit (all Roche Diagnostics, Basel, CH). The following antibodies were used: Cleaved Caspase-3 (CC-3): Cell Signaling #9664, RRID:AB_2070042 (1:100); GFP: Abcam #ab290, RRID:AB_303395 (1:500); Ki67: Abcam #ab15580, RRID:AB_443209 (1:100); MYC: Zeta Corporation #Z2734RL (1:25); Nestin: Abcam #ab221660, RRID:AB_2909415 (1:2000); NeuN: Merck #MAB377, RRID:AB_2298772 (1:50); OLIG2: Merck #AB9610, RRID:AB_570666 (1:200); SMARCA4: Abcam #ab110641, RRID:AB_10861578 (1:25); and SOX2: Abcam #92,494, RRID:AB_10585428 (1:200).

Immunofluorescence (IF) stainings

IF stainings of formalin-fixed paraffin-embedded (FFPE) tissue were performed manually after deparaffinization and antigen retrieval with citrate buffer. For IF staining of GCPs in vitro, cells were fixed with 4% formaldehyde

for 10 min. In case of BrdU stainings, acidic pre-treatment (4 N HCl and 0.1 M sodium borate for 10 min each) was performed before blocking with 10% NGS in 0.3% Triton X-100. The following primary antibodies were used for incubation at 4 °C over night: BrdU: Invitrogen #B35128, RRID:AB_2536432 (1:100); MYC: Cell Signaling #5605, RRID:AB_1903938 (1:800); GFP (mouse): Invitrogen #A11120, RRID:AB_221568 (1:100); GFP (rabbit): Invitrogen #A11122, RRID:AB_221569 (1:100); and SMARCA4: Abcam #ab110641, RRID:AB_10861578 (1:25). Secondary antibodies (1:500) and DAPI (1 µg/ml) were added for 1 h at room temperature on the next day: anti-mouse Alexa 488: Cell Signaling Technology #4408S, anti-mouse Alexa 555: Cell Signaling Technology #4409S, anti-rabbit Alexa 488: Cell Signaling Technology #4412S, and anti-rabbit Alexa 546: Invitrogen #A11035.

Image quantifications

IF stainings of GCPs were quantified automatically using the *Automatic Measurement* tool of the NIS-Elements (AR 5.11.03) software. The threshold for fluorescence intensity and cell size was adjusted separately for each fluorescence channel and was applied to all samples to retrieve cell counts. At least three representative images were analyzed for each sample and staining. DAB stainings of tumors (MYC and GFP) were quantified with Image J (v 1.48a). All evaluated stainings were performed with the automated Ventana system and within the same run to ensure comparability of detected signals. Five pictures were taken from different areas within the tumor, DAB color deconvolution was applied, and resulting images (Color 2) were converted into 8-bit format. Masks with the following black/white thresholds were applied before measuring the corresponding area fraction: high signal: 0-125, medium signal: 125-150, low signal: 150-175, no signal: 175-255.

Western blot

For Western blotting, 30 µg of protein per sample were separated by SDS-PAGE (4-10% gradient) and were transferred onto a nitrocellulose membrane. After blocking with 5% milk powder in TBS-Tween, the membrane was incubated with the primary antibody at 4 °C overnight. The following antibodies were used: β-tubulin: Sigma-Aldrich #T4026, RRID:AB_477577 (1:500); GAPDH: GeneTex #100,118, RRID:AB_1080976; MYC: Cell Signaling #5605, RRID:AB_1903938 (1:500); and SMARCA4: Abcam #ab110641, RRID:AB_10861578 (1:10,000). After washing, the secondary horse-radish peroxidase (HRP) coupled antibody was applied for 1 h at room temperature: Goat-anti-mouse-HRP: Dako #P0447 (1:10,000) or Goat-anti-rabbit-HRP: Dako #P0448 (1:10,000). The Clarity Western ECL Substrate (Bio-Rad

Laboratories Inc, CA, USA) was used for detection on X-ray films.

RNA sequencing analysis

RNA Isolation from FFPE tissue was performed using the Maxwell RSC RNA FFPE kit (Promega Corporation, WI, USA). Prior to sequencing, RNA concentration and integrity was determined on an RNA 6000 Nano Chip on the Agilent 2100 Bioanalyzer system (Agilent Technologies, CA, USA). At least 100 ng total RNA per sample were used for sequencing. Ribosomal RNA was depleted with the RiboCop Human/Mouse/Rat V2 kit before library preparation with the CORALL Total RNA-seq V2 kit (both Lexogen GmbH, Vienna, AT). Pooled libraries were sequenced on a NextSeq500 sequencing system (Illumina, CA, USA) by 1 × 75 bp single-end sequencing for 75 cycles, generating at least 30 Mio reads per sample. Raw fastq files of mouse samples were processed in usegalaxy.eu [1]. Low quality reads were detected using *FastQC* (Galaxy Version 0.73 + galaxy0), and reads with average quality < 20 were trimmed with *Trimmomatic* (Galaxy Version 0.38.1). Reads were aligned to the GRCm39 (mm39) mouse reference genome using *STAR aligner* (Galaxy Version 2.7.8a + galaxy1). Gene expression was quantified with *featureCounts* (Galaxy Version 2.0.1 + galaxy2), and VST-normalized files were generated by *DESeq2* (Galaxy Version 2.11.40.7 + galaxy2). Further processing of data was performed with R (4.2.1).

Differential gene expression analysis between mouse samples was performed using *limma* (3.52.2) [54]. Genes orthologous to humans were used for volcano plots generated with *ggplot2* (3.4.1) with genes considered differentially expressed if $\text{LogFC} \geq 2.5$ and False Discovery Rate (FDR) adjusted $p \leq 0.01$. For gene set enrichment analysis, all mouse genes with $\text{LogFC} \geq 1.5$ and FDR adjusted $p \leq 0.01$ were considered using multiple packages from *clusterProfiler* (4.4.4) visualized with in-built *clusterProfiler* plots.

Human gene expression data were obtained from a previously published pediatric brain tumor cohort (Sturm et al. 2016 [61]; GSE73038). To compare mouse and human gene expression data, 14,151 orthologous genes between both datasets were used, and data were batch corrected for species differences using an in-house pipeline. The previously identified 14,151 orthologous genes were used for differential gene expression analysis between human tumor subtypes using *limma* (3.52.2) [54]. The 6,000 most differentially expressed genes (or 5,000 for MB only) were selected using Benjamini-Hochberg correction for multiple testing and sorting by F-statistic. Visualizations were performed using RStudio packages *umap* (0.2.9.0)

[41] and *Complex Heatmap* (2.12.1) [18] using Euclidean distance and Ward.D2 linkage for clustering. For the distance plots, Euclidean distance was measured (*Stats*

4.1.2 package), and plots were generated with *Complex Heatmap*.

DNA methylation analysis

DNA from frozen tumor biopsies (tumors 3 + 4) was isolated using the NucleoSpin Tissue kit (Macherey-Nagel, Düren, DE), whereas DNA isolation from FFPE tissue (tumor 1) was performed using the Maxwell RSC DNA FFPE kit (Promega Corporation). At least 150 ng of total DNA were used for bisulfite conversion with the EZ DNA Methylation kit (Zymo Research, CA, USA). Then, samples were analyzed on the Infinium Mouse Methylation BeadChip array covering > 285,000 CpG sites within the mouse genome on an iScan array scanner (both Illumina). Human tumor samples were analyzed on the MethylationEPIC 850k BeadChip array (Illumina). The use of biopsy-specimens for research upon anonymization was always in accordance with local ethical standards and regulations at the University Medical Center Hamburg-Eppendorf.

Data processing and analysis was performed with R (4.1.2). For preprocessing of raw data and extraction of beta values, the *Minfi* package [3] was used for human data, whereas the *SeSAMe* package [70] was used for mouse data. Then, quantile normalization of data was performed. For a comparison of murine samples to human brain tumor DNA methylation profiles, previously published data by Capper et al. [9] and Sharma et al. [59] were combined with data generated in-house (in total $n = 228$). Within the human dataset including all brain tumor entities, the 15,000 most differentially methylated CpG sites were identified. Out of these, 491 CpGs that are orthologous between the human and mouse genome were chosen for further analysis. Human and mouse datasets were combined and again, quantile normalization was performed. UMAPs [41] as well as hierarchically clustered heatmaps (*Complex Heatmap* package [18]) were generated based on the differential methylation of the previously chosen 491 CpGs. For the generation of distance plots, Pearson correlation (*Stats* 4.1.2 package) was applied, and plots were generated with the *Complex Heatmap* package.

Statistical analysis

All statistical analysis was performed using the GraphPad Prism (9.4.1) or R (4.1.2) software. The statistical tests applied to the data shown are stated in the respective figure legends. For each comparison, at least $n = 3$ samples per group were used and/or $n = 3$ independent experiments were conducted. P-values were corrected for multiple testing. All graphs depict mean values \pm standard deviation.

Results

Loss of SMARCA4 or MYC overexpression does not increase proliferation of granule cell precursors (GCPs) in vitro

In a first step, we investigated the influence of both *Smarca4* and *Myc* alterations on cell behavior in vitro. To induce a loss of SMARCA4 in GCPs, *Math1-creER^{T2}::Smarca4^{fl/fl}* mice received a single dose of tamoxifen at P3, and GCPs were isolated from the cerebellum when pups reached an age of 7-8 days. Successful knockdown of SMARCA4 was detected in around 50% of cells as shown in Western Blot and IF stainings (Fig. 1A-C). Proliferation was significantly decreased

in SMARCA4-negative cells at day 1 in culture, while no significant difference in proliferation was observed at day 3 or 5 in culture (Fig. 1D). Next, we analyzed the effect of MYC overexpression in GCPs by transduction with a lentiviral MYC-GFP construct. Successful transduction was validated by the presence of MYC protein in Western Blot and by positive GFP IF stainings with mean transduction rates ranging between 15.5 and 22.6% (Fig. 1E-G). Overall proliferation of non-induced *Math1-creER^{T2}::Smarca4^{fl/fl}* GCPs after transduction with MYC virus showed no difference compared to proliferation of cells transduced with a Mock-GFP construct

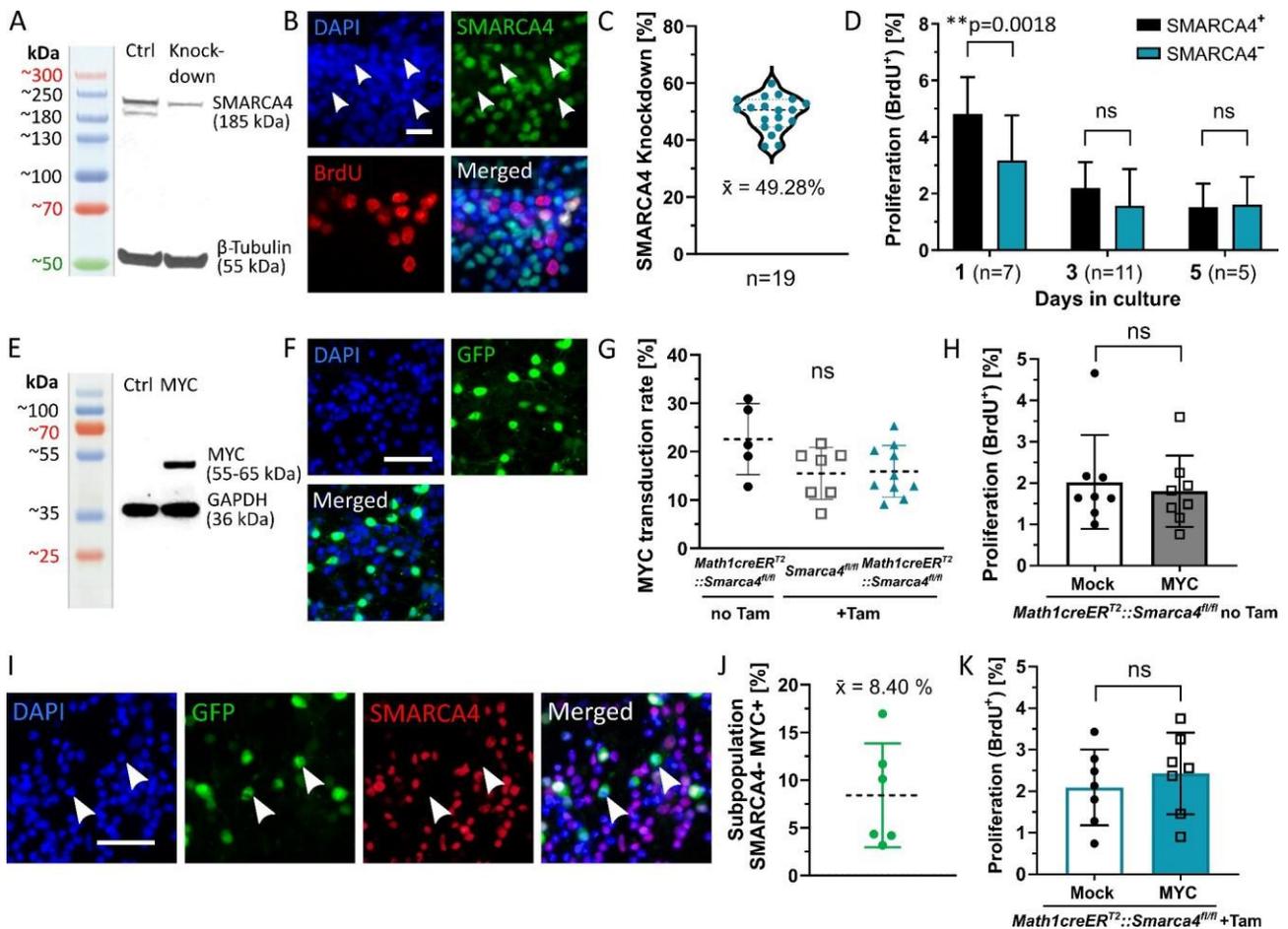


Fig. 1 Loss of SMARCA4 or MYC overexpression does not increase proliferation of GCPs in vitro. **(A)** Tamoxifen-induced knockdown of SMARCA4 is evident in Western Blot of P7/8 *Math1creER^{T2}::Smarca4^{fl/fl}* GCPs compared to controls (*Smarca4^{fl/fl}*) after tamoxifen injection at P3. Two SMARCA4 bands are detected as seen in previously published studies [19, 46]. **(B)** IF staining of knockdown GCPs at day 3 in culture shows loss of SMARCA4 protein and proliferation indicated by BrdU incorporation. White arrowheads mark SMARCA4-negative areas. **(C)** Evaluation of SMARCA4 knockdown in IF on day 3 in culture of 19 independent GCP cultures. **(D)** Proliferation as measured by BrdU incorporation in IF on day 1, 3, and 5 in culture, separately counted for SMARCA4-positive and -negative GCPs in knockdown cultures. Two-tailed paired t-tests were applied. **(E)** MYC expression is evident in Western Blot of wild-type P7/8 GCPs 72 h after transduction. **(F)** IF staining shows GFP signal 72 h after transduction of GCPs. **(G)** MYC transduction rates were evaluated in IF stainings of GCPs 72 h after transduction. The three groups include GCPs without tamoxifen (Tam) induction and GCPs of cre-negative (*Smarca4^{fl/fl}*) and cre-positive (*Math1creER^{T2}::Smarca4^{fl/fl}*) genotype after tamoxifen induction at P3. Tukey's multiple comparisons test was applied. **(H)** Overall proliferation as measured by BrdU incorporation in IF of *Math1creER^{T2}::Smarca4^{fl/fl}* GCPs without tamoxifen induction 72 h after transduction with Mock or MYC constructs. Paired two-tailed t-test was applied. **(I, J)** IF staining of tamoxifen-induced *Math1creER^{T2}::Smarca4^{fl/fl}* GCPs 72 h after transduction with MYC virus. The subpopulation with SMARCA4 protein loss and GFP signal (white arrowheads) constitutes around 8.4% of the whole cell culture. **(K)** Overall proliferation of tamoxifen-induced *Math1creER^{T2}::Smarca4^{fl/fl}* GCPs 72 h after transduction with Mock or MYC constructs. Paired two-tailed t-test was applied. Scale bar in B corresponds to 20 μ m, scale bars in F+I correspond to 50 μ m

(Fig. 1H). Subsequently, we combined both SMARCA4 loss and MYC overexpression by transducing tamoxifen-induced *Math1creER^{T2}::Smarca4^{fl/fl}* GCPs. As shown in Fig. 1I-J, the subpopulation of SMARCA4-deficient GFP-positive GCPs constituted around 8.4% of the whole cell culture. Again, overall proliferation was not significantly increased after MYC transduction (Fig. 1K). However, proliferation of the SMARCA4-deficient and successfully transduced subpopulation could not be analyzed separately since acidic pre-treatment required for BrdU stainings destroys GFP epitopes [7].

Loss of SMARCA4 and MYC overexpression cooperate to drive brain tumor formation in vivo

In a next step, we transplanted altered GCPs into immunodeficient *CD1^{nu/nu}* mice to further explore their tumorigenic potential in vivo. For this purpose, SMARCA4 knockdown GCPs were isolated from induced *Math1creER^{T2}::Smarca4^{fl/fl}* mice and were transduced with a lentiviral MYC construct as described above. On the next day, GCPs were dissociated and transplanted into the cerebella of *CD1^{nu/nu}* mice without pre-sorting

for recombined or transduced cells (Fig. 2A). Within a cohort of 19 transplanted mice, five mice developed a tumor in the cerebellum, presenting with neurological symptoms earliest four weeks and latest five months after transplantation (Fig. 2B). Histologically, tumors presented as a cell dense mass in HE stainings, with regions showing anaplastic features as well as apoptotic areas, consistent with large cell/anaplastic (LCA) histology frequently detected in MYC driven Group 3 MB (Fig. 2C-E) [13, 30]. IHC stainings revealed a loss of SMARCA4 in all tumor cells (Fig. 2F). The presence of recombined *Smarca4* in tumor biopsies was also verified by PCR, which confirmed that the loss of SMARCA4 was caused by genetic recombination (Fig. 2G). Furthermore, tumors stained positive for both GFP and MYC, thereby validating successful transduction with the MYC-GFP construct (Fig. 2H+I). Tumors were highly proliferative according to Ki67 signals and displayed a high degree of apoptosis as indicated by Cleaved Caspase-3 (CC3) staining (Fig. 2J+K). Staining for neural markers revealed scattered expression of SOX2 and Nestin, whereas no signal for NeuN or OLIG2 was detected (Fig. 2L-O). Altogether,

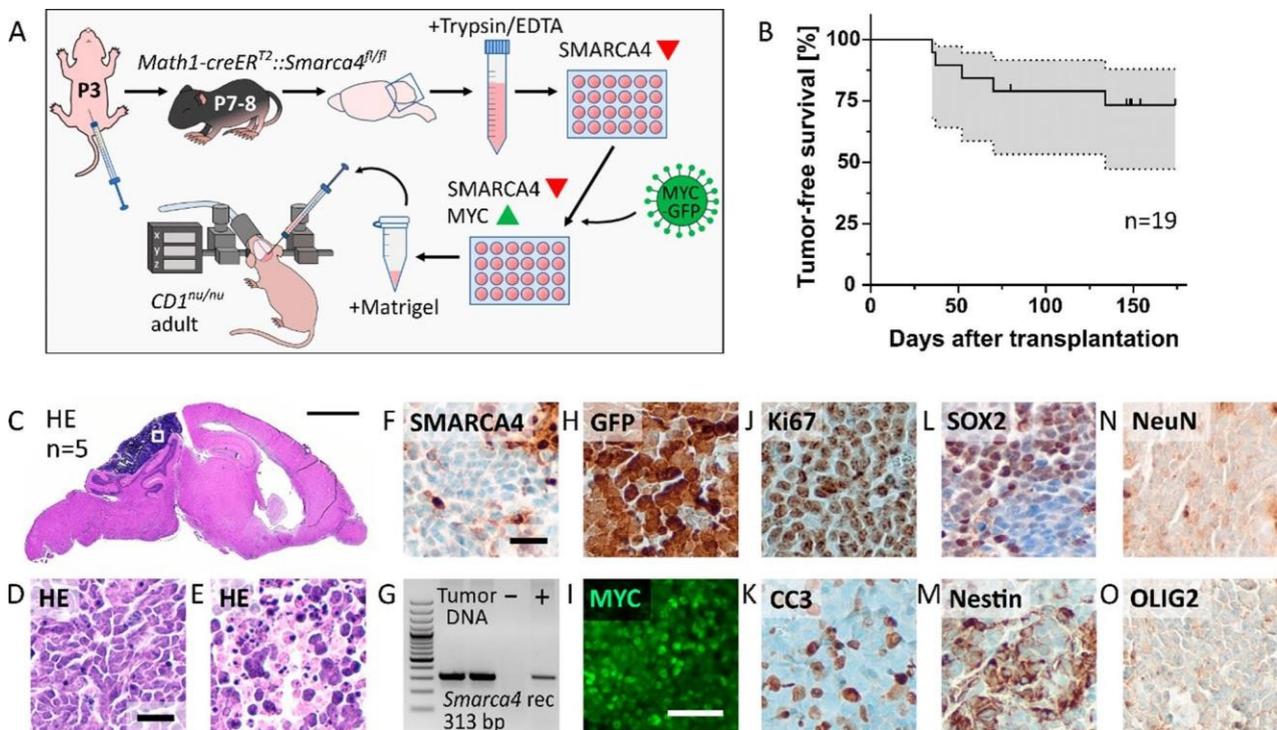


Fig. 2 Loss of SMARCA4 and MYC overexpression cooperate to drive brain tumor formation in vivo. (A) Schematic overview of the cell culture and transplantation protocol for the generation of SMARCA4-deficient MYC-overexpressing tumors. (B) Tumor-free survival of transplanted *CD1^{nu/nu}* mice; grey area represents the 95% confidence interval. Censored mouse at day 80 had to be sacrificed due to illness unrelated to tumor development. (C) Representative HE staining of tumors in the brains of n = 5 transplanted mice in sagittal brain section. (D,E) High-power HE stainings of distinct areas within the tumors showing (D) anaplastic or (E) apoptotic features. (F) Tumors show complete loss of SMARCA4 in IHC interspersed with SMARCA4-positive blood vessels. (G) PCR using DNA isolated from tumor biopsies confirms *Smarca4* recombination on a genetic level. (H-I) Tumors stain positive for (H) GFP and (I) MYC, confirming transduction with the MYC-GFP construct. (J) Tumors are highly proliferative as indicated by Ki67 stainings; (K) with a high degree of apoptosis according to Cleaved Caspase-3 (CC3) signals. (L-O) Tumors show scattered expression of (L) SOX2 and (M) Nestin but no signal for (N) NeuN or (O) OLIG2. Scale bar corresponds to 2 mm in C, to 25 μ m in D+F (also applicable to E, H, J-O), and to 50 μ m in I

these results affirmed the origin of detected tumors in the subpopulation (8.4%) of GCPs harboring both recombinated *Smarca4* and overexpressed MYC and showed proliferative capacity as well as undifferentiated nature of tumors.

