scientific reports

Check for updates

OPEN

Transcriptomic and proteomic spatial profiling of pediatric and adult diffuse midline glioma *H3 K27*-Altered

Sudarshawn Damodharan[1,4], Jack M. Shireman[2,4], Elliot Xie[3], Emily Distler[2], Christina Kendziorski[3] & Mahua Dey[2]✉

Diffuse midline glioma, *H3 K27*-altered (DMG) are highly aggressive malignancies of the central nervous system (CNS) that primarily affect the pediatric population. Large scale spatial transcriptomic studies have implicated that tumor microenvironmental landscape plays an important role in determining the phenotypic differences in tumor presentation and clinical course, however, data connecting overall transcriptomic changes to the protein level is lacking. The NanoString GeoMx™ Digital Spatial Profiler platform was used to determine the spatial transcriptomic and proteomic landscape in a cohort of both pediatric and adult *H3 K27*-altered DMG biopsy samples. Three fluorescently labeled antibodies targeting immune cells (CD45), epithelial cells (PanCK), tumor cells (*H3 K27M*) and a nucleic acid stain (SYTO-13) were used to establish regions of interest (ROI) for genomic and proteomic analysis. We found genetic alterations within the tumor which can be delineated across patient age and spatial location. We show that the H3 K27M mutation itself has a profound impact on tumor cells transcriptomics and interestingly we found limited fidelity between overall transcriptome and proteome. Our data also validate a previously described genomic signature at the proteomic level and reveal a special shift in the signature based on the local TME composition.

Diffuse midline glioma *H3 K27*-altered is a fatal central nervous system (CNS) malignancy that predominantly affects children of age 5–7 years. It is classified as a World Health Organization (WHO) grade 4 neoplasm with no curative treatments. The current standard approach to treatment is radiation therapy which is only palliative. The incidence rate for these tumors is difficult to accurately quantify due to the recent change in molecular classification as well as the rarity of cases. The most current incidence data collected by the World Health Organization (WHO) reports an incidence rate of 0.54 cases per million person years in adults, and a rate of 2.23 cases per million person years in people ≤ 20 years of age[1,2]. The prognosis of this neoplasm is extremely poor, especially in children with a 2-year survival rate of $< 10\%$[1,2]. Due to the rarity of the tumors in adults and the many subclassifications of DMG's, true prognosis in adults is still debated; however, it's generally agreed that adults have a better, while still overall poor, clinical prognosis[3–5]. These neoplasms occur primarily in midline structures within the CNS, with the brainstem and thalamus being the most prevalent and equally fatal anatomical locations[6–9]. Due to the critical anatomical locations of DMGs, surgical resection or needle biopsies were traditionally not performed, limiting the amount of pathological and molecular data obtained to date. However, with the development of sophisticated surgical technology over the last decade, biopsy has become routine resulting in a better understanding of DMGs at a genomic and molecular level.

These advanced molecular studies have led to further classification of DMGs into DMG H3K27-altered based on specific histone mutations[10,11] the most significant being highly recurrent histone mutations (*H3F3A* or *HIST1H3B/C*) which are present across pediatric and adult cases[7,12,13]. Liu et al. provided the first look into the genomic landscape of DMG at a spatial level across varying age and anatomical location[14]. In doing so, oligodendroglial lineage of precursor tumor cells were identified to be the majority in all clinical and anatomical groups and mesenchymal precursor signatures were increased in older patients with DMG[14]. Our understanding of the molecular and genetic landscape of these malignancies continues to grow dramatically, but primarily at

[1]Department of Pediatrics, Section of Hematology, Oncology & Stem Cell Transplantation, University of Chicago, Chicago, IL, USA. [2]Department of Neurosurgery, School of Medicine & Public Health, University of Wisconsin, UW Carbone Cancer Center, 600 Highland Ave, Madison, WI 53792, USA. [3]Department of Biostatistics and Medical Informatics, School of Medicine and Public Health, University of Wisconsin, Madison, WI, USA. [4]These authors contributed equally: Sudarshawn Damodharan and Jack M. Shireman. ✉email: dey@neurosurgery.wisc.edu

the mRNA and transcriptomic level[10–12,14], downstream proteomic concordance of the transcriptomic landscape has yet to be evaluated.

Although the mRNA landscape provides a good overview of the overall genomic alteration of a malignancy, it is the overall functional proteins that are responsible for the pathological behavior of the malignancy. Translation of mRNA into protein is a complex process with varying levels of translational fidelity between mRNA and protein[15–17]. Correlations between the differential expression of specific mRNAs and their corresponding proteins have been assessed in varying human disease processes with varying levels of concordance[18–23]. The fidelity between transcriptomic and proteomic landscape has never been explored in DMG.

To address these gaps in the literature we initiated this study in which we utilized spatial multi-omics profiling to comprehensively evaluate the molecular landscape of a cohort of pediatric and adult DMG tumor samples.

Methods

Clinical sample processing

Tissue samples (n = 8) and clinical data was collected from previously archived formalin-fixed-paraffin-embedded tissue (FFPE) samples of both pediatric (n = 4) and adult (n = 4) patients with confirmed pathological diagnosis of DMG, *H3 K27*-altered. IRB approval was obtained to utilize the archived pediatric and adult pathological samples for analysis at the University of Wisconsin-Madison (2022-0164-CP001). All patients analyzed within the study provided informed consent and all analyses were carried out according to UWSMPH/UW-Madison IRB direction and supervision. All samples were obtained from biopsies done pre-treatment at the time of initial diagnosis. From each sample, three cores were obtained to have each patient sample represented as a triplicate to evaluate the broader tumor landscape and proteogenomic profiles. (Fig. 1A, B, Table 1, Sup Fig. 1). All analysis conducted on human samples was performed in accordance with relevant guidelines and regulations.

Spatial genomic profiling

Transcriptomic and proteomic profiling was done on the above samples. The anatomical locations included the thalamus (n = 5) and pons (n = 3). To evaluate the intra-tumoral landscape and heterogeneity, the samples were stained with antibodies against the marker, PanCK (which in the brain can stain vasculature as well as broad cytoskeletal intermediate filaments likely to be present within the TME[24]; the leukocyte marker, CD45; mutant-tumor specific marker against H3 K27M; and the nuclear stain SYTO13. Staining colors are indicated in Fig. 1B.

