Alternative splicing in pediatric central nervous system tumors highlights oncofetal candidate *CLK1* exon 4

Ammar S. Naqvi[1,2,+], Patricia J. Sullivan[3,+], Ryan J. Corbett[3], Priyanka Sehgal[4], Karina L. Conkrite[5], Komal S. Rathi[1,2], Brian M. Ennis[1,2], Katharina E Hayer[6], Bo Zhang[1,2], Miguel A. Brown[1,2], Daniel P. Miller[1,2], Alex Sickler[3], Adam A. Kraya[1,2], Kaleem L. Coleman[3], Joseph M. Dybas[1,2], Zhuangzhuang Geng[1,2], Christopher Blackden[1,2], Shehbeel Arif[1,2], Antonia Chroni[1,2], Aditya Lahiri[6], Madison L. Hollawell[1,2], Phillip B. Storm[1,2], Dalia Haydar[3,7], Jessica B. Foster[5,8], Mateusz Koptyra[1,2], Peter J. Madsen[1,2], Sharon J. Diskin[5,6,8], Andrei Thomas-Tikhonenko[4,5,8,9], Adam C. Resnick[1,2], Jo Lynne Rokita[3,7,*]

[1] Center for Data-Driven Discovery in Biomedicine, Children's Hospital of Philadelphia, Philadelphia, PA, USA

[2] Division of Neurosurgery, Children's Hospital of Philadelphia, Philadelphia, PA, USA

[3] Center for Cancer and Immunology Research, Children's National Hospital, Washington, D.C., USA

[4] Division of Cancer Pathobiology, Children's Hospital of Philadelphia, Philadelphia, PA, USA

[5] Division of Oncology and Center for Childhood Cancer Research, Children's Hospital of Philadelphia, Philadelphia, PA

[6] Department of Biomedical and Health Informatics, Children's Hospital of Philadelphia, Philadelphia, PA, USA

[7] Department of Pediatrics, The George Washington University School of Medicine and Health Sciences, Washington, D.C., USA

[8] Department of Pediatrics, University of Pennsylvania, Philadelphia, PA, USA

[9] Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA, USA

[+]Authors contributed equally

*Correspondence:

Jo Lynne Rokita, Ph.D.

Children's National Hospital

111 Michigan Ave NW

Washington, DC 20010

E-mail: jrokita@childrensnational.org

Abstract

Background: Pediatric brain tumors are the leading cause of disease-related mortality in children, yet many aggressive tumors lack effective therapies. RNA splicing is a hallmark of cancer, but it has not yet been systematically studied in pediatric brain tumors.

Methods: We analyzed 729 pediatric brain tumors spanning histologies and molecular subtypes to quantify differential tumor splicing. We developed the *Splicing Burden Index (SBI)* to enable cross-sample comparisons and performed hierarchical clustering of highly variable splice events to define splicing-informed tumor groups. These were integrated with clinical outcomes, pathway activity, and proteogenomic data. Recurrent splice events were prioritized for predicted functional impact, and *in vitro* perturbation studies were performed targeting the splicing kinase *CDC-like kinase 1 (CLK1)*.

Results: SBI revealed substantial inter- and intra-histology heterogeneity. Clusters were enriched for histologies and molecular subtypes, several of which were independently associated with survival beyond histology and clinical covariates. Spliceosome pathway activity varied across clusters and was associated with worse survival, yet was not correlated with SBI, indicating distinct dimensions of splicing dysregulation. Functional prioritization identified a recurrent in *CLK1* exon 4, required for canonical kinase activity. *CLK1* exon 4 inclusion followed an oncofetal

pattern and showed context-dependent associations with outcome distinct from total *CLK1* expression. Pharmacologic inhibition and exon 4-specific perturbation of *CLK1* reduced tumor cell viability and disrupted cancer-relevant splicing and transcriptional programs.

Conclusions: This study systematically characterizes splicing in pediatric brain tumors, identifies splicing-informed subgroups, and prioritizes *CLK1* exon 4 as an oncofetal tumor-specific event, motivating further preclinical exploration.

Keywords: splicing, pediatric brain tumors, *CLK1*, splicing burden, therapeutic targeting

Key Points

- Splicing analysis of 729 pediatric CNS tumors identifies splicing-defined clusters.
- *CLK1* exon 4 inclusion is widespread and developmentally regulated.
- Exon-level *CLK1* regulation shows context-dependent links to prognosis in aggressive CNS tumors.

Importance of study

This study provides a comprehensive, cross-histology characterization of alternative splicing across 729 primary pediatric central nervous system (CNS) tumors, revealing widespread and heterogeneous splicing dysregulation with clinical relevance. By introducing the Splicing Burden Index and defining splicing-informed tumor clusters, we show that splicing patterns capture biologically and prognostically meaningful heterogeneity beyond traditional histologic classification. We further demonstrate that global splicing burden, spliceosome pathway activity, and exon-level regulation represent distinct dimensions of splicing biology. Functional prioritization of recurrent splice events identifies *CDC-like kinase 1 (CLK1)* exon 4 as a developmentally regulated, widely included splice event that produces the active kinase isoform

and shows context-dependent associations with outcome. Experimental perturbation of *CLK1* exon 4 reduces tumor cell viability and disrupts cancer-relevant transcriptional and splicing programs. Together, these findings establish alternative splicing as a fundamental and targetable feature of pediatric CNS tumor biology and provide a framework for splicing-directed therapeutic strategies in high-risk disease contexts.

Lay Summary

Brain tumors are the leading cause of cancer-related death in children, and many pediatric brain tumors remain difficult to treat. While most research has focused on genetic mutations, another important layer of regulation—how genes are "spliced" to make different versions of RNA and proteins—has been less well studied in childhood brain cancers. In this study, we analyzed RNA data from 729 pediatric brain tumors to understand how RNA splicing differs across tumor types. We found that tumors could be grouped into distinct categories based on their splicing patterns, and that these patterns were linked to differences in tumor behavior and patient outcomes. Some tumors showed widespread splicing changes that affected important cancer-related pathways. We also identified a splicing change in a gene called *CLK1* that is common in pediatric brain tumors and normally seen during early brain development. Changing how *CLK1* is spliced reduced tumor cell growth and disrupted cancer-related pathways in a representative patient-derived model. Importantly, drugs that target this pathway are already being tested in adults with cancer. Together, these findings show that RNA splicing plays a major role in pediatric brain tumor biology and may represent a new way to classify tumors and develop targeted treatments for children with brain cancer.

Introduction

Pediatric brain cancer is the number one cause of disease-related death in children[1]. Despite advances in genomic profiling, many aggressive pediatric central nervous system (CNS)

tumors lack effective targeted therapies, underscoring the need to identify additional regulatory mechanisms that drive tumor behavior and therapeutic vulnerability. One such mechanism is alternative pre-mRNA splicing, a fundamental post-transcriptional process that expands proteomic diversity and fine-tunes gene regulation, yet remains incompletely characterized in pediatric CNS tumors.

Prior studies have demonstrated that splicing alterations can contribute to pediatric brain tumor biology through diverse mechanisms. Rare, mutually exclusive mutations in spliceosome-associated factors such as *SF3B1* and *SF3B2* have been identified in pediatric high-grade gliomas (HGGs), disrupting processes involved in DNA replication, genome integrity, or transcriptional fidelity[2]. More recent work has shown that alternative splicing can activate oncogenic pathways in pediatric HGGs, including RAS/MAPK signaling, and is associated with worse clinical outcomes[3]. Functional relevance of splice isoforms has also been demonstrated *in vivo*, where a tumor-specific isoform of the neuronal cell adhesion molecule *NRCAM*, but not its canonical transcript, was essential for pediatric HGG xenograft growth[4]. Additional studies have implicated *OTX2*-driven splicing programs in Group 3 medulloblastoma (MB)[5], U1 snRNA splice site mutations in SHH MB[6], tumor-restricted splice-derived neoantigens in MB[7], and pathogenic *NF1* splice variants in Neurofibromatosis Type I–associated tumors[8].

Collectively, these studies establish alternative splicing as a biologically important contributor to pediatric brain tumor pathogenesis. However, they have largely focused on isolated mechanisms, specific tumor types, or individual splice events. As a result, the broader architecture of splicing dysregulation across pediatric CNS tumors—and how exon-level splicing integrates with tumor lineage, signaling pathways, and clinical outcome—remains poorly defined. Particularly, it is unclear which splicing alterations are merely correlative and which actively rewire protein function, signaling networks, or immune visibility to drive malignant behavior.

Alternative splicing is especially relevant in the developing brain, which exhibits the most complex and conserved splicing programs of any tissue[9,10]. These programs are tightly regulated by *trans*-acting RNA-binding proteins (RBPs), including Serine-rich Splicing Factors (SRSFs) and heterogeneous nuclear ribonucleoproteins (hnRNPs), whose dysregulation can profoundly alter cellular identity and function[11,12]. Perturbation of these regulatory networks, or of kinases that modulate splicing factor activity, has the potential to create lineage-specific dependencies in cancer.

Here, we perform a comprehensive, cross-histology analysis of transcriptome-wide alternative splicing across 729 pediatric CNS tumors spanning diverse histologic and molecular subtypes. We identify widespread alternative splicing alterations that define splicing-informed tumor subgroups, reveal heterogeneity within and across tumor lineages, and uncover recurrent splice events predicted to alter protein function. Using a functional prioritization framework, we identify CDC-like kinase 1 (*CLK1*) as a candidate regulator of oncogenic splicing programs in pediatric CNS tumors. Canonical *CLK1* activity requires inclusion of exon 4, whereas skipping of this exon produces catalytically inactive isoforms[13]. Prior studies have identified *CLK1* as a therapeutic target in multiple adult cancers[14,15,16–20], but none have examined its role in pediatric cancer. Here, we demonstrate that exon 4 inclusion in *CLK1* is widespread, developmentally-regulated, and functionally relevant in pediatric CNS tumors, supporting the concept that splice-defined states may represent actionable layers of tumor biology.

Materials and Methods

Study participants

Study participants include pediatric brain tumor patients whose genomic data were deposited into and obtained from the OpenPedCan[21] project. Histologies include atypical teratoid rhabdoid tumor (ATRT, N = 24), choroid plexus tumor (CPT, N = 20), craniopharyngioma (CPG,

N = 29), diffuse midline glioma (DMG, N = 29), ependymoma (N = 69), germ cell tumor (N = 8), low-grade glioma (LGG, N = 200), medulloblastoma (MB, N = 105), meningioma (N = 17), mesenchymal tumor (N = 17), mixed neuronal-glial tumor (GNT, N = 63), neurofibroma plexiform (N = 10), non-neoplastic tumor (N = 26), other CNS embryonal tumor (N = 8), other high-grade glioma (other HGG, N = 74), schwannoma (N = 13), and Rare CNS tumors (N = 17) (**Table S1**).

Primary data analyses

Somatic primary workflows were implemented by the Kids First Data Resource Center as described in the Open Pediatric Brain Tumor Atlas (OpenPBTA)[22] and OpenPedCan[Citation error] projects. The code for these workflows, including RNA-seq quantification, fusion identification, RNA splicing, and SNV, INDEL, CNV, SV calling, can be found at https://github.com/d3b-center/OpenPedCan-workflows. Sample-level data can be found through the Kids First Portal at https://kidsfirstdrc.org/. To detect alternative splicing, we ran rMATS turbo (v. 4.1.0)[23] with GENCODE v39 GFF annotations on single samples, as described by the Kids First RNA-Seq workflow (https://github.com/d3b-center/OpenPedCan-workflows). We filtered for alternative splicing events with ≥ 10 junction read counts. To compare RNA-Seq from *CLK1* exon 4 morpholino-treated cells vs control morpholino-treated cells, we ran rMATs with three biological replicates for each condition `--b1 --b2`. This grouped analysis calculated ΔPSI, p-values, and FDR statistics for each splice event. These results were then used for all downstream processing throughout the manuscript.

Cell Culture

The high-grade glioma patient-derived cell lines 7316-1763 and 7316-1769 were obtained by CBTN request, and the KNS-42 cell line was obtained from Accegen (ABC-TC0532). The pediatric HGG cell line KNS-42 was cultured in DMEM-F12 (GIBCO, 11320033) supplemented

with 10% FBS (GIBCO, 26140079), 2 mmol/L L-glutamine (GIBCO, 25030081), and 1X penicillin/streptomycin (GIBCO, 15140122) at 37°C and 5% CO_2. The cell line was authenticated by Guardian Forensic Sciences (Abington, PA) using the GenePrint 24 (Promega, B1870) short tandem repeat kit. Cells tested negative for mycoplasma using the EZ-PCR Mycoplasma Detection Kit (Biological Industries, 20-700-20) and were used for a maximum of 12 passages post thaw.

Morpholino Treatments

A Vivo-Morpholino ACTCTTCTGGAAACGTCAAGTGGGC (Gene Tools, LLC) targeting the intron 3-exon 4 splice junction was used to skip exon 4 in *CLK1*. Cells were treated with 1, 5, and 10 μM concentrations of *CLK1* morpholino and 10 μM of Control morpholino. 48 hours post-treatment, cells were harvested for PCR and immunoblots.

RNA Extraction and Quantitative Real-time PCR (qRT-PCR)

Total RNA was isolated and treated with DNAse using the Maxwell RSC simplyRNA Cells kit (Promega, AS1390) with the Maxwell RSC48 Instrument (Promega) per the manufacturer's instructions. Next, 2 μg of RNA were reverse-transcribed using SuperScript IV (Invitrogen, 18090010). Primers used for *CLK1* mRNA transcript quantification are listed in **Table S6K**. qRT-PCR was performed using PowerSYBR Green PCR Master Mix (Invitrogen, 4367659) on an Applied Biosystems Viia7 machine. The amplification was performed using the following settings: denaturation at 95°C for 10 min, followed by 40 cycles of denaturation at 95°C for 15 s and annealing at 60°C for 1 min. The comparative cycle threshold (CT) method was applied to quantify the expression levels of *CLK1*. The fold change of gene expression was calculated by the equation 2ΔΔCT, with *HPRT* (Thermo Fisher, 4453320, assay ID: Hs02800695_m1) used as the housekeeping gene.