Intratumoral heterogeneity of MYC amplification within Group 3 MB has been described as an important factor in metastasis and therapy resistance [51]. Therefore, we analyzed levels of MYC expression in different regions of our tumors, which revealed striking heterogeneity in between but also within samples (Additional File 1, Fig. S1A-C). All tumors contained areas with varying degrees of MYC signal including cells without any MYC signal. In contrast, GFP signals were uniformly high in all tumors, suggesting regulation of MYC expression independent from successful transduction with the MYC-GFP construct (Additional File 1, Fig. S1D-F).

Moreover, we examined brains and spines for leptomeningeal dissemination, which is detected in around 40% of human Group 3 MB and has also been recapitulated in other MYC-driven medulloblastoma models [33, 38]. In our model, we observed leptomeningeal spread within the brain in four out of five tumor-bearing mice, affecting the cerebral cortex, the midbrain, and the brain stem (Additional File 1, Fig. S1G-J). However, we did not detect any dissemination in the spines of affected mice.

Differential gene expression in MYC/SMARCA4 tumors

To characterize MYC/SMARCA4 tumors on a molecular level, we performed RNA sequencing using FFPE biopsy punches of four mouse tumors. As a control, we simultaneously sequenced FFPE-derived RNA of a previously established SHH MB mouse model (*Math1cre::Smo^{fl/wt}* [58]) and of *Math1cre^{ERT2}::Smarca4^{fl/fl}* P7 whole cerebella. The comparison of MYC/SMARCA4 tumors to *Math1cre^{ERT2}::Smarca4^{fl/fl}* cerebella revealed *Myc* as the most significantly upregulated gene in our model (Additional File 1, Fig. S2A; Additional File 2, Table S1). Gene set enrichment analysis revealed downregulation of terms associated with neuronal development and differentiation, while upregulated terms were mainly associated with ribosome biogenesis and ribosomal RNA (rRNA) synthesis and processing, a characteristic hallmark for MYC-driven cancers (Additional File 1, Fig. S2B,C) [66]. Comparison of gene expression profiles of MYC/SMARCA4 tumors to the established SHH MB mouse model again confirmed upregulation of *Myc*, while *MycN* as a target of SHH signaling was significantly downregulated (Fig. 3A; Additional File 2, Table S2). Other downregulated genes included *Atoh1* and *Barhl1*, both markers for granule cells, of which low levels of *BARHL1* have been associated with a less favorable prognosis in MB [50]. On the other hand, *Hoxa5* and *Fabp4*, both associated with increased malignancy in

gliomas, were upregulated in MYC/SMARCA4 tumors [11, 21]. Gene set enrichment analysis revealed downregulated GO terms mostly linked to neuronal development (Fig. 3B). Meanwhile, terms associated with transmembrane transport and synaptic signaling were upregulated in our model (Fig. 3C). Pathway analysis confirmed the downregulation of SHH signaling but also reduction of Notch and PI3K-Akt-mTOR signaling, whereas glycolysis/gluconeogenesis as well as G protein signaling pathways were upregulated in MYC/SMARCA4 tumors (Fig. 3D + E).

MYC/SMARCA4 tumors show molecular resemblance to human Group 3 MB

In a next step, we integrated our RNA sequencing data with previously published gene expression data to test comparability of our murine tumors to human brain tumors. An integration with a data set comprising several pediatric brain tumor entities (Sturm et al. 2016 [61]) revealed resemblance of our model to human MB in both UMAP and Euclidian clustering (Fig. 4A-B). While mouse SHH MB serving as a validation displayed unambiguous proximity to human SHH MB, MYC/SMARCA4 tumors showed similarity to both SHH MB and Group 3/4 MB in both approaches. A distance plot analysis considering mean values for each subgroup indicated closest proximity of both mouse SHH MB and our MYC/SMARCA4 tumors to human SHH MB (Fig. 4C). Based on these results, we further evaluated the similarity to specific MB subgroups by comparing our mouse model exclusively to MB samples. Within the human MB cohort, we again performed gene expression analysis to identify the most differentially expressed genes between MB subgroups. An integration of our mouse data resulted in closest similarity of MYC/SMARCA4 tumors to Group 3 MB in both UMAP and Euclidian clustering, whereas mouse SHH MB reliably clustered with human SHH MB (Fig. 4DE). In both approaches, tumor 3 formed an exception by clustering closely with SHH MB. However, we did not detect apparent differences to the other three samples in histological appearance and levels of MYC or SMARCA4 in this tumor. Distance plot analysis further confirmed closest proximity of MYC/SMARCA4 tumors to Group 3 MB (Fig. 4F).

Human brain tumors and biologically relevant tumor subgroups can be reliably classified according to their DNA methylation profile [9]. Therefore, we additionally isolated DNA of three mouse tumors (tumors 1, 3, and 4 from RNA sequencing analysis) and performed global DNA methylation analysis using the Mouse Methylation Bead Chip. These data were integrated with a human MB dataset comprising in-house analyzed samples and previously published cohorts [9, 59]. UMAP and Euclidian clustering according to differential methylation of 491

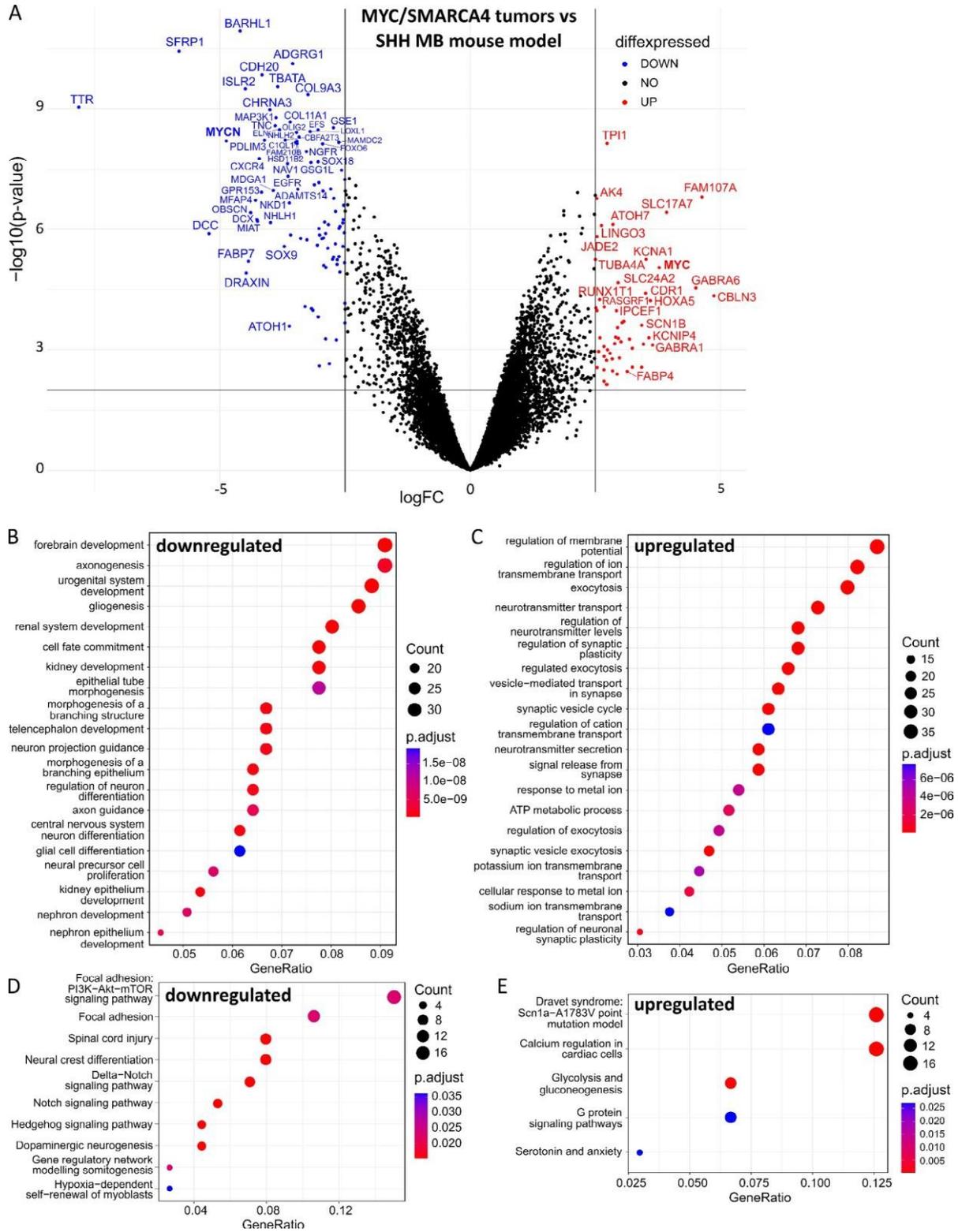


Fig. 3 Differential gene expression of MYC/SMARCA4 tumors compared to an established mouse SHH MB mouse model. (A) Volcano plot depicting differential gene expression between our MYC/SMARCA4 tumor model ($n = 4$) and the *Math1-cre::Smo^{fl/wt}* SHH MB mouse model ($n = 3$) as assessed by RNA sequencing analysis. Only genes orthologous in mice and humans were visualized, and differential expression with $\log_{2}FC \geq 2.5$ and $p \leq 0.01$ was considered significant (blue/red coloring) after Benjamini-Hochberg correction. A detailed list of differentially expressed genes is included in Additional File 2, Table S2. (B,C) Gene set enrichment analysis was performed based on significantly differentially expressed genes considering all mouse genes with $\log_{2}FC \geq 1.5$ and $p \leq 0.01$. (D,E) Deregulated wiki pathways considering differentially expressed genes across all mouse genes with $\log_{2}FC \geq 1.5$ and $p \leq 0.01$

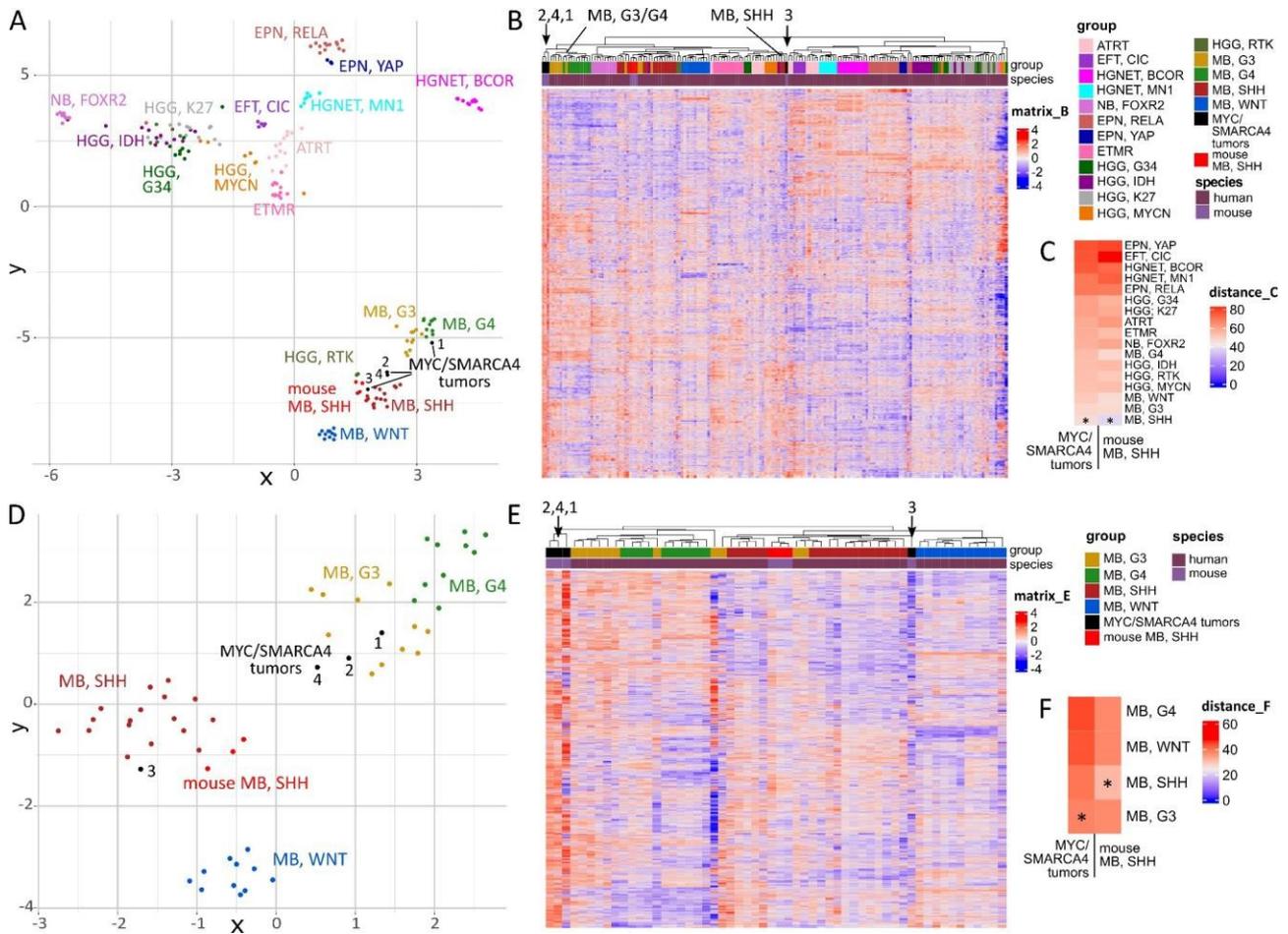


Fig. 4 MYC/SMARCA4 tumors show similarities to Group 3 MB in gene expression analysis. (A) UMAP clustering of mouse tumors profiled by RNA sequencing and published expression data of pediatric brain tumors (Sturm et al. 2016 [61]). Out of the 14,151 orthologous genes identified between both datasets, the 6,000 most differentially expressed genes within the human dataset were used for clustering. Mouse SHH MB show resemblance to their human counterpart, whereas MYC/SMARCA4 tumors display similarity to both SHH MB and Group 3/4 MB. (B) Hierarchical clustering according to differentially expressed genes shows proximity of MYC/SMARCA4 tumors to the Group 3/4 MB cluster for three samples, whereas tumor 3 clusters with a subset of SHH MB (black arrows). (C) Distance plot shows closest resemblance of both mouse tumor models to SHH MB. Asterisks mark shortest distance. (D) UMAP clustering of mouse tumors and human MB subgroups only (Sturm et al. 2016) according to the 5,000 most differentially expressed genes within the human MB dataset out of 14,151 orthologous genes. MYC/SMARCA4 tumors appear closest to Group 3 MB. (E) Hierarchical clustering confirms proximity of MYC/SMARCA4 tumors to the Group 3/4 MB cluster with exception of tumor 3 (black arrows). (F) Distance plot shows closest resemblance of MYC/SMARCA4 tumors to Group 3 MB. EFT, CIC = Ewing sarcoma family tumor with *CIC* alteration; HGNET, BCOR = High-grade neuroepithelial tumor with *BCOR* alteration; NB, FOXR2 = Neuroblastoma with *FOXR2* activation; EPN, RELA = Ependymoma with *RELA* fusion; EPN, YAP = Ependymoma with *YAP* fu- sion; ETMR = Embryonal tumor with multilayered rosettes; HGG, G34 = *H3F3A* G34 mutant high-grade glioma; HGG, IDH = *IDH* mutant high-grade glioma; HGG, K27 = *H3F3A* K27 mutant diffuse midline glioma; HGG, MYCN = *MYCN*-amplified high-grade glioma; HGG, RTK = *IDH/H3F3A* wild-type high-grade glioma of the receptor tyrosine kinase (RTK) subtype; MB, G3 = MB, Group 3; MB, G4 = MB, Group 4

orthologous CpG sites showed good separation of human MB subgroups, with MYC/SMARCA4 tumors clustering in close proximity to Group 3/4 MB (Fig. 5A-B). A distance plot confirmed highest resemblance of MYC/ SMARCA4 tumors to Group 3 MB (Fig. 5C).

Discussion

In this study, we successfully generated SMARCA4-deficient tumors in mice resembling human Group 3 MB both histologically and molecularly. Although SMARCA4 loss or MYC overexpression did not increase proliferation

of GCPs in vitro, the combination of both alterations induced tumor formation after orthotopic transplantation in vivo. An important role of altered SMARCA4 in MB development was suspected before since the overexpression of SMARCA4 wild-type represses tumor development in an OTX2/MYC Group 3 MB mouse model [5]. Our study now confirmed these assumptions by showing a selection for SMARCA4-deficient cells in all detected MYC/SMARCA4 tumors.

On its own, a loss of SMARCA4 in GCPs does not harbor tumorigenic potential as indicated by decreased

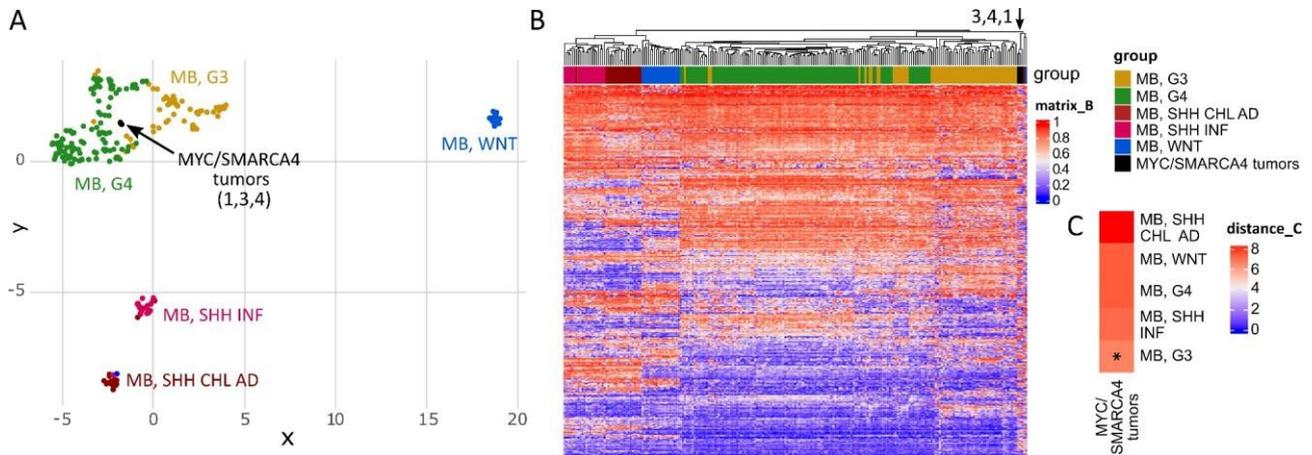


Fig. 5 MYC/SMARCA4 tumors show similarities to Group 3/4 MB in DNA methylation analysis. (A) UMAP clustering according to DNA methylation of mouse tumors (Mouse Methylation BeadChip) and human MB (Capper et al. 2018 [9], Sharma et al. 2019 [59], and in-house analyzed samples, n = 228) using 491 orthologous CpG sites out of the 15,000 most differentially methylated CpG sites within the human dataset. Mouse MYC/SMARCA4 tumors (n = 3) show most similarity to MB, Group 3/4. (B) Heatmap clustering according to DNA methylation of the same samples and CpG sites similarly shows proximity of the MYC/SMARCA4 tumors to MB, Group 3/4 (black arrow). (C) Distance Plot using the mean methylation values summarized for every sub- group shows lowest distance of MYC/SMARCA4 tumors to MB, Group 3. MB, G3 = MB, Group 3; MB, G4 = MB, Group 4; MB, SHH CHL AD = Medulloblastoma SHH-activated (children and adults); MB, SHH INF = Medulloblastoma SHH-activated (infants)

proliferation of SMARCA4-deficient GCPs in vitro. This observation could be attributed to the previously described failure of SMARCA4-deficient GCPs to respond to SHH protein, which is added to cell cultures as a mitogen [69]. Moreover, we have shown before that a postnatally induced loss of SMARCA4 in *Math1-creER^{T2}::Smarca4^{fl/fl}* mice delays migration of GCPs to the internal granular layer in vivo but does not affect the cerebellar phenotype seen later in development [26]. Similarly, overexpression of MYC alone did not increase proliferation of GCPs in vitro. In contrast, Pei et al. have shown higher proliferation and increased ability to form neurospheres after transducing cerebellar stem cells with a stabilized MYC^{T58A} construct [47]. Lentiviral transduction of SOX2-positive cerebellar progenitors with MYC^{T58A} is even sufficient to drive formation of Group 3-like MB in mice [64]. However, the choice of a wild-type MYC construct in our study could play a crucial role. Kawachi et al. did not detect development of MB after overexpression of wild-type MYC alone by *in utero* electroporation [33]. Moreover, Swartling et al. have shown that overexpression of stabilized MYC^{T58A} in neural stem cells results in the development of brain tumors, while overexpression of wild-type MYC does not [63]. Consequently, aberrant chromatin remodeling by the loss of SMARCA4 in our model might cause stabilization of wild-type MYC required for the development of tumors.