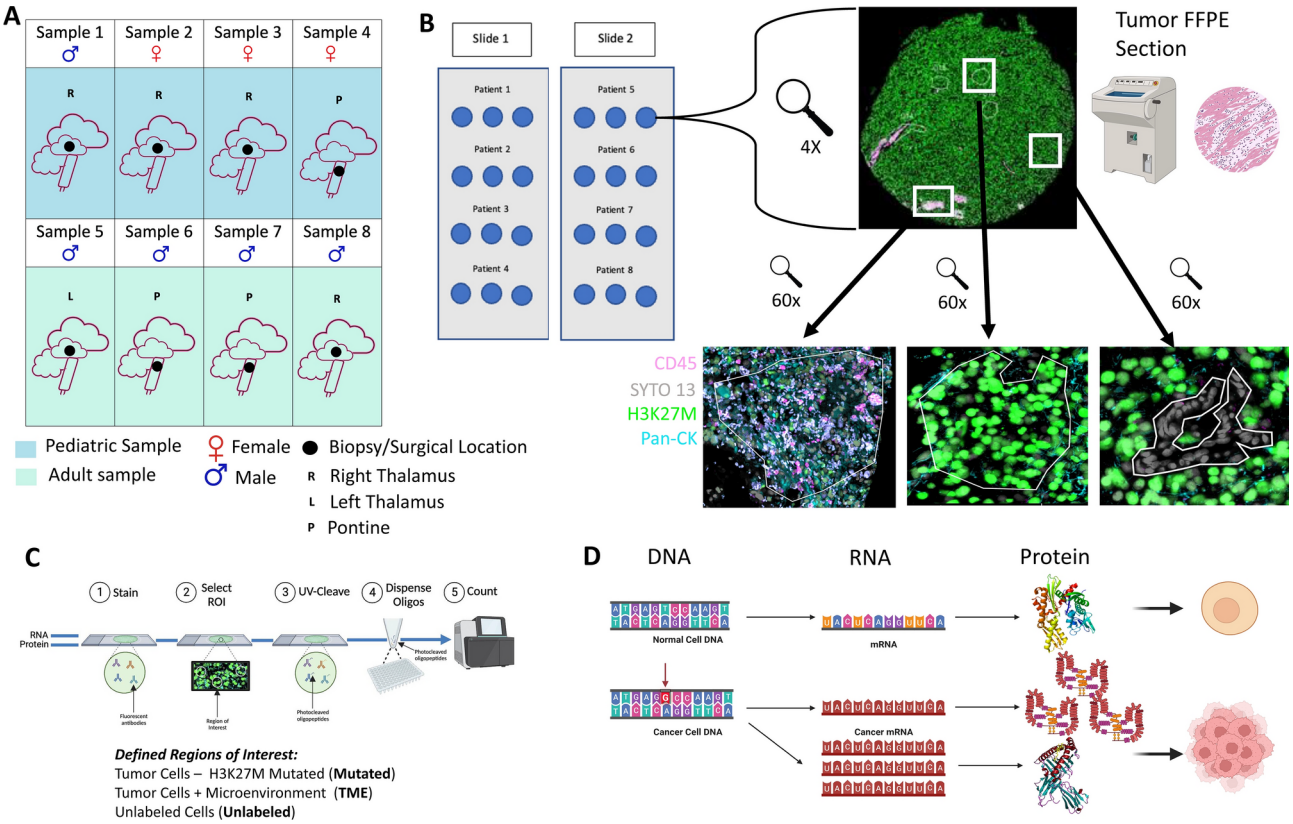

Fig. 1. Patient cohort and experimental workflow: (**A**) Graphic depicting the tumor location, patient age, and location of the tissue biopsy analyzed for the study. (**B**) Graphic depicting the slide layout of the Nanostring GeoMx slide used for sequencing. (**C**) Cartoon illustrating the experimental workflow of the project and how the ROIs were defined. (**D**) Cartoon depicting mismatch in RNA and Protein correlation levels in cancer cells compared to normal cells.

Patient	Age at diagnosis (years)	Sex	Race	Location	Treatment	OS (months)
1	6	Male	Caucasian	Right thalamus	XRT + TMZ, ONC201	14
2	11	Female	Hispanic	Right thalamus	XRT	9
3	14	Female	Black	Right thalamus	XRT	12
4	3	Female	Black	Pontine	XRT	2
5	43	Male	Caucasian	Pontine	XRT	6
6	46	Male	Hispanic	Left thalamus	None	1
7	48	Male	Caucasian	Pontine	XRT + TMZ, ONC201	26
8	25	Male	Black	Right thalamus	XRT + bevacizumab	24

Table 1. Clinical characteristics of pediatric and adult patients. *XRT* radiotherapy, *OS* overall survival, *TMZ* temozolomide.

We performed spatial transcriptomic and proteomic characterization of all samples via in situ sequencing utilizing the Nanostring digital spatial profiling (DSP) system and Illumina sequencing platforms (Fig. 1B, C). Across all samples, a total of 95 ROIs were marked for this experiment; divided between both populations of samples. Each ROI accounted for approximately 500uM of total space with the differing cells encompassed within this. The ROI categories utilized for this experiment were H3 K27M Ab predominant, H3 K27 ab + CD45 and areas predominantly with unlabeled cells (non-H3 K27M positive cells). Each ROI was assessed for the composition of mutant neoplastic cells, immune cells and interactions with one another in collaboration with a neuropathologist, pediatric hematologist, and Nanostring support staff. A layout of selection of ROI's in the context of a greater tissue section can be seen in Supplementary Fig. 1.

ROI collection and filtering

A total of 95 ROIs were initially collected for analysis (Fig. 1C) and subjected to default QC parameters as outlined by Nanostring. This QC process resulted in the dropping of 9 ROI's largely based on total nuclei count being below manufactures standard detection rate (gene detection rate larger than 0.05 and nuclei count larger than 100). For deconvolution analysis which requires a more in-depth analysis with more marker genes present and can be disproportionately affected by noise we imposed a stricter cutoff (gene detection rate larger than 0.15) this resulted in 49 total ROI's being analyzed for the deconvolution. Because of the stricter cutoff imposed we further validated the deconvolution analysis by applying two separate algorithms (Redeconve 1.1.2[25]) & CIBERSORTX[26]) and ensuring the results matched with results obtained from SpatialDecon (1.8.0).

Annotation of ROIs

Each ROI underwent several runs of manual annotation. This involved an in-depth examination of the corresponding-colored images and the screening of highly expressed protein markers characteristic of each region.

Correlation analysis between protein and RNA

For correlation analysis, proteins and RNA were matched according to their respective GeneIDs. A Pearson correlation coefficient was calculated for each pair to assess the degree of linear correlation between the protein and RNA expression levels in the ROIs. Hypothesis testing to identify differences between average correlations in adult vs. pediatrics (and in mutated vs. unlabeled vs. TME) was conducted using t-tests on Fisher's Z-transformed correlation values. Z transformation scores and the Pearson correlation coefficient averaged across all ROI's is included within the visualization (Fig. 1D).

Deconvolution analysis

Deconvolution of the spatial data was carried out using the SpatialDecon algorithm[27]. This method utilizes log-normal regression and integrates background modeling, thereby enhancing performance beyond that of traditional least-squares methods. SpatialDecon is adept at quantifying cell populations within ROIs and provides precise estimates of cell abundance. The reference matrix, crucial for this analysis, was sourced from a study focused on DMG single-cell RNA sequencing data[14]. The processing of this data was in strict accordance with the protocols and methodologies detailed in Liu et al.[14]. For each ROI, SpatialDecon (1.8.0) generated beta values, representing the cell abundance scores of the rolled-up/major cell types. These beta values were organized into a matrix with major cell types as rows and ROIs as columns. To assess differences in cell type abundances between groups, we applied the Wilcoxon rank-sum test to these cell abundance scores. Comparisons were made across different age groups and spatial location. The Wilcoxon rank-sum test was chosen for its robustness and, importantly, its lack of assumption regarding distribution of the data. All deconvolution analyses were repeated with two separate algorithms to ensure fidelity as described above.