Protein Extraction

Cultured cells were washed once in chilled D-PBS (pH 7.4) and lysed in RIPA buffer containing 50 mM Tris-HCl, pH 7.4, NP-40 (1%), deoxycholate (0.25%), 150 mM NaCl, 1 mM EDTA pH 8.0, 1x protease and phosphatase inhibitor cocktail (Pierce Halt Inhibitor Cocktail, Thermo Fisher Scientific, 78446), and SDS (0.1%). Total protein in the lysate was estimated by the DC Protein assay (BioRad Laboratories, 5000111).

Detection of Proteins Using Immunoblot Analysis

70 µg of total protein were mixed with 5X SDS loading dye (Biorad, 161-0374) and resolved on 10% SDS-polyacrylamide gel. The protein was transferred onto a PVDF membrane (Immobilin-P, Millipore, IPVH00010) and probed with α-CLK1 mouse monoclonal primary antibody (Santa Cruz, sc-515897) and HRP-conjugated secondary antibody (Cell Signaling Technology, 7076S). Bands were detected using enhanced chemiluminescence (Millipore, WBKLS0500) and captured by a Chemiluminescence imager (GE Healthcare). β-actin was used as the loading control and probed with α-β-actin rabbit monoclonal antibody (Cell Signaling Technology, 12262S).

pan-DYRK/CLK1 inhibitor Cirtuvivint (SM08502) experiments

The KNS-42 cell line was cultured in DMEM-F12 (Gibco, 11330032) supplemented with 10% FBS (Corning, MT3501CV, lot 003322001) and additional L-glutamine (Thermo Fisher, 25030081) to a final concentration of 4.5 mM. Dissociation was performed with Trypsin-EDTA (0.05%, Thermo 25300054) and counted on a DeNovix Cell Drop cell counter.

For growth kinetics, 10,000 (3 day assay) or 6,000 (6 day assay) cells were plated per well into a 96-well plate (Greiner Bio-One, 655098) in a 200 uL total volume per well. Plates were placed into an Incucyte SX5 device and scanned every 2 hours for several days to measure growth via a mask designed uniquely for this cell type. At the end point of the assay, cell viability

was analyzed with CellTiter Glo 2.0 reagent (Promega, G9242) by replacing half the media with reagent and reading on a Promega GloMax device.

Cirtuvivint (MedChem Express, HY-137435) was resuspended in 100% DMSO (Sigma, D2650-5X5ML) to 1 mM and stored in aliquots at -80 C. Dosing was optimized via serial dilution at a range of 20 uM to 0.02 uM against a vehicle control equivalent to the highest dosing of drug. Cells were plated and at 24 hours, 100 uL of media were removed from each well and replaced with drug media for a final dose range of 0.01, 0.05, 0.5, 0.1, 1, 5, and 10 uM. Cells were untouched for 3 days total while growth was monitored via Incucyte.

Cell Viability Assay

Cell viability was measured using the CellTiter-Glo (CTG) luminescent cell viability assay (Promega, G7570). Cells were seeded in white 96-well flat-bottom plates at a density of 24,000 cells per well and treated the following day with either 7.5 μM control or CLK1 exon 4 targeted morpholino. Luminescence was measured using a Biotek Synergy 2 plate reader at 24, 48, 72, and 96 hours.

Results

Pediatric brain tumors display heterogeneous global patterns of differential splicing

To investigate alternative splicing in pediatric brain tumors, we analyzed total RNA-Seq data from 729 diagnostic tumors from the Open Pediatric Cancer (OpenPedCan) project[21] (**Figure 1A**). Demographic and clinical characteristics for each patient and tumor are provided in **Table S1**. We quantified splicing at single-sample resolution using percent-spliced-in (PSI) values computed with rMATS turbo[23], which measure relative exon or junction usage independent of total gene expression, and identified skipped exon (SE), alternative 5' splice site (A5SS), alternative 3' splice site (A3SS), and retained intron (RI) events.

Given the biological and molecular diversity of pediatric CNS tumors, we asked whether distinct splicing patterns would be observed across histologies and molecular subtypes. Among recurrent (N ≥ 2) differential splice events, SE events were the most frequent (**Figure 1B**), consistent with a prior report in pediatric HGGs[3]. We next assessed whether these events were histology-specific or shared across tumor types. We observed both shared and histology-specific splicing events (**Figure S1A**), with medulloblastoma (MB), low-grade glioma (LGG), and non-DMG high-grade glioma (other HGG) tumors exhibiting the largest number of unique recurrent events, consistent with their larger representation in the cohort. Importantly, this pattern persisted after normalizing for the number of patients within each histology (**Figure 1C**), indicating that both shared and histology-enriched splicing programs are present across pediatric CNS tumors. A complete list of unique events per histology is reported in **Table S2A**.

To enable cross-sample comparisons of alternative splicing activity, we developed the splicing burden index (SBI): a sample-level metric quantifying the proportion of differential alternative splicing events in each tumor (**Online Methods**). Across the cohort, the median SBI was 0.0351 (3.51%). Among tumor types, LGGs exhibited the lowest median SBI (2.16%), while germ cell tumors (GCTs) had the highest (10.1%) (**Figure 1D**). Notably, SBI varied substantially within tumor histologies. Tumor groups with high SBI variance (> 3rd quartile variance) included other CNS embryonal tumors, GCTs, rare CNS tumors, neurofibroma plexiforms, and ATRTs, indicating marked inter-tumoral heterogeneity in splicing programs. Although overall SBI values were modest on a cohort-wide scale, this variability underscores the heterogeneity of splicing programs across pediatric CNS tumors.

We next asked whether tumors with a low tumor mutation burden (TMB) might exhibit an increased splicing burden as an alternate mechanism contributing to tumorigenesis. Instead, we observed a very weak positive correlation between TMB and SBI (Pearson's R = 0.13, p-value = 6.9e-4; **Figure S1B**), which persisted after excluding hyper-mutant tumors (R = 0.14, p-value = 3.4e-4, **Figure S1C**). Although statistically significant, these correlations explain little variance,

indicating that TMB and splicing burden are largely independent across the cohort. Taken together, these results suggest that TMB is not a major determinant of global splicing burden in pediatric CNS tumors. When stratified by histology, significant positive correlations were observed only in GCTs ($R = 0.79$, $p = 0.033$) and MB ($R = 0.21$, $p = 0.048$), whereas schwannomas showed a significant inverse relationship between TMB and SBI ($R = -0.78$, $p = 0.0049$) (**Figure S1D**). Schwannomas also exhibited the lowest variance in SBI across tumor types, consistent with tighter constraint on splicing burden in this lineage.

Splice events cluster pediatric brain tumor histologies and are associated with survival outcomes

To assess whether CNS tumors exhibit shared patterns of alternative splicing, we performed hierarchical clustering of samples using PSI values from the top 5,000 most variable splice events across all primary tumors (N = 729). This analysis identified 10 distinct clusters, each enriched for specific histologies and/or molecular subtypes (**Figure 2A-B,** and **S2A-D**). Notable enrichments included Cluster 4 with ependymoma (OR = 222.8), Cluster 1 with LGGs (OR = 47.9), and Cluster 9 with MB (OR = 115.1), with the latter driven predominantly by Group 4 tumors (OR = 94.4) (**Figure S2B**). These results are consistent with prior work demonstrating that the MB subgroups WNT, SHH, Group 3, and Group 4 can be delineated based on splicing patterns inferred by expression arrays[24]. HGGs, including DMGs, exhibited the greatest splicing heterogeneity, with samples distributed across 8 of the 10 clusters (**Figure S2C-D**). Cluster membership for all samples is provided in **Table S2B**.

To assess whether these clustering patterns primarily reflected differences in cellular composition rather than tumor-intrinsic splicing programs, we quantified the proportion of PSI variance explained by estimated cell-type composition inferred from bulk RNA-Seq gene expression data using BRETIGEA[25]. For most cluster-defining events, less than 30% of total PSI variance was attributable to cell-type composition; approximately 15% of events showed 30–50%

variance explained, and 10% exceeded 50% (**Figure 2C**). These findings indicate that while cellular composition contributes to splicing variability for a subset of events, it does not account for the majority of splicing differences that define the observed clusters.

We next examined overall survival (OS) across splicing-defined clusters. After adjusting for histology and additional clinical covariates including tumor resection and age, several clusters remained independently associated with OS relative to Cluster 1 (**Figure 2D**). Specifically, Cluster 9 showed improved OS, whereas Clusters 3, 4, 5, and 7 were associated with worse OS. These results indicate that splicing-defined clusters are associated with prognostically relevant heterogeneity beyond histologic classification alone. Given these associations, we evaluated SBI as a covariate in the survival models and tested for interactions between SBI and cluster membership. We identified significant SBI-cluster interactions in Clusters 5, 7, 8, and 10, indicating that the relationship between splicing burden and OS varies across splicing-defined contexts. In Clusters 5, 7, 8, and 10, higher SBI was associated with improved OS within those clusters. However, at the cluster level, membership in Clusters 7 and 8, as well as higher SBI values across the cohort, remained associated with poorer OS (**Figure 2E**). SBI distributions by cluster, stratified by histology, are shown in **Figure S2E**.

To explore the functional impact of splicing alterations, we identified differentially-expressed cancer-associated signaling pathways across the 10 splicing-defined clusters using GSVA (**Figure 2F**). The KEGG spliceosome pathway was significantly upregulated in Clusters 7, 8, and 9, which were associated with poorer OS (**Table S2C,** Bonf-adj $p < 0.05$; **Figure S3A**) and significantly downregulated in Clusters 2, 5, and 6, which were associated with more favorable outcomes (**Table S2,** Bonf-adj $p < 0.05$; **Figure S3A**). These results indicate marked heterogeneity in spliceosome pathway activity across splicing-defined tumor contexts.

We next assessed whether spliceosome activity, quantified by GSVA score, was associated with survival outcomes across the cohort using a multivariate cox proportional hazards model adjusting for tumor resection, glioma grade group, age at diagnosis, and cluster

membership. Tumor resection was a favorable prognostic factor (HR = 0.25 for gross/near total and 0.44 for partial, $p < 0.01$), whereas non-LGG tumors (HR = 13.60, $p < 0.01$), Cluster 7 membership (HR = 4.36, $p < 0.01$), and higher spliceosome GSVA scores (HR = 2.91, $p < 0.01$) were independently associated with worse OS (**Figure S3B**).

To evaluate whether increased KEGG spliceosome GSVA scores reflect elevated protein abundance, we integrated RNA expression with matched proteogenomic data (N = 122) from the Clinical Proteomic Tumor Analysis Consortium (CPTAC)[26]. Protein (**Figure S3C**) and RNA (**Figure S3E**) expression of KEGG spliceosome pathway genes (**Table S3A**) were visualized alongside GSVA scores, revealing a significant positive correlation between protein abundance and GSVA scores (R = 0.4, p = 2.6e-6; **Figure S3E**), supporting the use of transcript-based GSVA as a proxy for pathway activity. Notably, spliceosome GSVA scores were not correlated with splicing burden (**Figure S3F**), indicating that global spliceosome pathway activity and splicing burden capture distinct aspects of splicing dysregulation.

Widespread splicing alterations are associated with expression changes in splicing factors and recurrent splice events with predicted functional impact

To investigate potential mechanisms underlying widespread splicing alterations, we assessed recurrent somatic alterations across the cohort (N = 657, **Figure S4A**). While recurrent alterations were observed in canonical oncogenic drivers, recurrent predicted deleterious mutations in components of the splicing machinery were rare overall. Specifically, only 3.5% (23/657) of tumors harbored PolyPhen- and SIFT-predicted deleterious mutations in 50 of the 150 HUGO-annotated spliceosome genes (**Table S3B, S3D**) and 18.1% (119/657) harbored predicted deleterious mutations across 446 of the 1,350 known SF genes (**Tables S3C, S3E**). These findings indicate that widespread splicing alterations in pediatric CNS tumors are unlikely to be primarily driven by recurrent coding mutations in splicing machinery genes.

Prior studies have shown that in the absence of recurrent splicing factor mutations, changes in the expression of splicing regulators can influence downstream splicing programs and contribute to tumorigenesis[27–29]. Guided by this framework, we performed differential gene expression (DE) analysis between high and low SBI tumors, focusing on known splicing factors and related genes[30] (**Table S3C**). We found 65.3% (N = 881/1350) of these genes were significantly differentially-expressed (adjusted p-value < 0.05, **Figure 3A-B, Table S3F**). Notably, 88% (30/34) of genes encoding the SRSFs and hnRNPs – families of *trans*-acting splicing regulators with established roles in exon-associated splicing[31] – were significantly differentially expressed between high and low SBI tumors, suggesting that SBI reflects coordinated splicing factor–associated regulatory states.

To further assess whether global splicing burden reflects coordinated regulation of splicing machinery, we examined correlations between SBI and the expression of splicing factor (SF) genes across splicing-defined clusters. We generated a correlation heatmap of the top 40 SF genes most correlated with SBI, stratified by cluster (**Figure 3C**), with corresponding correlation statistics summarized by cluster and by histology (**Table S4A-B**). SBI showed significant positive correlations with SF gene expression in Clusters 2, 6, 7, 8, 9, and 10, whereas Clusters 1, 3, 4, and 5 exhibited predominantly negative correlations. Notably, Cluster 6 displayed a mixed pattern, with subsets of SF genes showing both positive and negative correlations with SBI. These findings indicate that increased splicing burden is associated with coordinated, but cluster-specific, regulation of splicing factor expression rather than uniform upregulation of the splicing machinery across all tumors.

To further prioritize recurrent splice changes with potential biological relevance across splicing-defined tumor states, we developed a robust and adaptable workflow to identify recurrent (N ≥ 2) differentially-spliced (ΔPSI z-score > |2|) events, ΔPSI > 0.2 (sample compared to cohort), and predicted functional impact (**Figure 3D**). Across the cohort, we identified 97,189 recurrent

differential splicing events. Of these, 34,953 events were prioritized based on predicted gain or loss of annotated UniProt functional features, including disulfide bonds (N = 3,627), localization signals (N = 1,551), amino acid modification sites (N = 6,997), and protein domains (N = 28,779) (**Table S5**).

To focus on potentially targetable splice alterations, we restricted this set to functional splice events occurring in kinases, yielding 1,129 events. We further refined this list by prioritizing kinases with established roles in splicing regulation, resulting in 37 candidate splice events across 11 genes: *CLK1, CLK2, CLK3, CLK4, FASTK, MARK2, PKN2, PRKDC, SMG1, SRPK1,* and *SRPK3*. Applying additional filters based on expression (TPM > 10) and recurrence (> 10 tumors), we identified *CLK1, CLK3, FASTK,* and *PRKDC* as top candidates comprising 11 unique splice events. We excluded seven RI events, which are known to have higher false-positive rates in short-read sequencing[32], and two additional RI events in *CLK3* that were annotated as A5SS and A3SS. After these exclusions, only two SE events remained: *CLK1* exon 4 and *PRKDC* exon 80. Comparison of *PRKDC* exon 80 inclusion across normal polyA-sequenced cohorts showed that our PBTA polyA data closely matched GTEx (**Figure S4B**), indicating that the stranded-specific signal may reflect a technical artifact. We therefore prioritized *CDC-like kinase 1 (CLK1)*, a well-established regulator of alternative splicing[11], as a candidate node within high-burden splicing states for further investigation. *CLK1* regulates SRSF proteins through hyper-phosphorylation of SR-rich domains, promoting cooperative RNA binding and splicing activity[33–35]. Canonical *CLK1* activity requires inclusion of exon 4[13]. We observed differential splicing of this exon, with 34 tumors showing significant skipping and a mean ΔPSI of 0.29 (**Table S5C)**. This event is illustrated using sashimi plots contrasting tumors with high versus low exon 4 inclusion (**Figure 3E**). Notably, the majority of tumors (n = 695 out of 729) showed high *CLK1* exon 4 inclusion (mean PSI = 0.790; **Table S5D**); **Figure S4C**), motivating further investigation of this splice event in the context of tumor biology and development.

Given the identification of *CLK1* exon 4 skipping as a recurrent splice event across the cohort, we next examined whether exon 4 inclusion (PSI) and total *CLK1* expression were associated with survival outcomes. Using multivariate Cox proportional hazards models adjusted for cluster membership, extent of resection, histology, and age at diagnosis, we found that higher *CLK1* exon 4 PSI was independently associated with improved overall survival (HR = 0.17, $p < 0.01$; **Figure 3F**). Tumor resection remained a favorable prognostic factor, whereas non-LGG tumors and Cluster 7 membership were associated with worse outcomes, consistent with prior analyses.

In contrast, when examining *CLK1* expression, we observed cluster-specific associations with survival. In Clusters 3 and 7—enriched for *BRAF* V600E–mutant low-grade gliomas (Cluster 3) and for other high-grade gliomas, including diffuse midline gliomas (Cluster 7) – either higher *CLK1* exon 4 containing transcript abundance or total *CLK1* transcript abundance – was associated with worse overall survival after adjusting for resection status, histology, and age (HR = 1.16 in Cluster 3 and 5.96 in Cluster 7, $p < 0.01$; **Figures 3G–H**). These findings indicate that exon-level splicing regulation and gene-level expression of *CLK1* have distinct and lineage-associated relationships with clinical outcome.

Consistent with these observations, analysis of spliceosome pathway activity within Cluster 7 revealed that lower GSVA scores were associated with improved overall and event-free survival, whereas higher scores were associated with poorer outcomes (**Figures 3I, S4D**). Together, these results suggest that while elevated *CLK1* expression and spliceosome pathway activity track with adverse prognosis in specific high-risk clusters, exon 4 inclusion represents a distinct splicing feature whose association with outcome is independent of overall expression levels.

To further resolve the relationships between *CLK1* exon-level splicing, gene expression, and global splicing burden, we examined correlations between *CLK1* transcript abundance (TPM), exon 4 PSI, and SBI across splicing-defined clusters and histologies (**Figure S5A-C**). *CLK1*

expression and SBI showed strong, cluster-dependent relationships, with positive correlations in Clusters 7, 8, and 9 and a strong inverse correlation in Cluster 1 (**Figure S5A**). When stratified by histology, *CLK1* expression was positively correlated with SBI in ATRT, DMG, medulloblastoma, meningioma, other HGGs, and schwannoma, but negatively correlated in glioneuronal tumors (**Figure S5B**). In contrast, *CLK1* exon 4 PSI was strongly and positively correlated with *CLK1* expression across all clusters and histologies (**Figure S5C**). Together, these analyses indicate that exon 4 inclusion is tightly coupled to *CLK1* expression, whereas the relationship between *CLK1* expression and global splicing burden is context dependent and varies across tumor histologies.

The uniform coupling between *CLK1* exon 4 inclusion and expression across tumor contexts, together with the lineage-specific relationship between *CLK1* expression and splicing burden, prompted us to next examine whether exon 4 inclusion reflects a developmentally regulated splicing state.

***CLK1* exon 4 inclusion marks an oncofetal, lineage-associated splicing state with context-specific sensitivity to CLK inhibition**

To contextualize our findings within normal neurodevelopment, we examined *CLK1* exon 4 inclusion in non-tumor brain tissues using RNA-seq data from the Evolutionary Developmental (Evo-Devo) Atlas (N = 59) and GTEx (N = 2,642). *CLK1* exon 4 inclusion was significantly higher in pediatric brain tumors compared to GTEx adult brain, was comparable to fetal Evo-Devo samples, and decreased with age across both Evo-Devo and the PBTA ($p < 0.05$; **Figure 4A; Table S5E**). This pattern is consistent with an oncofetal splicing program[36], and supports the model that pediatric CNS tumors re-engage developmentally-regulated *CLK1* splicing states that are normally attenuated during postnatal brain maturation. Reactivation of this state may impose distinct post-transcriptional regulatory requirements that are not uniformly buffered across lineages.

Next, we evaluated whether *CLK1* and its exon 4 inclusion is associated with selective sensitivity in pediatric brain tumors. Analysis of the Cancer Dependency Map (DepMap) revealed that CNS tumor cell lines with high expression of the exon 4-included *CLK1* transcript ENST00000321356 (used as a proxy for exon 4-containing isoform abundance) exhibit significantly greater CRISPR sensitivity, reflected by more negative Chronos gene effect scores (0 ~ nonessential, -1 ~ common essentials), compared to cell lines with low exon 4 inclusion (Wilcoxon $p = 0.034$; **Figure 4B**). Notably, this association was restricted to CNS and myeloid malignancies and was not observed across other cancer types (**Figure S6A**), indicating lineage-specific sensitivity to perturbation of *CLK1*.

To investigate molecular features associated with this context specificity, we performed a correlation-based analysis of gene expression with the *CLK1* exon 4-containing transcript across CNS and myeloid cell lines from DepMap. We identified genes whose expression levels were significantly correlated with *CLK1* within each lineage. While only 3.9% of *CLK1*-correlated genes were shared between the two lineages (**Figure S6B, Table S4C**), both the shared core and the substantially larger lineage-specific gene sets were significantly enriched for RNA splicing, spliceosome-associated pathways, and chromatin remodeling pathways (B-H adjusted $p < 0.05$; **Figure S6C, 4C**). These findings suggest that *CLK1* sensitivity arises in cellular contexts characterized by coordinated, regulated exon-level splicing programs embedded within broader transcriptional and epigenetic networks, rather than simply elevated spliceosome component expression or generalized spliceosome activation.

Among pediatric CNS tumor cell lines in DepMap, the diffuse hemispheric glioma line KNS-42 exhibited the strongest sensitivity to *CLK1* perturbation (**Figure 4D**) and was therefore selected for mechanistic interrogation. KNS-42 also demonstrates robust inclusion of *CLK1* exon 4, consistent with the splicing patterns observed in our primary tumor cohort. To extend beyond a single model, we additionally evaluated two patient-derived pediatric brain tumor cell lines (7316-1763 and 7316-1769) that similarly exhibit high *CLK1* exon 4 inclusion with short-read

RNA-Seq. Using Oxford Nanopore long-read RNA sequencing, we confirmed that these models express comparable ratios of *CLK1* isoforms that include or skip exon 4, validating the splice event initially identified by short-read RNA sequencing (**Figure 4E**).

Next, we tested the impact of CLK1 inhibition in KNS-42 cells using the pan-Dyrk/Clk inhibitor cirtuvivint (SM08502)[19]. Using the IncuCyte Live Cell Analysis System to monitor real-time proliferation, we observed a significant reduction in cell growth at multiple concentrations over a 6-day period (**Figure 4F**). Additionally, we observed a dose-dependent decrease in cell viability using CellTiter-Glo at three days (**Figure 4G**) and six days (**Figure S6D**) post-treatment of 0.5, 1, 5, and 10 μM Cirtuvivint. Although the reduction in proliferation/viability with cirtuvivint was dose-dependent and statistically significant, the magnitude was modest. Given that cirtuvivint is a pan-CLK/DYRK inhibitor, we next performed exon 4-specific perturbation experiments to more directly test the contribution of the catalytically active CLK1 isoform.

CLK1 regulates the SR family of splicing factor proteins through hyper-phosphorylation of the SR-rich peptide regions of SR proteins to induce cooperative RNA binding and increased activity[33–35]. We therefore postulated that exon 4 inclusion is required to produce a stable, full-length, catalytically active CLK1 protein isoform. To directly test this hypothesis, we modulated *CLK1* exon 4 splicing using targeted morpholino oligomers (see **Online Methods**), in which we forced exon 4 skipping in the KNS-42 cell line. We performed qRT-PCR and observed a near total loss of the *CLK1* exon 4 inclusion transcript at both 5 and 10 μM of exon 4 targeted morpholino, evidenced by reduced expression of the exon 3-4 junction. At these same concentrations, we observed increased *CLK1* exon 4 skipping using primers targeting the exon 3-5 junction (**Figure 4H**). Importantly, forced *CLK1* exon 4 skipping resulted in near complete loss of CLK1 protein at 5 and 10 μM (**Figure 4I**), corroborating previous work that *CLK1* exon 4 is required for full-length and catalytically active *CLK1*[37–39]. Next, we assessed the functional impact of *CLK1* exon 4 splicing using CellTiter-Glo and confirmed that cells with high *CLK1* exon 4 skipping (*CLK1* exon 4 targeting morpholino) exhibited significantly decreased viability compared to those with *CLK1*

exon 4 inclusion (non-targeting morpholino) at 24, 72, and 96 hours ($p \leq 0.01$, within-time Student's *t-test*, **Figure 4J**). Together, these data support a model in which *CLK1* exon 4 inclusion marks a lineage- and context-dependent splicing state associated with selective sensitivity to *CLK1* perturbation in the pediatric KNS-42 model tested.

To identify transcriptional and splicing changes associated with forced *CLK1* exon 4 splicing, we performed RNA-seq from KNS-42 cells treated with morpholino oligomers (N = 3 controls, N = 3 targeted to skip exon 4). We performed differential gene expression (DE) analysis and identified 1,322 genes with differential expression (569 upregulated, 753 downregulated) between cells treated with morpholino or non-targeting control (**Figure 5A, Table S6A**), including 78 oncogenes or tumor suppressor genes (TSGs, **Figure 5B**). Next, we identified a total of 4,001 unique differential splicing (DS) events (SE = 2,256; Mutually Exclusive [MXE] = 926; A5SS = 181; A3SS = 271; and RI = 367; **Table S6B-G**), including 267 oncogenes or TSGs (**Figure 5B**). There were 120 genes (2.2%) which were both DE and DS (**Figure 5D, Table S6H**), indicating these may impact total protein abundance in tumors. DS genes were significantly over-represented for mitotic spindle, E2F targets, G2M checkpoint, and nucleotide excision repair pathways (Bonf-adj $p < 0.05$, **Figure 5E**). To further investigate the impact on DNA repair and other pathways, we performed GSVA of DNA repair and cancer signaling pathways using DS oncogenes and TSGs and found that depletion of CLK1 was associated with upregulation of stress- and cytokine-associated signaling pathways, including TNFA/NFkB, PI3K/AKT/MTOR, IL6/JAK/STAT3, alongside downregulation of multiple DNA repair pathways, consistent with a possible metabolic and cell-state transition. (**Figures 5F-G and S5E-F**). Importantly, *CLK1* morpholino-mediated skipping resulted in significant downregulation of *CLK1* and known downstream targets, including *LRP5, AXIN2,* and *LEF1* ($Log_2FC < -0.1$, B-H p-adj < 0.05, **Table S6I**), accompanied by suppression of WNT signaling (**Figure 5F-G**). While prior studies have primarily examined pharmacologic *CLK1* inhibition, the affected pathways observed here are

directionally concordant with reported pathway-level effects of the pan-CLK/DYRK inhibitor cirtuvivint (SM08502)[19], suggesting that exon 4-dependent perturbation of *CLK1* engages similar downstream regulatory programs.

Finally, we asked whether *CLK1 exon 4* inclusion levels are associated with changes in essential oncogenes defined by the pediatric gene dependency maps of the Childhood Cancer Model Atlas[40]. We identified 20 such genes that also exhibit significant gene dependencies in PBTA cell lines (**Figure 5H-I, Table S6J**), including seven regulators of MAPK signaling: *BRAF, EZH2, RAF1, JUN, FGFR1, FGFR2,* and *SRC*. For instance, expression of mRNAs encoding proto-oncogene *SRC*[41,42] was higher in cells with high *CLK1* exon 4 inclusion (non-targeting morpholino), consistent with a role for *CLK1*-associated splicing states in modulating *SRC* expression. Differential splicing effects were more complex, as they involved multiple transcripts within individual genes; however, taken together, these data suggest that transcript-level changes associated with differential *CLK1* exon 4 inclusion may contribute to sensitivities in pathways such as DNA repair and MAPK signaling. This interpretation is consistent with established links between alternative splicing with cancer progression[43–45].

While extending functional validation to additional pediatric brain tumor models, non-malignant CNS counterparts, and *in vivo* systems will be critical for defining therapeutic windows and translational relevance, these findings position *CLK1* exon 4 as a candidate lineage-associated axis of splicing regulation with potential clinical relevance in pediatric CNS tumors.

Discussion

Pediatric brain tumors remain the leading cause of disease-related mortality in children. In this study, we perform a large-scale, cross-histology analysis of alternative splicing across pediatric CNS tumors and demonstrate that splicing dysregulation is pervasive, heterogeneous, and clinically informative. By integrating exon-level splicing, gene expression, pathway activity,

proteogenomics, and functional perturbation, our work provides a framework for uncovering how splicing programs contribute to tumor biology beyond canonical genomic alterations.

We introduce the Splicing Burden Index (SBI) as a quantitative, sample-level metric that enables comparison of differential splicing across tumors in a cohort without reliance on matched normal tissue, an important consideration when normal brain tissue is difficult to obtain. SBI varied across and within tumor histologies, highlighting substantial inter-tumoral heterogeneity in splicing programs. Contrary to our initial hypothesis, tumor mutational burden (TMB) showed only a very weak association with SBI, explaining little variance and arguing against splicing dysregulation serving primarily as a compensatory mechanism in genetically quiet tumors. Instead, global splicing burden appears to reflect context-dependent dysregulation of RNA processing programs.

Clustering tumors by highly variable splice events identified ten splicing-defined groups enriched for specific histologies and molecular subtypes. Several clusters remained independently associated with survival after adjustment for histology and clinical covariates, demonstrating that splicing captures prognostically-relevant heterogeneity beyond historical tumor classification alone. Importantly, the relationship between splicing burden and outcome was context dependent, with cluster-specific interactions indicating that the biological consequences of splicing dysregulation are not uniform across pediatric CNS tumors.

Pathway-level analyses further underscored this complexity. Spliceosome pathway activity varied across splicing-defined clusters and was independently associated with adverse outcomes in multivariate models. However, spliceosome GSVA scores were not correlated with SBI, indicating that pathway activation and splicing burden represent distinct dimensions of splicing dysregulation. Integration with matched proteogenomic data confirmed that transcript-based GSVA scores reflect underlying protein abundance, supporting their biological relevance.

Mechanistically, deleterious coding mutations in spliceosome components and splicing factor genes were rare, whereas expression changes in splicing regulators were widespread, including across SRSF and hnRNP family members. These findings support a model in which altered expression of splicing regulators, rather than recurrent genetic disruption of splicing machinery, contributes to widespread splicing alterations in pediatric CNS tumors.

To identify splice events with potential functional relevance, we prioritized recurrent differential splicing events predicted to alter annotated protein features. This approach highlighted *CDC-like kinase 1 (CLK1)* as a compelling candidate, based on a recurrent event affecting exon 4, which is required for canonical kinase activity. Although most tumors exhibited high exon 4 inclusion, a subset showed significant skipping, motivating further investigation. Integration with developmental transcriptomes further demonstrated that *CLK1* exon 4 inclusion follows an oncofetal pattern, supporting a model in which pediatric CNS tumors re-engage developmentally regulated splicing states that may create context-specific sensitivities.

Clinical analyses revealed that exon-level splicing and gene-level expression of *CLK1* have distinct and context-dependent associations with outcome. Higher exon 4 inclusion was independently associated with improved survival across the cohort, whereas elevated *CLK1* expression tracked with worse outcomes in specific high-risk clusters. Correlation analyses clarified these relationships: exon 4 inclusion was tightly coupled to *CLK1* expression across all contexts, while the relationship between *CLK1* expression and global splicing burden varied by cluster and histology. These findings highlight the importance of distinguishing isoform-level regulation from bulk expression and pathway activity.

Functional perturbation of *CLK1* exon 4 in a pediatric HGG model reduced overall *CLK1* protein abundance, impaired cell viability, and induced widespread transcriptional and splicing changes affecting cell cycle, DNA repair, and MAPK signaling pathways. While limited to a single

pediatric tumor model, these experiments provide proof-of-concept evidence that exon-level disruption of *CLK1* can perturb oncogenic programs. The modest phenotype observed with pan-CLK/DYRK inhibition highlights both the challenges of polypharmacologic targeting and the importance of isoform-resolved perturbation strategies when evaluating splicing-associated functional sensitivities.

Notably, *CLK* family kinases, including *CLK1*, are already under clinical investigation in multiple adult malignancies through early phase trials. The Pan-Clk/Dyrk Inhibitor cirtuvivint (SM08502) is being used in a phase 1 clinical trial in patients with acute myeloid leukemia (AML) and myelodysplastic syndromes (MDS)[46], and has shown preclinical efficacy in non-CNS solid tumors such as triple negative breast cancer, pancreatic cancer, castrate-resistant prostate cancer, colorectal cancer, endometrial cancer, and non-small cell lung cancer[14,15,16–20]. An ATP-competitive, macrocyclic inhibitor of the CLK family, BH-30236[47], is in a Phase 1/1b clinical trial for patients with AML and MDS[48]. Finally, the CLK1-specific inhibitor, CTX-712, is in Phase 1/2 trial for relapsed or refractory AML and high risk MDS[49], underscoring *CLK1* and its family as a target across diverse tumor histologies. Our study extends the rationale for CLK1 inhibition to pediatric brain tumors, particularly those with high splicing burden and exon 4 inclusion.

In summary, our work provides a systematic, integrative view of alternative splicing across pediatric CNS tumors, identifies splicing-informed tumor clusters with clinical relevance, and prioritizes *CLK1* exon 4 inclusion as a recurrent, developmentally-regulated splice event candidate for further preclinical testing. By distinguishing exon-level regulation from global splicing burden and pathway activation, this study highlights the importance of isoform-resolved analyses and provides a foundation for future splicing-directed therapeutic strategies in pediatric brain cancer.

Limitations of the study

In this study, initial splicing quantifications relied on short-read RNA-Seq technology, which limits resolution of full-length isoforms and complex multi-exon events. Future integration with long-read RNA-Seq will enable more comprehensive isoform-level analyses[50]. The absence of matched normal RNA for each tumor of origin tissue restricts the ability to define mutually exclusive or tissue-specific splicing events. For example, within histologies (eg: LGG), the primary site of the tumor can vary widely depending on diagnosis and it would be ideal to match each tumor to its tissue of origin. We mitigated these limitations by leveraging multiple non-tumor reference cohorts, developmental datasets, and the SBI metric. Finally, while functional validation supports a role for *CLK1* exon 4 in tumor cell fitness, extending these studies to additional pediatric models, non-malignant CNS counterparts, and *in vivo* systems will be essential to define therapeutic windows and translational relevance.

Funding

This work was funded in part by the National Institutes of Health (R03OD036498 to JLR), the Chad Tough Foundation, the Children's Hospital of Philadelphia Division of Neurosurgery, the Children's Brain Tumor Network, and the anonymous private investors to the Children's National Hospital Brain Tumor Institute. This research utilized the Common Fund Data Ecosystem (CFDE) Data Coordinating Center (DCC) data and the Velsera CAVATICA cloud platform provided by the CFDE Cloud Workspace Pilot (1OT2OD030162-01) and the Kids First Data Research Center Cloud Credit Program (U2C HD109731 – 08S1). Additional support was provided by Cure Search for Children's Cancer Foundation Acceleration Initiative Award (ATT) and Mildred L. Roeckle Endowed Chair in Pathology at Children's Hospital of Philadelphia (ATT).

Conflicts of Interest

The authors declare no conflicts of interest.

Author Contributions

Conceptualization: ASN, JLR, ACR, RJC, PJS

Methodology: ASN, JLR, RJC, ACR, KSR, AK, SA, KLC, PJS

Software: BE, ASN, RJC, PJS, JLR

Validation: PS, KC, RJC, KLC, KEH, ASN, AL, PS, PJS

Formal Analysis: ASN, JLR, RJC, KLC, KSR, BE, PJS, AS

Investigation: ASN, JLR, RJC, KSR, AK, PJS

Resources: JLR, ACR, ATT, PBS, CB, SJD, Kal LC, DH

Data Curation: JLR, ASN, RJC

Writing - Original Draft: ASN, JLR, KSR, PS, KLC, PJS

Writing - Review & Editing: ASN, JLR, RJC, SJD, KLC, JBF, JMD, PJS

Visualization: ASN, JLR, KSR, RJC, KLC, PJS

Supervision: JLR, ACR, ATT

Project Administration: JLR, ACR

Funding Acquisition: JLR, ACR, PBS, ATT, SJD

Ethics statement

This study utilized publicly available genomic data from the Open Pediatric Cancer (OpenPedCan) project, derived from the Children's Brain Tumor Network (CBTN) and the Pediatric Neuro-Oncology Consortium (PNOC). These consortia have established protocols for

the ethical collection and sharing of patient data. The patient-derived cell lines 7316-1763 and 7316-1769 were obtained via request from the CBTN, and the KNS-42 cell line was purchased commercially. While this study did not directly involve the collection of new human or animal subject data, the use of these pre-existing resources implies that the original data collection was conducted with appropriate institutional review board approval and informed consent procedures.

Data Availability

Patient genomic data

All pediatric brain tumor raw data are available upon request from the database of Genotypes and Phenotypes (dbGAP), accession number phs002517.v2.p2, and/or from the Children's Brain Tumor Network (https://cbtn.org) and the Pacific Pediatric Neuro-Oncology Consortium (pnoc.us) for data not immediately available in dbGaP. All processed data used in this study were derived from the OpenPedCan project[21] v13 data release at https://github.com/d3b-center/OpenPedCan-analysis. All code for the manuscript analyses and figures are openly available at https://github.com/rokitalab/clk1-splicing.

Non-tumor control RNA-Seq

RNA-Seq data from the GTEx project (dbGAP Accession phs000424) and the Evo-Devo atlas (Array Express Accession E-MTAB-6814) were harmonized with GENCODE v39 using the Kids First Data Resource Center workflow at https://github.com/kids-first/kf-rnaseq-workflow.

CLK1 morpholino RNA-Seq

RNA-sequencing data from the *CLK1* morpholino experiment has been deposited in GSE273841.

Merged primary and summary data

Merged primary matrices and summary files utilized in this manuscript were derived from are openly accessible via the download script in the https://github.com/rokitalab/clk1-splicing repository.

Acknowledgements

We graciously thank all of the patients and families of the Children's Brain Tumor Network (CBTN) and the Pediatric Neuro-oncology Consortium for donating tissue which enabled this research. The authors thank the following collaborators who contributed experiments, analyses, code, and/or code review that were ultimately not included in the manuscript: Poonam Sonawane, Cullen Wilson, Krutika S. Gaonkar, and Run Jin.

References

1. Thorbinson, C., and Kilday, J.-P. (2021). Childhood Malignant Brain Tumors: Balancing the Bench and Bedside. Cancers *13*. https://doi.org/10.3390/cancers13236099.

2. Mackay, A., Burford, A., Carvalho, D., Izquierdo, E., Fazal-Salom, J., Taylor, K.R., Bjerke, L., Clarke, M., Vinci, M., Nandhabalan, M., et al. (2017). Integrated Molecular Meta-Analysis of 1,000 Pediatric High-Grade and Diffuse Intrinsic Pontine Glioma. Cancer Cell *32*, 520–537.e5.

3. Siddaway, R., Milos, S., Vadivel, A.K.A., Dobson, T.H.W., Swaminathan, J., Ryall, S., Pajovic, S., Patel, P.G., Nazarian, J., Becher, O., et al. (2022). Splicing is an alternate oncogenic pathway activation mechanism in glioma. Nat. Commun. *13*, 588.

4. Sehgal, P., Naqvi, A.S., Higgins, M., Liu, J., Harvey, K., Jarroux, J., Kim, T., Mankaliye, B., Mishra, P., Watterson, G., et al. (2025). NRCAM variant defined by microexon skipping is a targetable cell surface proteoform in high-grade gliomas. Cell Rep. *44*, 116099.

5. Saulnier, O., Zagozewski, J., Liang, L., Hendrikse, L.D., Layug, P., Gordon, V., Aldinger, K.A., Haldipur, P., Borlase, S., Coudière-Morrison, L., et al. (2024). A group 3 medulloblastoma stem cell program is maintained by OTX2-mediated alternative splicing. Nat. Cell Biol. *26*, 1233–1246.

6. Suzuki, H., Kumar, S.A., Shuai, S., Diaz-Navarro, A., Gutierrez-Fernandez, A., De Antonellis, P., Cavalli, F.M.G., Juraschka, K., Farooq, H., Shibahara, I., et al. (2019).

Recurrent noncoding U1 snRNA mutations drive cryptic splicing in SHH medulloblastoma. Nature *574*, 707–711.

7. Rivero-Hinojosa, S., Grant, M., Panigrahi, A., Zhang, H., Caisova, V., Bollard, C.M., and Rood, B.R. (2021). Proteogenomic discovery of neoantigens facilitates personalized multi-antigen targeted T cell immunotherapy for brain tumors. Nat. Commun. *12*, 6689.

8. Koczkowska, M., Chen, Y., Xie, J., Callens, T., Gomes, A., Wimmer, K., and Messiaen, L.M. (2023). Analysis of 200 unrelated individuals with a constitutional NF1 deep intronic pathogenic variant reveals that variants flanking the alternatively spliced NF1 exon 31 [23a] cause a classical neurofibromatosis type 1 phenotype while altering predominantly NF1 isoform type II. Hum. Genet. *142*, 849–861.

9. Scotti, M.M., and Swanson, M.S. (2016). RNA mis-splicing in disease. Nat. Rev. Genet. *17*, 19–32.

10. Merkin, J., Russell, C., Chen, P., and Burge, C.B. (2012). Evolutionary Dynamics of Gene and Isoform Regulation in Mammalian Tissues. Science *338*, 1593–1599.

11. Chen, S., Yang, C., Wang, Z.-W., Hu, J.-F., Pan, J.-J., Liao, C.-Y., Zhang, J.-Q., Chen, J.-Z., Huang, Y., Huang, L., et al. (2021). CLK1/SRSF5 pathway induces aberrant exon skipping of METTL14 and Cyclin L2 and promotes growth and metastasis of pancreatic cancer. J. Hematol. Oncol. *14*, 60.

12. Fuentes-Fayos, A.C., Vázquez-Borrego, M.C., Jiménez-Vacas, J.M., Bejarano, L., Pedraza-Arévalo, S., L-López, F., Blanco-Acevedo, C., Sánchez-Sánchez, R., Reyes, O., Ventura, S., et al. (2020). Splicing machinery dysregulation drives glioblastoma development/aggressiveness: oncogenic role of SRSF3. Brain *143*, 3273–3293.

13. Shkreta, L., Delannoy, A., Toutant, J., and Chabot, B. (2024). Regulatory interplay between SR proteins governs CLK1 kinase splice variants production. RNA *30*, 1596–1607.

14. Zhang, C., Chen, Y., Zhu, S., Wang, Z., Lin, H., Lu, J., Zhang, H., Weng, Y., Huang, X., Li, G., et al. (2025). Targeting CLK1/SRSF7 axis-dependent alternative splicing sensitizes pancreatic ductal adenocarcinoma to chemotherapy and immunotherapy. Drug Resist. Updat. *83*, 101292.

15. Liu, N., van der Velde, J.J.S., Ramdjielal, S., Koedoot, E., van Overbeek, N.K., Batenburg, D., Vertegaal, A.C.O., van de Water, B., and Le Dévédec, S.E. (2025). Pharmacological CLK inhibition disrupts SR protein function and RNA splicing blocking cell growth and migration in TNBC. Breast Cancer Res. *27*, 140.

16. Tolcher, A., Babiker, H.M., Chung, V., Kim, E., Moser, J., Karim, R., Vandross, A., Sommerhalder, D., Scott, A.J., Fakih, M., et al. (2021). Abstract CT112: Initial results from a Phase 1 trial of a first-in-class pan-CDC-like kinase inhibitor (SM08502) with proof of mechanism in subjects with advanced solid tumors. Cancer Res. *81*, CT112–CT112.

17. ClinicalTrials.Gov https://www.clinicaltrials.gov/study/NCT05084859.

18. A. Scott, J. Call, S. Chandana, E. Borazanci, G. Falchook, R. Bordoni, S. Richey, A. Starodub, V. Chung, N. Lakhani, E. Lam, K. Schaffer, J. Wang, G. Shapiro, J. Sachdev, D. Beaupre and A. Tolcher (2022). Preliminary Evidence of Clinical Activity from Phase 1 and

1b Trials of the CLK/DYRK Inhibitor Cirtuvivint (CIRT) in Subjects with Advanced Solid Tumors.

19. Bossard, C., McMillan, E.A., Creger, E., Eastman, B., Mak, C.-C., Beaupre, D.M., and White, M.A. (2021). The Pan-Clk/Dyrk Inhibitor Cirtuvivint (SM08502) Exposes Mechanistic Underpinnings of Alternative Splicing As a Therapeutic Vulnerability in Heme Malignancies. Blood *138*, 2950.

20. Corr, B.R., Moroney, M.R., Woodruff, E., Watson, Z.L., Jordan, K.R., Danhorn, T., Bailey, C., Wolsky, R.J., and Bitler, B.G. (2023). Combination CDC-like kinase inhibition (CLK)/Dual-specificity tyrosine-regulated kinase (DYRK) and taxane therapy in CTNNB1-mutated endometrial cancer. bioRxivorg. https://doi.org/10.1101/2023.04.04.535570.

21. Geng, Z., Wafula, E., Corbett, R.J., Zhang, Y., Jin, R., Gaonkar, K.S., Shukla, S., Rathi, K.S., Hill, D., Lahiri, A., et al. (2025). The Open Pediatric Cancer Project. Gigascience *14*. https://doi.org/10.1093/gigascience/giaf093.

22. Shapiro, J.A., Gaonkar, K.S., Spielman, S.J., Savonen, C.L., Bethell, C.J., Jin, R., Rathi, K.S., Zhu, Y., Egolf, L.E., Farrow, B.K., et al. (2023). OpenPBTA: The Open Pediatric Brain Tumor Atlas. Cell Genomics *3*. https://doi.org/10.1016/j.xgen.2023.100340.

23. Shen, S., Park, J.W., Lu, Z.-X., Lin, L., Henry, M.D., Wu, Y.N., Zhou, Q., and Xing, Y. (2014). rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc. Natl. Acad. Sci. U. S. A. *111*, E5593–E5601.

24. Dubuc, A.M., Morrissy, A.S., Kloosterhof, N.K., Northcott, P.A., Yu, E.P., Shih, D., Peacock, J., Grajkowska, W., van Meter, T., Eberhart, C.G., et al. (2012). Subgroup-specific alternative splicing in medulloblastoma. Acta Neuropathol. *123*, 485–499.

25. McKenzie, A.T., Wang, M., Hauberg, M.E., Fullard, J.F., Kozlenkov, A., Keenan, A., Hurd, Y.L., Dracheva, S., Casaccia, P., Roussos, P., et al. (2018). Brain cell type specific gene expression and co-expression network architectures. Sci. Rep. *8*, 8868.

26. Petralia, F., Tignor, N., Reva, B., Koptyra, M., Chowdhury, S., Rykunov, D., Krek, A., Ma, W., Zhu, Y., Ji, J., et al. (2020). Integrated Proteogenomic Characterization across Major Histological Types of Pediatric Brain Cancer. Cell *183*, 1962–1985.e31.

27. Karni, R., de Stanchina, E., Lowe, S.W., Sinha, R., Mu, D., and Krainer, A.R. (2007). The gene encoding the splicing factor SF2/ASF is a proto-oncogene. Nat. Struct. Mol. Biol. *14*, 185–193.

28. Kaida, D., Schneider-Poetsch, T., and Yoshida, M. (2012). Splicing in oncogenesis and tumor suppression. Cancer Sci. *103*, 1611–1616.

29. Golan-Gerstl, R., Cohen, M., Shilo, A., Suh, S.-S., Bakàcs, A., Coppola, L., and Karni, R. (2011). Splicing factor hnRNP A2/B1 regulates tumor suppressor gene splicing and is an oncogenic driver in glioblastoma. Cancer Res. *71*, 4464–4472.

30. Sebestyén, E., Singh, B., Miñana, B., Pagès, A., Mateo, F., Pujana, M.A., Valcárcel, J., and Eyras, E. Large-scale analysis of genome and transcriptome alterations in multiple tumors unveils novel cancer-relevant splicing networks. Preprint, https://doi.org/10.1101/023010 https://doi.org/10.1101/023010.

31. Zhu, J., Mayeda, A., and Krainer, A.R. (2001). Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. Mol. Cell *8*, 1351–1361.

32. David, J.K., Maden, S.K., Wood, M.A., Thompson, R.F., and Nellore, A. (2022). Retained introns in long RNA-seq reads are not reliably detected in sample-matched short reads. Genome Biol. *23*, 240.

33. Ninomiya, K., Kataoka, N., and Hagiwara, M. (2011). Stress-responsive maturation of Clk1/4 pre-mRNAs promotes phosphorylation of SR splicing factor. J. Cell Biol. *195*, 27–40.

34. Dominguez, D., Tsai, Y.-H., Weatheritt, R., Wang, Y., Blencowe, B.J., and Wang, Z. (2016). An extensive program of periodic alternative splicing linked to cell cycle progression. Elife *5*. https://doi.org/10.7554/eLife.10288.

35. Aubol, B.E., Plocinik, R.M., Keshwani, M.M., McGlone, M.L., Hagopian, J.C., Ghosh, G., Fu, X.-D., and Adams, J.A. (2014). N-terminus of the protein kinase CLK1 induces SR protein hyperphosphorylation. Biochem. J *462*, 143–152.

36. Laurence, D.J., and Neville, A.M. (1972). Foetal antigens and their role in the diagnosis and clinical management of human neoplasms: a review. Br. J. Cancer *26*, 335–355.

37. Hillman, R.T., Green, R.E., and Brenner, S.E. (2004). An unappreciated role for RNA surveillance. Genome Biol. *5*, R8.

38. Uzor, S., Zorzou, P., Bowler, E., Porazinski, S., Wilson, I., and Ladomery, M. (2018). Autoregulation of the human splice factor kinase CLK1 through exon skipping and intron retention. Gene *670*, 46–54.

39. Duncan, P.I., Stojdl, D.F., Marius, R.M., and Bell, J.C. (1997). In vivo regulation of alternative pre-mRNA splicing by the Clk1 protein kinase. Mol. Cell. Biol. *17*, 5996–6001.

40. Sun, C.X., Daniel, P., Bradshaw, G., Shi, H., Loi, M., Chew, N., Parackal, S., Tsui, V., Liang, Y., Koptyra, M., et al. (2023). Generation and multi-dimensional profiling of a childhood cancer cell line atlas defines new therapeutic opportunities. Cancer Cell *41*, 660–677.e7.

41. Dehm, S.M., and Bonham, K. (2004). SRC gene expression in human cancer: the role of transcriptional activation. Biochem. Cell Biol. *82*, 263–274.

42. Georgescu, M.-M., Islam, M.Z., Li, Y., Circu, M.L., Traylor, J., Notarianni, C.M., Kline, C.N., and Burns, D.K. (2020). Global activation of oncogenic pathways underlies therapy resistance in diffuse midline glioma. Acta Neuropathol Commun *8*, 111.

43. Zhang, J., and Manley, J.L. (2013). Misregulation of pre-mRNA alternative splicing in cancer. Cancer Discov. *3*, 1228–1237.

44. Nomura, S., Yoshitomi, H., Takano, S., Shida, T., Kobayashi, S., Ohtsuka, M., Kimura, F., Shimizu, H., Yoshidome, H., Kato, A., et al. (2008). FGF10/FGFR2 signal induces cell migration and invasion in pancreatic cancer. Br. J. Cancer *99*, 305–313.

45. Lawrence, M.S., Stojanov, P., Mermel, C.H., Robinson, J.T., Garraway, L.A., Golub, T.R., Meyerson, M., Gabriel, S.B., Lander, E.S., and Getz, G. (2014). Discovery and saturation

analysis of cancer genes across 21 tumour types. Nature *505*, 495–501.

46. National Cancer Institute (.gov) (2016). https://www.cancer.gov/research/participate/clinical-trials-search/v?id=NCI-2024-05326&loc=0&q=SM08502&rl=1.

47. Cui, J.J., Jiang, P., Deng, W., Zhai, D., Ling, N., Li, D., Wang, Z., Hu, Y., Darjania, L., Whitten, J., et al. (2024). Abstract 5944: Discovery of BH-30236: A novel macrocyclic CLK inhibitor targeting alternative splicing in cancers. Cancer Res. *84*, 5944–5944.

48. National Cancer Institute (.gov) (2016). https://www.cancer.gov/research/participate/clinical-trials-search/v?id=NCI-2024-06415&loc=0&q=clk&rl=1.

49. National Cancer Institute (.gov) (2016). https://www.cancer.gov/research/participate/clinical-trials-search/v?id=NCI-2023-02412&r=1.

50. Chen, Y., Davidson, N.M., Wan, Y.K., Yao, F., Su, Y., Gamaarachchi, H., Sim, A., Patel, H., Low, H.M., Hendra, C., et al. (2025). A systematic benchmark of Nanopore long-read RNA sequencing for transcript-level analysis in human cell lines. Nat. Methods *22*, 801–812.

Figure Legends

Figure 1: Pediatric brain tumors display heterogeneous global patterns of differential splicing.

(A) Circos plot of CNS tumors used in this study, categorized by histology, tumor location, and reported gender. Non-neoplastic tumors consist of benign tumors and/or cysts. Rare tumors are CNS tumors with a small (<4) number of patient samples. **(B)** Lollipop plot illustrating the total number of recurrent differential splicing events across the cohort, classified by splicing type (SE: single exon, RI: retained intron, A5SS: alternative 5' splice site, A3SS: alternative 3' splice site, top panel created with BioRender.com). **(C)** Barplots of the number of histology-specific recurrent SE splicing events per patient. Histologies are reverse-ordered by the total number of unique events (skipping + inclusion). **(D)** Distribution plots of splicing burden index (SBI) of all splicing events and types by histology. Shaded regions represent high (blue; ≥ Quartile 3) and low (yellow; ≤ Quartile 1) SBI groups.

Figure 2: Splicing-based clustering reveals heterogeneous tumor groups with distinct histologies, pathway activities, and clinical outcomes.

(A) Hierarchical clustering heatmap of PSI values for top 5,000 most variable splice events across tumors. **(B)** Stacked barplot showing histology sample membership fraction for each cluster. **(C)** Stacked barplot showing the contribution of PSI variance for each cell-type composition. **(D)** Forest plot of cox proportional hazards multivariate interaction OS model for clusters relative to Cluster 1, including covariates for tumor resection and age at diagnosis. Black and white diamonds indicate statistically significant and not significant hazard ratios (HRs), respectively, with intervals denoting 95% confidence intervals. Gray diamonds indicate reference levels of factor covariates. **(E)** Forest plot of cox proportional hazards multivariate interaction OS model for clusters, including covariates for tumor resection, age at diagnosis, and cluster SBI. Black and white diamonds indicate statistically significant and not significant hazard ratios (HRs), respectively, with intervals denoting 95% confidence intervals. Gray diamonds indicate reference levels of factor covariates. **(F)** Heatmap of top cancer-related pathways by cluster (GSVA scores represented by blue/orange color).

Figure 3: Splicing state-associated event prioritization highlights *CLK1* exon 4 as a recurrent functional splice event.

(A) Volcano plot illustrating differentially-expressed splicing factor genes with high SBI compared to those with low SBI (NS = not significant, FC = fold change, colored dots represent $\log_2 FC > |.5|$ and/or Benjamini and Hochberg (B-H) adjusted p-value < 0.05) **(B)** Barplot showing significance of canonical trans-acting splicing factors. Fold-change directionality is annotated as + or -. **(C)** Correlation heatmap showing Pearson R between SBI and SF gene expression, stratified by cluster. Stars denote * = FDR < 0.05 and ** = FDR < 1.0e-5. **(D)** Workflow to prioritize candidate differential exon-level splicing events that alter UniProt-defined functional sites, created with BioRender.com. Stacked bar plots represent the fraction of exon inclusion, skipping, or mixed

splicing events categorized by predicted impact. **(E)** Sashimi plot of two representative tumor samples with either high (BS_30VC3R2Q) or low (BS_FYP2B298) *CLK1* exon 4 inclusion. Reads supporting exon 4 skipping are marked with an asterisk (*), and exon 4 is indicated by a hash (#) on the transcript plot. **(F)** Cox proportional hazards model forest plot of OS for cluster membership, including covariates for extent of tumor resection, histology group, *CLK1* exon 4 PSI, and age at diagnosis. Black and white diamonds indicate statistically significant and not significant hazard ratios, respectively, with intervals denoting 95% confidence intervals. Gray diamonds indicate reference levels of factor covariates. **(G)** Cox proportional hazards model forest plot of EFS for Cluster 3 and **(H)** OS for cluster 7 adjusted for exon 4-included *CLK1* transcript abundance, including covariates for extent of tumor resection, histology group, and age at diagnosis. "Collapsed" indicates combined histologies with N < 3 to retain these tumors in the model. Black and white diamonds indicate statistically significant and not significant hazard ratios, respectively, with intervals denoting 95% confidence intervals. Gray diamonds indicate reference levels of factor covariates. **(I)** Kaplan–Meier curve for overall survival (OS) in Cluster 7, stratified by high vs low GSVA score.

Figure 4. *CLK1* exon 4 is required for pediatric brain tumor cell line growth and viability.

(A) Boxplots of CLK1 exon 4 inclusion stratified by age in non-tumor RNA-seq brain tissue data from the Evolutionary Developmental (Evo-Devo) Atlas (N = 59) and GTEx (N = 2,642). **(B)** Boxplot of DepMap dependency scores stratified by high or low *CLK1* exon 4 containing transcript expression in brain tumor cell lines. Wilcoxon p-value shown. **(C)** Dot plot of Gene Ontology (GO) gene-set enrichment analysis for genes correlated with *CLK1* across CNS cell lines from DepMap. Dot size indicates gene count size, and color represents enrichment significance. **(D)** Ranked dotplot of DepMap dependency scores in brain tumor cell lines with pediatric line KNS-42 highlighted in black. **(E)** Table showing percent spliced in for all *CLK1* exon 4 transcripts in patient-derived cell lines (7316-1763 and 7316-1769 from the CBTN) and KNS-42 (commercial) derived

using short RNA-Seq or percent *CLK1* exon 4 transcripts using either long-read (ONT) sequencing. **(F)** Proliferation of KNS-42 cells treated with increasing concentrations of pan-DYRK/CLK1 inhibitor Cirtuvivint over six days. **(G)** Day 3 cell viability of KNS-42 cells treated with increasing concentrations of Cirtuvivint. Stars denote Bonferroni-adjusted p-values following pairwise Student's *t-tests*. **(H)** Barplot showing the RNA expression fold-change in cells treated with control morpholino or morpholino targeting the *CLK1* exon 3-4 junction or exon 3-5 junction. **(I)** Western blot of CLK1 with increasing morpholino treatment of 1, 5, and 10 μM. **(J)** Cell viability of cells treated with *CLK1* exon 4 morpholino or non-targeting morpholino. Stars denote within-time paired Student's *t-tests*.

Figure 5. Perturbing *CLK1* disrupts RNA splicing and impairs oncogenic transcriptional programs.

(A) Volcano plot illustrating genes differentially-expressed in KNS-42 cells treated with *CLK1* exon 4 targeting morpholino compared to cells treated with non-targeting morpholino. **(B)** Stacked barplot showing differential splicing in KNS-42 cells treated with *CLK1* exon 4 targeting morpholino compared to cells treated with non-targeting morpholino categorized by splicing type (SE, MXE, A3SS, A5SS, and RI). **(C)** Barplots displaying number of differentially expressed (DE) genes or differentially spliced (DS) genes affecting functional sites categorized by gene family.**(D)** Venn diagram depicting overlap of all DS and DE genes **(E)** Over-representation analysis of HALLMARK and KEGG pathways for DS cancer genes. Dot size indicates gene count size, and color represents enrichment significance. **(F)** Heatmap displaying single-sample HALLMARK GSVA scores for DS genes affecting functional sites in cells treated with *CLK1* exon 4 morpholino or non-targeting morpholino. **(G)** Barplots illustrate the mean GSVA score difference by treatment (n = 3 replicates per treatment). **(H)** Venn diagram depicting overlap of DS and DE genes and significant (Wald FDR < 0.05, z-score < -1.5) dependency genes (DG) identified in matched CBTN HGG cell lines through CRISPR dependency experiments from the Childhood Cancer Model Atlas

(CCMA v3). **(I)** Ranked dotplot of significant CRISPR gene dependency mean z-scores for pediatric HGG cell lines with *CLK1* expression and splicing-based target genes highlighted in red and blue, respectively.

Figure 1

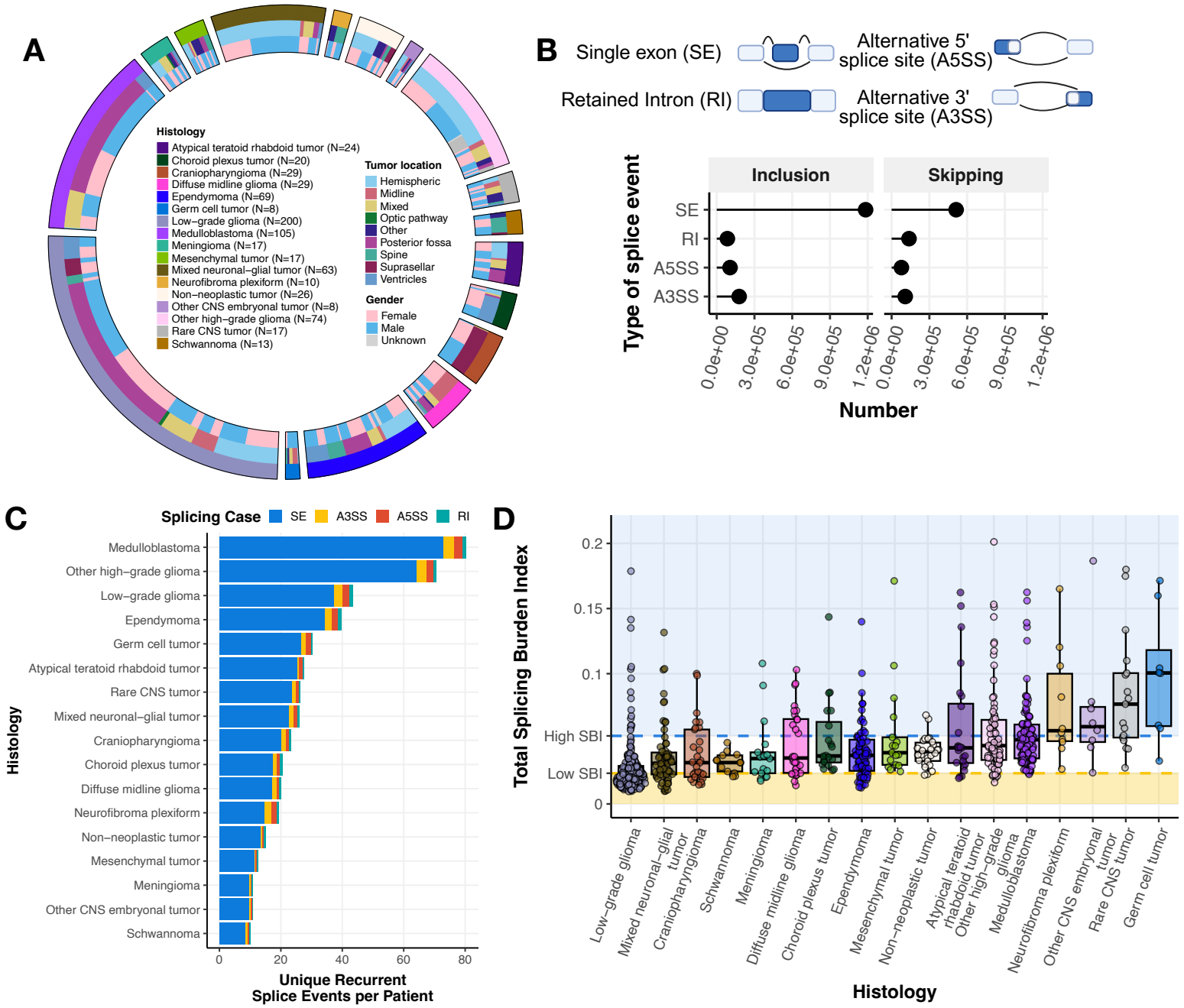

A
Histology
Atypical teratoid rhabdoid tumor (N=24)
Choroid plexus tumor (N=20)
Craniopharyngioma (N=29)
Diffuse midline glioma (N=29)
Ependymoma (N=69)
Germ cell tumor (N=8)
Low−grade glioma (N=200)
Medulloblastoma (N=105)
Meningioma (N=17)
Mesenchymal tumor (N=17)
Mixed neuronal−glial tumor (N=63)
Neurofibroma plexiform (N=10)
Non−neoplastic tumor (N=26)
Other CNS embryonal tumor (N=8)
Other high−grade glioma (N=74)
Rare CNS tumor (N=17)
Schwannoma (N=13)
Tumor location
Hemispheric
Midline
Mixed
Optic pathway
Other
Posterior fossa
Spine
Suprasellar
Ventricles
Gender
Female
Male
Unknown
B
Single exon (SE)
Alternative 5' splice site (A5SS)
Retained Intron (RI)
Alternative 3' splice site (A3SS)
Inclusion
Skipping
Type of splice event
SE
RI
A5SS
A3SS
0.0e+00
3.0e+05
6.0e+05
9.0e+05
1.2e+06
Number
C
Splicing Case
SE
A3SS
A5SS
RI
Histology
Medulloblastoma
Other high−grade glioma
Low−grade glioma
Ependymoma
Germ cell tumor
Atypical teratoid rhabdoid tumor
Rare CNS tumor
Mixed neuronal−glial tumor
Craniopharyngioma
Choroid plexus tumor
Diffuse midline glioma
Neurofibroma plexiform
Non−neoplastic tumor
Mesenchymal tumor
Meningioma
Other CNS embryonal tumor
Schwannoma
0
20
40
60
80
Unique Recurrent Splice Events per Patient
D
Total Splicing Burden Index
0.2
0.15
0.1
High SBI
Low SBI
0
Low−grade glioma
Mixed neuronal−glial tumor
Craniopharyngioma
Schwannoma
Meningioma
Diffuse midline glioma
Choroid plexus tumor
Ependymoma
Mesenchymal tumor
Non−neoplastic tumor
Atypical teratoid rhabdoid tumor
Other high−grade glioma
Medulloblastoma
Neurofibroma plexiform
Other CNS embryonal tumor
Rare CNS tumor
Germ cell tumor
Histology

Figure 2

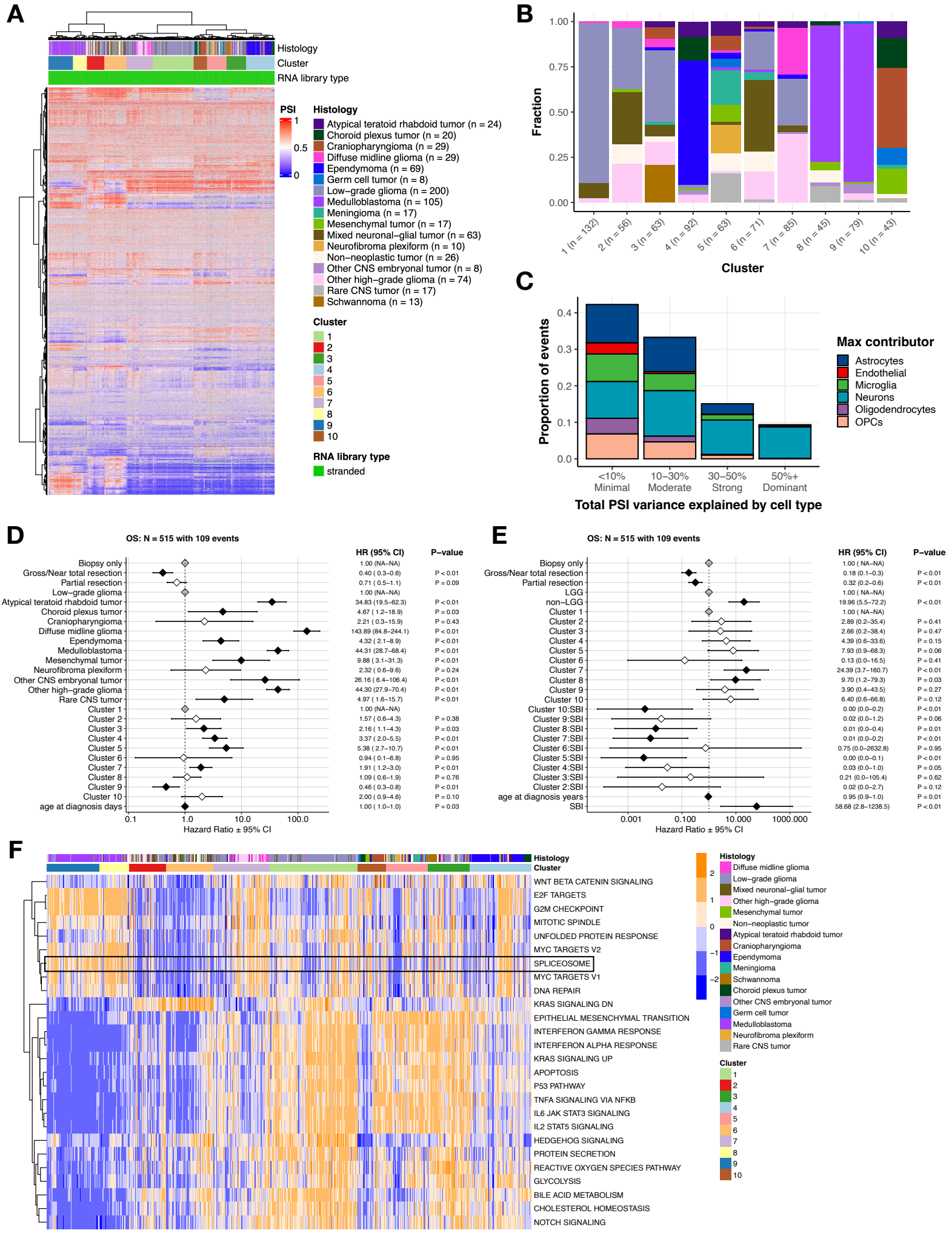

A
Histology
Cluster
RNA library type
PSI
1
0.5
0
Histology
Atypical teratoid rhabdoid tumor (n = 24)
Choroid plexus tumor (n = 20)
Craniopharyngioma (n = 29)
Diffuse midline glioma (n = 29)
Ependymoma (n = 69)
Germ cell tumor (n = 8)
Low−grade glioma (n = 200)
Medulloblastoma (n = 105)
Meningioma (n = 17)
Mesenchymal tumor (n = 17)
Mixed neuronal−glial tumor (n = 63)
Neurofibroma plexiform (n = 10)
Non−neoplastic tumor (n = 26)
Other CNS embryonal tumor (n = 8)
Other high−grade glioma (n = 74)
Rare CNS tumor (n = 17)
Schwannoma (n = 13)
Cluster
1
2
3
4
5
6
7
8
9
10
RNA library type
stranded
B
Fraction
1.00
0.75
0.50
0.25
0.00
1 (n = 132)
2 (n = 56)
3 (n = 63)
4 (n = 92)
5 (n = 63)
6 (n = 71)
7 (n = 85)
8 (n = 45)
9 (n = 79)
10 (n = 43)
Cluster
C
Proportion of events
0.4
0.3
0.2
0.1
0.0
<10% Minimal
10−30% Moderate
30−50% Strong
50%+ Dominant
Total PSI variance explained by cell type
Max contributor
Astrocytes
Endothelial
Microglia
Neurons
Oligodendrocytes
OPCs
D
OS: N = 515 with 109 events
HR (95% CI)
P−value
Biopsy only 1.00 (NA−NA)
Gross/Near total resection 0.40 (0.3−0.6) P < 0.01
Partial resection 0.71 (0.5−1.1) P = 0.09
Low−grade glioma 1.00 (NA−NA)
Atypical teratoid rhabdoid tumor 34.83 (19.5−62.3) P < 0.01
Choroid plexus tumor 4.67 (1.2−18.9) P = 0.03
Craniopharyngioma 2.21 (0.3−15.9) P = 0.43
Diffuse midline glioma 143.89 (84.8−244.1) P < 0.01
Ependymoma 4.32 (2.1−8.9) P < 0.01
Medulloblastoma 44.31 (28.7−68.4) P < 0.01
Mesenchymal tumor 9.88 (3.1−31.3) P < 0.01
Neurofibroma plexiform 2.32 (0.6−9.6) P = 0.24
Other CNS embryonal tumor 26.16 (6.4−106.4) P < 0.01
Other high−grade glioma 44.30 (27.9−70.4) P < 0.01
Rare CNS tumor 4.97 (1.6−15.7) P < 0.01
Cluster 1 1.00 (NA−NA)
Cluster 2 1.57 (0.6−4.3) P = 0.38
Cluster 3 2.16 (1.1−4.3) P = 0.03
Cluster 4 3.37 (2.0−5.5) P < 0.01
Cluster 5 5.38 (2.7−10.7) P < 0.01
Cluster 6 0.94 (0.1−6.8) P = 0.95
Cluster 7 1.91 (1.2−3.0) P < 0.01
Cluster 8 1.09 (0.6−1.9) P = 0.76
Cluster 9 0.46 (0.3−0.8) P < 0.01
Cluster 10 2.00 (0.9−4.6) P = 0.10
age at diagnosis days 1.00 (1.0−1.0) P = 0.03
0.1
1.0
10.0
100.0
Hazard Ratio ± 95% CI
E
OS: N = 515 with 109 events
HR (95% CI)
P−value
Biopsy only 1.00 (NA−NA)
Gross/Near total resection 0.18 (0.1−0.3) P < 0.01
Partial resection 0.32 (0.2−0.6) P < 0.01
LGG 1.00 (NA−NA)
non−LGG 19.96 (5.5−72.2) P < 0.01
Cluster 1 1.00 (NA−NA)
Cluster 2 2.89 (0.2−35.4) P = 0.41
Cluster 3 2.66 (0.2−38.4) P = 0.47
Cluster 4 4.39 (0.6−33.6) P = 0.15
Cluster 5 7.93 (0.9−68.3) P = 0.06
Cluster 6 0.13 (0.0−16.5) P = 0.41
Cluster 7 24.39 (3.7−160.7) P < 0.01
Cluster 8 9.70 (1.2−79.3) P = 0.03
Cluster 9 3.90 (0.4−43.5) P = 0.27
Cluster 10 6.40 (0.6−66.8) P = 0.12
Cluster 10:SBI 0.00 (0.0−0.2) P < 0.01
Cluster 9:SBI 0.02 (0.0−1.2) P = 0.06
Cluster 8:SBI 0.01 (0.0−0.4) P = 0.01
Cluster 7:SBI 0.01 (0.0−0.2) P < 0.01
Cluster 6:SBI 0.75 (0.0−2632.8) P = 0.95
Cluster 5:SBI 0.00 (0.0−0.1) P < 0.01
Cluster 4:SBI 0.03 (0.0−1.0) P = 0.05
Cluster 3:SBI 0.21 (0.0−105.4) P = 0.62
Cluster 2:SBI 0.02 (0.0−2.7) P = 0.12
age at diagnosis years 0.95 (0.9−1.0) P = 0.01
SBI 58.68 (2.8−1238.5) P < 0.01
0.001
0.100
10.000
1000.000
Hazard Ratio ± 95% CI
F
Histology
Cluster
WNT BETA CATENIN SIGNALING
E2F TARGETS
G2M CHECKPOINT
MITOTIC SPINDLE
UNFOLDED PROTEIN RESPONSE
MYC TARGETS V2
SPLICEOSOME
MYC TARGETS V1
DNA REPAIR
KRAS SIGNALING DN
EPITHELIAL MESENCHYMAL TRANSITION
INTERFERON GAMMA RESPONSE
INTERFERON ALPHA RESPONSE
KRAS SIGNALING UP
APOPTOSIS
P53 PATHWAY
TNFA SIGNALING VIA NFKB
IL6 JAK STAT3 SIGNALING
IL2 STAT5 SIGNALING
HEDGEHOG SIGNALING
PROTEIN SECRETION
REACTIVE OXYGEN SPECIES PATHWAY
GLYCOLYSIS
BILE ACID METABOLISM
CHOLESTEROL HOMEOSTASIS
NOTCH SIGNALING
2
1
0
−1
−2
Histology
Diffuse midline glioma
Low−grade glioma
Mixed neuronal−glial tumor
Other high−grade glioma
Mesenchymal tumor
Non−neoplastic tumor
Atypical teratoid rhabdoid tumor
Craniopharyngioma
Ependymoma
Meningioma
Schwannoma
Choroid plexus tumor
Other CNS embryonal tumor
Germ cell tumor
Medulloblastoma
Neurofibroma plexiform
Rare CNS tumor
Cluster
1
2
3
4
5
6
7
8
9
10

Figure 3

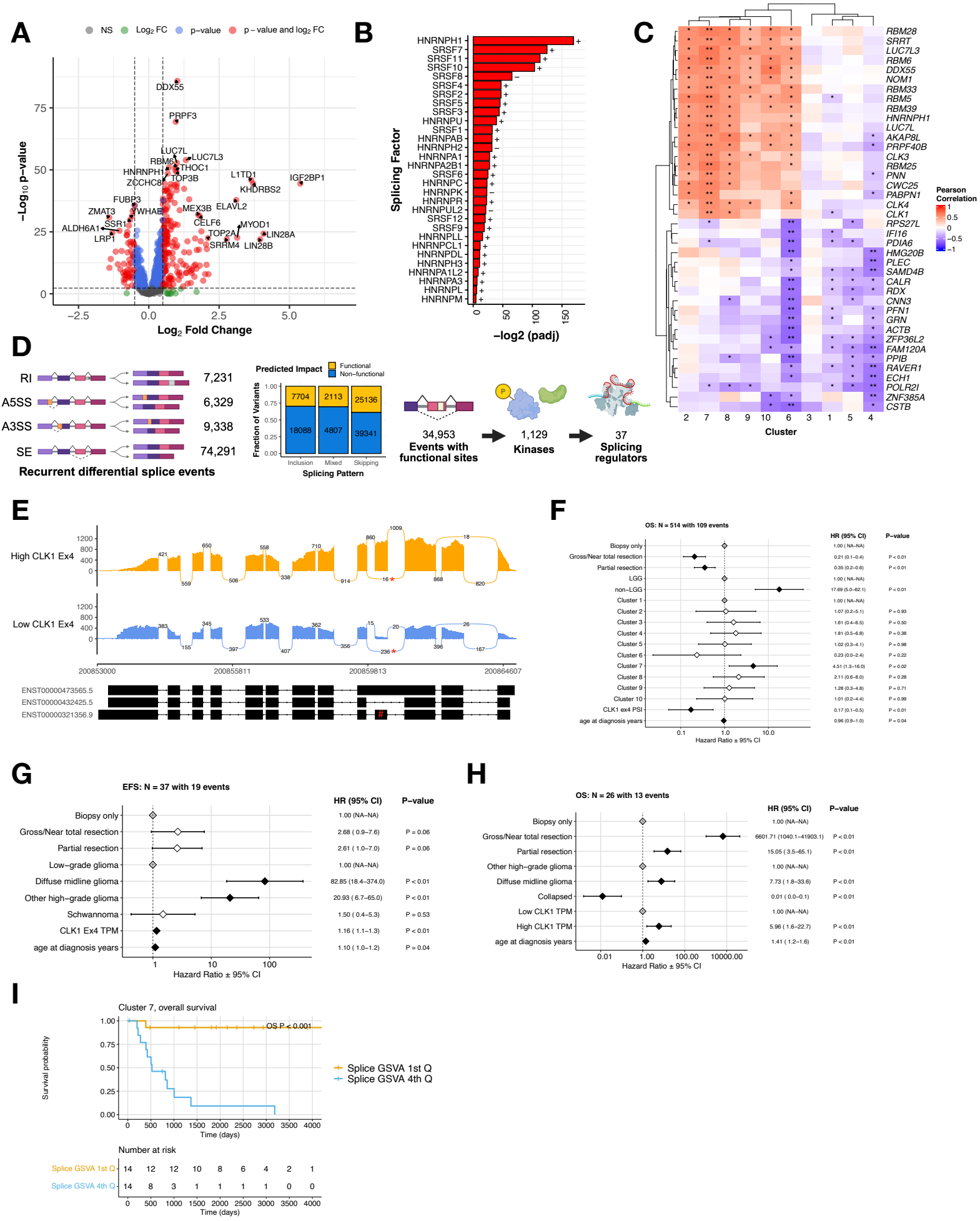
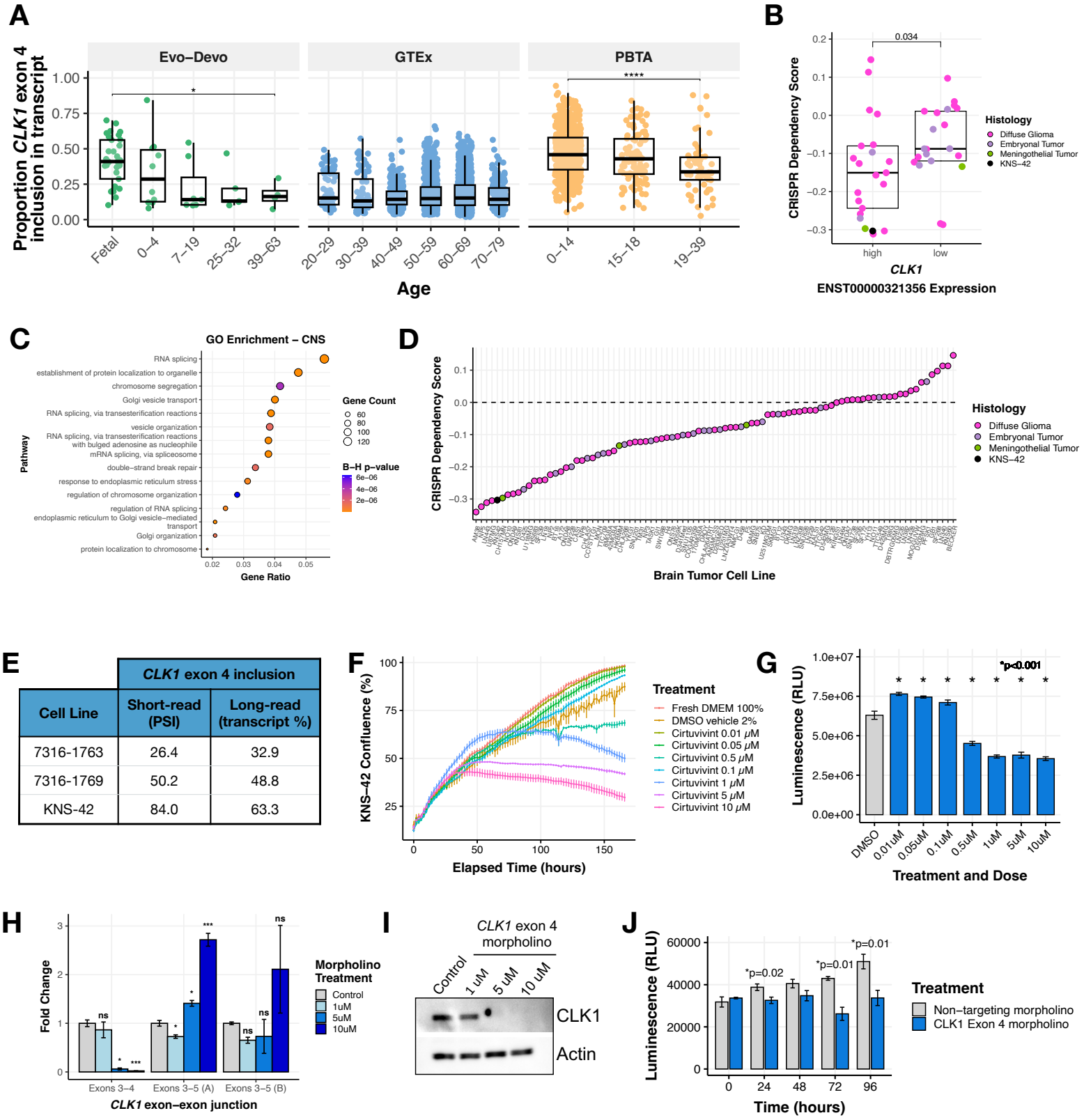
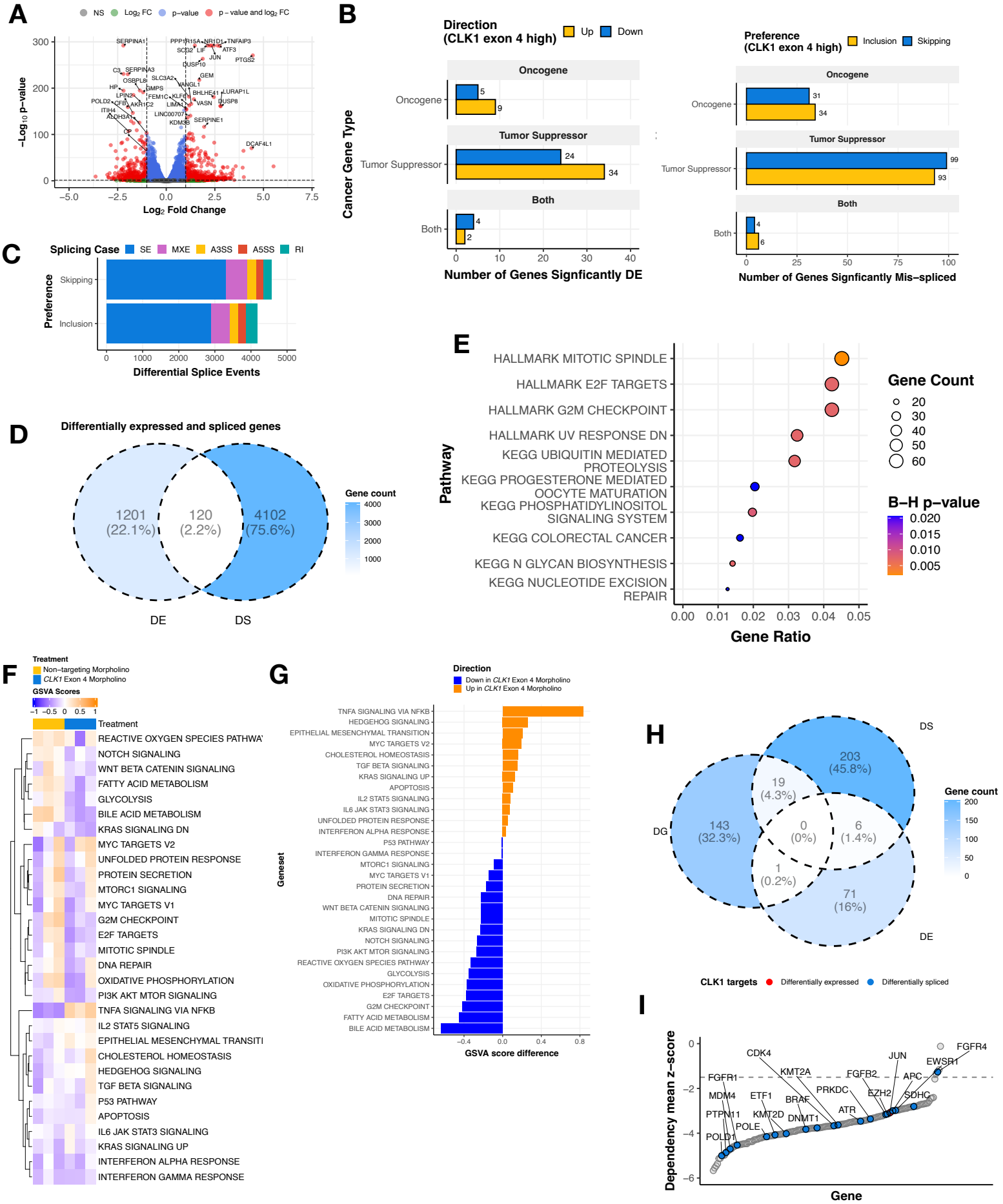

A
NS
Log2 FC
p-value
p - value and log2 FC
−Log10 p-value
Log2 Fold Change
DDX55
PRPF3
LUC7L
LUC7L3
RBM6
THOC1
HNRNPH1
TOP3B
ZCCHC8
L1TD1
IGF2BP1
KHDRBS2
FUBP3
ZMAT3
YWHAE
MEX3B
ELAVL2
SSR1
CELF6
ALDH6A1
MYOD1
LRP1
TOP2A
LIN28A
SRRM4
LIN28B
B
Splicing Factor
−log2 (padj)
HNRNPH1
SRSF7
SRSF11
SRSF10
SRSF8
SRSF4
SRSF2
SRSF5
SRSF3
HNRNPU
SRSF1
HNRNPAB
HNRNPH2
HNRNPA1
HNRNPA2B1
SRSF6
HNRNPC
HNRNPK
HNRNPR
HNRNPUL2
SRSF12
SRSF9
HNRNPLL
HNRNPCL1
HNRNPDL
HNRNPH3
HNRNPA1L2
HNRNPA3
HNRNPL
HNRNPM
C
Pearson Correlation
Cluster
RBM28
SRRT
LUC7L3
RBM6
DDX55
NOM1
RBM33
RBM5
RBM39
HNRNPH1
LUC7L
AKAP8L
PRPF40B
CLK3
RBM25
PNN
CWC25
PABPN1
CLK4
CLK1
RPS27L
IFI16
PDIA6
HMG20B
PLEC
SAMD4B
CALR
RDX
CNN3
PFN1
GRN
ACTB
ZFP36L2
FAM120A
PPIB
RAVER1
ECH1
POLR2I
ZNF385A
CSTB
D
RI
A5SS
A3SS
SE
7,231
6,329
9,338
74,291
Recurrent differential splice events
Predicted Impact
Functional
Non−functional
Fraction of Variants
Inclusion
Mixed
Skipping
Splicing Pattern
34,953
Events with functional sites
1,129
Kinases
37
Splicing regulators
E
High CLK1 Ex4
Low CLK1 Ex4
ENST00000473565.5
ENST00000432425.5
ENST00000321356.9
F
OS: N = 514 with 109 events
HR (95% CI)
P−value
Biopsy only
Gross/Near total resection
Partial resection
LGG
non−LGG
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Cluster 10
CLK1 ex4 PSI
age at diagnosis years
Hazard Ratio ± 95% CI
G
EFS: N = 37 with 19 events
HR (95% CI)
P−value
Biopsy only
Gross/Near total resection
Partial resection
Low−grade glioma
Diffuse midline glioma
Other high−grade glioma
Schwannoma
CLK1 Ex4 TPM
age at diagnosis years
Hazard Ratio ± 95% CI
H
OS: N = 26 with 13 events
HR (95% CI)
P−value
Biopsy only
Gross/Near total resection
Partial resection
Other high−grade glioma
Diffuse midline glioma
Collapsed
Low CLK1 TPM
High CLK1 TPM
age at diagnosis years
Hazard Ratio ± 95% CI
I
Cluster 7, overall survival
OS P < 0.001
Survival probability
Time (days)
Splice GSVA 1st Q
Splice GSVA 4th Q
Number at risk

Figure 4



Cell Line	*CLK1* exon 4 inclusion	
	Short-read (PSI)	Long-read (transcript %)
7316-1763	26.4	32.9
7316-1769	50.2	48.8
KNS-42	84.0	63.3

Figure 5


A
NS
Log2 FC
p–value
p – value and log2 FC
−Log10 p–value
Log2 Fold Change
B
Direction (CLK1 exon 4 high)
Up
Down
Cancer Gene Type
Oncogene
Tumor Suppressor
Both
Number of Genes Signficantly DE
Preference (CLK1 exon 4 high)
Inclusion
Skipping
Number of Genes Signficantly Mis–spliced
C
Splicing Case
SE
MXE
A3SS
A5SS
RI
Preference
Skipping
Inclusion
Differential Splice Events
D
Differentially expressed and spliced genes
1201 (22.1%)
120 (2.2%)
4102 (75.6%)
Gene count
DE
DS
E
HALLMARK MITOTIC SPINDLE
HALLMARK E2F TARGETS
HALLMARK G2M CHECKPOINT
HALLMARK UV RESPONSE DN
KEGG UBIQUITIN MEDIATED PROTEOLYSIS
KEGG PROGESTERONE MEDIATED OOCYTE MATURATION
KEGG PHOSPHATIDYLINOSITOL SIGNALING SYSTEM
KEGG COLORECTAL CANCER
KEGG N GLYCAN BIOSYNTHESIS
KEGG NUCLEOTIDE EXCISION REPAIR
Pathway
Gene Ratio
Gene Count
B–H p–value
F
Treatment
Non–targeting Morpholino
CLK1 Exon 4 Morpholino
GSVA Scores
REACTIVE OXYGEN SPECIES PATHWAY
NOTCH SIGNALING
WNT BETA CATENIN SIGNALING
FATTY ACID METABOLISM
GLYCOLYSIS
BILE ACID METABOLISM
KRAS SIGNALING DN
MYC TARGETS V2
UNFOLDED PROTEIN RESPONSE
PROTEIN SECRETION
MTORC1 SIGNALING
MYC TARGETS V1
G2M CHECKPOINT
E2F TARGETS
MITOTIC SPINDLE
DNA REPAIR
OXIDATIVE PHOSPHORYLATION
PI3K AKT MTOR SIGNALING
TNFA SIGNALING VIA NFKB
IL2 STAT5 SIGNALING
EPITHELIAL MESENCHYMAL TRANSITI
CHOLESTEROL HOMEOSTASIS
HEDGEHOG SIGNALING
TGF BETA SIGNALING
P53 PATHWAY
APOPTOSIS
IL6 JAK STAT3 SIGNALING
KRAS SIGNALING UP
INTERFERON ALPHA RESPONSE
INTERFERON GAMMA RESPONSE
G
Direction
Down in CLK1 Exon 4 Morpholino
Up in CLK1 Exon 4 Morpholino
Geneset
GSVA score difference
H
DS
DG
DE
203 (45.8%)
19 (4.3%)
143 (32.3%)
0 (0%)
6 (1.4%)
1 (0.2%)
71 (16%)
Gene count
CLK1 targets
Differentially expressed
Differentially spliced
I
Dependency mean z–score
Gene
FGFR4
EWSR1
CDK4
JUN
KMT2A
FGFR2
APC
FGFR1
PRKDC
EZH2
SDHC
MDM4
ETF1
BRAF
ATR
PTPN11
KMT2D
DNMT1
POLE
POLD1