The fact that MYC/SMARCA4 tumors did not only show high resemblance to the transcriptome of Group 3 MB but also displayed similarities to SHH MB could be attributed to the cellular origin of our tumors. SHH

MB are derived from GCPs as previously demonstrated in several mouse models and confirmed by comparisons to single-cell RNA sequencing data of murine and human cell populations [4, 58, 60, 67]. In our model, we specifically targeted *Math1*-positive GCPs by tamoxifen-induced *Smarca4* recombination at P3. GCPs are among many other neural progenitor populations that have been used before to model Group 3 MB in mice [33, 38, 64]. This fits to the fact that the exact cellular origin of Group 3 MB cannot be clearly assigned to a single murine cell population in the brain [67]. Indeed, recently published work provides evidence for both Group 3 and 4 MB originating from a distinct cell population in the subventricular zone of the human rhombic lip that does not exist in mice [22, 35, 60]. This divergence from previously used cells of origin should be considered in future attempts at modeling Group 3 MB in mice.

Nevertheless, SMARCA4-deficient MB mouse models could provide a valuable platform to explore targeted therapeutic options for affected patients. For now, the limited penetrance of our tumor model restricts its suitability for such studies. Pre-sorting for successfully transduced cells before transplantation could increase the fraction of MYC-overexpressing SMARCA4-deficient cells within the injected mixture, possibly also enhancing engraftment. However, this would also entail one more day of *in vitro* culture before transplantation for the GFP signal to be detectable. Consequently, fragile SMARCA4-deficient GCP cultures might show reduced viability and proliferative capacity by then. Alternative approaches include the introduction of an additional SMARCA4 deficiency in a recently developed transgenic MYC

driven MB mouse model or the use of other promoters such as *B1bp-cre* or *GFAP-cre* to drive earlier deletion of *SMARCA4* [38].

In comparing gene expression profiles of our *MYC*/*SMARCA4* tumors to an established *SHH* MB model, we identified upregulation of G protein signaling and glucose metabolism in our tumor model. Tao et al. have previously shown altered glucose metabolic pathways in a *MYC* driven MB mouse model and were successful in treating tumor cells with specific inhibitors of upregulated lactate dehydrogenase A [64]. Furthermore, several studies have suggested histone deacetylase (HDAC) inhibitors for treating *MYC*-driven Group 3 MB with efficacy shown both in cell lines in vitro and in mouse models in vivo [14, 16, 36, 48, 49]. It might be of great interest to explore similar treatment regimens in *SMARCA4*-

deficient MB, especially since the response could differ

significantly. For example, Romero et al. have shown that *SMARCA4*-deficient lung cancer cells do not respond to HDAC inhibition but in contrast are sensitive to inhibition of the demethylases *KDM6A/B*, even if *MYC* is concurrently amplified [56]. This observation emphasizes the importance of considering alternative treatment options for *SMARCA4*-deficient MB.

Conclusions

For the first time, we showed cooperative effects between *MYC* overexpression and a *SMARCA4* loss in driving tumorigenesis in cerebellar precursors. Tumors displayed histological and molecular resemblance to Group 3 MB with a distinct selection for *SMARCA4*-deficient cells. Taken together, these findings provide evidence for a tumor-promoting role of a *SMARCA4* deficiency in Group 3 MB. Consequently, our observations pave the way for further investigations on *SMARCA4*-deficient MB mouse models with the potential to identify therapeutic targets specific to these frequently occurring alterations.

List of Abbreviations

ATRT	Atypical/teratoid rhabdoid tumor
BAF	BRG1 / BRM-associated factor
BrdU	Bromodeoxyuridine
BRG1	BRAHMA related gene 1
CC3	Cleaved-Caspase 3
FDR	False Discovery Rate
FFPE	Formalin-fixed paraffin-embedded
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GCP(s)	Granule cell precursor(s)
IF	Immunofluorescence
IHC	Immunohistochemistry
LCA	Large cell/anaplastic
MB	Medulloblastoma
NSCLC	Non-small cell lung cancer
OLIG2	Oligodendrocyte transcription factor 2
rRNA	Ribosomal RNA
SCCOHT	Small cell carcinoma of the ovary, hypercalcemic type
SHH	Sonic Hedgehog

SMARCA4	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily A, member 4
SOX2	Sex determining region Y
Tam	Tamoxifen
UMAP	Uniform Manifold Approximation and Projection for Dimension Reduction
WNT	Wingless/Int-1

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40478-023-01654-2>.

Additional File 1

Additional File 2

Acknowledgements

We thank Vanessa Thaden, Jacqueline Tischendorf, Margarethe Gregersen (all University Medical Center Hamburg-Eppendorf, Germany), and Kristin Hartmann (Mouse Pathology Facility, University Medical Center Hamburg-Eppendorf) for excellent technical support. We further thank Dr Pierre Chambon (IGBMC, Illkirch-Graffenstaden, France) for providing the *Smarca4* mice and Dr Daisuke Kawauchi (National Cancer Center Hospital, Tokyo, Japan) for providing the *MSCV-MYC-IRES-RFP* construct. Finally, we acknowledge the support of the Next Generation Sequencing platform and the small animal models core facility of the Leibniz Institute of Virology, Hamburg, Germany.

Authors' contributions

Conceptualization: C.G., D.H., U.S.; Investigation: C.G.; Formal analysis: C.G., S.G., M.S.; Methodology: C.G., S.G., M.S., D.H., C.K., C.L., J.N.; Software: S.G., M.S.; Visualization: C.G., S.G., M.S.; Writing - original draft: C.G.; Writing - review & editing: C.G., S.G., M.S., D.H., C.K., C.L., J.N., U.S.; Supervision: U.S., Project administration: U.S., Funding acquisition: U.S.

Funding

This work was supported by the Deutsche Krebshilfe, Bonn, Germany (#70113754). Ulrich Schüller was additionally supported by the Fördergemeinschaft Kinderkrebszentrum Hamburg, Hamburg, Germany. Open Access funding enabled and organized by Projekt DEAL.

Data Availability

The datasets generated and analyzed within this study are available in the GEO repository, <https://www.ncbi.nlm.nih.gov/geo/>, accession numbers GSE235625 (RNA sequencing data) and GSE235924 (DNA methylation data). Other data and material described in this study are available from the corresponding author upon request.

Declarations

Ethics approval and consent to participate

All experimental procedures on animals were approved by the Government of Hamburg, Germany (N113/16, N050/2018, N099/2019) and were performed according to national regulations.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 12 July 2023 / Accepted: 17 September 2023

Published online: 02 November 2023

References

1. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Cech M, Chilton J, Clements D, Coraor N, Grünig BA et al (2018) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res* 46:537–544. <https://doi.org/10.1093/nar/ekv379>

2. Alfert A, Moreno N, Kerl K (2019) The BAF complex in development and disease. *Epigenetics Chromatin* 12:19. <https://doi.org/10.1186/s13072-019-0264-y>
3. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA (2014) Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* 30:1363–1369. <https://doi.org/10.1093/bioinformatics/btu049>
4. Ayrault O, Zhao H, Zindy F, Qu C, Sherr CJ, Roussel MF (2010) Atoh1 inhibits neuronal differentiation and collaborates with Gli1 to generate medulloblastoma-initiating cells. *Cancer Res* 70:5618–5627. <https://doi.org/10.1158/0008-5472.Can-09-3740>
5. Ballabio C, Anderle M, Giancesello M, Lago C, Miele E, Cardano M, Aiello G, Piazza S, Caron D, Gianni Fet et al (2020) Modeling medulloblastoma in vivo and with human cerebellar organoids. *Nat Commun* 11:583. <https://doi.org/10.1038/s41467-019-13989-3>
6. Bogershausen N, Wollnik B (2018) Mutational landscapes and phenotypic spectrum of SWI/SNF-Related Intellectual Disability Disorders. *Front Mol Neurosci* 11:252. <https://doi.org/10.3389/fnmol.2018.00252>
7. Boulanger JJ, Staines WA, LeBlanc V, Khoo E-L, Liang J, Messier C (2016) A simple histological technique to improve immunostaining when using DNA denaturation for BrdU labelling. *J Neurosci Methods* 259:40–46. <https://doi.org/10.1016/j.jneumeth.2015.11.006>
8. Bultman S, Gebuhr T, Yee D, La Mantia C, Nicholson J, Gilliam A, Randazzo F, Metzger D, Chambon P, Crabtree G et al (2000) A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. *Mol Cell* 6:1287–1295. [https://doi.org/10.1016/s1097-2765\(00\)00127-1](https://doi.org/10.1016/s1097-2765(00)00127-1)
9. Capper D, Jones DTW, Sill M, Hovestadt V, Schrimpf D, Sturm D, Koelsche C, Sahm F, Chavez L, Reuss DE et al (2018) DNA methylation-based classification of central nervous system tumours. *Nature* 555: 469–474. <https://doi.org/10.1038/nature26000>
10. Cavalli FMG, Remke M, Rampasek L, Peacock J, Shih DJH, Liu B, Garzia L, Torchia J, Nor C, Morrissy AS et al (2017) Intertumoral Heterogeneity within Medulloblastoma Subgroups. *Cancer Cell* 31: 737–754. <https://doi.org/10.1016/j.ccell.2017.05.005>
11. Cimino PJ, Kim Y, Wu HJ, Alexander J, Wirsching HG, Szulzewsky F, Pitter K, Ozawa T, Wang J, Vazquez J et al (2018) Increased HOXA5 expression provides a selective advantage for gain of whole chromosome 7 in IDH wild-type glioblastoma. *Genes Dev* 32:512–523. <https://doi.org/10.1101/gad.312157.118>
12. Clapier CR, Iwasa J, Cairns BR, Peterson CL (2017) Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol* 18:407–422. <https://doi.org/10.1038/nrm.2017.26>
13. Eberhart CG, Kepner JL, Goldthwaite PT, Kun LE, Duffner PK, Friedman HS, Strother DR, Burger PC (2002) Histopathologic grading of medulloblastomas: a Pediatric Oncology Group study. *Cancer* 94:552–560. <https://doi.org/10.1002/cncr.10189>
14. Ecker J, Oehme I, Mazitschek R, Korshunov A, Kool M, Hielscher T, Kiss J, Selt F, Konrad C, Lodrini M (2015) targeting class I histone deacetylase 2 in MYC amplified group 3 medulloblastoma. *Acta Neuropathol Commun* 3: 22. <https://doi.org/10.1186/s40478-015-0201-7>
15. Fernando TM, Piskol R, Bainer R, Sokol ES, Trabucco SE, Zhang Q, Trinh H, Maund S, Kschonsak M, Chaudhuri S et al (2020) Functional characterization of SMARCA4 variants identified by targeted exome-sequencing of 131,668 cancer patients. *Nat Commun* 11:5551. <https://doi.org/10.1038/s41467-020-19402-8>
16. Furchert SE, Lanvers-Kaminsky C, Jürgens H, Jung M, Loidl A, Frühwald MC (2007) Inhibitors of histone deacetylases as potential therapeutic tools for high-risk embryonal tumors of the nervous system of childhood. *Int J Cancer* 120:1787–1794. <https://doi.org/10.1002/ijc.22401>
17. Gajjar A, Robinson GW, Smith KS, Lin T, Merchant TE, Chintagumpala M, Mahajan A, Su J, Bouffet E, Bartels U et al (2021) Outcomes by clinical and molecular features in children with medulloblastoma treated with risk-adapted therapy: results of an International Phase III Trial (SJMB03). *J Clin Oncol* 39:822–835. <https://doi.org/10.1200/JCO.20.01372>
18. Gu Z, Eils R, Schlesner M (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32:2847–2849. <https://doi.org/10.1093/bioinformatics/btw313>
19. Habel N, El-Hachem N, Soysouvanh F, Hadhiri-Bziouche H, Giuliano S, Nguyen S, Horák P, Gay AS, Debayle D, Nottet Net et al (2021) FBXO32 links ubiquitination to epigenetic reprogramming of melanoma cells. *Cell Death Differ* 28:1837–1848. <https://doi.org/10.1038/s41418-020-00710-x>
20. Hasselblatt M, Nagel I, Oyen F, Bartelheim K, Russell RB, Schuller U, Juncker-storff R, Rosenblum M, Alassiri AH, Rossi Set et al (2014) SMARCA4-mutated atypical teratoid/rhabdoid tumors are associated with inherited germline alterations and poor prognosis. *Acta Neuropathol* 128:453–456. <https://doi.org/10.1007/s00401-014-1323-x>
21. He ZC, Liu Q, Yang KD, Chen C, Zhang XN, Wang WY, Zeng H, Wang B, Liu YQ, Luo M (2022) HOXA5 is amplified in glioblastoma stem cells and promotes tumor progression by transcriptionally activating PTPRZ1. *Cancer Lett* 533: e215605. <https://doi.org/10.1016/j.canlet.2022.215605>
22. Hendrikse LD, Haldipur P, Saulnier O, Millman J, Sjoboen AH, Erickson AW, Ong W, Gordon V, Coudière-Morrison L, Mercier AL et al (2022) Failure of human rhombic lip differentiation underlies medulloblastoma formation. *Nature* 609: 1021–1028. <https://doi.org/10.1038/s41586-022-05215-w>
23. Hodges HC, Stanton BZ, Cermakova K, Chang CY, Miller EL, Kirkland JG, Ku WL, Veverka V, Zhao K, Crabtree GR (2018) Dominant-negative SMARCA4 mutants alter the accessibility landscape of tissue-unrestricted enhancers. *Nat Struct Mol Biol* 25:61–72. <https://doi.org/10.1038/s41594-017-0007-3>
24. Holdhof D, Schoof M, Hellwig M, Holdhof NH, Niesen J, Schuller U (2020) hGFAP-Positive stem cells depend on Brg1 for proper formation of cerebellar and cerebellar structures. *Cereb Cortex* 30:1382–1392. <https://doi.org/10.1093/cercor/bhz173>
25. Holdhof D, Schoof M, Al-Kersh S, Spohn M, Kresbach C, Göbel C, Hellwig M, Indenbirken D, Moreno N, Kerl K et al (2021) Brahma-related gene 1 has time-specific roles during brain and eye development. *Development* 148:e196147. <https://doi.org/10.1242/dev.196147>
26. Holdhof D, On JH, Schoof M, Göbel C, Schüller U (2021) Simultaneous Brg1 knockout and MYCN overexpression in cerebellar Granule Neuron Precursors is insufficient to drive tumor formation but temporarily enhances their proliferation and delays their Migration. *Cerebellum* 20:410–419. <https://doi.org/10.1007/s12311-020-01219-2>
27. Holdhof D, Johann PD, Spohn M, Bockmayr M, Safaei S, Joshi P, Masliah-Planchon J, Ho B, Andrianteranagna M, Bourdeau Fet et al (2021) Atypical teratoid/rhabdoid tumors (ATRTs) with SMARCA4 mutation are molecularly distinct from SMARCB1-deficient cases. *Acta Neuropathol* 141:291–301. <https://doi.org/10.1007/s00401-020-02250-7>
28. Indra AK, Dupé Vr, Bornert J-M, Messaddeq N, Yaniv M, Mark M, Chambon P, Metzger D (2005) Temporally controlled targeted somatic mutagenesis in embryonic surface ectoderm and fetal epidermal keratinocytes unveils two distinct developmental functions of BRG1 in limb morphogenesis and skin barrier formation. *Development* 132:4533–4544. <https://doi.org/10.1242/dev.02019>
29. Jones DT, Jäger N, Kool M, Zichner T, Hutter B, Sultan M, Cho YJ, Pugh TJ, Hovestadt V, Stütz AM et al (2012) Dissecting the genomic complexity underlying medulloblastoma. *Nature* 488:100–105. <https://doi.org/10.1038/nature11284>
30. Juraschka K, Taylor MD (2019) Medulloblastoma in the age of molecular subgroups: a review. *J Neurosurg Pediatr* 24:353–363. <https://doi.org/10.3171/2019.5.PEDS18381>
31. Kadoch C, Hargreaves DC, Hodges C, Elias L, Ho L, Ranish J, Crabtree GR (2013) Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat Genet* 45:592–601. <https://doi.org/10.1038/ng.2628>
32. Kawauchi D, Robinson G, Uziel T, Gibson P, Reh J, Gao C, Finkelstein D, Qu C, Pounds S, Ellison DW et al (2012) A mouse model of the most aggressive subgroup of human medulloblastoma. *Cancer Cell* 21:168–180. <https://doi.org/10.1016/j.ccr.2011.12.023>
33. Kawauchi D, Ogg RJ, Liu L, Shih DJH, Finkelstein D, Murphy BL, Reh JE, Korshunov A, Calabrese C, Zindy Fet et al (2017) Novel MYC-driven medulloblastoma models from multiple embryonic cerebellar cells. *Oncogene* 36:5231–5242. <https://doi.org/10.1038/ncr.2017.110>
34. Korshunov A, Okonechnikov K, Stichel D, Schrimpf D, Delaidelli A, Tonn S, Mynarek M, Sievers P, Sahm F, Jones DTW et al (2022) Gene expression profiling of Group 3 medulloblastomas defines a clinically tractable stratification based on KIRREL2 expression. *Acta Neuropathol* 144:339–352. <https://doi.org/10.1007/s00401-022-02460-1>
35. Leto K, Arancillo M, Becker EB, Buffo A, Chiang C, Ding B, Dobyns WB, Dusart I, Haldipur P, Hatten ME et al (2016) Consensus Paper: Cerebellar Development. *Cerebellum* 15: 789–828. <https://doi.org/10.1007/s12311-015-0724-2>
36. Li XN, Shu Q, Su JM, Perlaky L, Blaney SM, Lau CC (2005) Valproic acid induces growth arrest, apoptosis, and senescence in medulloblastomas by increasing histone hyperacetylation and regulating expression of p21Cip1, CDK4, and CMYC. *Mol Cancer Ther* 4:1912–1
37. Machold R, Fishell G (2005) Math1 is expressed in temporally discrete pools of cerebellar rhombic-lip neural progenitors. *Neuron* 48:17–24. <https://doi.org/10.1016/j.neuron.2005.08.028>
38. Mainwaring OJ, Weishaupt H, Zhao M, Rosén G, Borgenvik A, Breinschmid L, Verbaan AD, Richardson S, Thompson D, Clifford SC et al (2023) ARF suppression by MYC but not MYCN confers increased malignancy of aggressive pediatric brain tumors. *Nat Commun* 14:1221. <https://doi.org/10.1038/s41467-023-36847-9>
39. Mashtalir N, D'Avino AR, Michel BC, Luo J, Pan J, Otto JE, Zullow HJ, McKenzie ZM, Kubiak RL, St Pierre Ret et al (2018) Modular Organization and Assembly of SWI/SNF family chromatin remodeling complexes. *Cell* 175:1272–1288. <https://doi.org/10.1016/j.cell.2018.09.032>
40. Matsumoto S, Banine F, Struve J, Xing R, Adams C, Liu Y, Metzger D, Chambon P, Rao MS, Sherman LS (2006) Brg1 is required for murine neural stem cell