Gene regulatory network reconstruction and cell-state identification

To comprehensively characterize the transcriptional state of each ROI, we utilized SCENIC (0.12.1)[28], a computational approach for concurrent gene regulatory network reconstruction and cell-state identification. The SCENIC methodology involves three steps: initially, GRNboost2 identifies gene sets co-expressed with transcription factors. Subsequently, these gene sets are refined through cis-regulatory motif analysis using

RcisTarget, leading to the formation of regulons by retaining modules with significant motif enrichment. The activity of each regulon in individual ROIs is quantified using the AUCell algorithm, yielding an AUC score that reflects the subnetwork activity. This score is then utilized to assess the characteristics of each ROI.

Over-representation analysis

Over-representation analysis (ORA) was done using webgestalt[29–31] with default parameters on functional database "gene ontology" or biological pathways "KEGG". Enrichment statistics are calculated using a hypergeometric test to evaluate the significance of enrichment and calculate a p-value, which is then adjusted for multiple comparisons (Bonferroni), as well as a false discovery rate (FDR)[29–31]. For this study FDR is reported for all ORA comparisons made with significance being considered if FDR < 0.05 or trending significance at FDR < 0.01.

Results

Patient cohort and clinical characteristics

Patient characteristics along with treatment details are listed in Table 1 and Fig. 1A along with experimental details and schematics (Fig. 1B–D). All patients in the study had mutations within the *H3F3A* gene which encodes histone H3.3 and typically correlates with a more severe phenotype in this tumor[5]. For the pediatric patients, the median age at diagnosis was 8.5 (range: 3–14 years); median overall survival (OS) was 9.2 months (range: 2–14 months). For the adult patients, the median age of diagnosis was 40.5 (range 25–48); median overall survival (OS) was 14.2 months (range: 1–26 months).

SCENIC profiling reveals spatially differential regulon enrichment

SCENIC was used to reconstruct gene regulatory networks and estimate cellular states for each ROI. The combined adult and pediatric samples were classified spatially by grouped ROIs into 3 categories: H3K27M mutated tumor cells (Mutated), Non H3K27M mutated cells (Unlabeled/Grey), and tumor cells + surrounding microenvironment (TME). These ROIs were confirmed using IHC staining on our tissue samples, allowing for anchoring and targeting of the resulting transcriptomic and proteomic analysis done by the GEOMX system. Mutated ROIs contained only cells expressing H3K27M marks (green), while TME ROIs contained H3K27M tumor cells (green), immune cells (pink), and cytoskeletal/cytokeratin cells (blue) (Fig. 2A). Because of the nature of these biopsies (fine needle aspiration targeted to tumor bulk) and reported low surrounding cell invasion within these tumors[14] ROIs containing cells with no staining (Unlabeled) were hypothesized to contain mostly DMG tumor cells that do not carry the H2K27M mutation which is seen in the tumor bulk according to recent reports[8,32]. To investigate this hypothesis, the top 25 significantly enriched genes/regulons identified using SCENIC were subject to over representation analysis (ORA) and compared to an equal number of enriched genes derived from a dataset of normal thalamic and pontine tissue derived from hits across both the Allen Human Brain Cell Atlas and the Human Protein atlas, which was filtered for midbrain location specific expression. This comparison demonstrated highly enriched signaling for cancer associated genes and tumor suppressors with enriched terms such as Transcriptional Regulation in Cancer and Pathways in Cancer (FDR < 0.05). Conversely, the same ORA analysis on the normal thalamic and pontine tissue failed to show enrichment for any cancer related pathways instead demonstrating insignificant enrichment for cell developmental processes and neuronal signatures with enrichment terms for regulation of synaptic signaling and mononuclear cell proliferation (FDR > 0.05) (Sup Fig. 2A, B). These data highlight the altered transcriptional profile of these unlabeled cells compared to typical human midbrain cells suggesting they could be tumor cells lacking the H2K27M mutation or resident brain cells with altered transcriptional profiles due to their proximity to DMG tumor cells.

Across our spatial classifications, SCENIC identified significantly different regulons driven by ARID3C and MSX1 in the TME, TBP and ELK1 in Mutated, and NFATC2 and Stat6 in Unlabeled ROIs (Fig. 2B,C). These regulons were enriched for cytokine and interleukin signaling in the TME ROIs (FDR < 0.05), limited significance cancer induced senescence and TP53 gene regulation in the Mutated ROIs (FDR < 0.10), and TNF and WNT pathway signaling in Unlabeled ROIs (FDR < 0.05) (Fig. 2C). These results highlight that spatially distinct regions within the tumor rely on differential transcriptional programs especially with regards to invading immune cells and surrounding vasculature. Furthermore, the H3K27 mutation itself seems to impact the transcriptional landscape significantly with unlabeled cells displaying more developmental like signaling pathways as opposed to mutated tumor cells displaying more canonical cancer signaling networks.

DIPG genomic signature varies across age and spatial location

Molecular and genomic classification of DMG remains in its infancy due to the extremely limited number of available patient samples compared to other glioma types. In a recently published seminal study[14] the authors describe a ubiquitous oligodendrocyte precursor cell (OPC) like cellular state across both pediatric and adult DMG. Other cell states identified in the tumor include Mesenchymal (Mes) like and Astrocytic (AC) like, as well as cycling cells. Interestingly, neuronal signatures are not seen within DMGs from either population. We validated this in our dataset as well where we only saw enrichment for neuronal signatures during ORA analysis of normal midbrain controls (Sup Fig. 2A, B). Utilizing the genomic signatures derived from their published dataset[14], we compared our samples across age and spatial location and were able to corroborate the baseline signatures seen in adults and pediatrics. In our data examining Mutated ROIs we found pediatric tumors enriching for AC like signatures (Wilcoxon rank test, pediatric vs adults; p-adj < 0.05) as well as a trend towards OPC like (specifically OPC like 3) signatures (Fig. 3A). Interestingly, when we examined TME ROIs we saw a shift in the observed signatures with adult tumors now showing increased OPC like signature enrichment (OPC like 1) (Wilcoxon rank test; pediatric vs adults, p-adj < 0.01) while pediatric samples displayed a higher Mes like (Wilcoxon rank test; pediatric vs adults, p-adj < 0.01) and Microglial signature (Wilcoxon rank test; pediatric vs