- maintenance and gliogenesis. *Dev Biol* 289:372–383. <https://doi.org/10.1016/j.ydbio.2005.10.044>
41. McInnes L, Healy J, Saul N, Großberger L (2018) UMAP: Uniform Manifold approximation and projection. *J Open Source Softw* 3:e861. <https://doi.org/10.21105/joss.00861>
 42. Merk DJ, Ohli J, Merk ND, Thatikonda V, Morrissy S, Schoof M, Schmid SN, Harrison L, Filser S, Ahlfeld J (2018) Opposing Effects of CREBBP Mutations Govern the Phenotype of Rubinstein-Taybi Syndrome and Adult SHH Medulloblastoma. *Dev Cell* 44: 709–724 e706 <https://doi.org/10.1016/j.devcel.2018.02.012>
 43. Moreno N, Schmidt C, Ahlfeld J, Pöschl J, Dittmar S, Pfister SM, Kool M, Kerl K, Schuller U (2014) Loss of Smar proteins impairs cerebellar development. *J Neurosci* 34:13486–13491. <https://doi.org/10.1523/jneurosci.2560-14.2014>
 44. Northcott PA, Buchhalter I, Morrissy AS, Hovestadt V, Weischenfeldt J, Ehrenberger T, Gröbner S, Segura-Wang M, Zichner T, Rudneva VA et al (2017) The whole-genome landscape of medulloblastoma subtypes. *Nature* 547:311–317. <https://doi.org/10.1038/nature22973>
 45. Pantelouris EM (1968) Absence of thymus in a mouse mutant. *Nature* 217:370–371. <https://doi.org/10.1038/217370a0>
 46. Park YK, Lee JE, Yan Z, McKernan K, O'Haren T, Wang W, Peng W, Ge K (2021) Interplay of BAF and MLL4 promotes cell type-specific enhancer activation. *Nat Commun* 12:1630. <https://doi.org/10.1038/s41467-021-21893-y>
 47. Pei Y, Moore CE, Wang J, Tewari AK, Eroshkin A, Cho YJ, Witt H, Korshunov A, Read TA, Sun JL et al (2012) An animal model of MYC-driven medulloblastoma. *Cancer Cell* 21: 155–167 <https://doi.org/10.1016/j.ccr.2011.12.021>
 48. Pei Y, Liu KW, Wang J, Garancher A, Tao R, Esparza LA, Maier DL, Udaqa YT, Murad N, Morrissy Set al et al (2016) HDAC and PI3K antagonists cooperate to inhibit growth of MYC-Driven Medulloblastoma. *Cancer Cell* 29:311–323. <https://doi.org/10.1016/j.ccr.2016.02.011>
 49. Perla A, Fratini L, Cardoso PS, Nör C, Brunetto AT, Brunetto AL, de Farias CB, Jaeger M, Roesler R (2020) Histone deacetylase inhibitors in Pediatric Brain Cancers: Biological Activities and therapeutic potential. *Front Cell Dev Biol* 8. <https://doi.org/10.3389/fcell.2020.00546>
 50. Pöschl J, Lorenz A, Hartmann W, von Bueren AO, Kool M, Li S, Peraud A, Tonn JC, Herms J, Xiang M (2011) expression of BARHL1 in medulloblastoma is associated with prolonged survival in mice and humans. *Oncogene* 30: 4721–4730 <https://doi.org/10.1038/ncr.2011.173>
 51. Qin N, Paisana E, Langini M, Picard D, Malzkorn B, Custódia C, Cascão R, Meyer FD, Blümel L, Göbbels S (2022) intratumoral heterogeneity of MYC drives medulloblastoma metastasis and angiogenesis. *Neuro Oncol* 24: 1509–1523 <https://doi.org/10.1093/neuonc/noac068>
 52. Ramos P, Karnezis AN, Hendricks WP, Wang Y, Tembe W, Zismann VL, Legendre C, Liang WS, Russell ML, Craig DW et al (2014) Loss of the tumor suppressor SMARCA4 in small cell carcinoma of the ovary, hypercalcemic type (SCCOHT). *Rare Dis* 2:e967148. https://doi.org/10.4161/21675_49X.2014.967148
 53. Reyes JC, Barra J, Muchardt C, Camus A, Babinet C, Yaniv M (1998) Altered control of cellular proliferation in the absence of mammalian brahma (SNF2alpha). *EMBO J* 17:6979–6991. <https://doi.org/10.1093/emboj/17.23.6979>
 54. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015) Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43:e47. <https://doi.org/10.1093/nar/gkv007>
 55. Robinson G, Parker M, Kranenburg TA, Lu C, Chen X, Ding L, Phoenix TN, Hedlund E, Wei L, Zhu X et al (2012) Novel mutations target distinct subgroups of medulloblastoma. *Nature* 488:43–48. <https://doi.org/10.1038/nature11213>
 56. Romero OA, Vilarubi A, Albuquerque-Bejar JJ, Gomez A, Andrades A, Trastulli D, Pros E, Setien F, Verdura S, Farre Let et al (2021) SMARCA4 deficient tumours are vulnerable to KDM6A/UTX and KDM6B/JMJD3 blockade. *Nat Commun* 12:4319. <https://doi.org/10.1038/s41467-021-24618-3>
 57. Salmon P, Trono D (2006) Production and titration of Lentiviral vectors. *Curr Protoc Neurosci* 37. 4.21.21–24.21.24
 58. Schüller U, Heine VM, Mao J, Kho AT, Dillon AK, Han YG, Huillard E, Sun T, Ligon AH, Qian Y et al (2008) Acquisition of granule neuron precursor identity is a critical determinant of progenitor cell competence to form shh-induced medulloblastoma. *Cancer Cell* 14:123–134. <https://doi.org/10.1016/j.ccr.2008.07.005>
 59. Sharma T, Schwalbe EC, Williamson D, Sill M, Hovestadt V, Mynarek M, Rutkowski S, Robinson GW, Gajjar A, Cavalli F et al (2019) Second-generation molecular subgrouping of medulloblastoma: an international meta-analysis of Group 3 and Group 4 subtypes. *Acta Neuropathol* 138:309–326. <https://doi.org/10.1007/s00401-019-02020-0>
 60. Smith KS, Bihanic L, Gudenus BL, Haldipur P, Tao R, Gao Q, Li Y, Aldinger KA, Iskusnykh IY, Chizhikov VV et al (2022) Unified rhombic lip origins of group 3 and group 4 medulloblastoma. *Nature* 609:1012–1020. <https://doi.org/10.1038/s41586-022-05208-9>
 61. Sturm D, Orr BA, Toprak UH, Hovestadt V, Jones DTW, Capper D, Sill M, Buchhalter I, Northcott PA, Leis J et al (2016) New brain tumor entities emerge from molecular classification of CNS-PNETs. *Cell* 164:1060–1072. <https://doi.org/10.1016/j.cell.2016.01.015>
 62. Sumi-Ichinose C, Ichinose H, Metzger D, Chambon P (1997) SNF2beta-BRG1 is essential for the viability of F9 murine embryonal carcinoma cells. *Mol Cell Biol* 17:5976–5986. <https://doi.org/10.1128/MCB.17.10.5976>
 63. Swartling Fredrik J, Savov V, Persson Anders I, Chen J, Hackett Christopher S, Northcott Paul A, Grimmer Matthew R, Lau J, Chesler L, Perry A et al (2012) Distinct neural stem cell populations give rise to disparate brain tumors in response to N-MYC. *Cancer Cell* 21:601–613. <https://doi.org/10.1016/j.ccr.2012.04.012>
 64. Tao R, Murad N, Xu Z, Zhang P, Okonechnikov K, Kool M, Rivero-Hinojosa S, Lazarski C, Zheng P, Liu Y et al (2019) MYC drives Group 3 Medulloblastoma through Transformation of Sox2(+) astrocyte progenitor cells. *Cancer Res* 79:1967–1980. <https://doi.org/10.1158/0008-5472.CAN-18-1787>
 65. Taylor MD, Northcott PA, Korshunov A, Remke M, Cho YJ, Clifford SC, Eberhart CG, Parsons DW, Rutkowski S, Gajjar A et al (2012) Molecular subgroups of medulloblastoma: the current consensus. *Acta Neuropathol* 123:465–472. <https://doi.org/10.1007/s00401-011-0922-z>
 66. van Riggelen J, Yetil A, Felsher DW (2010) MYC as a regulator of ribosome biogenesis and protein synthesis. *Nat Rev Cancer* 10:301–309. <https://doi.org/10.1038/nrc2819>
 67. Vladoiu MC, El-Hamamy I, Donovan LK, Farooq H, Holgado BL, Sundaravada-nam Y, Ramaswamy V, Hendrikse LD, Kumar S, Mack SC et al (2019) Child-hood cerebellar tumours mirror conserved fetal transcriptional programs. *Nature* 572:67–73. <https://doi.org/10.1038/s41586-019-1158-7>
 68. Wiczeorek D, Bogershausen N, Beleggia F, Steiner-Haldenstatt S, Pohl E, Li Y, Milz E, Martin M, Thiele H, Altmüller J et al (2013) A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. *Hum Mol Genet* 22:5121–5135. <https://doi.org/10.1093/hmg/ddt366>
 69. Zhan X, Shi X, Zhang Z, Chen Y, Wu JI (2011) Dual role of Brg chromatin remodeling factor in sonic hedgehog signaling during neural development. *PNAS* 108:12758–12763. <https://doi.org/10.1073/pnas.1018510108>
 70. Zhou W, Triche TJ Jr, Laird PW, Shen H (2018) SeSAMe: reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions. *Nucleic Acids Res* 46:e123–e123. <https://doi.org/10.1093/nar/gky691>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliation.

4.3. HGG-MYCN amplified project in original format



Mouse models of pediatric high-grade gliomas with MYCN amplification reveal intratumoral heterogeneity and lineage signatures

Received: 2 April 2023

Accepted: 14 November 2023

Published online: 24 November 2023

Check for updates

Melanie Schoof^{1,2}, Shweta Godbole³, Thomas K. Albert⁴, Matthias Dottermusch^{3,5}, Carolin Walter⁶, Annika Ballast⁴, Nan Qin^{7,8,9,10}, Marlina Baca Olivera^{7,8,9,10}, Carolin Göbel^{1,2}, Sina Neyazi^{1,2}, Dörthe Holdhof^{1,2}, Catena Kresbach^{1,2,5,11}, Levke-Sophie Peter^{1,2}, Gefion Dorothea Epplen^{1,2}, Vanessa Thaden^{1,2}, Michael Spohn¹, Mirjam Blattner-Johnson^{12,13}, Franziska Modemann^{11,14}, Martin Mynarek^{2,11}, Stefan Rutkowski², Martin Sill^{12,15}, Julian Varghese⁶, Ann-Kristin Afflerbach^{1,2}, Alicia Eckhardt^{1,2,16}, Daniel Münter⁴, Archana Verma⁴, Nina Struve^{11,16}, David T. W. Jones^{12,13}, Marc Remke^{7,8,9,10}, Julia E. Neumann^{3,5}, Kornelius Kerl^{4,17} & Ulrich Schüller^{1,2,5,17} ✉

Pediatric high-grade gliomas of the subclass MYCN (HGG-MYCN) are highly aggressive tumors frequently carrying *MYCN* amplifications, *TP53* mutations, or both alterations. Due to their rarity, such tumors have only recently been identified as a distinct entity, and biological as well as clinical characteristics have not been addressed specifically. To gain insights into tumorigenesis and molecular profiles of these tumors, and to ultimately suggest alternative treatment options, we generated a genetically engineered mouse model by breeding *hGFAP-cre::Trp53^{FUF}::Isl-MYCN* mice. All mice developed aggressive forebrain tumors early in their lifetime that mimic human HGG-MYCN regarding histology, DNA methylation, and gene expression. Single-cell RNA sequencing revealed a high intratumoral heterogeneity with neuronal and oligodendroglial lineage signatures. High-throughput drug screening using both mouse and human tumor cells finally indicated high efficacy of Doxorubicin, Irinotecan, and Etoposide as possible therapy options that children with HGG-MYCN might benefit from.

Recently, tumors originally diagnosed as primitive neuroectodermal tumors (PNETs) have been reclassified into multiple different pediatric brain tumor classes¹. One clearly distinguishable class of these tumors are “pediatric high-grade gliomas with *MYCN* amplification”, hereafter called “HGG-MYCN”. This recently described tumor entity was further characterized by Korshunov et al. and Tauziède-Espariat et al. as being

highly aggressive pediatric gliomas with a poor prognosis and a median overall survival of only 14 months^{2–4}. When compared to other pediatric glioma, HGG-MYCN present with the worst survival of subtypes³. Unlike other glioma, HGG-MYCN histologically present with undifferentiated, densely packed cell nuclei and highly circumscribed tumors without typical glial features. Molecularly, these tumors often

A full list of affiliations appears at the end of the paper. ✉ e-mail: u.schueller@uke.de

carry amplification of the *MYCN* gene and somatic or constitutional mutations in *TP53*^{2,4,5}. Usually, these tumors are treated according to protocols for high-grade glioma with no specificities or targeted therapies available.

MYCN is a transcription factor of the *MYC* family, which consists of three paralogues: cellular (*c-*), lung-carcinoma derived (*l-*), and neuroblastoma-derived (*n-*)-myelocytomatosis (*myc*). Even though all three share a common structure containing so-called *MYC* boxes and a basic helix-loop-helix domain, they show a distinct expression profile in normal tissue and disease. *MYCN* is expressed in neural tissues and is essential for normal central nervous system (CNS) development^{6,7}. It was discovered in 1983 as an amplified gene in neuroblastoma, a childhood tumor of the neural crest, with similarities to the *c-MYC* gene^{8,9}. Shortly after its discovery, *MYCN* was associated with high-risk neuroblastoma and poor prognosis¹⁰⁻¹².

MYCN has also been described to be involved in a number of further tumor entities, spanning hematologic malignancies^{13,14}, lung cancers¹⁵, and malignancies of the nervous system¹⁶. In CNS tumors, *MYCN* is mainly involved in tumorigenesis of retinoblastoma¹⁷, medulloblastoma^{18,22}, and glioblastoma²³⁻²⁵.

Although a few mouse tumor models driven by alterations of *MYCN* have been described, none of them have developed gliomas with similarities to human HGG-*MYCN*. The most prominent *MYCN*-driven tumor models mimic human neuroblastoma, in which *MYCN* expression, driven by tyrosine-hydroxylase (TH)- or dopamine- β -hydroxylase- (*Dbh*)-*cre*, leads to the development of aggressive tumors, modeling the human disease^{26,27}. Forcing *MYCN* expression in the hindbrain by the *Glil*-promoter induced the development of tumors, which resemble human medulloblastoma (MB) with the highest similarity to group 3 MB²⁸. This mouse model is well established as a group 3 MB model although *MYCN* is usually associated with group 4 MB²⁹. Furthermore, mutated *MYCN* is able to transform different types of neural stem cells (NSCs) leading to medulloblastoma- and glioblastoma-like tumors in hind- and forebrain, respectively³⁰. On the other hand, forced expression of wild-type *MYCN* in the hindbrain or entire CNS does not necessarily lead to brain tumor development^{31,32}. Given the lack of existing mouse models for HGG-*MYCN* on one hand and the urgent need for alternative treatment modalities for this aggressive disease on the other hand, we here aimed to develop a murine model for HGG-*MYCN* mimicking their human counterparts. As HGG-*MYCN* often carry both alterations, *MYCN* amplifications and *TP53* mutations, we generated a mouse model with combined expression of wild-type human *MYCN* and a loss of *Trp53*, which, on its own, induces glioma development in mice with a long latency and reduced penetrance³³.

In this work, we successfully generate a mouse model for HGG-*MYCN* by inducing *MYCN* expression and simultaneous *Trp53* deletion in *hGFAP-cre* expressing cells (*hGFAP-cre::Trp53^{Fl/Fl}::Isl-MYCN*), and show that these mice develop large forebrain tumors with 100% penetrance within 90 days. These tumors recapitulate human HGG-*MYCN* histologically and molecularly. We use a multi-omic approach to dissect the tumor biology of murine gliomas as well as high-throughput drug screening to identify alternative treatment options for these aggressive tumors.

Results

HGG-*MYCN* tumors represent a distinct tumor entity and frequently carry alterations in *TP53* and *MYCN*

Studies published on the recently discovered rare HGG-*MYCN* entity describe a group of aggressive pediatric gliomas, which form a distinct cluster in global DNA methylation analysis^{1,2,34,35}. We collected respective data from published and five in-house cases of HGG-*MYCN* diagnosed by DNA methylation profiling and used a reference set of pediatric brain tumors to confirm the distinctiveness of these tumors by Uniform Manifold Approximation and Projection (UMAP). We

included the most common brain tumors as well as potential differential diagnoses and analyzed the global DNA methylation of 2514 tumors, including 47 HGG-*MYCN* (Fig. 1a). Clinical information of the HGG-*MYCN* patients revealed a median age of 8 years with only 4 out of 89 patients older than 20 years (range: 1–56 years, Fig. 1b). Sex distribution was almost balanced (Fig. 1c). DNA sequencing information was available for 47 cases, and we found *TP53* mutations in 68% and *MYCN* amplifications in 60% of those cases (Fig. 1d, e). Thirty-six percent of the analyzed cases carried both a loss of *TP53* and an amplification of *MYCN* (Fig. 1f). HGG-*MYCN* were detected throughout the entire brain with a preponderance of tumors occurring in the frontal and temporal lobe (Fig. 1g). The tumors presented with multiple chromosomal aberrations and, in part, a clearly visible *MYCN* amplification ($n = 19$) as visualized in a heatmap of the copy number variations of 47 tumors as well as by Fluorescence-in situ hybridization (FISH) or immunohistochemistry (IHC). IHC showing accumulated p53 indicating a mutated *TP53* (Fig. 1h-l).

hGFAP-cre driven loss of p53 and expression of *MYCN* induce brain tumor formation in mice

In order to generate an appropriate mouse model for HGG-*MYCN*, we exploited the Cre-LoxP system to simultaneously induce the deletion of *Trp53* and force expression of human wild-type *MYCN* under the control of the *hGFAP* promoter. This promoter is active from embryonic day (E) 13.5 onwards and targets radial glia, which later differentiates to distinct cell types of the CNS, including neurons, oligodendrocytes, astrocytes, and adult neural stem cells³⁶. Loss of p53 is accomplished by an allele carrying loxP-sites after exon 1 and 10 in the *Trp53* gene, leading to a large deletion in the gene upon recombination. The expression of *MYCN* is achieved by inserting an allele into the ubiquitously expressed *Rosa26* locus, where a stop codon flanked by two loxP sites is preceding the *MYCN* open-reading-frame (Fig. 2a).

All *hGFAP-cre::Trp53^{Fl/Fl}::Isl-MYCN* animals ($n = 24$) developed neurological symptoms within 90 days of life with the first animals being symptomatic around postnatal day 40 (Fig. 2b). The animals abruptly presented with hydrocephalus and akinesia and were sacrificed as soon as symptoms appeared. Upon necropsy, large forebrain tumors and enlarged ventricles were macroscopically found (Fig. 2c). Midbrain and hindbrain as well as the spinal cord and other organs appeared macro- and microscopically normal. This was of particular interest, since *MYCN* amplifications and/or *TP53* mutations have also been observed in a subset of spinal ependymoma as well as in pediatric medulloblastoma^{21,22,37}. In this model, none of the microscopically analyzed mice ($n = 17$) showed overexpression of *MYCN* in these regions.

Histological examination of the murine brains revealed large, cell-dense tumors of the forebrain with a heterogeneous cell morphology (Fig. 2d, S1). To determine the timing of tumor initiation, we sacrificed animals at postnatal day (P) 7 and 37 prior to the establishment of any symptoms. At P7, no tumor lesion was detected by histological examination of the brains of six mice, whereas at P37, tumor lesions were detected in all analyzed mice ($n = 5$). These small tumor lesions presented with densely packed, highly proliferating cells in the outermost layer of the olfactory bulb (OB) (Fig. S1).

Genetic analysis of the tumors developing in *hGFAP-cre::Trp53^{Fl/Fl}::Isl-MYCN* mice showed a recombined stop codon as well as a recombined *Trp53* allele as expected (Fig. 2e). We also analyzed potential copy number variations (CNVs) of three representative mouse tumors (Fig. 2f). Due to the genetic engineering, all three mouse tumors revealed an amplification of the *Rosa26*-locus with the inserted *MYCN* on chromosome 6 (asterisk in Fig. 2f). Moreover, they showed chromosomal aberrations private to one tumor as well as recurrent chromosomal aberrations appearing in at least two of three tumors at chromosomes 7, 14, or 16 (Fig. 2f). This suggests that the combination of *MYCN* and *Trp53*

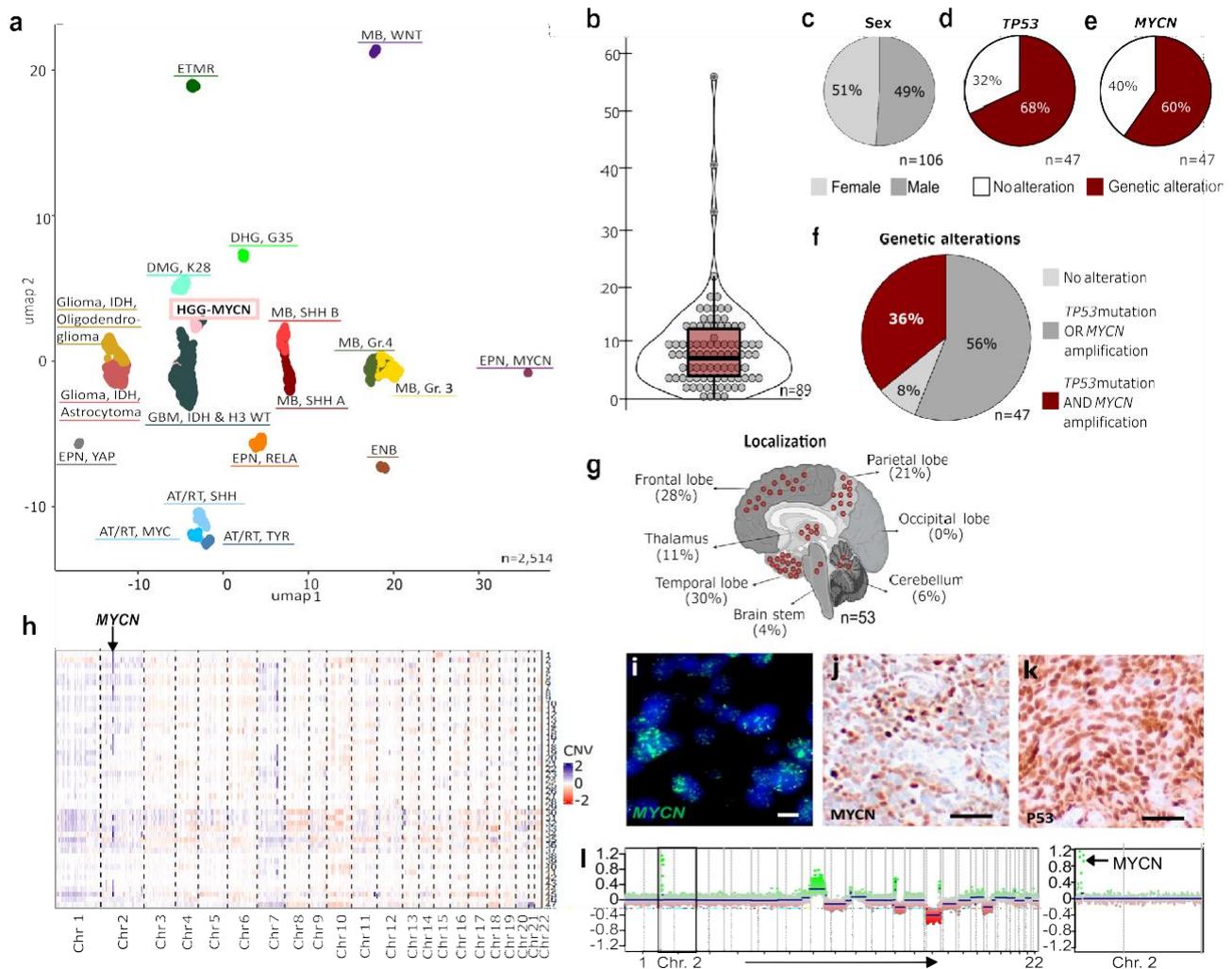


Fig. 1 | Tumors of the DNA methylation class HGG-MYCIN carry MYCN amplifications and TP53 mutations. **a** UMAP of global DNA methylation profiling of 2514 cases of multiple brain tumor entities including the most common brain tumors as well as potential differential diagnoses including 47 HGG-MYCIN using the 10,000 most differentially methylated CpG sites. **b** Age distribution of $n = 89$ HGG-MYCIN tumors with a median age of 8 years (bounds of box = 3–13, whiskers = min:1, max:22). **c** Sex distribution of 106 HGG-MYCIN. **d** 47 HGG-MYCIN were screened for TP53 mutations, of which 68% carried a mutation. **e** The same 47 cases as in (d) were analyzed for their MYCN status. **f** Of the same 47 cases shown in (d) and (e), only 8% carried no TP53 or MYCN alteration, whereas 36% of those carry both alterations. **g** Tumors can be found throughout the entire brain with the majority of cases

located in the temporal and frontal lobes. **h** Heatmap showing copy number variations of 47 HGG-MYCIN. The copy number profile of such tumors is imbalanced with a clearly visible *MYCN* amplification (Chr. 2), highlighted by the arrow. *MYCN* amplifications can also be detected by FISH analysis (representative case shown in (i), three independent tumors showing this amplification were analyzed), while IHC may serve as a surrogate marker (representative case in (j), the same three tumors were analyzed). **k** Nuclear p53 accumulation indicating impaired p53 function can be detected by IHC (representative case, three independent tumors were analyzed). **l** Representative CNV plot of a HGG-MYCIN with a magnification of chromosome 2 with the *MYCN* amplification (CNV plots of 47 tumors were generated). Scale bar in (i) = 5 μm , in (j) & (k) = 50 μm . Source data are provided as a Source Data file.

alterations induce further genomic changes that may be needed for tumor development.

We next characterized the mouse tumors by microscopy to see whether they resemble their human counterparts. Tumors of both species showed irregularly shaped, densely packed cell nuclei (Fig. 2g, n) and were positive for MYCN (Fig. 2h, o). They were also highly proliferative (Fig. 2i, p) and expressed Nestin (Fig. 2j, q), SOX2 (Fig. 2k, r), OLIG2 (Fig. 2l, s), and GFAP (Fig. 2m, t). Other histological markers were also similarly expressed in mouse as well as human tumors: Neurofilament, NeuN, and Synaptophysin were not expressed but tumors of both, mouse and human, expressed TubB3 (Fig. S2).

So far, the cell-of-origin as well as time and place of tumor onset for HGG-MYCIN is unknown. Therefore, we decided to investigate tumor development in other cell populations of the developing CNS by breeding animals with the same genetic alterations in other target

cell populations. We bred *Sox2-cre::Trp53^{Fl/Fl}::Isl-MYCIN* and *Blbp-cre::Trp53^{Fl/Fl}::Isl-MYCIN* mice, which initiate recombination upon E6.5 or E9.5, respectively. The *Sox2*-mediated recombination was embryonically lethal (Fig. S3a) with severely underdeveloped animals at E16.5 (Fig. S3b–d). The *Blbp*-mediated recombination also led to prenatal lethality except for two animals surviving until P0 and two until P18, on which they presented with hydrocephalus (Fig. S3e). Histological examination of the P18 brains showed no signs of tumor development, suggesting rather a developmental defect as the cause for the hydrocephalus (Fig. S3f–h).

Mouse HGG-MYCIN molecularly resemble their human counterparts

We next investigated whether murine and human tumors shared similarities regarding their DNA methylation profiles. We used bead

chip arrays to detect the methylation status of 285,000 CpG sites of

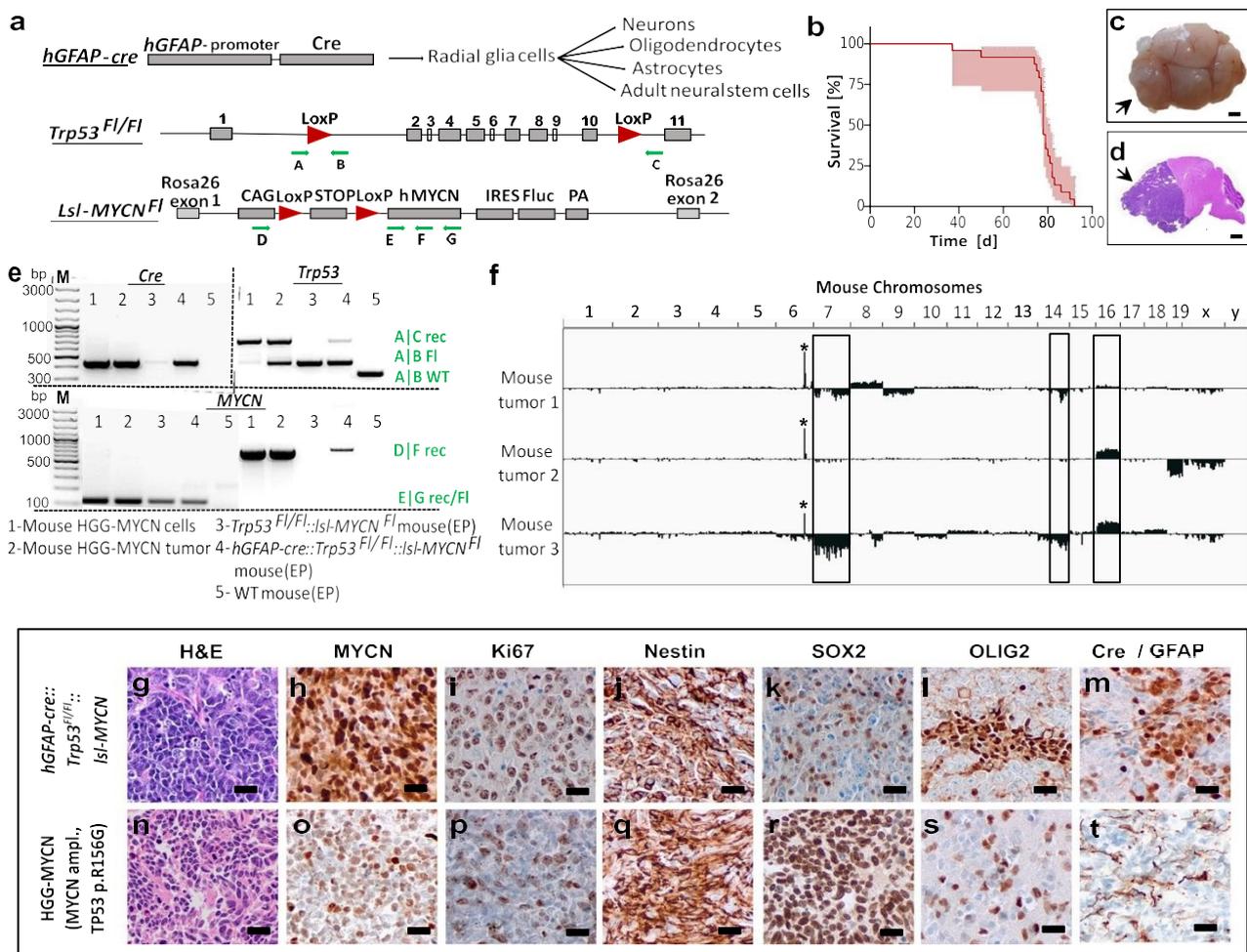


Fig. 2 | Mouse HGG-MYCNC develop within the first 100 days of life and match their human counterparts histologically. **a** Genetics of the HGG-MYCNC mouse model. Cre is expressed under control of the *hGFAP* promoter, which targets different cell populations shown on the right. The floxed alleles of *Trp53* and *MYCN* are depicted, loxP sites are shown as red arrowheads, and primers for genotyping and proof of recombination are shown as green arrows. **b** Kaplan-Meier survival curve of mice with HGG-MYCNC ($n = 24$) as percent survival, light red area showing the asymmetrical 95% confidence interval. **c** Macroscopic image of a mouse brain carrying an HGG-MYCNC (arrow). **d** Hematoxylin and Eosin (H&E) stained brain with a large forebrain tumor (arrow). **e** Representative PCR result of genotyping and the detection of allele recombination. Results are shown for cultured mouse tumor cells (1), fresh mouse tumor tissue (2), ear biopsy of a mouse not carrying the Cre recombinase (3), ear biopsy of a mouse with an HGG-MYCNC (4), and (5) a wild-type

mouse carrying none of the transgenes. Bands result from the primers indicated by green arrows in (a). The same PCR was performed for all animals generated in the study including the $n = 24$ animals included in the Kaplan-Meier survival analysis. **f** Copy number variation plots of three mouse HGG-MYCNC. An amplification of the *Rosa26*-locus, in which the *MYCN* is inserted, is visible in all three samples (marked by the star). Other recurrent copy number changes are observed in at least two of the tumors (marked by the rectangle). **g-t** Immunohistochemical comparability of mouse and a human HGG-MYCNC. The pictures show representative micrographs, all stainings were performed independently on at least three samples. EP ear punch, Fl floxed allele, M marker, Rec recombined allele, WT wild-type allele. The scale bar in c and d corresponds to 2 mm, and the scale bar in (g-t) corresponds to 20 μ m. Source data are provided as a Source Data file.

the mouse genome. Of these sites, 141 CpG sites were identical to the human 850k (EPIC) array. The beta-values of these 141 sites were sufficient to distinguish different human brain tumor entities according to the visualization via UMAP. When clustering the methylation data of the murine tumors with the most common pediatric brain tumors as well as potential differential diagnosis, mouse HGG-MYCNC (hot pink) showed the highest similarity to the human HGG-MYCNC (light pink, Fig. 3a). As a control, we included tumors of *Math1-cre::SmoM2^{Fl/wt}* mice, which are a well-known murine model for sonic hedgehog medulloblastoma (SHH-MB)³⁸ and which were located nearest to human SHH MB, as expected (Fig. 3a).

Next, we generated transcriptomic profiles of the murine tumors and compared them to gene expression data of human HGG-MYCNC and other pediatric brain tumors¹. The visualization of gene expression profiles by UMAP analyses revealed the highest simi-

larity between mouse HGG-MYCNC (hot pink) to their human

counterparts (light pink, Fig. 3b). Again, SHH-MBs served as an internal control with high similarities between murine (red) and human SHH-MB tumors (dark red). We further analyzed the similarity of gene expression profiles by calculating the Euclidean distance of averaged transcriptome data per group and performed an agreement of differential expression (AGDEX³⁹) analysis. Both analyses confirmed the high similarity between mouse and human HGG-MYCN (Fig. 3c, d, Euclidean distance: HGG-MYCN - mouse HGG-MYCN = 24.4, AGDEX: HGG-MYCN - mouse HGG-MYCN, AGDEX cos: 0.190659811, p -value = 0.007). When we compared *MYCN* expression in human and mouse HGG-MYCN relative to other human gliomas or healthy mouse tissue of the OB and cerebellum and the SHH-MB mouse model, and performed gene set enrichment analysis (GSEA) for *MYCN* target genes, we found that *MYCN* itself as well as its target gene sets are significantly enriched in tumors of both species, human and mouse (Fig. 3e, f, two-sided Welch's- t -test,

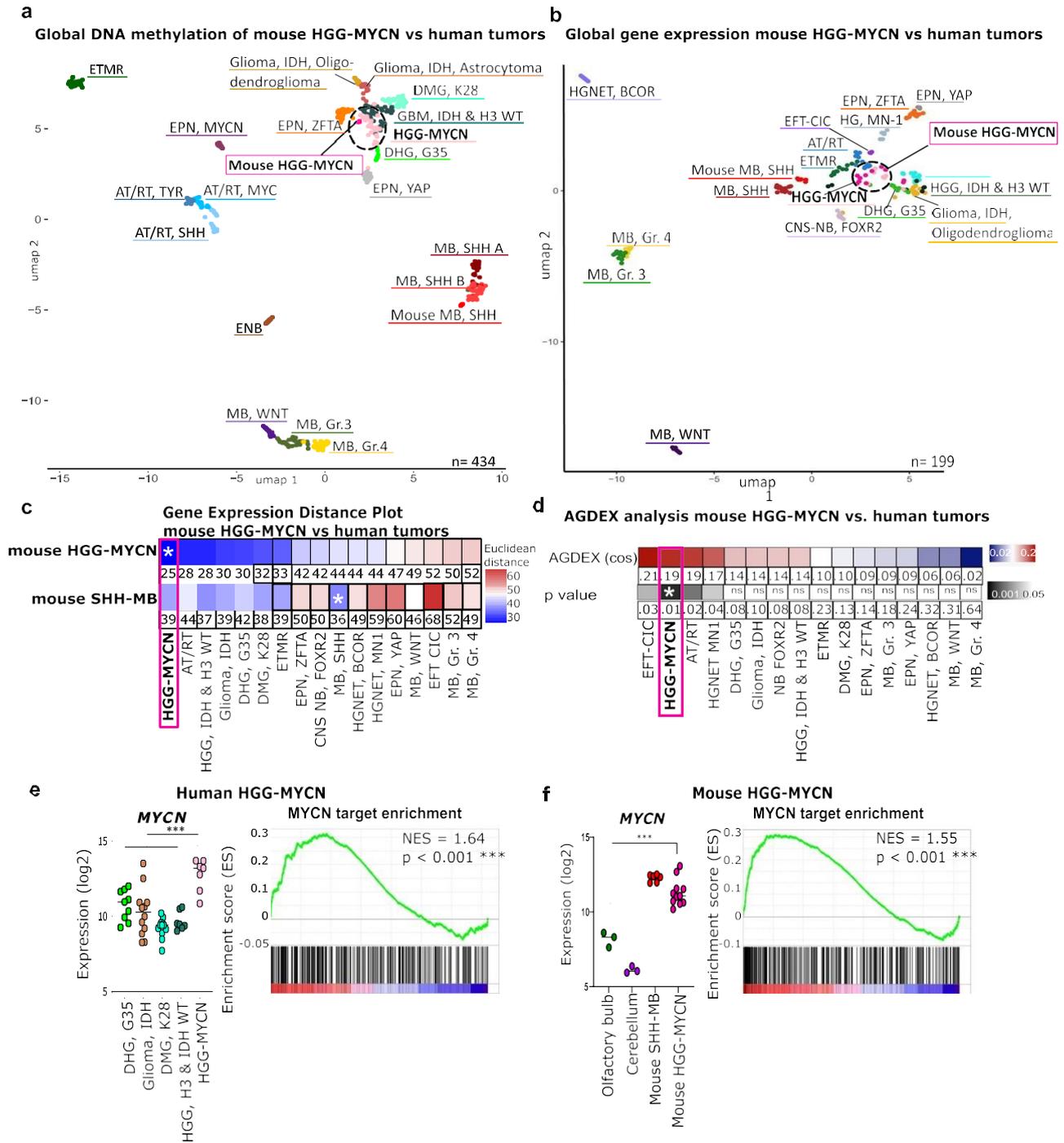


Fig. 3 | Mouse HGG-MYCNCN match their human counterparts molecularly. a UMAP of global DNA methylation. Betavalues of 141 identical CpG sites between mouse

and human were used for comparison of similarity. Mouse methylome data were generated with the Illumina Mouse Methylation bead chip array and compared to human data generated with the EPIC array. The three mouse tumors (in hot pink) show most similarity to the human HGG-MYCNCN group (pink). b UMAP of global gene expression data. Eleven mouse tumors were profiled by RNA sequencing and their gene expression profile was compared to published gene expression data of human „CNS PNET“ tumors 1. Data were normalized for interspecies differences by employing RNA Seq. data of *Math1-cre::SmoM2^{Fl/wt}* mouse tumors and human SHH medulloblastoma. c Distance plot of mouse HGG-MYCNCN and mouse SHH-MBs and human tumors based on the 500 most significantly expressed genes. Mouse tumors showed most similarity to human HGG-MYCNCN and human SHH-MB, respectively. The values of the Euclidean distance are displayed, and the asterisk marks the smallest values and thereby highest similarity. d AGDEX analysis of mouse HGG-MYCNCN

also shows the high similarity between murine and human tumors (analysis based on 14,416 orthologous genes). The values of the AGDEX analysis with their respective p-values are given. The asterisk marks the smallest p-value and thereby highest similarity. e, f Human and mouse HGG-MYCNCN show significantly higher MYCN expression compared to other glioma entities or control tissue (from mouse olfactory bulb (OB) or cerebellum) and comparable expression to the SHH-MB mouse model as well as highly significant enrichment for MYCN target genes. Human tumors: non-mycn, n = 40, HGG-MYCNCN n = 7. Mouse data: OB tissue n = 3, mouse HGG-MYCNCN n = 11., two-sided Welch's t-test, human: p = 0.0002, 95% confidence interval = 2.001 to 4.320, murine: p = 0.0002, 95% confidence interval = 2.143 to 4.085. Source data are provided as a Source Data file.

and human were used for comparison of similarity. Mouse methylome data were generated with the Illumina Mouse Methylation bead chip array and compared to human data generated with the EPIC array. The three mouse tumors (in hot pink) show most similarity to the human HGG-MYCNCN group (pink). b UMAP of global gene expression data. Eleven mouse tumors were profiled by RNA sequencing and their gene expression profile was compared to published gene expression data of human „CNS PNET“ tumors 1. Data were normalized for interspecies differences by employing RNA Seq. data of *Math1-cre::SmoM2^{Fl/wt}* mouse tumors and human SHH medulloblastoma. c Distance plot of mouse HGG-MYCNCN and mouse SHH-MBs and human tumors based on the 500 most significantly expressed genes. Mouse tumors showed most similarity to human HGG-MYCNCN and human SHH-MB, respectively. The values of the Euclidean distance are displayed, and the asterisk marks the smallest values and thereby highest similarity. d AGDEX analysis of mouse HGG-MYCNCN

human: $p = 0.0002$, 95% confidence interval = 2.001 to 4.320, murine: $p = 0.0002$, 95% confidence interval = 2.143 to 4.085).