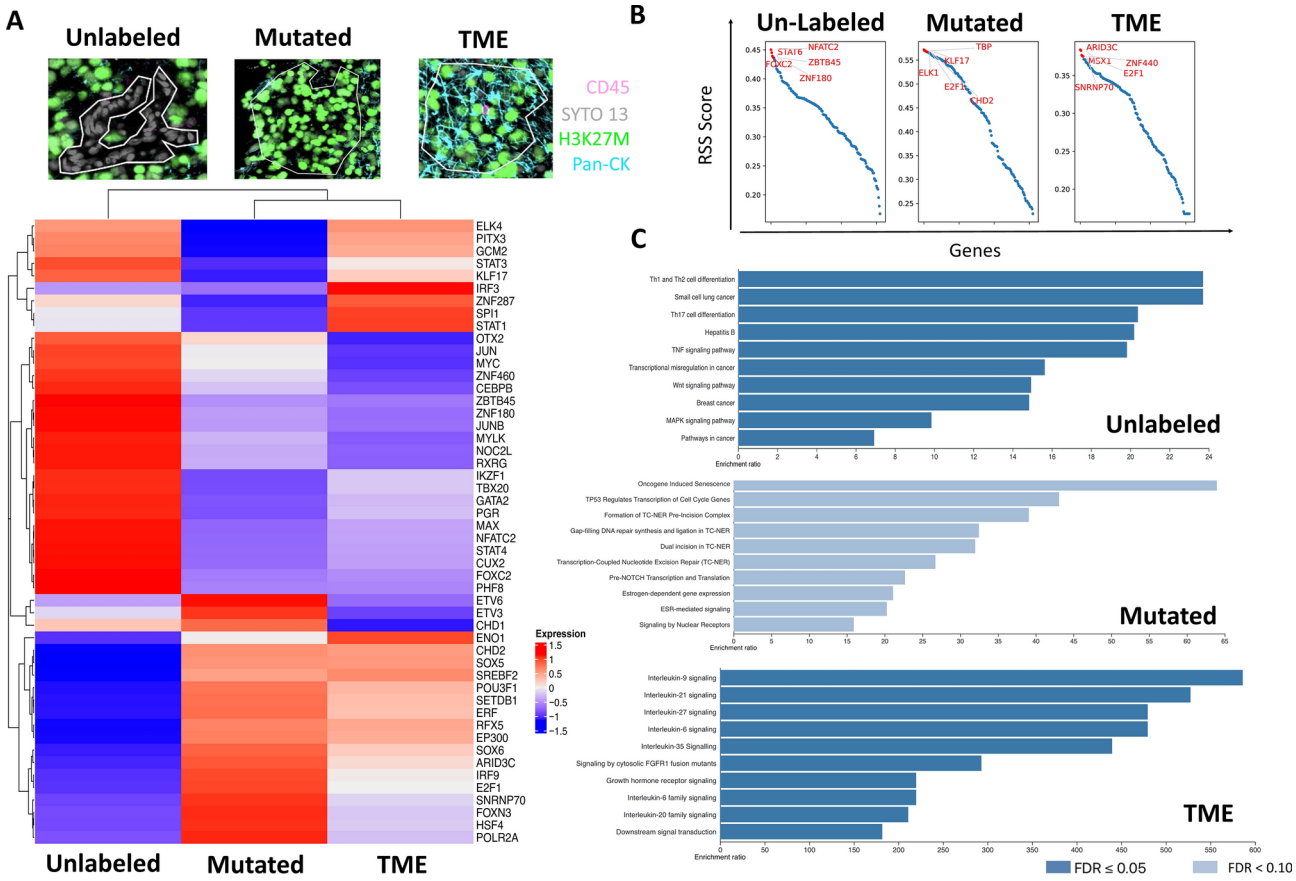

Fig. 2. SCENIC profiling reveals spatially differential regulon enrichment: (**A**) SCENIC regulon enrichment across our selected ROIs visualized using a heatmap displaying the set of transcription factors and their AUCell score across ROI's. (**B**) Regulon Specificity Score (RSS) across the analyzed ROIs with the top genes highlighted in red. (**C**) Over representation analysis of the genes contained within the SCENIC enrichment for each ROI analyzed.

adults, p-adj < 0.05) (Fig. 3B). To further validate these regional gene signature shifts seen with the SpatialDecon algorithm we applied two different deconvolution algorithms CIBERSORTX and ReDeconve and saw similar trends with pediatric Mes like enrichment and adult OPC like enrichment (Sup Fig. 3A, B).

These genomic signatures were originally derived from RNA data. However, our dataset includes a subset of protein targets (147 probes total) allowing for partial validation of the expression pattern for one of the component genes of the OPC like signature, EGFR (Sup Fig. 4A,B) at the protein level. These results indicate that even though our study has a relatively limited sample size, it still captures and validates a previously published genomic classification across patient age using both RNA and protein expression data. Our results also highlight a spatially driven shift in these signatures when more complex cell compositions are examined. This could indicate a TME driven shift in the tumor cells' biology or simply be reflective of the more complex TME present along outer edges of tumor compared to the tumor cell dense cores.

Correlation between RNA and Protein varies both spatially and across patient age

A large portion of the current transcriptomics work assumes that increased mRNA abundance results in increased protein abundance. While this is often the case, it is well known that the amount of translated protein can vary widely from the amount of detected mRNA[33–35]. Leveraging the unique Nanostring GeoMx technology we detected both RNA and protein probes across our entire range of samples. This allowed for comparison of matched RNA to protein pairs to assess transcription to translation fidelity and characterize the extent to which it varies across our samples. We first visualized the correlation between mRNA and protein abundance (ranging from -1 being 100% negative correlation to 1 being 100% positive correlation) across the adult and pediatric tumor samples. We found similar rates of correlation on average in adults (0.19 + /− 0.27 SD) and pediatrics (0.33 + /- 0.25 SD) (Sup Fig. 5A,B) but differences were observed among the top 10 gene pairs that are most (and least) correlated (Fig. 4A). Spatially, similar correlation results were observed with no statistical difference in correlation rates (mutated vs unlabeled vs TME: 0.28 + /− 0.31SD, 0.24 + /− 0.26SD, 0.29 + /− 0.27 SD; p > 0.05) (Sup Fig. 4B), however, differences among the specific gene pairs that were correlated vs. anticorrelated were seen (Fig. 4B). Across adult and pediatric samples positive correlation values between RNA and protein were similar for some genes (OLIG2, Vimentin, MAP2) but differential correlation was also seen in adults and pediatric

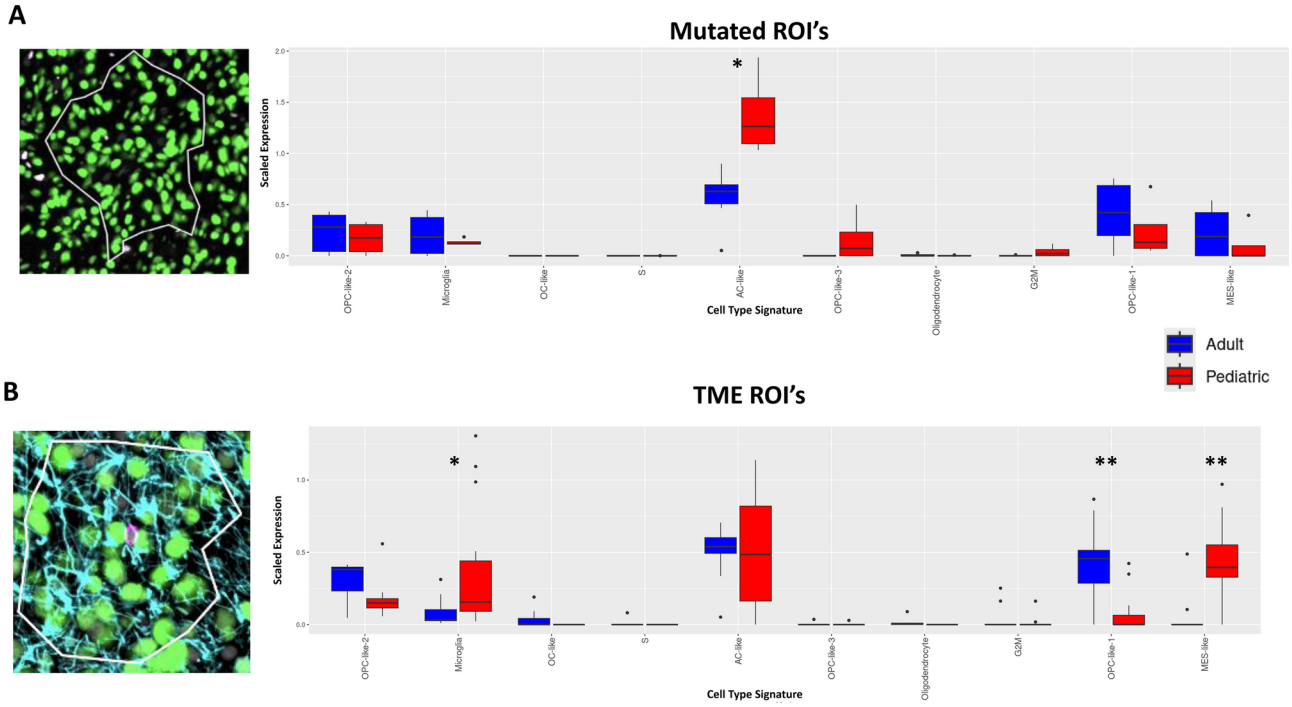

Fig. 3. DIPG genomic signature varies across age and spatial location: (**A**) Visual representation of the Mutated ROI used for deconvolution analysis as well as a histogram of the results of analysis comparing the adult (red) and pediatric (blue) populations within the Mutated ROI. (**B**) Visual representation of the TME ROI used for deconvolution analysis as well as a histogram of the results of analysis comparing the adult and pediatric populations within the TME ROI. *p-adj < 0.05, **p-adj < 0.01; Wilcoxon ranked test.

samples. Spatially, similar proteins were observed positively correlated across regions that were also observed across ages (OLIG2, MBP, MAP2). In terms of negative correlation CTLA4 was a top negative correlation across both age and spatial ROIs, an interesting finding due to the relevance of CTLA4 treatments across many other cancer types[36–38]. Overall, these results show that not only are different transcriptional programs being activated in tumors across age and regions, but these transcriptional programs do not always faithfully lead to the corresponding proteins being translated (Fig. 4C). If translational fidelity is not considered, predictions biased towards RNA content, but not true protein translation, may result.

Common genomic therapeutic targets show differing RNA to protein correlation

Much of the upcoming targeted and immunotherapeutic interventions that are being studied for DMG have come about based upon previously established transcriptomic information. However, downstream proteomic data of clinically relevant therapeutic targets have not been readily studied. To examine this in our dataset we found RNA–Protein pairs that represented promising or already tried therapeutic targets in this group of malignancy and quantified their RNA and protein levels across all samples. Expression of receptor tyrosine kinase (RTK) has been shown to be prevalent on these malignancies with on-going and previously completed studies examining EGFR as a relevant target both as a small molecule inhibitor and in combination with immunotherapeutic agents[39–41]. We observed high concordance for EGFR between RNA and protein, with protein levels higher than would have been predicted by mRNA abundance (Fig. 5). Checkpoint blockade has also been tested both as the monotherapy and combinatorial immunotherapy for this group of malignancy[42–44], thus we examined PD1, PDL-1 and CTLA-4 expression correlation in our dataset. We observed low concordance between RNA and protein across all the above markers with CTLA-4 especially enriched for negative correlation when examined in our TME ROIs (see Fig. 4E). For these molecules, the amount of RNA detected was more than the actual amount of protein detected which could lead to skewed optimistic prediction of outcome for therapeutic targeting in clinical settings. B7-H3 and BRAF have been well characterized across tumors and within our dataset we found close concordance for RNA and protein levels for B7-H3 but not BRAF. To further explore more clinically relevant targets with our data we filtered GSEA enrichment results using Enrichr[45] for genes involved in pathways that resulted in molecular secretions. The results demonstrated the top enriched pathways belonging to IL6 and ATF2 signaling events with glucocorticoid receptor signaling and calcium signaling also being present, likely related to the very neuronally active locations of these tumors (Sup Fig. 6). Together, these results highlight the crucial point that extensive targeting solely based on transcriptomics may result in false positives as well as false negatives and integrating proteomic data may strengthen early-stage clinical drug screening.

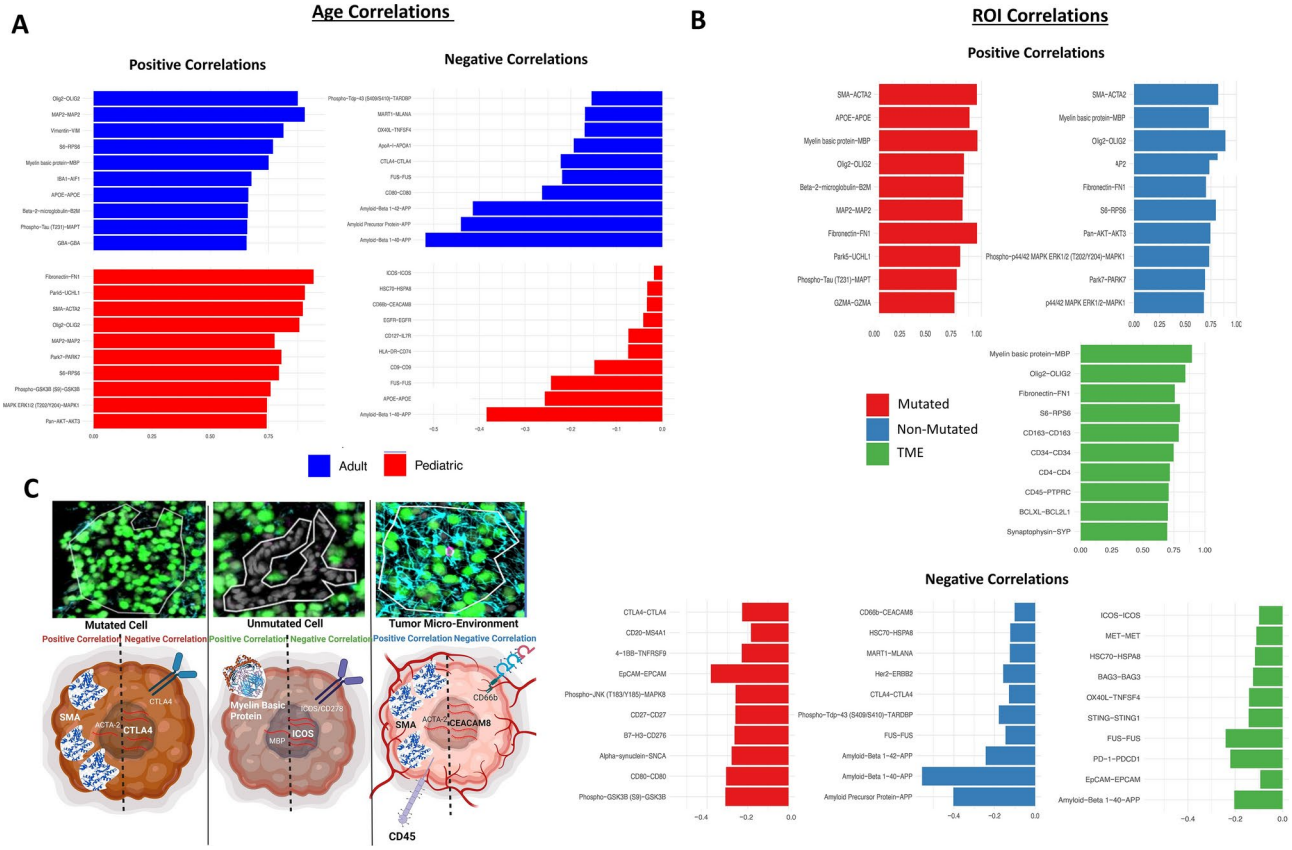

Fig. 4. Correlation between RNA and Protein varies both spatially and across patient age: (**A**) Histogram representations of the top 10 positively or negatively correlated RNA/protein pairs across adults (blue) or pediatric samples (red). (**B**) Histogram representations of the top 10 positively or negatively correlated RNA/protein pairs across all 3 ROIs contained in our sample population. (**C**) Cartoon representation of translational fidelity using the top 3 positive or negatively correlated RNA/protein pairs across ROI's.

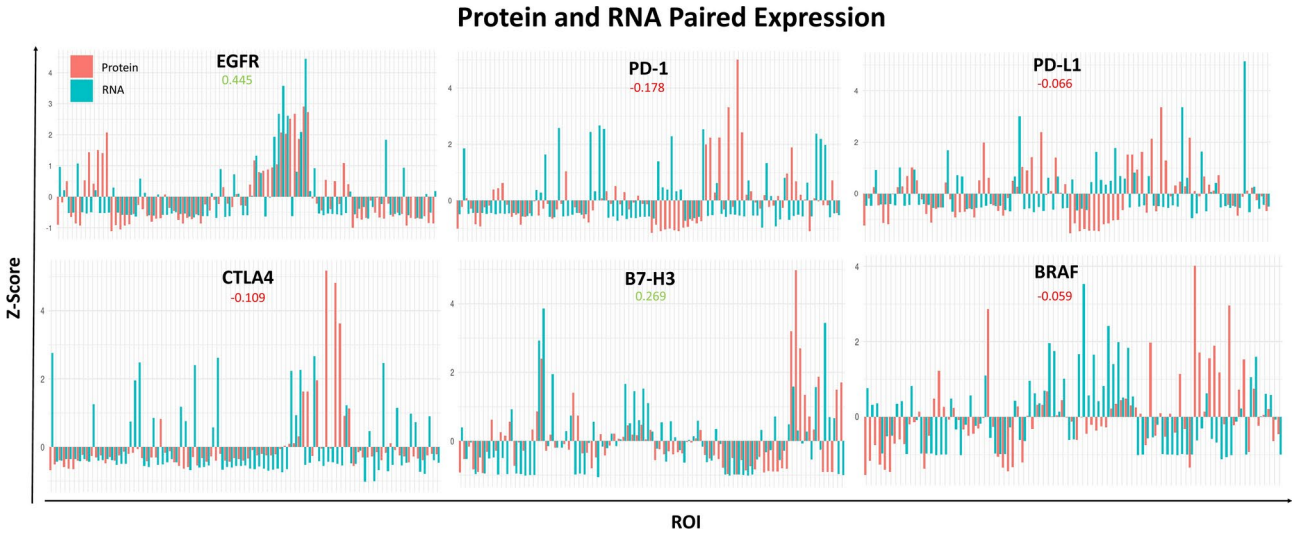

Fig. 5. Common genomic therapeutic targets show differing RNA to protein correlation: Histogram representation of expression of RNA (blue) and protein (red) levels across all samples for common genomic targets of DMG. The y-axis displays z-transformed expression levels, allowing for standardized comparison across different targets. The x-axis represents individual ROIs. For each pair, Pearson correlation coefficients between RNA and protein expression across all ROIs are reported. These correlations are color-coded: green for positive correlations and red for negative correlations.

Discussion

DMG continues to be a devastating disease with limited therapeutic options, however, in the last decade neurosurgical and neuro-oncological techniques have advanced considerably allowing for biopsy and sample collection from these tumors. This in turn has led to several novel clinical trials currently underway based on the genomic targets discovered from comprehensive transcriptomic studies. Our study aims to push the field forward by taking the next step in genomic screening by examining spatial aspects of the tumor architecture as well as proteomic profiles. Spatially, we found that tumor ROIs that contain portions of TME display distinct gene expression signatures compared to regions with dense tumor cell populations. Using proteomics, we demonstrate that transcription to translation fidelity is not always maintained in both adult and pediatric DMG. Finally, we interrogated common genetic targets proposed both for these tumors as well as broadly used in other gliomas and reveal the shortfalls of predicting therapeutic targeting efficacy based on RNA data alone. Continued research in this field will be crucial in advancing both the molecular and clinical understanding of these rare and lethal tumors.

A recent publication by Liu and colleagues demonstrated an increased tendency for pediatric DMG to be dominated by an OPC like cell type while adult DMG enriched for a mesenchymal cell phenotype[14]. In our cohort we saw similar patterns across adult and pediatric tumors, however when examining the spatial landscape, we observed a change in the signature in ROIs composed of both tumor and TME cells. Also of note, in both our cohorts of data there is little to no neuronal signature expressed within any sample set. This is interesting because these tumors arise in very neuronally dense and eloquent tissue but seem to not be as intermixed with their surroundings like glioblastomas (GBM), the most common primary CNS malignancy in adults, which are known to integrate into neuronal synapses and proliferate more in response to neuronal firing[46–49]. This may indicate that DMGs interacts more with other brain resident cells such as microglia, astrocytes, or oligodendrocytes as opposed to neurons. It's also worth noting that there was little to no outside immune cell infiltration into the tumor bulk itself which is consistent with reports of DMG being an immune desert[50,51]. Finally, due to the eloquent nature of the brain regions these tumors colonize, samples obtained anti-mortem are often limited to fine needle aspiration biopsies which likely enrich for tumor bulk and not tumor surroundings.