Single-cell transcriptomics of mouse HGG-MYCN revealed a highly heterogeneous population of oligodendroglial tumor cells

As knowledge about the cellular architecture and the cell-of-origin of HGG-MYCN is lacking and no single cell RNA sequencing (scRNA seq) data of HGG-MYCN is available, we employed scRNA seq to dissect the cellular composition of these murine tumors. We profiled the single cell transcriptomes of seven mouse tumors using 10X genomics technology. The cells were harvested from OB/tumor tissue on P43 ($n = 2$), P70 ($n = 2$), P77 ($n = 2$), and P92 ($n = 1$). ScRNA seq. resulted in 23 distinct clusters (Fig. 4a). The tumor cell clusters were clearly identified based on their increased expression of human *MYCN* and the firefly luciferase (*FLUC*) reporter gene (Fig. 4b). The remaining clusters were unequivocally assigned as non-malignant cells of different types by a combination of differential gene expression (DGE) analysis and comparison with known highly specific marker genes of distinct tumor microenvironment (TME) cell types (Fig. 4c, S4).

Based on their similarities in the gene expression patterns, the tSNE algorithm divided the tumor cells into distinct clusters (Fig. 4a-c). These tumor cell clusters were split into three larger areas: clusters 0 and 13 were composed only of cells of the most mature tumor (P92 tumor), clusters 1/2/3/4/9/15/18 were composed of cells resembling oligodendrocytic-lineage (OL) cells and cluster 5 expresses a neuronal-like signature.

In order to define the gene expression signatures of the different tumor cell populations, we employed a combination of DGE and functional gene network analysis as well as gene ontology (GO) annotation to identify the differences and similarities of distinct clusters within the murine HGG-MYCN (Fig. 4d). These analyses clearly show an oligodendrocytic tumor cell population, a neuronal tumor cell population, and a third tumor population, which is transcriptionally clearly distinct and expresses markers of neither cell lineage. Tumor cells of this cluster were solely harvested from the most mature tumor of a 92-day-old mouse.

In order to determine the potential cellular origin of the murine tumor cells, we compared gene expression profiles of the different tumor cell clusters to a reference dataset of the ventricular zone/subventricular zone (VZ/SVZ) of the mouse. We used logistic regression to identify similarities between the reference clusters and our tumor cells. This analysis reveals a high similarity of HGG-MYCN tumor cells with adult neural stem (aNSC) and transit-amplifying neural precursor cells (TAC). Some tumor cell clusters (5,0,13,3) also show similarities to neuroblasts (NB) and some clusters to cells of the oligodendrocytic lineage (1,15,0,13,3). These results suggest that murine HGG-MYCN may originate from multipotent stem or progenitor cells in the SVZ that differentiate into both glial and neuronal lineages.

To understand the tumor development further, we investigated the tumor cells of the different stages of tumor development. The comparison of the TME and the tumor cell clusters at P43, P70, P77, and P92 revealed a reduction of TME cells throughout tumor development and a change in tumor cell populations (Fig. 4f). The tumor cells of the youngest mice are mostly OL-like. At P70, the NB-like tumor cell population appears which is also detectable at P77. At P92, only a small proportion of tumor cells can be assigned to the OL- or NB-like type but instead most cells belong to a unique cell population which

we therefore called P92-tumor.

Sensitivity of HGG-MYCN cells to Doxorubicin, Etoposide and Irinotecan in an in vitro drug screen

Next, we used our mouse model to identify improved therapeutic options to treat HGG-MYCN, since standard treatment is still inefficient. For this, we performed a high-throughput drug screening of 639 compounds in a human HGG-MYCN cell line (pbt-04) and a cell line isolated from our mouse model (pn003, Fig. 5a, Supplementary data 1). To expedite clinical use, we analyzed drug response using a clinical anticancer library comprising nearly 80% of FDA-approved drugs. The chosen drugs encompass both standard chemotherapy and targeted therapy for various types of cancer. The selection doesn't limit to the blood-brain barrier permeability since, due to the recent development of materials science and nanotechnology, multiple strategies could be used to deliver drugs across BBB.

We determined the drug response by comparing the normalized area under the curve (AUC) of a cell viability assay. We then focused on the 100 compounds, that were most efficient in reducing cell viability in both, mouse and human cells. Among these 100 compounds, we found 18 drug classes represented by more than two drugs, suggesting a specific mechanism of action against these tumors (Fig. 5b). Thirty of these drugs are FDA approved, and for 14 of those, CNS permeability has been described (Fig. 5c). By focusing on the FDA approved drugs, we identified 12 compounds that are already used for the treatment of pediatric brain tumor patients and can therefore be considered relatively safe to use also in HGG-MYCN patients (Fig. 5d, e).

The top three compounds, that are FDA approved and used to treat pediatric brain tumor patients, are Doxorubicin, Etoposide, and SN38 as the active metabolite of Irinotecan. All three compounds showed high efficacy against our HGG-MYCN cell lines, mixed response in other adult and pediatric glioblastoma reference cells and had almost no effect on healthy fibroblasts (Fig. 5f-h).

Discussion

We show that combined loss of p53 and forced expression of MYCN in neural precursor cells is sufficient to drive brain tumor formation. Mice carrying these alterations in cells targeted by *hGFAP*-driven recombination develop large forebrain tumors. These murine tumors show a similar histology and marker expression as human tumors of the methylation class "pediatric high-grade glioma MYCN". In addition, we found the highest similarity between mouse tumors and HGG-MYCN compared to other aggressive pediatric brain tumors in DNA methylation and gene expression profiles.

We investigated the single cell transcriptomic landscape of the tumors and could show a tumor evolution of the mouse tumors by analyzing different mouse ages and stages of tumor development. We included mouse samples of early tumor onset at P43, further developed tumors at P70 and tumors of symptomatic mice at P77 and P92. During tumor development, we observed changes in the TME but also a change in the tumor cell populations. At P92, a unique tumor cell population is detected. This mouse was exceptional in survival, as only 4.4% of animals survive until P90. This might explain the difference of the observed tumor cell population from the other six mouse tumors but as we could only generate data of this one exceptionally long surviving mouse, we cannot draw definite conclusions.

The age of disease onset differed between mice and humans, with children having a median age of disease onset of eight years and mice becoming symptomatic in adolescence. This difference might be due

to species differences in maturation of neural cells or depending on environmental influences. As the age of disease onset is before adulthood in both species and the molecular biology of mouse and human tumors is very similar, we consider our model as a reliable model for the human disease.

Alterations in *MYCN* and *TP53* also occur in other human brain tumor entities, such as medulloblastoma. To investigate whether the same transgenes will lead to different brain tumors in other target cells and further characterize the cell-of-origin of our tumors, we used different promoters to drive Cre expression in the developing brain. We employed the *Sox2* promoter, which is expressed earlier in embryonic development (E6.5 instead of E13.5), and the *Blb* promoter, which is expressed from E9.5 onwards. In fact, Swartling et al. described that

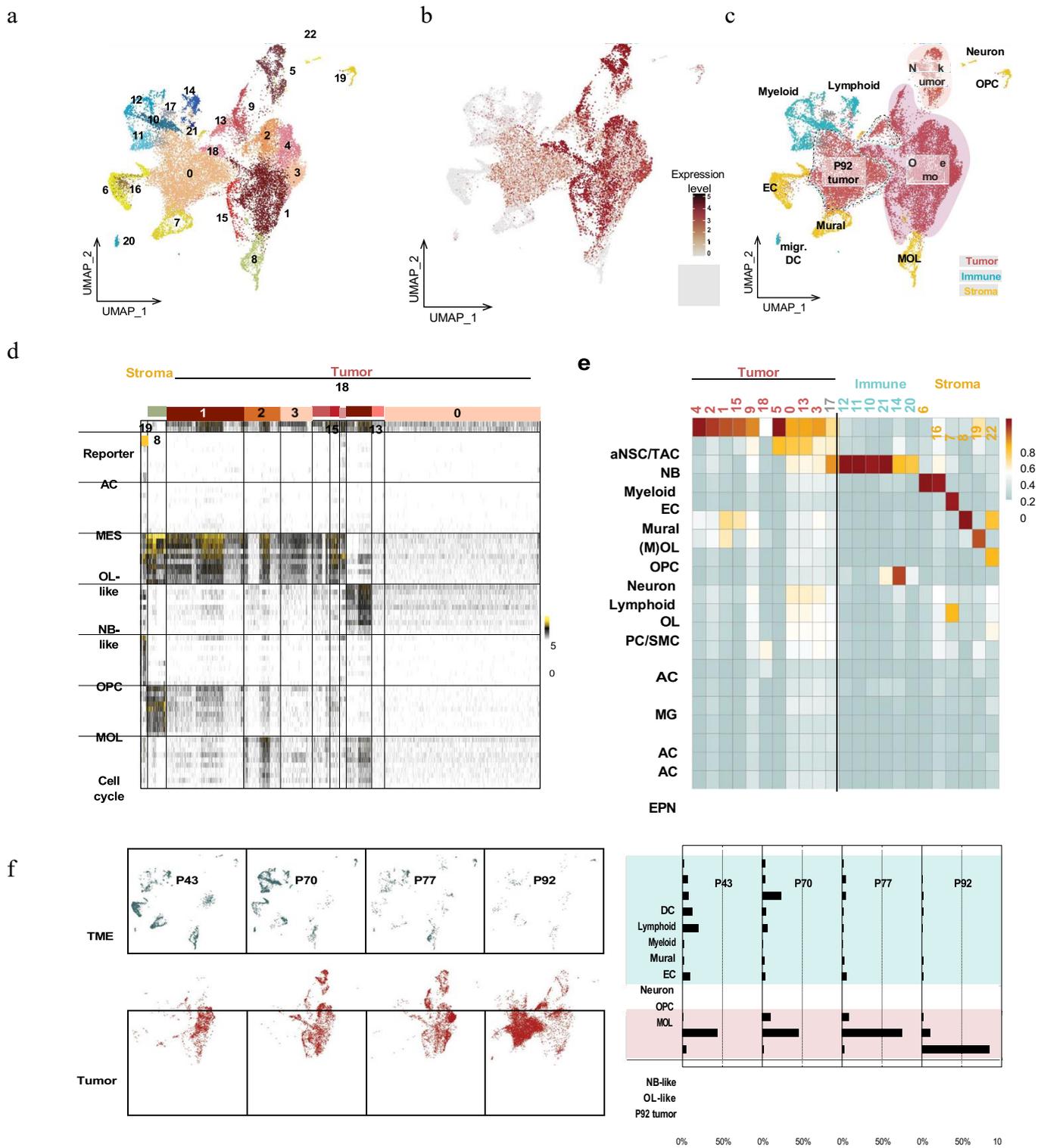


Fig. 4 | Mouse HGG-MYCN reveal a high intratumoral heterogeneity with oligodendroglial and neural cell populations and a time-resolved change in tumor composition. **a** UMAP of single cell RNA sequencing data of seven mouse HGG-MYCN (2xP43, 2xP70, 2xP77 and 1x P92) including 24,938 cells. 23 cell clusters were identified by the Seurat algorithm. **b** To identify tumor cell clusters, expression of human *MYCN* and Luciferase (*FLUC*), were plotted. **c**, **d** Cell clusters were annotated by analysis of marker gene expression. This revealed immune as well as stromal cell clusters and three main superclusters of tumor cells. Tumor superclusters consist of an oligodendroglial-like, a neuronal-like, and a cluster which was only detected in the most mature mouse tumor. **e** Cell clusters were compared to a reference atlas of the VZ/SVZ of the mouse by logistic regression. Similarity in gene expression is displayed in red, less similarity in blue. Mouse tumor

cell clusters show similarity to precursor cells of the stem cell niche, suggesting a tumor origin in this region. **f** UMAP and bargraph depicting the changes in cell composition of samples of different mouse ages. The TME content is reduced in later tumor stages. Early tumors are only OL-like, during tumor development, an NB-like population appears. At P92, a unique tumor cell population is detected, showing neither OL nor NB-like features. AC astrocyte, aNSC adult neural stem cell, DC dendritic cell, EC endothelial cells, EPN ependymal cells, MES mesenchymal, MG microglia, migr. DC migratory dendritic cells, MOL mature oligodendrocyte, mural/fibroblasts, pericytes, smooth muscle cells etc., NB neuroblast, OL oligodendrocytic, OPC oligodendrocytic precursor cell, PC pericytes, SMC smooth muscle cells, TAC transit amplifying cell, TME tumor microenvironment, UMAP uniform manifold approximation and projection.

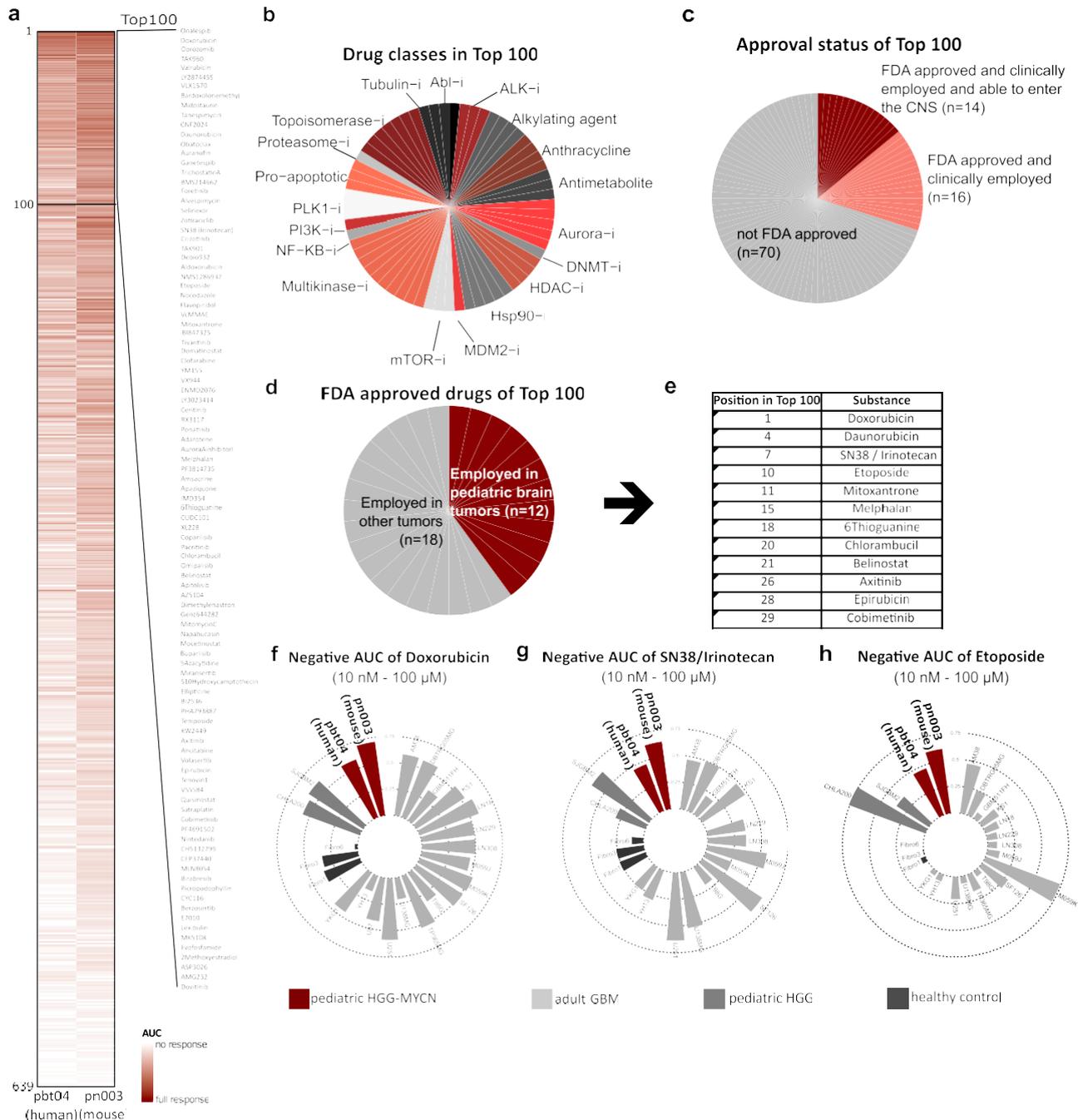


Fig. 5 | High throughput drug screen indicates efficacy of Doxorubicin, Etoposide, and Irinotecan for the treatment of HGG-MYCN. a Heatmap of the AUC in a drug screen of 639 substances in human (pbt-04) and murine (pn003) HGGMYCN cell lines. A darker color indicates a stronger response. **b** Among the top 100 substances of the drug screen, we detected 18 drug classes with at least two substances. The most prominent were anthracyclins, aurora inhibitors, Hsp90- inhibitors, multikinase inhibitors, and topoisomerase inhibitors. **c** Of the top 100 most effective substances in the screen, 30 were FDA approved. Of those, 14 have

been described to be delivered into the CNS. **d** Of the 30 FDA-approved substances, 12 have been used to treat pediatric brain tumor patients as depicted in (e). **f-h** The top three substances Doxorubicin, SN-38 (active metabolite of Irinotecan) and Etoposide are efficiently impairing growth of HGG-MYCN cells, while showing a mixed effect on other glioma cell lines and almost no effect on healthy fibroblasts. Depicted is the negative AUC (1 minus the respective AUC). AUC area under the curve. AUC values are supplied in Supplementary data 1. Source data are provided as a Source Data file.

MYCN drives different brain tumor types depending on the cell of origin³⁰. However, although both promoters have previously been described to drive medulloblastoma formation in mice^{40,41}, none of the generated animals developed any brain tumors. The embryonal lethality of MYCN expression and TP53 loss in Sox2- as well as B1bp-expressing cell populations implies that the physiologic expression of at least one of the two genes is essential during normal CNS devel-

opment in this timeframe. This indicates a similar function in human

CNS development which could be a hint to the time of human HGG-MYCN development.

We particularly examined the cerebellum and the spine of our *hGFAP-cre::Trp53^{Fl/Fl}::Isl-MYCN* mice for signs of tumor development. Since no tumors were detected we conclude that loss of p53 and overexpression of MYCN, despite being expressed there, is either not sufficient for the development of medulloblastoma or spinal ependy- moma or other promoters may be needed. However, we cannot rule

out that *hGFAP-cre::Trp53^{fl/fl}::Isl-MYC*N mice died from their forebrain tumors prior to the development of further lesions. Together with the scRNAseq data, we can deduce that the cell of origin for the murine HGG-MYC*N* is most likely a relatively undifferentiated precursor cell that is able to differentiate into the oligodendroglial and neuronal lineage as we identified cells differentiating along both lineages in our tumors.

We neither detected any glioma in the cerebellum of the mice, although in humans, at least a few cerebellar cases of HGG-MYC*N* have been described in the literature⁴². In our mice, tumors appear to initiate in the OB, which may be caused by differences in the sensory input of mice and humans. Brain tumor development and especially glioblastoma formation is known to be dependent on sensory input, and in humans, other brain regions, such as visual and hearing systems, might be more prone to tumor development^{43,44}. In addition, tumor development in our mouse model coincides with the area of adult neurogenesis in mice, which is highest in the SVZ, from where cells migrate into the OB. In humans, the highest rate of adult neurogenesis can be found in the striatum and almost no neurons are added to the OB after initial development⁴⁵. This species difference in neurogenesis might also contribute to differences in tumor location.

Concluding, our mouse model of HGG-MYC*N* has limitations in modeling the human disease. The genetic alteration driving the mouse tumors are observed in humans, but the combination of both is only found in 36% of human tumors. Therefore, our mouse model is genetically only mimicking this subset of human tumors. In addition, also the age of tumor onset, the localization of tumors, as well as the exact copy number alterations are different between mouse and human HGG-MYC*N*. We show similarities between the two species on a molecular level including gene expression and DNA methylation, but these methods are also not a definite proof for the translational relevance of our model. However, all employed methods for comparing mouse and human tumors including histology, transcriptomic, and epigenomics indicate a resemblance of human HGG-MYC*N* by mouse HGG-MYC*N*. The observed differences most likely depend on species-dependent differences in neurogenesis and external stimuli as discussed above. A genetically engineered mouse model has inherent limitations as mice are different from human in many aspects. Therefore, conclusions drawn from mouse models can only answer some questions of human tumor development. Anyways, modeling a tumor in an intact organism with a functional immune system and a complete organ system, will help understand aspects of the tumor. Employing a mouse model together with other techniques of modeling human tumors will help increasing the treatability and survival of patients. MYC*N* was discovered in 1983 by its tumor-driving role in neuroblastoma^{8,9}, a tumor of the sympathetic nervous system, and mouse models, in which tumors were initiated by amplification or forced expression of MYC*N*, were neuroblastoma models. Apart from these murine neuroblastoma, a few MYC*N*-driven brain tumor models have been described. One model, generated by transduction of NSCs with a mutationally stabilized MYC*N* and transplantation of respective cells into nude mice, generated forebrain tumors with features of glioma. However, at that time, techniques were limited to thoroughly compare the molecular landscapes of mouse and human tumors³⁰.

We combined MYC*N* amplification with an *hGFAP*-mediated loss of *Trp53*. The latter alone has been described to induce formation of IDHwt/H3wt glioblastoma with a penetrance of 100%, although tumor formation occurs significantly later in life than tumor development in our HGG-MYC*N* mice³³. So, neither a *hGFAP*-mediated MYC*N* amplification³² nor a *hGFAP*-mediated *TP53* loss³³ is sufficient to drive HGG-MYC*N* formation. Only the combination of both alterations cooperate to induce such tumors. A possible mechanism of how these two alterations co-act to induce tumor formation is that MYC*N* drives cell proliferation, but only in cells that are genetically unstable for example by a loss of *Trp53*. This increased proliferation can lead to

tumor formation, whereas cells with a functional p53 will activate cell cycle checkpoints and prevent uncontrolled proliferation.

We show that cells derived from our mouse model can be cultivated in vitro and be employed for preclinical testing, which is required as a preparation for the in vivo application of potential drugs.

We employed our mouse model and a human cell line to find potential therapies in a high-throughput drug screen. Thereby, we identified Etoposide, Doxorubicin, and Irinotecan as potential treatment options in patients with HGG-MYC*N*. All these substances are used in the treatment of pediatric brain tumors and can therefore be considered safe to use. We think that using these substances instead or in addition to the current treatment with radiation and optionally Temozolomide might substantially improve patient survival, although further in vivo studies are required. These in vivo studies could include the comparison of the current treatment with radiation and Temozolomide and Irinotecan, Etoposide, Doxorubicin and any combination of those substances. In addition, more targeted treatments described for other MYC*N* driven tumors like for example AURKA-, CDK2/9- and DNA replication repair inhibitors could be tested. In addition, it might be useful to employ our mouse model as well as PDX models of HGG-MYC*N* in such treatment studies. Our here described mouse model mimics the human disease regarding histology, DNA methylation as well as gene expression. In summary, we believe that it can serve as a reliable preclinical model for this poorly understood brain tumor entity.

Methods

All work performed in this project complies with all relevant ethical regulations including animal protection laws of the state of Hamburg and the local ethical standards and regulations at the University Medical Center Hamburg-Eppendorf (§12 HmbKHG). Only anonymized human data were included in the project and informed consent of all human patients was obtained.

Transgenic animals

The generation of *hGFAP-cre*³⁶, *Blbp-cre*⁴⁶, *Sox2-cre*⁴⁷, *Trp53^{fl/fl}*⁴⁸, and *Isl-MYC*N**²⁶ as well as *Math1-cre::SmoM2^{fl/wt}*³⁸ transgenic mouse lines has been described previously. All animal procedures were performed in accordance with applicable animal protection laws and approved by the state of Hamburg (Reference N2019/99). Genotyping was performed by polymerase-chain reaction with primer pairs described in the original publications (Supplementary Table 1). For all analyses, mice of both sexes were used in equal numbers. The mice were kept at 22 °C and 45–65% humidity on a 12 h light/dark cycle and water and food was available ad libitum. The animals were monitored daily by experienced personnel. All animals were sacrificed upon detectable neurological symptoms including ataxia and hydrocephalus as approved by the state of Hamburg. The termination criteria defined were never exceeded in the study.

Human	cell	line	culture
Pbt04 cells (SCRI, Brain tumor resource lab) were grown in NeuroCultNS-A Basal Medium (Human) supplemented with NeuroCult NS-A Proliferation Supplements—Human, Epidermal Growth Factor (EGF); Fibroblast Growth Factor (FGF) and penicillin/streptomycin (P/S) in laminin coated flasks.			

Primary murine tumor cell culture

For the establishment of murine tumor cell lines, fresh tumors were dissected and dissociated with Accutase and DNase. After dissociation, cells were seeded in mouse NSC medium (DMEM/F12 + Glutamax, HEPES, MEM-Non Essential Amino Acids, P/S, B27 Supplement, EGF, FGF) and grown in a humidified incubator at 37 °C. As soon as large

spheres formed, they were dissociated with Accutase and re-seeded.

High-throughput drug screening

High-throughput drug screening using an in-house semi-automated platform was performed. For this approach, we measured the drug responses of pbt-04 and a murine HGG-MYCN cancer cell line to 639 small-molecule compounds (including FDA-approved, phase I-IV, and experimental drugs). All compounds were dispensed using a D300e Digital printer (Tecan) in a 6–8-point serial dilution covering a concentration range from 0.0043 to 25 μM in 1536-well plates (Corning). After 72 h treatment, cell viability was detected by CellTiter-Glo Luminescent cell viability assay (Promega) using a Spark plate reader (Tecan). Dose-response to compounds was measured based on a normalized area under the curve (AUC). The AUC values for all compounds can be found in Supplementary data 1. AUC data of additional glioma cell lines as well as healthy control fibroblasts were generously provided for comparison by Marc Remke and Nan Qin.

Histology, immunohistochemistry, and FISH-analysis

For hematoxylin and eosin (H&E) and immunohistochemistry (IHC) stains, brain tissue was fixed in 4% paraformaldehyde/PBS for at least 12 h. The tissue was dehydrated, embedded in paraffin, and sectioned at 4 μm according to standard protocols. All IHC stains were performed on a Ventana System (Roche) using standard protocols. The following antibodies were employed: MYCN (Cell Signaling, 517053, 1:1000), P53 (Dako, M7001, 1:800), Ki67 (Abcam, ab16667, 1:100), Nestin (Abcam, ab221660, 1:2000), SOX2 (Abcam, ab97959, 1:200), OLIG2 (Millipore, AB9610, 1:200), Cre (Covance / DCS-diagnostics, PRB-106P, 1:100) and GFAP (Dako, M0761, 1:200). FISH analyses were performed to detect possible amplifications at the MYCN locus using standard procedures and a SPEC MYCN/2q11 dual color probe (Zytovision, Germany).

RNA Sequencing (RNA Seq)

We employed 11 mouse HGG-MYCN tumor samples (6 male and 5 female) as well as 6 mouse SHH-MBs (sex not determined) and three control samples (2 male and 1 female) of the cerebellum and three of the olfactory bulb (2 male and 1 female). After isolation of total RNA with TRIzol (Invitrogen), RNA integrity was analyzed with the RNA 6000 Nano Chip on an Agilent 2100 Bioanalyzer (Agilent Technologies). From total RNA, mRNA was extracted using the NEBNext Poly(A) mRNA Magnetic Isolation module (New England Biolabs) and RNA-Seq libraries were generated using the NEXTFLEX Rapid Directional qRNA-Seq Kit (Bioo Scientific) as per the manufacturer's recommendations. Concentrations of all samples were measured with a Qubit 2.0 Fluorometer (Thermo Fisher Scientific), and distribution of fragment lengths of the final libraries was analyzed with the DNA High Sensitivity Chip on an Agilent 2100 Bioanalyzer (Agilent Technologies). All samples were normalized to 2 nM and pooled equimolarly. The library pool was sequenced on the NextSeq500 (Illumina) with 1 \times 75 bp read length and 16.1 to 18.6 Mio reads per sample.

For each sample, sufficient quality of the raw reads was confirmed by FastQC v0.11.8⁴⁹. Afterwards, the reads were aligned to the mouse reference genome GRCh38 with STAR v2.6.1c⁵⁰ and simultaneously counted per gene by employing the *-quantmode GeneCounts* option. Counts are based on the Ensembl annotation release 95. Differentially expressed genes were estimated with DESeq2 v1.22.2⁵¹.

RNA Sequencing analysis

Raw fastq files of mouse samples were processed in usegalaxy.eu⁵². Low quality reads were detected using FastQC (Galaxy Version 0.73+galaxy0). Trimmomatic (Galaxy Version 0.38.1) was used for trimming poor quality reads (reads with average quality <20). Reads were aligned to the mm39.ncbiRefseq.gtf.gz using STAR aligner (Galaxy Version 2.7.8a+galaxy1) and gene expression was quantified using featureCounts (Galaxy Version 2.0.1+galaxy2). DESeq2 (Galaxy Version 2.11.40.7+galaxy2) was used for generating VST-normalized

files for all samples. Human gene expression data was obtained from GSE73038¹. Mouse samples were measured in two different batches, and the VST-normalized files were combined and corrected for batch effect using ComBat from sva package (3.44.0) in Rstudio (4.2.1). For comparing mouse (11 tumors with MYCN amp and *Trp53* mutation and 6 SHH tumors) and human data (173 samples), only orthologous genes were used ($n = 14,416$). In order to correct for species-specific batch effect, ComBat was applied after combining the mouse and human data (GSE73038¹). Average tumor subgroup-specific gene expression was used for calculating euclidean distance. Sample-sample distance plot was visualized using ComplexHeatmap (2.12.1). Limma (3.52.2) was used for performing differential expression analysis. Top 500 differentially expressed genes (adjusted for multiple testing using Bonferroni-Hochberg correction and sorted by F-statistic) were visualized using umap (0.2.9.0) and ComplexHeatmap (2.12.1) in Rstudio. AGDEX was performed in C++ as described previously⁵³. AGDEX was based on 14,416 orthologous genes with human sonic hedgehog medulloblastoma (SHH MB) and mouse SHH MB as reference group for differential expression of tumor samples in each species, respectively. The statistical significance of the observed AGDEX value was determined via permutation as described^{39,53}. The significance of changes in expression of MYCN was determined by Welch's *t*-test of non-MYCN and HGG-MYCN human tumors and mouse OB vs. mouse HGG-MYCN. Gene set enrichment analysis (GSEA) was performed using the software GSEA v4.2.0 of the Broad Institute^{54,55}. The gene set comprised 344 genes with transcription factor binding evidence in the MYCN-21190229-SHEP-21N-HUMAN profile from the CHEA Transcription Factor Binding Site Profiles dataset^{56,57}. Genes were ranked by signal-to-noise ratio and statistical significance was determined via 1000 gene set permutations. $P < 0.05$ was considered significant.

Global DNA-methylation analysis

Three mouse HGG-MYCN tumors were employed to methylation array (two male and one female). After isolation of genomic DNA, 200–500 ng DNA was used for bisulfite conversion by the EZ DNA Methylation Kit (Zymo Research). Then, the Mouse Methylation Bead Chip Array (Illumina) covering 285,000 CpG sites on the mouse genome was used on an iScan device (Illumina).

In order to compare the global DNA-methylation profile of mouse tumors with that of human samples, we collected human methylation profiles generated with the Illumina EPIC array, consisting of 850,000 CpG sites. The use of biopsy-specimens for research upon anonymization was always in accordance with local ethical standards and regulations at the University Medical Center Hamburg-Eppendorf. Apart from data generated for diagnostic purposes in-house and by cooperation partners, we included data generated and published by Capper et al. in 2018 (GSE73801)⁴² for the overview of human brain tumor entities. The raw data were preprocessed and beta values were extracted with minfi⁵⁸. The 10,000 most differentially methylated CpG sites were used to generate a UMAP with the UMAP package⁵⁹.

For the comparison of mouse and human tumors, only EPIC data were employed and $n = 431$ human brain tumors were included. Beta values were extracted from human idat files using minfi and the entire human dataset was quantile normalized. Mouse data were processed with SeSAMe⁶⁰ (v1.12.9), beta values were extracted and the data were quantile normalized. 141 identical CpG sites were extracted from both datasets and human and mouse datasets were combined. After combination, the mixed mouse-human dataset was quantile normalized and an UMAP was generated.

The heatmap of the copy number variation (CNV) of 47 HGG-MYCN was created from idat files. CNV data was generated with the Conumee package (v1.28.0)⁶¹. CNV values of the bin signals were calculated and depicted as a heatmap using ComplexHeatmap. Mouse CNV plots were created similarly by employing the conumee pipeline

with a few modifications including using an in-house generated reference set of normal mouse tissue.

Single-cell RNA sequencing analysis

Mouse tumors of seven *hGFAP-cre::Trp53^{Fl/Fl}::Isl-MYC^{Fl/wt}* mice (43–92 days old, 5 male and 2 females) were isolated and minced. Samples were digested for 30 min at 37 °C in a freshly prepared solution of papain (Worthington) in pre-warmed DMEM/F12 medium (Thermo Fisher) with DNase (Worthington) and passed through a 40 µm cell strainer (Corning). Red blood cells were depleted using ACKlysis buffer (Thermo Fisher), and single-cell suspensions were cryopreserved. After thawing, non-vital cells that stained positive for 7-Aminoactinomycin D (eBioscience) were removed using a BD FACS Aria II cell sorter (BD Biosciences). Approximately 10,000 vital cells were used as input for scRNA-seq. Single-index libraries were generated with Chromium Single Cell 3' v3.1 technology (10x Genomics) and sequenced using a NextSeq 2000 sequencing instrument (high-throughput kit, 100 cycles) at the Genomics Core Facility (University Hospital Münster, Germany) after quality control using a TapeStation 2200 (Agilent Technologies). The samples were analyzed with the 10x Genomics Cell Ranger v6.0.2 pipeline⁶² and Seurat R package v4.0.5⁶³. Raw data were converted to fastq format with the Cell Ranger mkfastq function and then aligned against the murine reference transcriptome mm10 v2020-A with Cell Ranger count and default values. Seurat objects were generated for the samples based on the following filter criteria: at least three cells, a minimum feature count of 200, and cells with <25% of mitochondrial genes. Outlier cells with a high nCount_RNA value were classified as doublets and removed (threshold: 40,000–60,000). The filtered data were then normalized, integrated, and clustered with Seurat, using a resolution parameter of 0.5. Feature plots, UMAPs, dotplots, and heatmap visualizations were created with Seurat functions; a cluster-based cell type annotation was conducted based on the expression of characteristic marker genes per cell type. Lists of differentially expressed genes per cluster were calculated with Seurat's findMarkers function, using the MAST test "hinter" using a resolution parameter of 0.5. The genes for each signature depicted in Fig. 4 are shown in Supplementary Fig. 4.

Finally, a logistic regression analysis based on the approach of Young et al.⁶⁴ was used to identify similarities between the murine scRNA-seq clusters and a reference dataset by Mizrak et al.⁶⁵ (GSE134918). Briefly, a multinomial regression model was trained on the chosen reference dataset with the R package glmnet⁶⁶, and cell types were predicted for the original mouse data based on this model. Probabilities per reference cell type were aggregated on cluster level using R's mean function, and visualized as heatmaps with the package pheatmap⁶⁷.

Statistical information

Animal survival was depicted as a Kaplan–Meier curve, made and assessed using GraphPad Prism 9 software. Significant *p* values are reported for appropriate figures in the figure legend. The sample size (*n*) is given in the figure legend or in the respective figure panel. Error bars represent the standard error of mean. The cut-off values for bioinformatic analyses are noted in the respective methods section. Further information is included in the Reporting summary.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The DNA methylation and RNA sequencing data including the raw data generated in this study have been deposited in the GEO database under accession code GSE227413. The scRNA sequencing data has

been deposited in the GEO database under accession code GSE237237. The human gene expression data was obtained from GSE73038. Human DNA methylation data was obtained from GSE73801 and GSE215240. Mouse scRNA sequencing data publicly available from GSE134918 was employed in this study. AUC data from other glioma cell lines were provided by Marc Remke and Nan Qin. Source data are provided with this paper.

References

1. Sturm, D. et al. New brain tumor entities emerge from molecular classification of CNS-PNETs. *Cell* 164, 1060–1072 (2016).
2. Korsunov, A. et al. H3-IDH-wild type pediatric glioblastoma is comprised of molecularly and prognostically distinct subtypes with associated oncogenic drivers. *Acta Neuropathol.* 134, 507–516 (2017).
3. Sturm, D. et al. Multiomic neuropathology improves diagnostic accuracy in pediatric neuro-oncology. *Nat. Med.* 29, 917–926 (2023).
4. Schoof, M. et al. Malignant gliomas with H3F3A G34R mutation or MYCN amplification in pediatric patients with Li Fraumeni syndrome. *Acta Neuropathol.* 142, 591–593 (2021).
5. Guerrini-Rousseau, L. et al. Pediatric high-grade glioma MYCN is frequently associated with Li-Fraumeni syndrome. *Acta Neuropathol. Commun.* 11, 3 (2023).
6. Knoepfler, P. S., Cheng, P. F. & Eisenman, R. N. N-myc is essential during neurogenesis for the rapid expansion of progenitor cell populations and the inhibition of neuronal differentiation. *Genes Dev.* 16, 2699–2712 (2002).
7. van Bokhoven, H. et al. MYCN haploinsufficiency is associated with reduced brain size and intestinal atresias in Feingold syndrome. *Nat. Genet.* 37, 465–467 (2005).
8. Schwab, M. et al. Amplified DNA with limited homology to myc cellular oncogene is shared by human neuroblastoma cell lines and a neuroblastoma tumour. *Nature* 305, 245–248 (1983).
9. Kohl, N. E. et al. Transposition and amplification of oncogene-related sequences in human neuroblastomas. *Cell* 35, 359–367 (1983).
10. Schwab, M. et al. Chromosome localization in normal human cells and neuroblastomas of a gene related to c-myc. *Nature* 308, 288–291 (1984).
11. Seeger, R. C. et al. Association of multiple copies of the N-myc oncogene with rapid progression of neuroblastomas. *N. Engl. J. Med.* 313, 1111–1116 (1985).
12. Campbell, K. et al. Association of heterogeneous MYCN amplification with clinical features, biological characteristics and outcomes in neuroblastoma: a report from the Children's Oncology Group. *Eur. J. Cancer* 133, 112–119 (2020).
13. Hirvonen, H., Hukkanen, V., Salmi, T. T., Pelliniemi, T.-T. & Alitalo, R. L-myc and N-myc in hematopoietic malignancies. *Leuk. Lymphoma* 11, 197–205 (1993).
14. van Lohuizen, M., Breuer, M. & Berns, A. N-myc is frequently activated by proviral insertion in MuLV-induced T cell lymphomas. *EMBO J.* 8, 133–136 (1989).
15. Nau, M. M. et al. Human small-cell lung cancers show amplification and expression of the N-myc gene. *Proc. Natl Acad. Sci. USA* 83, 1092–1096 (1986).
16. Rickman, D. S., Schulte, J. H. & Eilers, M. The expanding world of N-MYC-driven tumors. *Cancer Discov.* 8, 150–163 (2018).
17. Lee, W.-H., Murphree, A. L. & Benedict, W. F. Expression and amplification of the N-myc gene in primary retinoblastoma. *Nature* 309, 458–460 (1984).
18. Aldosari, N. et al. MYCC and MYCN oncogene amplification in medulloblastoma. A fluorescence in situ hybridization study on paraffin sections from the Children's Oncology Group. *Arch. Pathol. Lab. Med.* 126, 540–544 (2002).