Our study is the first to include spatial proteomic profiling and we demonstrate that while translational fidelity was measured at similar levels across adults and pediatric samples, as well as across the ROIs, there were subtle differences in the types of genes that were either faithfully or not faithfully translated from mRNA. These results have significant implications for cancer therapeutics in general which to date has relied mostly on screening mRNA due to difficulties in obtaining proteomic data. Our examination of common genomic targets discovered using mRNA alone shows that although this method can result in true proteomic hits, there is also room for error including both false positive and false negative hits being prioritized (such as EGFR false negative vs. PD-1/PDL-1 false positive). Although there are some studies suggesting mRNA, specifically differentially expressed mRNA, may be a faithful readout of true protein synthesis[52], this has yet to be widely examined in the context of brain malignancies, especially in the pediatric population. Kosti and colleagues describe RNA to protein correlation profiling conducted across a wide variety of normal and cancer tissues by integrating data between the Human Proteome Map project and RNA-Seq measurements from the Genotype-Tissue Expression project[53]. In this study the average correlation ranged from 0.36 to 0.5, slightly higher than our reported correlations. Although this study correlated across many more tissue sites and samples than our study, they did not examine any brain tissue correlation either in the normal or malignant setting. Furthermore, these correlations come from an across sample design (integrating two separate datasets) whereas our correlations come from the same samples. Saratsis et al. do provide a study of RNA and protein in DIPG data, however, they report only concordance of expression levels (54%) rather than actual RNA to protein correlation[54]. These samples are all also obtained postmortem which depending on the interval between death and analysis can introduce significant noise into the proteomics data[55]. Overall, we believe further study is required, specifically on anti-mortem proteomic analysis of malignant and normal brain tissue, to truly understand this effect.

One of the limitations of our study is that the number of protein probes available is far less than the number of RNA probes, and so a genome-wide comparison of RNA to protein was not possible. It's also necessary because of GeoMX technical limitations to compare protein probes and RNA probs across slides rather than only from the same slide possibly introducing noise. We also did not have a large enough number of samples to include statistically significant sex comparisons. In addition, due to the technical limitations of the Nanostring platform regarding the number of antibodies that can be used for defining ROIs, we had a subset of cells that were unstained. Since all our samples were targeted needle biopsies aimed at obtaining the most pathological tumor tissue, we hypothesized that these were most likely tumor cells that lacked the H3K27m mutation. In validating this using ORA analysis, we discovered a unique genetic program within these cells compared to H3K27m tumor cells involving some immune signatures as well as cancer and developmental gene enrichment. Although we only visualize a limited number of immune cells in our samples, there were samples with strong enrichment for immune related signaling, especially when TME ROIs were considered. This may be a result of tumor cells hijacking immune signaling methods (seen in other gliomas[56–58]) or could result from immune cells sitting mostly on the peripheral edges of the tumor (which would not be captured during biopsy) inducing inflammation. An in depth understanding of the immune response and its resulting inflammation in this context may prove critical as we try to advance therapies, particularly CAR-T cell based therapies which are currently undergoing clinical trials[59–61].

Although DMG remains a 100% lethal tumor with incredibly limited survival, novel research is starting to illuminate its genetic background and expose possible therapeutic vulnerabilities. Further studies across multiple institutions to increase sample size will be the key in validating and advancing effective therapies and ultimately making a difference in patient survival.

Data availability

All sequence data collected and analyzed for the manuscript will be deposited into a GEO repository to be released publicly upon acceptance and publication of the article. Any valid requests for additional data will be honored and should be directed to the corresponding author.

Received: 20 March 2024; Accepted: 16 September 2024
Published online: 30 September 2024

References

1. Louis, D. N. *et al.* The 2021 WHO classification of tumors of the central nervous system: A summary. *Neuro-Oncol.* **23**, 1231–1251 (2021).
2. *Central Nervous System Tumours*. vol. 5 (International Agency for Research on Cancer, 2021).
3. Dono, A., Takayasu, T., Ballester, L. Y. & Esquenazi, Y. Adult diffuse midline gliomas: Clinical, radiological, and genetic characteristics. *J. Clin. Neurosci.* **82**, 1–8 (2020).
4. Meyronet, D. *et al.* Characteristics of H3 K27M-mutant gliomas in adults. *Neuro-Oncol.* **19**, 1127–1134 (2017).
5. Lu, V. *et al.* Impact of the H3K27M mutation on survival in pediatric high-grade glioma: A systematic review and meta-analysis. *J. Neurosurg. Pediatr.* **23**, 308–316 (2019).
6. Mackay, A. *et al.* Integrated molecular meta-analysis of 1000 pediatric high-grade and diffuse intrinsic pontine glioma. *Cancer Cell* **32**, 520-537.e5 (2017).
7. Khuong-Quang, D.-A. *et al.* K27M mutation in histone H3.3 defines clinically and biologically distinct subgroups of pediatric diffuse intrinsic pontine gliomas. *Acta Neuropathol.* **124**, 439–447 (2012).
8. Jones, C. & Baker, S. J. Unique genetic and epigenetic mechanisms driving paediatric diffuse high-grade glioma. *Nat. Rev. Cancer* **14**, 651–661 (2014).
9. Wu, G. *et al.* Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. *Nat. Genet.* **44**, 251–253 (2012).
10. Huang, T. *et al.* Detection of histone H3 K27M mutation and post-translational modifications in pediatric diffuse midline glioma via tissue immunohistochemistry informs diagnosis and clinical outcomes. *Oncotarget* **9**, 37112–37124 (2018).
11 Wang, J. *et al.* Epigenomic landscape and 3D genome structure in pediatric high-grade glioma. *Sci. Adv.* **7**, eabg4126 (2021).
12 Schwartzentruber, J. *et al.* Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* **482**, 226–231 (2012).
13. Tune, B. G. *et al.* From pediatric to adult brain cancer: Exploring histone H3 mutations in Australian brain cancer patients. *Biomedicines* **11**, 2907 (2023).
14. Liu, I. *et al.* The landscape of tumor cell states and spatial organization in H3–K27M mutant diffuse midline glioma across age and location. *Nat. Genet.* **54**, 1881–1894 (2022).
15. Kapur, M. & Ackerman, S. L. mRNA translation gone awry: Translation fidelity and neurological disease. *Trends Genet.* **34**, 218–231 (2018).
16. Chung, C. *et al.* The fidelity of transcription in human cells. *Proc. Natl. Acad. Sci.* **120**, e2210038120 (2023).
17. Gamba, P. & Zenkin, N. Transcription fidelity and its roles in the cell. *Curr. Opin. Microbiol.* **42**, 13–18 (2018).
18. Asplund, A. *et al.* Expression profiling of microdissected cell populations selected from basal cells in normal epidermis and basal cell carcinoma. *Br. J. Dermatol.* **158**, 527–538 (2008).
19. Dhanasekaran, S. M. *et al.* Delineation of prognostic biomarkers in prostate cancer. *Nature* **412**, 822–826 (2001).
20. Greenbaum, D., Colangelo, C., Williams, K. & Gerstein, M. Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biol.* **4**, 117 (2003).
21. Fu, N. *et al.* Comparison of protein and mRNA expression evolution in humans and chimpanzees. *PLoS One* **2**, e216 (2007).
22. Shankavaram, U. T. *et al.* Transcript and protein expression profiles of the NCI-60 cancer cell panel: An integromic microarray study. *Mol. Cancer Ther.* **6**, 820–832 (2007).
23. Gry, M. *et al.* Correlations between RNA and protein expression profiles in 23 human cell lines. *BMC Genom.* **10**, 365 (2009).
24. Pathology Outlines - Cytokeratins (CK) - general. https://www.pathologyoutlines.com/topic/stainsckgeneral.html.
25. Zhou, Z., Zhong, Y., Zhang, Z. & Ren, X. Spatial transcriptomics deconvolution at single-cell resolution using Redeconve. *Nat. Commun.* **14**, 7930 (2023).
26 Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **12**, 453 (2015).
27. Danaher, P. *et al.* Advances in mixed cell deconvolution enable quantification of cell types in spatial transcriptomic data. *Nat. Commun.* **13**, 385 (2022).
28. Aibar, S. *et al.* SCENIC: Single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).
29. Wang, J., Vasaikar, S., Shi, Z., Greer, M. & Zhang, B. WebGestalt 2017: A more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucl. Acids Res.* **45**, W130–W137 (2017).
30. Liao, Y., Wang, J., Jaehnig, E. J., Shi, Z. & Zhang, B. WebGestalt 2019: Gene set analysis toolkit with revamped UIs and APIs. *Nucl. Acids Res.* **47**, W199–W205 (2019).
31. Zhang, B., Kirov, S. & Snoddy, J. WebGestalt: An integrated system for exploring gene sets in various biological contexts. *Nucl. Acids Res.* **33**, W741–W748 (2005).
32. Larson, J. D. *et al.* Histone H3.3 K27M accelerates spontaneous brainstem glioma and drives restricted changes in bivalent gene expression. *Cancer Cell* **35**, 140-155.e7 (2019).
33. de Sousa Abreu, R. *et al.* Global signatures of protein and mRNA expression levels. *Mol. Biosyst.* **5**, 1512–1526 (2009).
34. Sonneveld, S., Verhagen, B. M. P. & Tanenbaum, M. E. Heterogeneity in mRNA Translation. *Trends Cell Biol.* **30**, 606–618 (2020).
35. Edfors, F. *et al.* Gene-specific correlation of RNA and protein levels in human cells and tissues. *Mol. Syst. Biol.* **12**, 883 (2016).
36. He, X. & Xu, C. Immune checkpoint signaling and cancer immunotherapy. *Cell Res.* **30**, 660–669 (2020).
37. Sampson, J. H., Gunn, M. D., Fecci, P. E. & Ashley, D. M. Brain immunology and immunotherapy in brain tumours. *Nat. Rev. Cancer* **20**, 12–25 (2020).
38. Havel, J. J., Chowell, D. & Chan, T. A. The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. *Nat. Rev. Cancer* **19**, 133–150 (2019).
39. Dong, X. *et al.* Anti-VEGF therapy improves EGFR-vIII-CAR-T cell delivery and efficacy in syngeneic glioblastoma models in mice. *J. Immunother. Cancer* **11**, e005583 (2023).
40. Wang, X. *et al.* EGFR is a master switch between immunosuppressive and immunoactive tumor microenvironment in inflammatory breast cancer. *Sci. Adv.* **8**, eabn7983 (2022).
41. Weller, M. *et al.* Rindopepimut with temozolomide for patients with newly diagnosed, EGFRvIII-expressing glioblastoma (ACT IV): A randomised, double-blind, international phase 3 trial. *Lancet Oncol.* **18**, 1373–1385 (2017).
42. Cacciotti, C. *et al.* Immune checkpoint inhibition for pediatric patients with recurrent/refractory CNS tumors: A single institution experience. *J. Neuro-Oncol.* **149**, 113–122 (2020).
43. Ausejo-Mauleon, I. *et al.* TIM-3 blockade in diffuse intrinsic pontine glioma models promotes tumor regression and antitumor immune memory. *Cancer Cell* **41**, 1911-1926.e8 (2023).

44. Kline, C. *et al.* Reirradiation and PD-1 inhibition with nivolumab for the treatment of recurrent diffuse intrinsic pontine glioma: A single-institution experience. *J. Neuro-Oncol.* **140**, 629–638 (2018).
45. Chen, E. Y. *et al.* Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinform.* **14**, 128 (2013).
46. Venkatesh, H. S. *et al.* Electrical and synaptic integration of glioma into neural circuits. *Nature* **573**, 1–7 (2019).
47. Venkatesh, H. S. *et al.* Neuronal activity promotes glioma growth through neuroligin-3 secretion. *Cell* **161**, 803–816 (2015).
48. Monje, M. *et al.* Roadmap for the emerging field of cancer neuroscience. *Cell* **181**, 219–222 (2020).
49. Venkatesh, H. S. *et al.* Targeting neuronal activity-regulated neuroligin-3 dependency in high-grade glioma. *Nature* **549**, 533–537 (2017).
50. Lin, G. L. *et al.* Non-inflammatory tumor microenvironment of diffuse intrinsic pontine glioma. *Acta Neuropathol. Commun.* **6**, 51 (2018).
51. Lieberman, N. A. P. *et al.* Characterization of the immune microenvironment of diffuse intrinsic pontine glioma: Implications for development of immunotherapy. *Neuro-Oncol.* **21**, 83–94 (2019).
52. Koussounadis, A., Langdon, S. P., Um, I. H., Harrison, D. J. & Smith, V. A. Relationship between differentially expressed mRNA and mRNA-protein correlations in a xenograft model system. *Sci. Rep.* **5**, 10775 (2015).
53. Kosti, I., Jain, N., Aran, D., Butte, A. J. & Sirota, M. Cross-tissue analysis of gene and protein expression in normal and cancer tissues. *Sci. Rep.* **6**, 24799 (2016).
54. Saratsis, A. M. *et al.* Comparative multidimensional molecular analyses of pediatric diffuse intrinsic pontine glioma reveals distinct molecular subtypes. *Acta Neuropathol.* **127**, 881–895 (2014).
55. Machaalani, R., Gozal, E., Berger, F., Waters, K. A. & Dematteis, M. Effects of post-mortem intervals on regional brain protein profiles in rats using SELDI-TOF-MS analysis. *Neurochem. Int.* **57**, 655–661 (2010).
56. Razavi, S.-M. *et al.* Immune evasion strategies of glioblastoma. *Front. Surg.* **3**, 11 (2016).
57. Gangoso, E. *et al.* Glioblastomas acquire myeloid-affiliated transcriptional programs via epigenetic immunoediting to elicit immune evasion. *Cell* https://doi.org/10.1016/j.cell.2021.03.023 (2021).
58 van Weverwijk, A. & de Visser, K. E. Mechanisms driving the immunoregulatory function of cancer cells. *Nat. Rev. Cancer* https://doi.org/10.1038/s41568-022-00544-4 (2023).
59. Majzner, R. G. *et al.* GD2-CAR T cell therapy for H3K27M-mutated diffuse midline gliomas. *Nature* **603**, 934–941 (2022).
60. Vitanza, N. A. *et al.* Intraventricular B7–H3 CAR T cells for diffuse intrinsic pontine glioma: Preliminary first-in-human bioactivity and safety. *Cancer Discov.* **13**, 114–131 (2022).
61. Thomas, B. C. *et al.* CAR T cell therapies for diffuse midline glioma. *Trends Cancer* **9**, 791–804 (2023).

Author contributions

SD/JMS/EX/CK/MD wrote and edited the manuscript. SD/EX/JMS/ED constructed figures and tables. JMS/EX completed data analysis. CK/MD supervised the project and supplied funding. SD/JMS contributed equally to the manuscript and are co-first authors.

Funding

This work was supported by the NIH K08NS092895 grant (MD). JMS is partly supported by NIH/NINDS T32 NS105602.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-73199-w.

Correspondence and requests for materials should be addressed to M.D.