19. Gessi, M. et al. MYCN amplification predicts poor outcome for patients with supratentorial primitive neuroectodermal tumors of the central nervous system. *Neuro Oncol.* 16, 924–932 (2014).
20. Northcott, P. A. et al. The whole-genome landscape of medulloblastoma subtypes. *Nature* 547, 311–317 (2017).
21. Jones, D. T. et al. Dissecting the genomic complexity underlying medulloblastoma. *Nature* 488, 100–105 (2012).
22. Kool, M. et al. Genome sequencing of SHH medulloblastoma predicts genotype-related response to smoothed inhibition. *Cancer Cell* 25, 393–405 (2014).
23. Bjerke, L. et al. Histone H3.3 mutations drive pediatric glioblastoma through upregulation of MYCN. *Cancer Discov.* 3, 512–519 (2013).
24. Brennan, C. et al. Glioblastoma subclasses can be defined by activity among signal transduction pathways and associated genomic alterations. *PLoS ONE* 4, e7752 (2009).
25. Perry, A. et al. Malignant gliomas with primitive neuroectodermal tumor-like components: a clinicopathologic and genetic study of 53 cases. *Brain Pathol.* 19, 81–90 (2009).
26. Althoff, K. et al. A Cre-conditional MYCN-driven neuroblastoma mouse model as an improved tool for preclinical studies. *Oncogene* 34, 3357–3368 (2015).
27. Weiss, W. A., Aldape, K., Mohapatra, G., Feuerstein, B. G. & Bishop, J. M. Targeted expression of MYCN causes neuroblastoma in transgenic mice. *EMBO J.* 16, 2985–2995 (1997).
28. Swartling, F. J. et al. Pleiotropic role for MYCN in medulloblastoma. *Genes Dev.* 24, 1059–1072 (2010).
29. Taylor, M. D. et al. Molecular subgroups of medulloblastoma: the current consensus. *Acta Neuropathologica* 123, 465–472 (2012).
30. Swartling, F. J. et al. Distinct neural stem cell populations give rise to disparate brain tumors in response to N-MYC. *Cancer Cell* 21, 601–613 (2012).
31. Holdhof, D., On, J. H., Schoof, M., Göbel, C. & Schüller, U. Simultaneous Brg1 knockout and MYCN overexpression in cerebellar granule neuron precursors is insufficient to drive tumor formation but temporarily enhances their proliferation and delays their migration. *Cerebellum* 20, 410–419 (2021).
32. Fielitz, K. et al. Characterization of pancreatic glucagon-producing tumors and pituitary gland tumors in transgenic mice overexpressing MYCN in hGFAP-positive cells. *Oncotarget* 7, 74415–74426 (2016).
33. Li, Y. et al. Murine models of IDH-wild-type glioblastoma exhibit spatial segregation of tumor initiation and manifestation during evolution. *Nat. Commun.* 11, 3669 (2020).
34. Tauziède-Espariat, A. et al. The pediatric supratentorial MYCN-amplified high-grade gliomas methylation class presents the same radiological, histopathological and molecular features as their pontine counterparts. *Acta Neuropathol. Commun.* 8, 104 (2020).
35. Tauziède-Espariat, A. et al. An integrative radiological, histopathological and molecular analysis of pediatric pontine histone-wildtype glioma with MYCN amplification (HGG-MYCN). *Acta Neuropathol. Commun.* 7, 87 (2019).
36. Zhuo, L. et al. hGFAP-cre transgenic mice for manipulation of glial and neuronal function in vivo. *Genesis* 31, 85–94 (2001).
37. Ghasemi, D. R. et al. MYCN amplification drives an aggressive form of spinal ependymoma. *Acta Neuropathol.* 138, 1075–1089 (2019).
38. Schüller, U. et al. Acquisition of granule neuron precursor identity is a critical determinant of progenitor cell competence to form Shh-induced medulloblastoma. *Cancer Cell* 14, 123–134 (2008).
39. Pounds, S. et al. A procedure to statistically evaluate agreement of differential expression for cross-species genomics. *Bioinformatics* 27, 2098–2103 (2011).
40. Ahlfeld, J. et al. Sox2 requirement in sonic hedgehog-associated medulloblastoma. *Cancer Res.* 73, 3796–3807 (2013).
41. Gibson, P. et al. Subtypes of medulloblastoma have distinct developmental origins. *Nature* 468, 1095–1099 (2010).
42. Capper, D. et al. DNA methylation-based classification of central nervous system tumours. *Nature* 555, 469–474 (2018).
43. Pan, Y. et al. NF1 mutation drives neuronal activity-dependent initiation of optic glioma. *Nature* 594, 277–282 (2021).
44. Chen, P. et al. Olfactory sensory experience regulates gliomagenesis via neuronal IGF1. *Nature* 606, 550–556 (2022).
45. Ernst, A. & Frisén, J. Adult neurogenesis in humans—common and unique traits in mammals. *PLoS Biol.* 13, e1002045 (2015).
46. Hegedus, B. et al. Neurofibromatosis-1 regulates neuronal and glial cell differentiation from neuroglial progenitors in vivo by both cAMP- and Ras-dependent mechanisms. *Cell Stem Cell* 1, 443–457 (2007).
47. Hayashi, S., Lewis, P., Pevny, L. & McMahon, A. P. Efficient gene modulation in mouse epiblast using a Sox2Cre transgenic mouse strain. *Mech. Dev.* 119, S97–S101 (2002).
48. Jonkers, J. et al. Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer. *Nat. Genet.* 29, 418–425 (2001).
49. Andrews, S. FASTQC, A quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc> (2010).
50. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21 (2012).
51. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550 (2014).
52. Afgan, E. et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 46, W537–W544 (2018).
53. Pöschl, J. et al. Genomic and transcriptomic analyses match medulloblastoma mouse models to their human counterparts. *Acta Neuropathol.* 128, 123–136 (2014).
54. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* 102, 15545–15550 (2005).
55. Mootha, V. K. et al. PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* 34, 267–273 (2003).
56. Rouillard, A. D. et al. The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database (Oxford)* 2016, <https://doi.org/10.1093/database/baw100> (2016).
57. Lachmann, A. et al. ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. *Bioinformatics* 26, 2438–2444 (2010).
58. Aryee, M. J. et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* 30, 1363–1369 (2014).
59. McInnes, L. & Healy, J. UMAP: uniform manifold approximation and projection for dimension reduction. Preprint at <https://arxiv.org/abs/1802.03426> (2020).
60. Zhou, W., Triche, T. J. Jr, Laird, P. W. & Shen, H. SeSAMe: reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions. *Nucleic Acids Res.* 46, e123–e123 (2018).
61. Hovestadt, V. & Zapatka, M. conumee: Enhanced copy-number variation analysis using Illumina DNA methylation arrays. R package version 1.9.0. <http://bioconductor.org/packages/conumee/> (2015).
62. Zheng, G. X. Y. et al. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* 8, 14049 (2017).
63. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* 184, 3573–3587. e3529 (2021).
64. Young, M. D. et al. Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science* 361, 594–599 (2018).

65. Mizrak, D. et al. Single-cell profiling and SCOPE-Seq reveal lineage dynamics of adult ventricular-subventricular zone neurogenesis and NOTUM as a key regulator. *Cell Rep.* 31, 107805 (2020).
66. Friedman, J. H., Hastie, T. & Tibshirani, R. Regularization paths for generalized linear models via coordinated descent. *J. Stat. Softw.* 33, 1–22 (2010).
67. Kolde, R. pheatmap: Pretty Heatmaps. R package version 1.0.12. CRAN. *R-project.org/package=pheatmap* (2019).

Acknowledgements

We are thankful for the excellent technical support by Jacqueline Tischendorf, Celina Soltwedel, and Daniela Jeising. We like to thank the animal facility of the LIV Hamburg, Daniela Indenbirken, and the sequencing facility at the LIV Hamburg and the Fred Hutchinson Cancer Research Center for providing us with the pbt-04 cell line. A.K.A. and A.E. are thankful for the support within the interdisciplinary graduate school “Innovative Technologies in Cancer Diagnostics and Therapy” funded by the City of Hamburg. US received support from the Wilhelm-Sander Stiftung and from the Fördergemeinschaft Kinderkrebszentrum Hamburg. K.K. received support by the Kinderkrebshilfe Münster e. V. A.B. (Annika Ballast) was funded by the Medizinerkolleg (MedK) Muenster.

Author contributions

Conceptualization, M.S. (Melanie Schoof), U.S., T.K.A., J.E.N., M.R. and K.K.; Performed experiments, M.S., S.G., T.K.A., M.D., C.W., A.B., N.Q., M.B.O., C.G., S.N., D.H., C.K., L.P., G.D.E., V.T., F.M., A.A., A.E., D.M. and M.Sp. (Michael Spohn), A.V.; Resources, M.B., M.M., S.R., Ma.S., J.V., N.S., D.T.W.J., M.R., J.E.N., K.K. and U.S.; Writing, M.S., T.K.A., U.S. and K.K.; Funding acquisition, U.S. and K.K. All authors provided critical feedback, helped shape the research and manuscript, and commented on the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

F.M. received support for meeting attendance from Servier, AbbVie, Incyte, Gilead, Jazz Pharmaceuticals, Novartis, Teva, Pfizer and Amgen; received support for medical writing from Servier and Springer Verlag; received research grants from Apis Technologies, Daiichi Sankyo and

received speaker honoraria from Servier, Jazz Pharmaceuticals, AbbVie, Astellas Pharma, Bristol-Myers Squibb, MSD Sharp & Dohme, Novartis, Pfizer, and Stemline Therapeutics. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-43564-w>.

Correspondence and requests for materials should be addressed to Ulrich Schüller.

Peer review information *Nature Communications* thanks Karisa Schreck, Adam Green and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions 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.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

¹Research Institute Children's Cancer Center, Hamburg, Germany. ²Department of Pediatric Hematology and Oncology, University Medical Center, Hamburg-Eppendorf, Hamburg, Germany. ³Center for Molecular Neurobiology (ZMNH), University Medical Center, Hamburg-Eppendorf, Hamburg, Germany. ⁴Pediatric Hematology and Oncology, University Children's Hospital Muenster, Muenster, Germany. ⁵Institute of Neuropathology, University Medical Center, Hamburg-Eppendorf, Hamburg, Germany. ⁶Institute of Medical Informatics, University of Muenster, Muenster, Germany. ⁷German Cancer Consortium (DKTK), Partner Site Essen/Düsseldorf, Düsseldorf, Germany. ⁸Department of Pediatric Oncology, Hematology, and Clinical Immunology, Medical Faculty, Heinrich Heine University, University Hospital Düsseldorf, Düsseldorf, Germany. ⁹Institute of Neuropathology, Heinrich Heine University, University Hospital Düsseldorf, Düsseldorf, Germany. ¹⁰High-Throughput Drug Screening Core Facility, Medical Faculty, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany. ¹¹Mildred Scheel Cancer Career Center HaTriCS4 University Medical Center Hamburg-Eppendorf, Hamburg, Germany. ¹²Hopp Children's Cancer Center (KiTZ), Heidelberg, Germany. ¹³Pediatric Glioma Research Group, German Cancer Research Center (DKFZ), Heidelberg, Germany. ¹⁴Department of Oncology, Hematology and Bone marrow transplantation, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. ¹⁵Division of Pediatric Neurooncology, German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany. ¹⁶Department of Radiotherapy & Radiation Oncology, Hubertus Wald Tumorzentrum-University Cancer Center Hamburg (UCCH), University Medical Center Hamburg-Eppendorf, Hamburg, Germany. ¹⁷These authors contributed equally: Kornelius Kerl, and Ulrich Schüller. ✉ e-mail: u.schueller@uke.de

5. Summary in German

Diese Studie liefert umfassende Einblicke in die molekulare Charakterisierung von Medulloblastomen (MB) und pädiatrischen hochgradigen Gliomen (HGG) sowie die Ähnlichkeiten gentechnisch veränderter Mausmodelle von HGG MYCN-amplified und Group3 MB mit menschlichen Gegenstücken und hilft bei der Identifizierung von mögliche therapeutische Ziele und personalisierte Behandlungsstrategien.

Zunächst konzentrierten wir uns auf die multiomische Profilierung von MB mit dem Ziel, klinisch relevante Subtypen und alternative Behandlungsziele zu identifizieren. Durch eine groß angelegte Proteomanalyse von FFPE-MB-Tumoren, die mit öffentlich verfügbaren Daten harmonisiert wurde, wurden sechs stabile Proteomcluster entdeckt, die jeweils mit unterschiedlichen DNA-Methylom- und Transkriptom-Subtypen assoziiert sind. Die Überlebensanalyse ergab, dass pG3myc MB die geringste Gesamtüberlebensrate aufwies, während pWNT-Patienten die beste Überlebensrate aufwiesen. Darüber hinaus enthüllte eine gruppenspezifische Korrelationsanalyse von DNA-Methylom- und Proteomdaten subtypspezifische Korrelationen, wobei PALMD als potenzieller Biomarker für den Hochrisiko-Subtyp pG3myc identifiziert wurde. Darüber hinaus deutete die N-Glykan-Profilierung auf unterschiedliche Ziele der Immuntherapie bei Hochrisiko-MB-Subtypen hin.

Anschließend untersuchten wir Mausmodelle von HGG mit MYCN-Amplifikation und betonten deren Ähnlichkeit mit menschlichen Gegenstücken. DNA-Methylierung und transkriptomische Profile von murinen Tumoren zeigten eine hohe Ähnlichkeit mit menschlichen HGG-MYCN-Tumoren. Der Vergleich von SHH-Mausmodellen mit ihrem menschlichen Gegenstück bestätigte erwartete Ähnlichkeiten und weist auf die Nützlichkeit von Mausmodellen bei der Untersuchung menschlicher Hirntumoren und der Entwicklung therapeutischer Strategien hin.

Schließlich haben wir uns mit der MYC/SMARCA4-gesteuerten MB-Bildung bei Mäusen befasst, diese Tumoren auf molekularer Ebene charakterisiert und ihre Ähnlichkeit mit menschlichen MB der Gruppe 3 aufgeklärt. Die RNA-Sequenzierung zeigte eine Hochregulierung von Myc und eine Herunterregulierung von Genen, die mit der neuronalen Entwicklung verbunden sind. Die Integration mit menschlichen MB-Datensätzen ergab eine molekulare Ähnlichkeit mit MB der Gruppe 3 basierend auf Genexpressions- und DNA-Methylierungsmustern.

Insgesamt liefert die Studie ein umfassendes Verständnis von MB- und HGG-MYCN-amplifizierten Tumoren, wobei der Schwerpunkt auf molekularen Subtypen, potenziellen Biomarkern und therapeutischen Zielen liegt. Diese Erkenntnisse ebnen den Weg für personalisierte Behandlungsstrategien und verbesserte Patientenergebnisse bei diesen verheerenden Hirntumoren bei Kindern.

6. Summary English

This study provides comprehensive insights into the molecular characterization of medulloblastoma (MB) and pediatric high-grade gliomas (HGG), along with the similarities of genetically engineered mouse models of HGG MYCN-amplified and Group3 MB to human counterparts, aiding in the identification of potential therapeutic targets and personalized treatment strategies.

Firstly, we focused on multi-omic profiling of MB, aiming to identify clinically relevant subtypes and alternative treatment targets. Through large-scale proteome analysis of FFPE MB tumours, harmonized with publicly available data revealed six stable proteomic clusters, each associated with distinct DNA methylome and transcriptome subtypes. Survival analysis highlighted pG3myc MB with least overall survival, while pWNT patients exhibited the best survival rate. Moreover, group-specific correlation analysis of DNA methylome and proteome data unveiled subtype-specific correlations, with PALMD as a potential biomarker identified for high-risk pG3myc subtype and TNC as a biomarker for the pWNT MB subtype with a good survival. Additionally, N-glycan profiling suggested differential immune therapy targets across high-risk MB subtypes.

Then we investigated MYC/SMARCA4-driven MB formation in mice, characterizing these tumours on a molecular level and elucidating their similarity to human Group 3 MB. RNA sequencing highlighted upregulation of Myc and downregulation of genes associated with neuronal development. Integration with human MB datasets revealed molecular resemblance to Group 3 MB based on gene expression and DNA methylation patterns.

Finally, we investigated mouse models of HGG with MYCN amplification, emphasizing their resemblance to human counterparts. DNA methylation and transcriptomic profiles of murine tumours demonstrated high similarity to human HGG-MYCN tumours. Comparison of SHH mouse models with its human counterpart confirmed expected

similarities, indicating the utility of mouse models in studying human brain tumours and developing therapeutic strategies

Overall, the study provides a comprehensive understanding of MB and HGG MYCN amplified tumours, emphasizing molecular subtypes, potential biomarkers, and therapeutic targets. These findings pave the way for personalized treatment strategies and improved patient outcomes in these devastating pediatric brain tumours.

7. Statement of own contribution to the publications

Shweta Godbole⁺, Hannah Voß⁺, Antonia Gocke, Simon Schlumbohm, Yannis Schumann, Bojia Peng, Martin Mynarek, Stefan Rutkowski, Matthias Dottermusch, Mario M. Dorostkar, Andrey Korshunov, Thomas Mair, Stefan M. Pfister, Marcel Kwiatkowski, Madlen Hotze, Philipp Neumann, Christian Hartmann, Joachim Weis, Friederike Liesche-Starnecker, Yudong Guan, Manuela Moritz, Bente Siebels, Nina Struve, Hartmut Schlüter, Ulrich Schüller, Christoph Krisp, Julia E. Neumann[#] – **Nature Communications (Accepted, Under editorial revision)**

“Multiomic profiling of medulloblastoma reveals subtype-specific targetable alterations at the proteome and N-glycan level”

Contributions: Conducted experiments, analyses of DNA methylome, proteome and transcriptome data, integration of the individual omics with publicly available data and multi-omic analysis. Wrote and reviewed the original draft of the manuscript.

Carolin Göbel, **Shweta Godbole**, Melanie Schoof, Dörthe Holdhof, Carolin Loose, Julia Neumann and Ulrich Schüller
Acta Neuropathologica Communications (Published)

“MYC overexpression and SMARCA4 loss cooperate to drive medulloblastoma formation in mice” (DOI: [10.1186/s40478-023-01654-2](https://doi.org/10.1186/s40478-023-01654-2))

Contributions: Analyses of the mouse transcriptome data, batch correction, Integration with human data.

Melanie Schoof, **Shweta Godbole**, Thomas K. Albert, Matthias Dottermusch, Carolin Walter, Annika Ballast, NanQin, Marlina Baca Olivera, Carolin Göbel, Sina Neyazi, Catena Kresbach, Levke-Sophie Peter, Dörthe Holdhof, Gefion Dorothea Eppelen, Vanessa Thaden, Michael Spohn, Mirjam Blattner-Johnson, Franziska Modemann, Martin Mynarek, Stefan Rutkowski, Martin Sill, Julian Varghese, Ann-Kristin Afflerbach, Alicia Eckhardt, Daniel Münter, Archana Verma, Nina Struve, David T.W. Jones,

Kornelius Kerl, Marc Remke, Julia E. Neumann & Ulrich Schüller
Nature Communications (Published)

“Mouse models of pediatric high-grade gliomas with MYCN amplification reveal intratumoral heterogeneity and lineage signatures”

(DOI: [10.1038/s41467-023-43564-w](https://doi.org/10.1038/s41467-023-43564-w))

Contributions: Analyses of the mouse transcriptome data, batch correction, Integration with human data.

8. Acknowledgments

First and foremost, I would like to thank my PhD supervisor, Prof. Dr. med. Julia E Neumann. I wouldn't be where I am had she not believed in me. She always supported me, encouraged me to work on new projects, to try new methods and gave me sufficient time to get acquainted with a variety of projects. I will always be grateful for all the exposure I got in this field and I owe it all to you, Julia. Further, I could never get through these four years had I not met my colleagues, friends and our "PhD support group": Jelena Navolic, Sina Brakopp, Niagr Mammadova and Tasja Lempertz. They were always there to hold me when I fell down with constant supply of sweets, coffee, encouragement, insightful discussions on our projects and immense support both in and outside of the lab. I would especially like to express my appreciation and would like to thank Dr. med. Matthias Dottermusch, Yannis Schumann, Anotnia Gocke for contributing great ideas and support for all my projects, our fruitful discussions in our lab meetings helped me a lot and paved for my research further. Thanks to entire AG Neumann!

I could not have completed this project without the support of my PhD committee supervisors Prof. Dr.med. Stefan Rutkowski and Prof. Dr. rer. nat. Hartmut Schlüter. Our PhD committee meetings, discussions helped me develop my project as I had experts from diverse fields. I would like to express my gratitude to Prof. Dr. med. Ulrich Schüller, who has also contributed immensely to our project not just by helping us gather samples for these rare tumours, but also by giving very constructive feedback to every presentation and also helping us through the massive revision process for our paper along with Prof. Rutkowski. I would like to thank Dr. Melanie Schoof and Dr. Carlin Göbel for trusting me with their data and let me contribute to their project.

I am particularly thankful to Dr. rer. nat. Hannah Voss, without whom the Medulloblastoma project wouldn't be complete. Her hardwork on this project, her support and help to get me into it, helped us together to achieve great results. From her lab, I would also like to thank Sönke Harder and Dr. rer.nat Christoph Krisp, who helped me a lot in the proteomics lab, to get into the sample extractions and troubleshooting our failed experiments. I would also like to thank Hannah and Simon Schlumbohm, for already developing HarmonizR, a tool which helped us all a lot in our projects.

I would also like to thank the Neuropathology diagnostics department, especially, Tasja, Kristin Hartmann, Ulrike Rumpf, Karin Gehlken, Carolina Janko, Celina Soltwedel, Nicole Borgardt and Helena Gladkov for helping us cut the samples, stainings and generation of the methylation data. Without your technical support, it would be difficult to finish our projects in time. I would like to thank the entire team and colleagues at Neuropathology, Kinderkrebs centrum who have supported us in this project by giving your valuable inputs.

Lastly, I dedicate my thesis to my family without whom I would not have the strength to get through anything in life- my parents Ruta and Satish Godbole, my husband, Arjun, you have been the three firm pillars of my life, helping me keep my fourth pillar (my career) stable. I am always grateful to my brother, Chinmay, Meghna and my mother and father-in-law who have cheered me on always! All of you have given me lots of love and freedom to dream big, guided me on my path so that I can achieve all my dreams. Thank you everyone, I am eternally grateful, always.

9. Eidesstattliche Versicherung

Ich versichere ausdrücklich, dass ich die vorliegende Arbeit mit dem Titel (Titel der Dissertation) selbständig und ohne unerlaubte fremde Hilfe, insbesondere ohne entgeltliche Hilfe von Vermittlungs- und Beratungsdiensten, angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel (einschließlich „Chatbots“ / KI) benutzt habe. Alle wörtlichen oder sinngemäßen Entlehnungen aus anderen Arbeiten sind an den betreffenden Stellen als solche kenntlich gemacht und im entsprechenden Verzeichnis aufgeführt, das gilt insbesondere auch für alle Informationen aus Internetquellen.

Ich erkläre zudem, dass ich die an der Medizinischen Fakultät Hamburg geltende „Satzung zur Sicherung guter wissenschaftlicher Praxis und zur Vermeidung wissenschaftlichen Fehlverhaltens an der Universität Hamburg“ in der jeweils gültigen Fassung eingehalten habe.

Ferner versichere ich, dass ich die Dissertation bisher nicht einem Fachvertreter oder einer Fachvertreterin an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.

Ich erkläre mich einverstanden, dass meine Dissertation vom Dekanat der Medizinischen Fakultät mit einer gängigen Software zur Erkennung von Plagiaten überprüft werden kann.

Unterschrift: