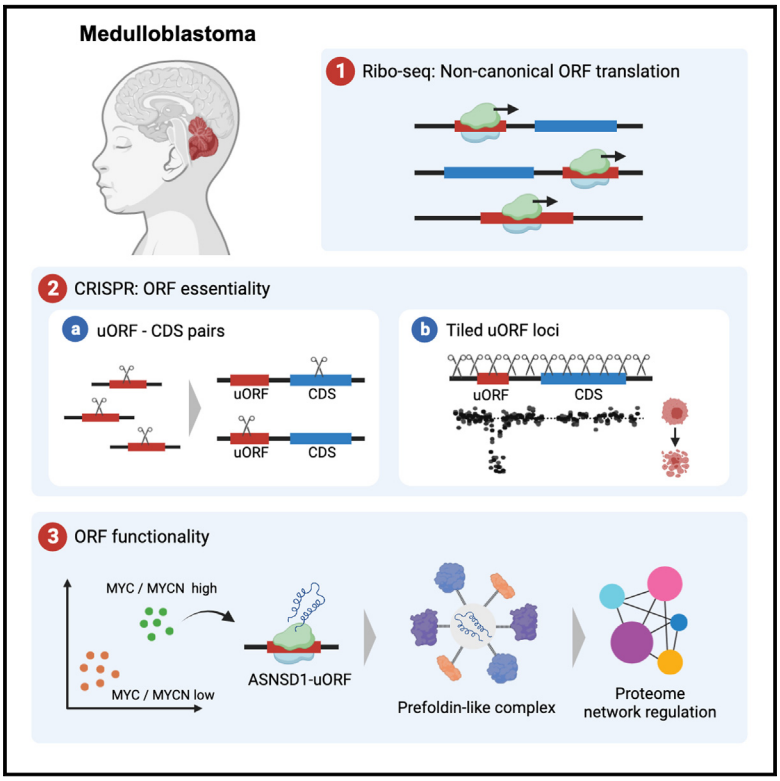
Molecular Cell

Article

Translation of non-canonical open reading frames as a cancer cell survival mechanism in childhood medulloblastoma

Graphical abstract



Highlights

- Ribo-seq reveals widespread translation of non-canonical ORFs in medulloblastoma
- High-resolution CRISPR tiling reveals uORF functions in medulloblastoma
- The ASNSD1-uORF microprotein controls downstream pathways with the prefoldin-like complex
- The ASNSD1-uORF microprotein is necessary for medulloblastoma cell survival

Authors

Damon A. Hofman, Jorge Ruiz-Orera, Ian Yannuzzi, ..., Pratiti Bandopadhayay, Sebastiaan van Heesch, John R. Prensner

Correspondence

s.vanheesch@prinsesmaximacentrum.nl (S.v.H.), prensner@umich.edu (J.R.P.)

In brief

Hofman et al. reveal a novel survival mechanism in high-risk childhood medulloblastoma through the translation of non-canonical open reading frames (ORFs). The study uncovers the critical role of these often-overlooked ORFs, particularly the ASNSD1-uORF, in medulloblastoma cell survival, offering new avenues for targeted therapies in pediatric cancer.

Hofman et al., 2024, Molecular Cell *84*, 261–276
January 18, 2024
https://doi.org/10.1016/j.molcel.2023.12.003

CellPress

Molecular Cell

CellPress
OPEN ACCESS

Article

Translation of non-canonical open reading frames as a cancer cell survival mechanism in childhood medulloblastoma

Damon A. Hofman,[1,15] Jorge Ruiz-Orera,[2,15] Ian Yannuzzi,[3] Rakesh Murugesan,[3] Adam Brown,[3,13] Karl R. Clauser,[3] Alexandra L. Condurat,[3,4] Jip T. van Dinter,[1] Sem A.G. Engels,[1] Amy Goodale,[3] Jasper van der Lugt,[1] Tanaz Abid,[3] Li Wang,[3] Kevin N. Zhou,[4,14] Jayne Vogelzang,[5,6] Keith L. Ligon,[5,6,7] Timothy N. Phoenix,[8] Jennifer A. Roth,[3] David E. Root,[3] Norbert Hubner,[2,9,10] Todd R. Golub,[3,4,11] Pratiti Bandopadhayay,[3,4,11] Sebastiaan van Heesch,[1,*] and John R. Prensner[12,16,*]

[1]Princess Máxima Center for Pediatric Oncology, Heidelberglaan 25, 3584 CS Utrecht, the Netherlands
[2]Cardiovascular and Metabolic Sciences, Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC), 13125 Berlin, Germany
[3]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
[4]Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, MA 02215, USA
[5]Department of Pathology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02215, USA
[6]Department of Pathology, Brigham and Women's Hospital, Boston, MA 02215, USA
[7]Department of Pathology, Boston Children's Hospital, Boston MA 02115, USA
[8]Division of Pharmaceutical Sciences, James L. Winkle College of Pharmacy, University of Cincinnati, Cincinnati, OH 45229, USA
[9]Charité-Universitätsmedizin, 10117 Berlin, Germany
[10]German Centre for Cardiovascular Research, Partner Site Berlin, 13347 Berlin, Germany
[11]Division of Pediatric Hematology/Oncology, Boston Children's Hospital, Boston, MA 02115, USA
[12]Department of Pediatrics, Division of Pediatric Hematology/Oncology and Biological Chemistry, University of Michigan Medical School, Ann Arbor, MI 48109, USA
[13]Present address: Arbor Biotechnologies, Cambridge, MA 02140, USA
[14]Present address: Kaiser Permanente Bernard J. Tyson School of Medicine, Pasadena, CA 91101, USA
[15]These authors contributed equally
[16]Lead contact
*Correspondence: s.vanheesch@prinsesmaximacentrum.nl (S.v.H.), prensner@umich.edu (J.R.P.)
https://doi.org/10.1016/j.molcel.2023.12.003

SUMMARY

A hallmark of high-risk childhood medulloblastoma is the dysregulation of RNA translation. Currently, it is unknown whether medulloblastoma dysregulates the translation of putatively oncogenic non-canonical open reading frames (ORFs). To address this question, we performed ribosome profiling of 32 medulloblastoma tissues and cell lines and observed widespread non-canonical ORF translation. We then developed a stepwise approach using multiple CRISPR-Cas9 screens to elucidate non-canonical ORFs and putative microproteins implicated in medulloblastoma cell survival. We determined that multiple lncRNA-ORFs and upstream ORFs (uORFs) exhibited selective functionality independent of main coding sequences. A microprotein encoded by one of these ORFs, ASNSD1-uORF or ASDURF, was upregulated, associated with *MYC*-family oncogenes, and promoted medulloblastoma cell survival through engagement with the prefoldin-like chaperone complex. Our findings underscore the fundamental importance of non-canonical ORF translation in medulloblastoma and provide a rationale to include these ORFs in future studies seeking to define new cancer targets.

INTRODUCTION

High-risk medulloblastoma remains one of the most recalcitrant pediatric cancers, and children with *MYC*-amplified disease frequently succumb to relapsed disease.[1–4] Besides *MYC* amplification, in-depth analyses of the medulloblastoma coding genome have identified and characterized additional somatic events in subsets of patients. Still, most tumors lack targetable mutations and do not yield insights regarding their aggressive behavior.[5–7] At the same time, medulloblastoma is known to exhibit extensive rewiring of RNA translational control[8,9] both through genetic mutation of the *DDX3X* RNA helicase in the WNT and SHH subtypes,[7,10–12] as well as in group 3/4 tumors through activation of the *MYCN* or *MYC* transcription factors,

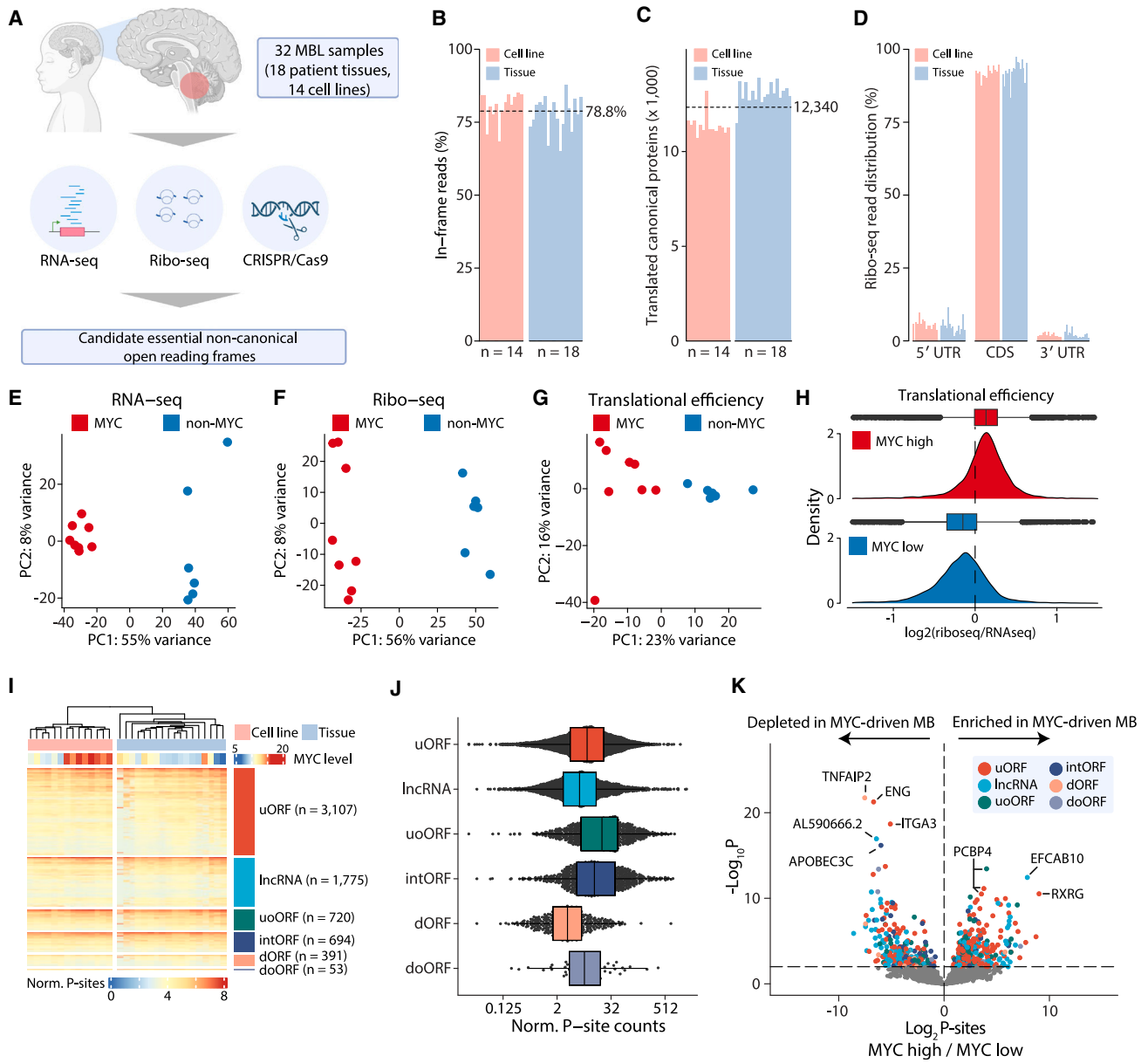


Figure 1. Comprehensive profiling of non-canonical ORF translation in medulloblastoma

(A) Schematic depiction of experimental approach.

(B) Bar plot showing the percentage of in-frame Ribo-seq reads across all 14 cell line samples and 18 tissue samples.

(C) Bar plot showing the number of translated canonical coding sequences (defined as P-sites per million > 1) across all samples.

(D) Bar plots showing percentages of reads mapping to coding sequences (CDSs) and untranslated regions (5′ UTR and 3′ UTR) of protein coding sequences across all samples.

(E) A principal component analysis (PCA) showing MYC-driven and non-MYC-driven cell lines using RNA-seq data.

(F) A PCA showing MYC-driven and non-MYC-driven cell lines using Ribo-seq data.

(G) A PCA separating MYC-driven from non-MYC-driven cell lines using translation efficiency values. Each dot represents one sample.

(H) A density plot showing the distribution of translational efficiency values for each gene in MYC-driven and non-MYC-driven medulloblastoma cell line subgroups. Boxplots show lower quartile, median, and upper quartile values, with whiskers extending to highest and lowest observations.

(I) Heatmap showing translation levels of translated non-canonical ORFs (rows) across all samples (columns). Rows and columns were clustered in an unsupervised manner within sample type (tissue and cell line) and ORF biotype groups. Samples are annotated by MYC translation levels. Translation levels are calculated as transformed normalized P-site counts. uORF, upstream open reading frame; uoORF, upstream overlapping open reading frame; intORF, internal open reading frame; dORF, downstream open reading frame; doORF, downstream overlapping open reading frame; lncRNA, long non-coding RNA open reading frame.

(legend continued on next page)

where recent genetic evidence indicates that control of RNA translation may be the most critical aspect of MYC function during tumorigenesis.[13–15] This deregulation of RNA translational control in medulloblastoma leads not only to a wide discrepancy between RNA and proteomic signatures[16,17] but also to a distinctive reliance on RNA translation factors[18] and potential therapeutic options.[19,20]

While translation of known proteins has been the focal point for prior research in medulloblastoma, as well as other childhood brain cancers, the human genome also contains thousands of non-canonical open reading frames (ORFs).[21] These previously understudied ORFs are ubiquitous regions of ribosome translation that occur separately from the known protein-coding sequences and have the capacity to influence gene activity or to encode proteins with distinct biological functions.[22–25] For example, individual cancer-associated ORFs may generate novel cancer targets that influence cell phenotypes,[26,27] whereas other classes of ORFs are critical effectors of oncogene-induced gene regulation.[28] However, the overall potential impact of such ORFs across and within cancers has not been determined.

Here, we have investigated the functional impact of translation of non-canonical ORFs in medulloblastoma. We demonstrate that these ORFs are commonly translated in medulloblastoma model systems and patient tumors, with translational control influenced by disease subtype. Using genome-wide CRISPR screens and ORF-specific saturation mutagenesis with CRISPR, we found that non-canonical ORFs are frequently essential for cell survival in medulloblastoma and describe widespread reliance on upstream ORFs (uORFs) in particular. From these, we identify a uORF in the *ASNSD1* gene that produces a microprotein, which is selectively upregulated and required for maintenance of cell survival by coordinating the function of the prefoldin-like complex, a poorly understood complex implicated in post-translational control.[29–31] Together, our findings demonstrate that oncogenic uORFs can act as critical disease mediators both in medulloblastoma and, by extension, human cancers more broadly.

RESULTS

Comprehensive translational profiling of medulloblastoma highlights biological subtypes

To characterize signatures of RNA transcription and translation in medulloblastoma, we profiled 32 unique patients/cell lines (14 medulloblastoma cell lines and 18 tumor samples; see STAR Methods) using RNA-seq and ribosome profiling[32] (Figure 1A; Tables S1A–S1F). Samples reflected major histological and molecular subtypes, including large cell/anaplastic and desmoplastic nodular, and MYC-driven subtypes (Table S1A). In total, we sequenced and mapped over 1.3 billion ribosome footprints across 32 samples (Table S1A; Figures S1A–S1C). For this, we further optimized the Ribo-seq procedure to capture high-quality ribosome footprints from low-input tumor samples down to 3 mg per sample (range: 3–75 mg). Ribosome profiling achieved an average of 78.8% in-frame reads (range 64.7%–84.8%) with an average of 12,340 translated known protein-coding sequences (CDSs) quantified per sample (range 10,712–13,868 CDSs) (Figures 1B–1D, S1D, and S1E). Tissue samples and cell lines exhibited similar performance metrics, with tumor samples yielding a higher number and thus greater diversity of detected CDSs (Figure 1C).

Clustering of cell lines by mRNA expression levels, as well as ribosome profiling, demonstrated distinct biological signatures between MYC-driven and non-MYC-driven cell lines (Figures 1E and 1F). Given prior proteogenomic data demonstrating discrepant RNA and protein signatures in medulloblastoma,[16,17] we next determined mRNA translational efficiency scores by comparing ribosome profiling and RNA-seq data (see STAR Methods and Table S1G) and observed clustering of MYC-driven compared with non-MYC cell lines, indicative of stark differences in translational control between medulloblastoma subtypes driven by MYC activity (Figure 1G). Indeed, compared with non-MYC-driven cells, MYC-driven cell lines exhibited a significantly increased mRNA translational efficiency overall (Figure 1H; Wilcoxon test; $p < 2.2 \times 10^{-16}$). We additionally confirmed differential translation signatures associated with MYC expression levels in patient samples, consistent with our observations in cell lines (Figures S1I–S1M and STAR Methods). Consistent with these results, Gene Ontology and gene set enrichment analyses highlighted pathways related to ribosome biogenesis, translation initiation and elongation, and neuronal differentiation as distinctive between subtypes depending on MYC activity (Figure S1G; Table S1H). Together, these data support prior observations that dysregulated RNA translational control is widespread in medulloblastoma and reflects underlying differences in tumor subtype biology.[16,17]

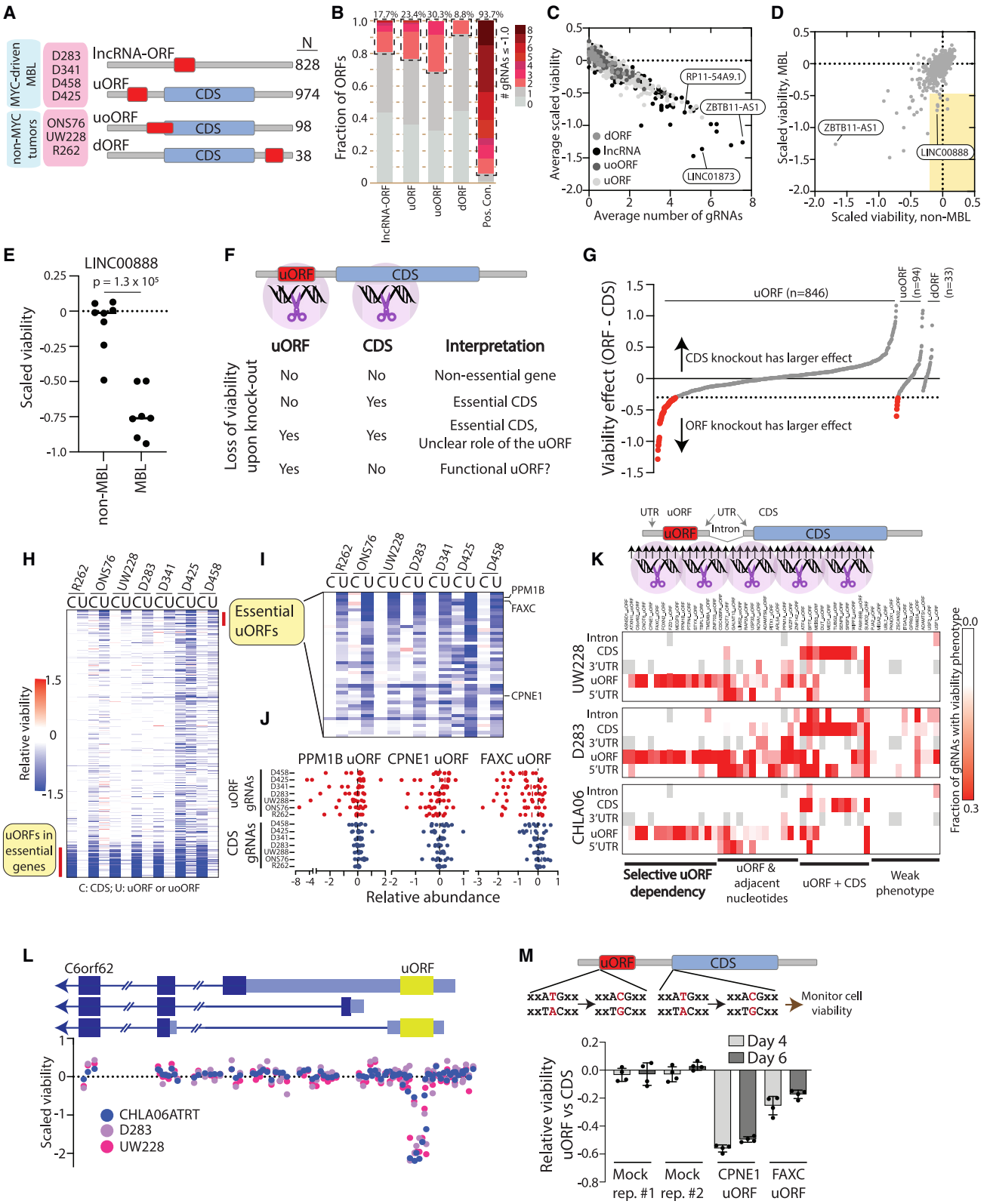
Translation of non-canonical ORFs is common in medulloblastoma

Motivated by increasing reports of functional non-canonical ORFs detected through translational profiling,[25,26,33,34] we next sought to quantify the contribution of these ORFs to the medulloblastoma translatome. We assessed the translation of 8,008 non-canonical ORFs derived from our previous analyses,[26] as well as a recently compiled human consensus ORF[21] catalog using our tissue and cell line ribosome profiling datasets. We observed the translation for 7,530 non-canonical ORFs in at least 1 sample and 6,740 in at least 5 samples (Figures 1I and 1J; Tables S1K and S1N). Among these, requiring detection by

(J) Boxplots showing distributions of translation levels of translated non-canonical ORFs, separated by ORF biotype. Each dot represents the mean translation level of one ORF across all samples. Boxplots show lower quartile, median, and upper quartile translation levels for each ORF biotype. Translation levels are calculated as normalized P-site counts. x axis reflects a $\log_2$ scale.

(K) Volcano plot of changes in translation levels between MYC-driven and non-MYC driven medulloblastoma cell lines. Each dot reflects a single non-canonical ORF, colored by ORF biotype. Dots above the dashed horizontal line have a false discovery rate (FDR) < 0.01. Labels for top 5 upregulated ($\log_2$ fold change > 2) ORFs with lowest p_{adj} and top 5 downregulated ORFs ($\log_2$ fold change < −2) with lowest p_{adj} are shown.

See also Figure S1 and Table S1.



(legend on next page)

Ribo-seq in at least 5 samples, translation of uORFs was most commonly and reproducibly found (n = 3,107), followed by the translation of lncRNA-ORFs (n = 1,775), upstream overlapping ORFs (uoORFs, n = 720), internal ORFs (intORFs, n = 694), downstream ORFs (dORFs, n = 391), and downstream overlapping ORFs (doORFs, n = 53). There was no statistically significant difference in the average number of detected non-canonical ORFs in cell lines (mean = 4,999) compared with the tissue samples (mean = 4,638; p value = 0.245, Student's t test) (Figure S1C). Importantly, translational efficiency analysis of non-canonical ORFs recapitulated disease clusters, similar to annotated CDSs, indicating subtype-specific control of non-canonical ORF translation (Figure S1H; Tables S1I–S1M). Overall, 717 non-canonical ORFs displayed differential translation levels between subtypes ($p_{adj} < 0.01$), with 268 ORFs showing increased translation in MYC-high medulloblastoma ($p_{adj} < 0.01$, $\log_2$ fold change > 2) (Figure 1K; Tables S1O and S1P).

This indicates that the medulloblastoma translatome is populated by thousands of diverse non-canonical ORFs and that translation of non-canonical ORFs is a characteristic feature of medulloblastoma disease subtypes.

Non-canonical ORFs are essential and specific in medulloblastoma cell survival

Non-canonical ORFs are increasingly recognized as serving key roles in cancer cell biology, in some cases through the generation of a stable bioactive protein.[23,26,27,35] Given their frequent and subtype-specific translation in medulloblastoma, we next sought to nominate non-canonical ORFs with key functional roles in this disease. We designed a CRISPR guide RNA library targeting 2,019 non-canonical ORFs and conducted loss-of-function knockout screening in seven medulloblastoma cell lines (four MYC driven and three non-MYC driven) in order to nominate non-canonical ORFs implicated in medulloblastoma cancer cell survival (Figures 2A and S2A–S2C; Table S2A). Performance metrics of the CRISPR screens were similar across cell lines and demonstrated high biological reproducibility (Figures S2D–S2J; Tables S2B–S2E).

In aggregate, 387 ORFs (21.1%) demonstrated an essentiality phenotype in at least one cell line, with 121 out of 387 of ORFs displaying an effect on cell survival in at least 2 independent cell lines (Figure 2B; Tables S2E–S2G). ORFs were considered essential if they exhibited at least two gRNAs with a normalized loss-of-function score of ≤ -1.0 (STAR Methods). Overall, upstream overlapping ORFs (uoORFs) and uORFs had higher rates of essentiality, although this observation was likely influenced by proximity to annotated CDSs and gene promoters (Figure S2K). dORFs, located in the 3′ UTRs of protein-coding mRNAs, exhibited the lowest rates of essentiality (Figure 2B), consistent with their generally lower translation rates (Figure 1J).

Across all cell lines, the strongest loss-of-function phenotypes were observed by the known pan-lethal effect of *ZBTB11-AS1*, which we previously characterized as an 88-amino-acid microprotein, as well as several other pan-lethal lncRNA-ORFs in *LINC01873* and *RP11-54A9.1* (Figure 2C).[26] A direct comparison of 528 ORFs screened in our current cohort of 7 medulloblastoma cell lines and our prior cohort of 8 non-medulloblastoma cell lines (comprising melanoma and carcinomas of the breast, lung, cervix, colon, liver, and prostate)[26] revealed 14 ORFs, the knockout of which had a significantly increased loss-of-viability phenotype in the medulloblastoma cohort (Figure 2D; Table S2H). Among these, we observed particularly pronounced medulloblastoma-specific viability effects for *LINC00888*, which encodes a microprotein whose translation is particularly elevated in MYC-driven medulloblastoma samples (Figures 2E and S2L). Thus, medulloblastoma

Figure 2. Non-canonical ORFs are frequently essential genes in medulloblastoma

(A) A schematic description of the cell lines and numbers of non-canonical ORFs evaluated by CRISPR screening.

(B) A bar plot showing frequency of essentiality among different classes of non-canonical ORFs. At least two gRNAs had to score as depleted to nominate an essential non-canonical ORF.

(C) A scatter plot showing the relationship between the average ORF knockout phenotype across cell lines compared with the average number of gRNAs with a viability score of ≤ -0.5 across cell lines. Previously identified ORFs from Prensner et al.,[26] are indicated.

(D) A scatter plot showing the correlation of ORF knockout phenotypes across a previously published panel of eight non-medulloblastoma cancer cell lines[26] and the current dataset of seven medulloblastoma cell lines. Medulloblastoma-specific effects are highlighted in the yellow box.

(E) The impact of knockout of an ORF in *LINC00888* in medulloblastoma and non-medulloblastoma cancer cell lines.[26] Each dot reflects an individual cell line. The y axis reflects the overall loss-of-viability phenotype of *LINC00888* knockout. p value by a two-tailed Student's t test.

(F) A schematic reflecting the knockout strategy to identify uORFs and uoORFs with putative functional consequences in medulloblastoma cell viability.

(G) A line graph showing the scaled loss of viability when comparing knockout of a uORF, uoORF, or dORF with knockout of the associated parental coding sequence (CDS) for that gene. The y axis shows the differential in viability effect. The x axis reflects each individual ORF.

(H) A heatmap showing scaled loss of viability for each pair of a parental CDS and a uORF or uoORF across all tested cell lines. Pan-essential CDSs are indicated. C, parental CDS; U, uORF or uoORF.

(I) An expanded view of the heatmap in (H), focusing on cases in which knockout of a uORF or uoORF resulted in substantially more loss of viability compared with knockout of the parental CDS.

(J) Individual gRNA level data for three essential uORFs. Here, each dot represents a gRNA to either the indicated uORF or the associated CDS. The y axis shows the cell line for the data points. The x axis shows the scaled loss of viability associated with the gRNA.

(K) Top, a schematic showing the tiling saturation gRNA library design. Bottom, a heatmap showing the fraction of gRNAs for the given genomic region of the indicated ORF that scored as displaying a loss-of-viability phenotype. ORFs are organized along the x axis according to whether they exhibited a selective knockout phenotype, a phenotype in conjunction with other gRNAs, or a weak phenotype.

(L) Individual gRNA-level data from the tiling saturation screen for the *C6orf62* uORF. Each dot represents a gRNA. The y axis shows the loss of viability associated with each gRNA. gRNAs are ordered along the x axis to align with the schematic of the *C6orf62* gene and uORF.

(M) Base editing of the *CPNE1* and *FAXC* uORF start codons or the start codons of their associated parental CDSs in D425 medulloblastoma cells. The barplot displays the differential in viability for uORF compared with CDS gRNA.

See also Figure S2 and Table S2.

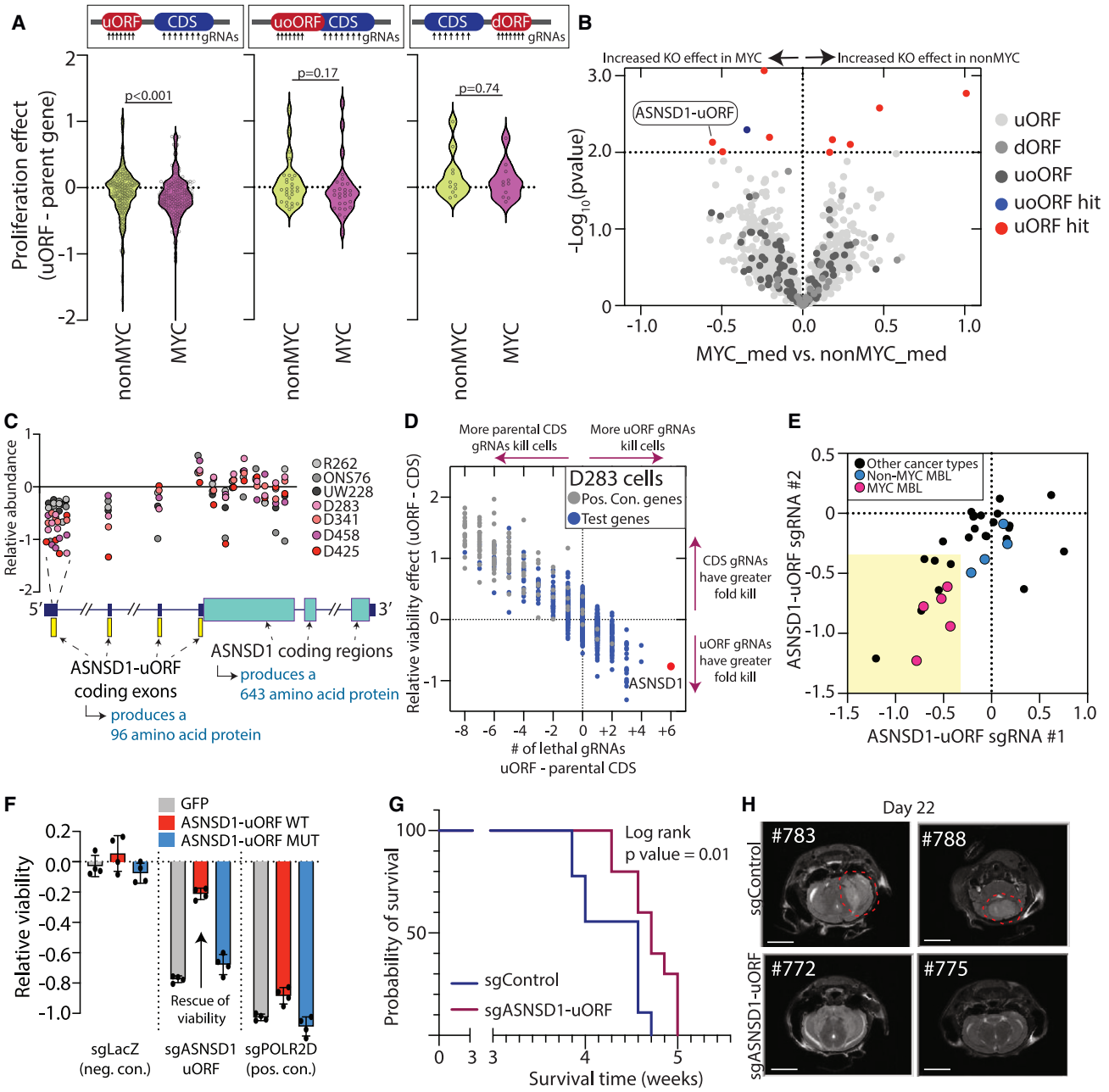


Figure 3. ASNSD1-uORF drives medulloblastoma cell survival

(A) Violin plots showing the differential viability phenotype in MYC- or non-MYC-driven medulloblastoma cells for knockout of uORFs, uoORFs, or dORFs that scored as hits in the CRISPR screen. p values by a Mann-Whitney U test.

(B) A volcano plot showing the differential viability phenotype of knockout of uORFs, uoORFs, and dORFs in MYC-driven and non-MYC-driven cell lines. Hits are indicated with the shown colors. p values are by a two-tailed Student's t test.

(C) Individual gRNA-level data for *ASNSD1-uORF* and *ASNSD1* parental CDS in the primary CRISPR screen. Each dot reflects a gRNA; dot colors reflect the indicated cell lines. The y axis shows scaled viability after knockout with each gRNA. The x axis reflects the genomic position of the gRNA relative to the *ASNSD1* gene structure shown below.

(D) A scatter plot comparing the magnitude of viability phenotype of uORF knockout relative to parental CDS knockout in D283 cells. The x axis shows the number of gRNAs inducing a loss-of-viability phenotype for the uORF minus that number for the parental CDS. The y axis shows the average loss-of-viability phenotype of the four most effective gRNAs for the uORF minus that number of the parental CDS. Positive control genes are shown in gray and other uORF genes are shown in blue.

(E) A scatter plot showing the degree of loss of viability for *ASNSD1-uORF* knockout using two gRNAs across 33 cell lines. MYC-driven medulloblastoma cells are shown in pink and nonMYC medulloblastoma cells are shown in blue. Other cell lines are shown in black.

(legend continued on next page)

may possess a unique landscape of non-canonical ORF functions.

Selective gene dependency for upstream open reading frames in medulloblastoma

While functionality of ORFs in some lncRNAs has been well established,[26,33,34,36] we were intrigued to note the abundant uORFs with an essentiality phenotype upon knockout (Figure 2B). Because most uORFs and uoORFs are conventionally thought to be regulatory sequences for adjacent canonical CDSs,[22,37] recent studies have indicated that some uORFs contain sequence variants[38,39] and encode protein products[40–44] that contribute to disease and function independent of the canonical CDS encoded by the same gene. We therefore sought to determine whether any uORFs or uoORFs harbored a selective cancer dependency phenotype that might suggest unique biological relevance of the ORF. To do this, we performed matched knockout of the uORF or uoORF and knockout of the adjacent CDS in 964 cases (>90% with at least 7 gRNAs per ORF) and compared the knockout phenotypes (Figure 2F; Table S2A).

We observed that 69 (7.2%) of uORFs or uoORFs exhibited a substantial loss-of-viability phenotype upon knockout that was not recapitulated by knockout of the adjacent CDS (Figures 2G–2J; Table S2F), of which 29/69 (42.0%) represented pan-lethal effects observed in at least 6 cell lines. To probe this observation further, we generated a custom tiling gRNA library that saturated 50 of the 69 mRNAs (median 79.5 gRNAs per gene, range 68–112) in which the uORF exhibited a lethality phenotype and performed loss-of-function screens in three cell lines (1 non-MYC MBL, 1 MYC MBL, and 1 atypical teratoid/rhabdoid tumor) (Figures S2M and S2N; Tables S2I–S2K). In total, 15 uORFs exhibited a knockout phenotype only when uORF-targeting gRNAs were used, corroborating the above-mentioned effects at a high resolution and indicating precise selective dependency relative to the CDS (Figures 2K and S2O–S2Q; Table S2L), as exemplified by the *C6orf62* uORF tiling knockout results (Figure 2L). Using two additional examples of uORFs located in the *CPNE1* and *FAXC* genes, we also verified that uORF translation was the critical feature for dependency through base editing of the uORF start codon (Figure 2M). These results indicate that a subset of uORFs may have unique roles in medulloblastoma cell viability.

Identification of a uORF in ASNSD1 as a genetic dependency in medulloblastoma

By comparing MYC- and non-MYC-driven cell lines, we were intrigued to observe that MYC-driven medulloblastoma cells exhibited enhanced essentiality phenotypes with uORF knockout ($p < 0.001$, Mann-Whitney U test) but not for uoORFs or dORFs (Figure 3A). While most differential uORF essentiality phenotypes were modest in magnitude, we focused attention a small subset of uORFs enriched in MYC-driven cells (Figure 3B). While several candidates revealed inconsistent effects of targeting the uORFs upon inspection (Figure S3A), the uORFs in the *TBPL1* and *ASNSD1* genes confirmed an effect of uORF knockout (Figure S3B). We therefore directed a focused effort on a uORF in the *ASNSD1* gene, which exhibited particular strength as a vulnerability gene in MYC-driven medulloblastoma (Figure 3B). This gene also demonstrated among the most differential phenotypes between uORF knockout and main CDS knockout, with a highly selective phenotype (Figures 3D and S3C).

This uORF encodes a conserved 96-amino-acid sequence that spans four exons of the *ASNSD1* 5′-UTR and has recently been observed and annotated in prior non-canonical ORF discovery efforts (Figures 3C, S3D, and S3E).[45–47] In humans, *ASNSD1* transcript expression is enriched in the cerebellum, with preferential expression during early development, consistent with the location and onset of childhood medulloblastoma (Figures S3F and S3G).

To confirm its role in medulloblastoma cell viability, we performed CRISPR-Cas9 knockout validation experiments for *ASNSD1-uORF* across 5 MYC-driven and 4 non-MYC-driven medulloblastoma cell lines, as well as a larger set of 24 non-medulloblastoma cell lines. Loss of cell viability following knockout of *ASNSD1-uORF* was prominent in MYC-driven medulloblastoma cell lines, whereas 18/24 (75.0%) of non-MBL cell lines did not show a consistent phenotype (Figure 3E; Table S3A). Moreover, re-expression of the wild-type ORF, but not a start-site mutant, rescued this phenotype (Figures 3F, S3H, and S3I), confirming the necessity of a protein-coding *ASNSD1-uORF* cDNA. In support of these observations, ectopic expression of ASNSD1-uORF led to a small but statistically significant increase in neural stem cell growth (9.8 vs. 7.9 doublings at 120 h; $p < 0.001$, two-tailed Student's t test; Figures S3J and S3K).

We next investigated the role for ASNSD1-uORF in medulloblastoma *in vivo*. Consistent with its importance in medulloblastoma cell viability *in vitro*, knockout of *ASNSD1-uORF* prolonged overall survival for mice with orthotopic xenografts of D425 MYC-driven medulloblastoma cells (Figures 3G, 3H, and S3L). While editing efficiency was limited for *in vivo* knockouts, we observed that the knockout allele fraction decreased, consistent with the outgrowth of cells lacking allele knockout (Figure S3M). To probe a role in autochthonous medulloblastoma tumorigenesis, we performed *in utero* electroporation of *ASNSD1-uORF* cDNA in conjunction with cDNAs for cMYC and a dominant-negative p53 (DNp53) into the developing murine cerebellum. However, addition of ASNSD1-uORF to cMYC and DNp53 in this model did not alter mouse survival (Figures S3N–S3P).

(F) A barplot showing the loss of viability for *ASNSD1-uORF* knockout in D341 cells stably overexpressing GFP, ASNSD1-uORF, or AUG-mutant ASNSD1-uORF. Black dots indicate individual data points. Error bars represent standard deviation.
(G) Overall survival for mice with D425 orthotopic xenografts in the murine cerebellum. sgControl mice (n = 9) are shown in blue and sg*ASNSD1-uORF* mice (n = 10) are shown in red. p value is by a log-rank test.
(H) Brain MRIs at day 22 post-injection for mice with sgControl orthotopic xenografts (#783 and #788) or sg*ASNSD1-uORF* orthotopic xenografts (#772 and #775). Scale bars represent 5 mm.
See also Figure S3 and Table S3.

Elevated ASNSD1-uORF protein levels in medulloblastoma

Given the importance of ASNSD1-uORF in high-risk medulloblastoma, we next asked whether its abundance was increased in this disease. Indeed, ASNSD1-uORF displayed higher levels of RNA translation in MYC-driven cell lines by Ribo-seq ($p = 0.0013$, Figure 4A). Moreover, using targeted mass spectrometry with size selection, we observed a significant upregulation of ASNSD1-uORF protein level, but not other small proteins, in 10 MYC-driven compared with 5 non-MYC-driven medulloblastoma cell lines ($p = 0.001$, Figures 4B and S4A). To validate these findings in patients, we leveraged publicly available mass spectrometry data for 45 pediatric medulloblastoma samples.[17] In this historical dataset, we noted that ASNSD1-uORF appeared correlated with MYC in group 3 tumors, although the analysis was underpowered (Figure S4B). Across all samples, high ASNSD1-uORF was also observed in samples in MYCN-high group 4 tumors, where high MYC and high MYCN are mutually exclusive (Figures S4C and S4D). These results are consistent with the well-known overlap in MYC and MYCN function[48] because both may bind the same DNA motifs,[49] dimerize with Max,[50] and control similar downstream cellular programs.[51] Therefore, we performed a merged analysis of ASNSD1-uORF protein levels in patient tumors with high levels of either MYC or MYCN, which revealed strong correlation between this uORF and the *MYC*-family transcription factors (Pearson $R = 0.47$, $p = 0.0009$) (Figure 4C).[17] Consistent with this finding, *MYCN* overexpression in ONS76 medulloblastoma cells (MYC low subtype) resulted in upregulation of ASNSD1-uORF protein levels (Figures S4E and S4F; Table S4A). This regulation appeared, at least in part, to be from post-transcriptional effects because ASNSD1-uORF protein levels did not correlate with *ASNSD1* mRNA levels or MYC/MYCN level in patient tissues (Figures S4G and S4H).

We also measured ASNSD1-uORF protein levels across 23 non-medulloblastoma cell lines with matched CRISPR knockout data (as in Figure 3E; Table S3A) and observed that, while some cell lines lacking an essentiality phenotype expressed ASNSD1-uORF protein, medulloblastoma cell lines displayed both prominent protein expression and a loss-of-viability knockout phenotype (Figure S4I). Lastly, a reanalysis of mass spectrometry data for 504 solid tumor, non-medulloblastoma cancer cell lines[52] demonstrated the greatest abundance of ASNSD1-uORF protein in MYCN-amplified neuroblastoma cell lines, consistent with MYC/MYCN regulation (Figure S4J). Taken together, our findings indicate that *ASNSD1-uORF* is a genetic dependency in high-risk medulloblastoma, which may be associated with its upregulation at the protein level in MYC- or MYCN-driven pediatric cancers.

ASNSD1-uORF functions coordinately with the prefoldin-like complex in medulloblastoma

To identify molecular mechanisms of ASNSD1-uORF protein in medulloblastoma, we pursued three strategies: protein-protein interactions, correlation of proteomic and genetic knockout signatures, and downstream molecular networks. First, we performed co-immunoprecipitation experiments for ectopically expressed ASNSD1-uORF followed by mass spectrometry (Figure 4D). Consistent with a prior report,[53] we observed a striking enrichment for multiple members of the prefoldin complex, which we validated with western blots (Figures 4E and 4F). We further validated this interaction by using co-immunoprecipitation of endogenous prefoldin subunit 6 (encoded by the *PFDN6 gene*) in D425 cells (Figures 4G and 4H), which confirmed enrichment of endogenous ASNSD1-uORF protein (Figures 4I and S4K; Table S4B).

Next, we sought to distinguish whether ASNSD1-uORF primarily operated in conjunction with the canonical prefoldin complex (PFD) or the more obscure prefoldin-like complex (PFDL) variant. The PFD is an evolutionarily conserved, hexameric protein chaperone complex, which is thought to play an important role in the stability of nascent proteins.[30,31] Several clinicopathological studies have associated PFD components with cancer,[54–56] including recent data suggesting that PFD proteins may be dysregulated in medulloblastoma.[20] While the canonical PFD is embryonic lethal in mouse knockout models, the non-canonical PFDL, which retains only two of the six components of the PFD complex (*PFDN2* and *PFDN6*), may have only subtle murine knockout phenotypes (Figure S4L).

To place ASNSD1-uORF in the context of PFD or PFDL, we first used the Archer et al. medulloblastoma mass spectrometry dataset[17] to correlate PFD or PFDL complex members to ASNSD1-uORF protein abundance. We observed that the PFDL-specific complex members are among the most highly correlated proteins with high statistical significance (p53 and DNA damage regulated 1 (PDRG1), URI1 prefoldin-like chaperone (URI1), and ubiquitously expressed prefoldin-like chaperone (UXT); Pearson correlations 0.756–0.826, Q values $< 10^{-12}$ Figure 4J; Table S4C). By contrast, PFD complex-specific proteins were not significantly correlated with ASNSD1-uORF abundance. PFDL proteins were significantly upregulated in MYC/MYCN-driven medulloblastomas, similar to ASNSD1-uORF (Figure S4M), and PFDL mRNAs exhibited an increased translational efficiency rate in cell line models as well (Figure S4N). Next, we established that genetic knockout of PFDL proteins recapitulated the phenotype of *ASNSD1-uORF* knockout. Specifically, we used pooled cell culture to knockout *ASNSD1-uORF* in >400 barcoded PRISM cancer cell lines for dropout screening[26] and compared its pattern of genetic dependency with those of PFD and PFDL protein knockout in the same cell lines in the DepMap database (www.depmap.org) (Figures 4K and S4O–S4P; Tables S4D–S4G). We found that members of the PFD and PFDL complexes readily clustered based upon the Pearson correlation of their knockout phenotype across the cell lines and that *ASNSD1-uORF* was strongly associated with the PFDL but not the PFD complex (Figure 4L). A focused analysis of PFD and PFDL members in medulloblastoma cell lines represented in the DepMap database confirmed a relative increase in viability effects from PFDL knockout in MYC-high cell lines (Figure S4Q). Lastly, purified ASNSD1-uORF tagged with a glutathione S-transferase (GST) tag showed direct interactions with the PFDL complex (Figure S4R).

Knockout of *ASNSD1-uORF* or multiple prefoldin members did not impact the abundance of cytoskeletal proteins such as actin and tubulin, which have previously been suggested[57] as downstream targets (Figure S5A). We therefore profiled

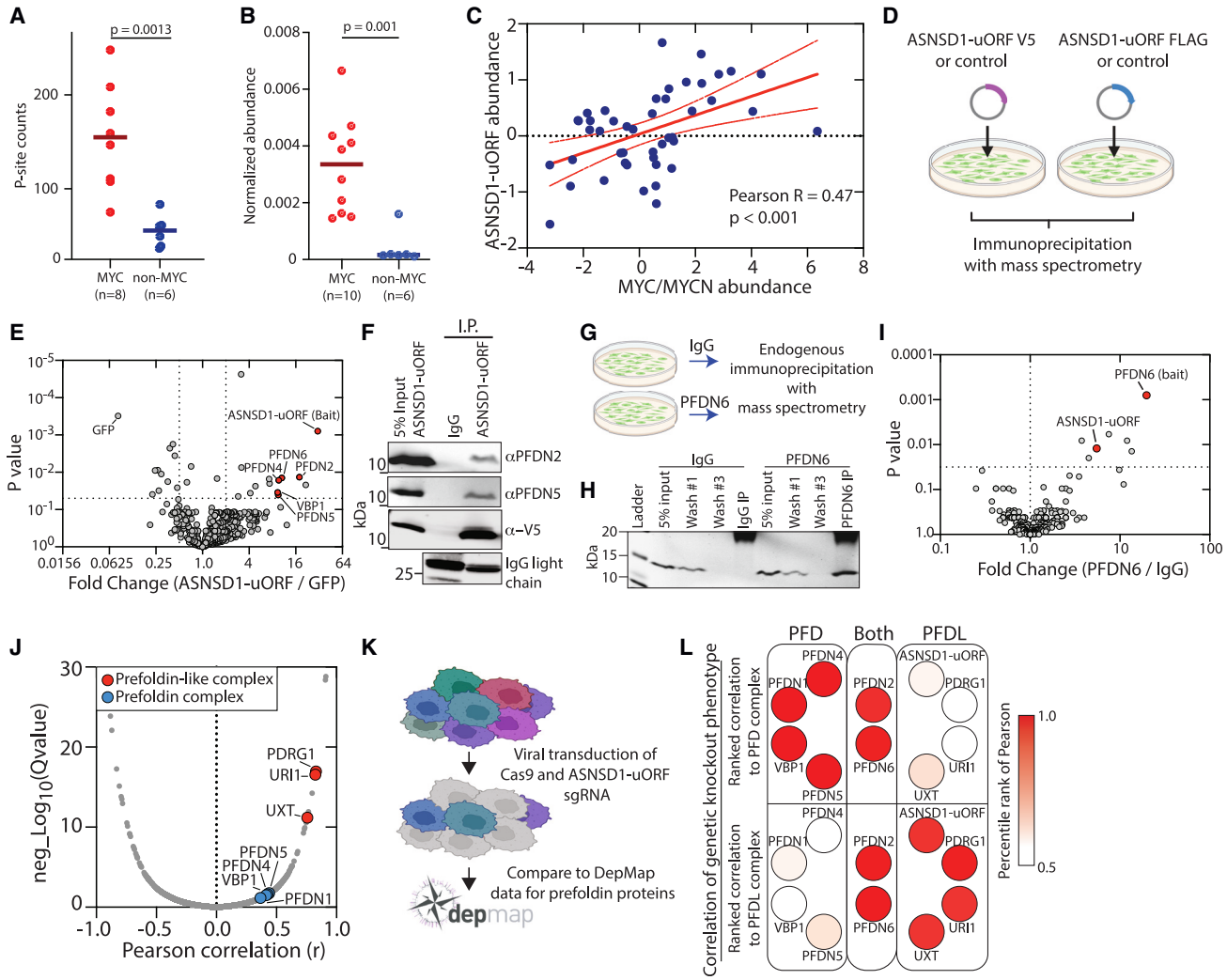


Figure 4. ASNSD1-uORF cooperates with the prefoldin-like complex in medulloblastoma

(A) Abundance of *ASNSD1-uORF* translation across medulloblastoma cell lines using Ribo-seq data. Each dot reflects a cell line. p value by a two-tailed Student's t test.

(B) Protein abundance of ASNSD1-uORF in a cohort of MYC-driven (n = 10) or non-MYC (n = 6) medulloblastoma cell lines. p value by a two-tailed Student's t test.

(C) A scatter plot correlating ASNSD1-uORF protein abundance to protein abundance of MYC and MYCN in medulloblastoma patient samples (n = 46) from the reanalyzed Archer et al. dataset.[17] Correlation and p values were determined by a Pearson R.

(D) A schematic showing the experimental design for ASNSD1-uORF co-immunoprecipitation from exogenous expression in HEK293T cells.

(E) A volcano plot showing enrichment of prefoldin and prefoldin-like complex proteins in ectopic ASNSD1-uORF co-immunoprecipitation in HEK293T cells. The x axis shows fold change of pull-down on a $\log_2$ scale. The y axis shows the p value by a two-tailed Student's t test.

(F) A western blot showing validation of PFDN2 and PFDN5 pull-down with ASNSD1-uORF co-immunoprecipitation.

(G) A schematic showing the experimental design for endogenous co-immunoprecipitation with PFDN6.

(H) Western blot validation of PFDN6 pull-down in D425 cells.

(I) Mass spectrometry analysis of interacting partners with endogenous PFDN6 co-immunoprecipitation. The x axis shows fold change of pull-down on a log10 scale. The y axis shows the p value by a two-tailed Student's t test.

(J) The correlation between ASNSD1-uORF protein abundance and prefoldin or prefoldin-like complex proteins from the reanalyzed Archer et al. medulloblastoma tissue samples (n = 46).[17] The x axis shows the Pearson correlation to ASNSD1-uORF. The y axis shows the adjusted Q value.

(K) A schematic showing the experimental design for correlating *ASNSD1-uORF* knockout phenotypes with knockout phenotypes of prefoldin proteins.

(L) A heatmap showing the percentile rank of the Pearson correlation coefficient for loss of viability across 484 cancer cell lines following *ASNSD1-uORF* knockout or prefoldin/prefoldin-like gene knockouts.

See also Figure S4 and Table S4

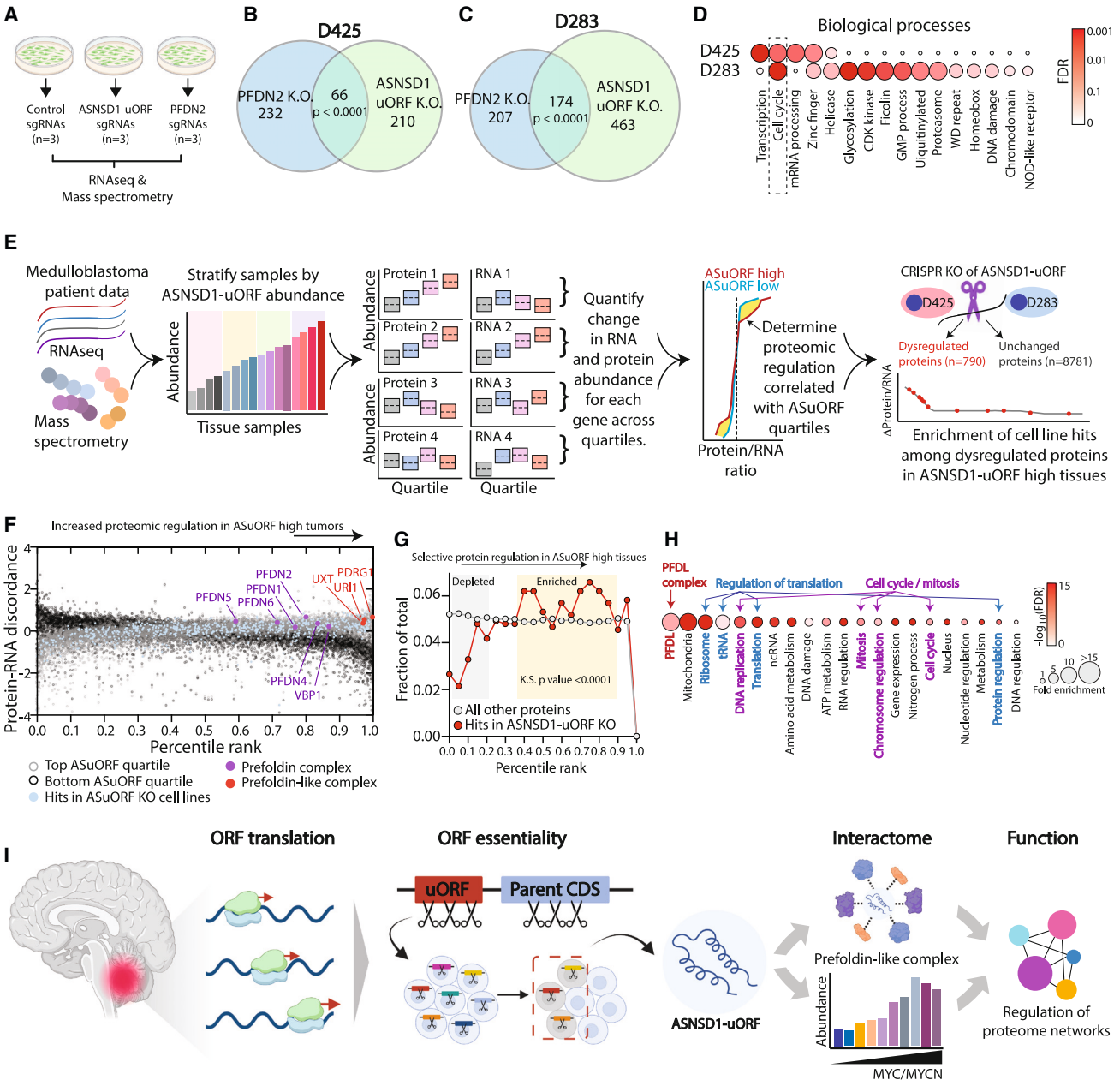


Figure 5. ASNSD1-uORF associates with increased RNA-protein dyssynchrony in medulloblastoma

(A) A schematic showing the experimental design for RNA-seq and mass spectrometry experiments to functionally characterize ASNSD1-uORF.

(B) Overlapping signatures of regulated proteins in mass spectrometry data for *ASNSD1-uORF* and *PFDN2* knockout in D425. p value by a Fisher's exact test.

(C) Overlapping signatures of regulated proteins in mass spectrometry data for *ASNSD1-uORF* and *PFDN2* knockout in D283. p value by a Fisher's exact test.

(D) Enriched biological processes identified in D425- or D283-signatures of proteins regulated by both *PFDN2* and *ASNSD1-uORF* in mass spectrometry datasets.

(E) A schematic illustrating the analytical approach used to define protein-RNA dyssynchrony in medulloblastoma and integrate *in vitro ASNSD1-uORF* knockout data.

(F) A scatter plot showing the protein-RNA discordance in medulloblastoma samples with the highest (top quartile) or lowest (bottom quartile) ASNSD1-uORF protein abundance. The x axis shows the percentile rank for proteins with the most to least protein-RNA discordance. Gray and black dots show protein-RNA discordance for the top and bottom quartiles, respectively. Blue dots indicate hits in the *ASNSD1-uORF* knockout proteomics dataset. Purple dots and red dops indicate prefoldin and prefoldin-like complex proteins, respectively.

(G) The distribution of the protein-RNA discordance score in medulloblastoma tissue samples for proteins identified in the *ASNSD1-uORF* knockout experiment in D283 and D425 cells. The fraction of proteins identified *in vitro* as either hits (N = 790) or non-hits (N = 8,781) are quantified across the percentile rank for

(legend continued on next page)

transcriptomic and proteomic changes following the knockout of *ASNSD1-uORF* or *PFDN2* in D425 and D283 cells (Figure 5A; Tables S5A–S5D). Importantly, the protein abundance of the ASNSD1 parent CDS was not depleted by these gRNAs (Figure S5B). For both cell lines, we observed an overlapping proteomic signature of co-regulated proteins (Figures 5B and 5C), which demonstrated minimal change by RNA-seq (Figures S5C and S5D), confirming a post-transcriptional role for the prefoldin complex. Probing these sets of proteins further revealed consistent biological functional groups, with proteins related to cell cycle showing prominently (Figure 5D; Table S5E). These proteins additionally showed concordant upregulation in ASNSD1-uORF-high tissues (Figure S5E).

Because the PFDL complex imparts post-transcriptional control, we next reasoned that tumors with high abundance of this complex would exhibit increased discordance between the levels of target proteins and their corresponding mRNAs (Figure 5E). By integrating matched RNA-seq and proteomics from medulloblastoma patient samples from Archer et al.[17] with the differentially regulated proteins from our cell line knockout experiments, we stratified the degree of protein-RNA dyssynchrony for each protein by ASNSD1-uORF abundance and assessed for bias in the distribution of the proteins significantly altered by *ASNSD1-uORF* knockout in either D425 or D283 cells (Figures 5E and 5F; Tables S5F–S5H). We found that proteins altered by *ASNSD1-uORF* knockout were significantly skewed to show increased protein-to-RNA dyssynchrony and depleted for genes that showed highly concordant protein-to-RNA ratios (Figure 5G; Table S5I). Consistent with our identification of cell cycle as a putative protein network regulated by *ASNSD1-uORF* and *PFDN2* knockout *in vitro*, proteins with highly skewed protein-RNA ratios in ASNSD1-uORF-high samples were significantly overrepresented for mitotic and cell-cycle proteins, along with translation-associated signatures (Figures 5H, S5F, and S5G; Tables S5J–S5M). Interestingly, the most enriched signature was for other members of the prefoldin-like complex (Figure 5H), which indicates coordinated post-transcriptional regulation of this complex and is consistent with our observation that several prefoldin-like proteins are upregulated in MYC(N)-high tissue samples (Figure S4M), as well as ONS76 cells with ectopic MYCN expression (Figures S5H and S5I). Lastly, in support of this association between ASNSD1-uORF and cell cycle, we found that *ASNSD1-uORF* knockout *in vitro* decreased S-phase incorporation of bromouridine (BrdU) (Figure S5J).Collectively, these data support a role for ASNSD1-uORF within the PFDL complex in mediating cancer cell viability by coordinating downstream signatures of proteome regulation related to cell cycle and mitosis, which may be relevant for medulloblastoma.

DISCUSSION

Here, we present a comprehensive analysis of the medulloblastoma translatome, generating matched Ribo-seq and RNA-seq data of 32 patient tissues and cell lines to enable the investigation of translated ORFs in this disease. We show that medulloblastoma reproducibly translates over 6,700 non-canonical ORFs, which represent a previously unstudied layer of biology in this embryonal brain cancer (Figure 5I). Using multiple CRISPR-Cas9 approaches to knock out over 2,000 ORFs, we broadly interrogate the contribution of non-canonical ORFs in cell survival across seven medulloblastoma cell lines. Overall, our results provide strong support for the growing community-wide interest in non-canonical ORFs as biological actors in both basic cell biology[24,25,34,58–60] and cancer pathophysiology.[26,35,61] As such, our data argue for the inclusion of non-canonical ORFs in cancer genomics studies.

We particularly observe that a subset of uORFs function to maintain cancer cell survival. While early literature on uORFs has emphasized their importance only as regulators of mRNA translation,[22,37,62] our efforts indicate that a sizable number of uORFs may operate as discrete biological actors. We are further able to pinpoint genetic dependency of 15 uORFs using high-density CRISPR-tiling approaches, which provides high-resolution genetic evidence for uORF functionality in these cases. These data support the hypothesis that some uORFs are specific genetic dependencies in cancer, although the annotated, adjacent protein-coding CDS is not. Indeed, this hypothesis would suggest that some genes found to be dependencies by RNA interference screening—in which a full mRNA is downregulated—fail to score in CRISPR knockout data targeting the CDS. *ASNSD1* points toward this: MYC-amplified medulloblastoma cell lines D458, D425, and D341 are among the most prominent hits in DEMETER shRNA data[63] for *ASNSD1* but do not score in the CRISPR-based DepMap (Figure S5K).

At the same time, we report the first example of molecular subtype-specific non-canonical ORF activity in childhood cancer. We focus on the role of the *MYC*-family transcription factors, which we find may drive non-canonical ORF translation. Here, we establish a specific role for the *ASNSD1-uORF* as a medulloblastoma cancer dependency, and its activity is linked to the MYC-family protein activity. Given the prominent role for *MYC* transcription factors in other cancer types, our observations that transcription factor amplification activates certain uORFs may have broader implications in cancer. To this end, we note that the example of ASNSD1-uORF protein is also more abundant in high-risk neuroblastoma cell lines, which may be due to impact on RNA translation by *MYCN* amplification.[64,65]

protein-RNA discordance score. The y axis reflects the fraction of total proteins in each group. The x axis reflects the percentile rank quantized into 0.05 increments. p value by a Kolmogorov-Smirnov test.

(H) The Gene Ontology terms that are most enriched in proteins (N = 383 with a two-tailed t test p value of <0.01) with a differential protein-RNA discordance score between tissue samples in the top and bottom quartiles of ASNSD1-uORF abundance. The color of the circles represents the degree of statistical significance for the Gene ontology (GO) term and the size of the size represents the fold enrichment of proteins in that GO term.

(I) A general model of non-canonical ORF translation in medulloblastoma.

See also Figure S5 and Table S5.

Lastly, we describe a mechanism for ASNSD1-uORF within the poorly understood prefoldin-like complex, which is thought to play a role in protein homeostasis similar to that of the prefoldin complex, a related but distinct entity.[30,57] As such, our data reinforce a prior observation association ASNSD1-uORF with the prefoldin-like complex as well as emerging evidence that protein homeostasis via the prefoldin complex is dysregulated in medulloblastoma.[20] While precise functions of the prefoldin-like complex remain incompletely understood, we observe that its impact on proteome regulation associates with specific, cancer-relevant biological functions, such as cell cycle. As a post-transcriptional mechanism of protein regulation, ASNSD1-uORF and the prefoldin-like complex lend additional evidence to the observation that the medulloblastoma proteome deviates substantially from the transcriptome.[16,17]

In summary, our findings exploit the known disease biology of medulloblastoma subtypes to provide cancer relevancy to the growing field of non-canonical ORFs and microproteins, providing context- and oncogene-specific consequences of non-canonical ORF translation. As such, our work provides additional rationale to investigate non-canonical ORFs and their translation as putative cancer target genes in medulloblastoma and other diverse malignancies.

Limitations of the study

This study has several important limitations. First, we were unable to acquire sufficient *MYC*-amplified medulloblastoma tissue samples to enable detailed analyses of this molecular subtype in patients and therefore perform some analyses only in cell lines. Because some of the tissue samples were acquired over 5 years ago, the assigned clinical molecular subtype did not universally use DNA methylation, which is currently the gold standard technique.[66] For functional analyses, we used a broad set of previously designed non-canonical ORFs,[21] not all of which were detected, and we profiled only a portion of them by CRISPR-Cas9 screens due to the complexity of those experiments. Because of the genotoxic nature of Cas9, some nominated ORFs may be complicated by Cas9-mediated effects of targeting the genome rather than the ORF itself. For *ASNSD1-uORF*, we observe that its biological effects are enriched in MYC-driven medulloblastoma but not exclusive, and there are other cell lines that show dependency on this gene, which we have not investigated here. Regarding the relationship of ASNSD1-uORF to MYC/MYCN, we have demonstrated a mechanism by which MYC/MYCN increases the translational efficiency of the ASNSD1-uORF CDS, but we have not evaluated the role of protein stability, which may also contribute to ASNSD1-uORF levels in medulloblastoma. The influence of factors such as ancestry, race, ethnicity, socioeconomic status, and gender on ASNSD1-uORF function could not be assessed due to unavailable data. Lastly, we did not observe that the addition of *ASNSD1-uORF* expression to *MYC* and *DNp53* resulted in altered medulloblastoma aggressiveness in mice. The reasons for this include the possibility that ASNSD1-uORF is insufficient to alter cancer aggressiveness *in vivo* or that we did not test the correct genetic background because p53 mutations are highly rare in MYC-driven medulloblastoma.[5,6]

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Mouse models
  - Murine orthotopic xenograft experiments
  - Murine magnetic resonance imaging
  - Murine *in utero* electroporation experiments
  - Cell lines and reagents
  - Primary tissue samples
- METHOD DETAILS
  - Immunoblot Analysis
  - RNA isolation and cDNA synthesis
  - Ribosome profiling
  - RNA sequencing
  - RNAseq sample clustering and pathway analysis
  - Obtaining gene-level RNA-seq read counts
  - Ribo-seq read alignment and processing
  - Calculating translational efficiency values
  - Quantifying ORF-level P-sites
  - Identifying differentially translated ORFs
  - ORF-level translational efficiency analysis
  - Tissue sample RNA-seq and Ribo-seq analyses
  - Lentiviral transduction for CRISPR screens
  - Primary and validation CRISPR screens
  - Analysis of CRISPR screening data
  - Assessment of Cas9 toxicity at gene promoters
  - Defining medulloblastoma-specific CRISPR hits
  - Analysis of CRISPR validation screen
  - Base editing
  - Nomination of ASNSD1-uORF
  - Prefoldin in medulloblastoma DepMap data
  - ASNSD1-uORF evolutionary analysis
  - ASNSD1 gene expression analysis
  - ASNSD1-uORF overexpression experiments
  - ASNSD1-uORF knockout experiments
  - Analysis of cell line knockout data
  - Pooled ASNSD1-uORF knockout
  - Analysis of pooled ASNSD1-uORF knockout data
  - CRISPR-seq
  - ASNSD1-uORF protein level in cell lines
  - ASNSD1-uORF in proteomics and RNA-seq
  - ASNSD1-uORF immunoprecipitation
  - PFDN6 co-immunoprecipitation
  - Expression of GST constructs
  - GST-Pull Down Assay
  - Identification of downstream targets
  - Stratifying ASNSD1-uORF proteomic targets
  - Calculation of protein-RNA discordance score
  - Bromouridine (BrdU) quantification

- Prefoldin complex lethality in murine embryos
- Comparing CRISPR data with Project Achilles

- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2023.12.003.

ACKNOWLEDGMENTS

We thank Greg Newby and David Liu from the Broad Institute for their gracious insights with base editing experiments and providing the ABE8e-NRCH base editor. We thank Edmond Chan (Columbia University) for insights into chromatin immunoprecipitation experiments. We thank Ross Tomiano at the Taplin Mass Spectrometry Facility and Julian Mintseris at the Thermo-Fisher Center for Multiplexed Proteomics at the Harvard Medical School for assistance with mass spectrometry experiments. We thank Maura Berkeley and Zachary Herbert at the Dana-Farber Cancer Institute Molecular Biology Core Facility for assistance with next-generation sequencing. We thank the Dana-Farber Cancer Institute Center for Patient Derived Models for cell line support. We thank the Boston Children's Hospital biobank and the DFHCC Neurooncology Program Tissue and Data Bank for biobanking support. We thank the PRISM team at the Broad Institute for assistance with PRISM cell line screening and PRISM cell library sample preparation. We thank Joelle Straehla for sharing Med2112-mCherry-Luc and Med411-GFP-Luc cells. We thank Quang-De Nguyen, Amy Cameron, and Murry Morrow at the Lurie Family Animal Imaging Center at the Dana-Farber Cancer Institute. J.R.P. acknowledges funding from the National Institutes of Health – National Cancer Institute (K08-CA263552-01A1), the Alex's Lemonade Stand Foundation Young Investigator Award (#21-23983), the St. Baldrick's Foundation Scholar Award (#931638), the DIPG/DMG Research Funding Alliance, the Cure ATRT Now Fund, the Musella Foundation for Brain Tumor Research, and a Collaborative Pediatric Cancer Research Awards Program/Kids Join the Fight award (#22FN23). T.R.G. acknowledges funding from the National Cancer Institute (1 R35 CA242457–01). K.L.L. and J.V. acknowledge support from the Pediatric Brain Tumor Foundation, the National Brain Tumor Society, and 3000 Miles to the Cure. S.v.H. acknowledges funding from Fonds Cancers (FOCA, Belgium), Stichting Reggeborgh (the Netherlands), and Bergh in het Zadel (the Netherlands). N.H. acknowledges funding from the European Union Horizon 2020 Research and Innovation Program(AdG788970), the Deutsche Forschungsgemeinschaft (SFB-1470 – B03), the Chan Zuckerberg Foundation (2019-202666), the Leducq Foundation (16CVD03), the British Heart Foundation, and the Deutsches Zentrum für Herz-Kreislauf-Forschung (BHF/DZHK: SP/19/1/34461). P.B. acknowledges funding from the Isabel V. Marxuach Fund for Medulloblastoma Research and the Jared Branfman Sunflowers for Life Fund for Pediatric Brain and Spinal Cancer Research.

AUTHOR CONTRIBUTIONS

Conceptualization, J.R.P., S.v.H., J.R.-O., and D.A.H.; methodology, J.R.P., D.A.H., S.v.H., J.R.-O., K.R.C., D.E.R., and P.B.; validation, J.R.P., D.A.H., J.R.-O., R.M., and I.Y.; formal analysis, J.R.P., D.A.H., J.R.-O., and K.R.C.; investigation, J.R.P., D.A.H., J.R.-O., R.M., I.Y., K.N.Z., A.L.C., A.B., K.R.C., S.A.G.E., J.v.d.L., T.N.P., A.H., T.A., J.A.R., D.E.R., L.W., and J.v.D.; resources, J.R.P., P.B., T.R.G., K.L.L., J.V., S.v.H., and N.H.; data curation, J.R.P., D.A.H., and J.R.-O.; writing – original draft, D.A.H., J.R.-O., J.R.P., and S.v.H.; writing – review & editing, D.A.H., J.R.-O., J.R.P., and S.v.H. with input from all authors; visualization, J.R.P., J.R.-O., and D.A.H.; supervision, J.R.P., S.v.H., P.B., T.R.G., and N.H.; project administration, S.v.H. and J.R.P.; funding acquisition, J.R.P., T.R.G., N.H., K.L.L., P.B., and S.v.H.

DECLARATION OF INTERESTS

K.L.L. reports the following interests: equity in Travera; research funds from Bristol Myers Squibb, SEngine Precision Medicine, Multiple Myeloma Research Foundation, and Eli Lilly and Company; and being a consultant or on the scientific advisory board for Bristol Myers Squibb, Travera, and IntegraGen. P.B. receives grant funding from Novartis Institute of Biomedical Research, and has received grant funding from Deerfield Therapeutics, both for unrelated projects. P.B. has also served on a paid advisory board for qed Therapeutics, unrelated to this work. D.E.R. receives research funding from members of the Functional Genomics Consortium (AbbVie, BMS, Jannsen, Merck, and Vir) and is a director of Addgene, Inc.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: May 16, 2023
Revised: August 30, 2023
Accepted: December 1, 2023
Published: January 3, 2024

REFERENCES

1. Hill, R.M., Richardson, S., Schwalbe, E.C., Hicks, D., Lindsey, J.C., Crosier, S., Rafiee, G., Grabovska, Y., Wharton, S.B., Jacques, T.S., et al. (2020). Time, pattern, and outcome of medulloblastoma relapse and their association with tumour biology at diagnosis and therapy: a multicentre cohort study. Lancet Child Adolesc. Health *4*, 865–874.
2. Schwalbe, E.C., Lindsey, J.C., Nakjang, S., Crosier, S., Smith, A.J., Hicks, D., Rafiee, G., Hill, R.M., Iliasova, A., Stone, T., et al. (2017). Novel molecular subgroups for clinical classification and outcome prediction in childhood medulloblastoma: a cohort study. Lancet Oncol. *18*, 958–971.
3. Ramaswamy, V., Remke, M., Bouffet, E., Bailey, S., Clifford, S.C., Doz, F., Kool, M., Dufour, C., Vassal, G., Milde, T., et al. (2016). Risk stratification of childhood medulloblastoma in the molecular era: the current consensus. Acta Neuropathol. *131*, 821–831.
4. Pfister, S., Remke, M., Benner, A., Mendrzyk, F., Toedt, G., Felsberg, J., Wittmann, A., Devens, F., Gerber, N.U., Joos, S., et al. (2009). Outcome prediction in pediatric medulloblastoma based on DNA copy-number aberrations of chromosomes 6q and 17q and the MYC and MYCN loci. J. Clin. Oncol. *27*, 1627–1636.
5. Jones, D.T.W., Jäger, N., Kool, M., Zichner, T., Hutter, B., Sultan, M., Cho, Y.J., Pugh, T.J., Hovestadt, V., Stütz, A.M., et al. (2012). Dissecting the genomic complexity underlying medulloblastoma. Nature *488*, 100–105.
6. Pugh, T.J., Weeraratne, S.D., Archer, T.C., Pomeranz Krummel, D.A., Auclair, D., Bochicchio, J., Carneiro, M.O., Carter, S.L., Cibulskis, K., Erlich, R.L., et al. (2012). Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. Nature *488*, 106–110.
7. Gröbner, S.N., Worst, B.C., Weischenfeldt, J., Buchhalter, I., Kleinheinz, K., Rudneva, V.A., Johann, P.D., Balasubramanian, G.P., Segura-Wang, M., Brabetz, S., et al. (2018). The landscape of genomic alterations across childhood cancers. Nature *555*, 321–327.
8. Delaidelli, A., Jan, A., Herms, J., and Sorensen, P.H. (2019). Translational control in brain pathologies: biological significance and therapeutic opportunities. Acta Neuropathol. *137*, 535–555.
9. Dey, A., Robitaille, M., Remke, M., Maier, C., Malhotra, A., Gregorieff, A., Wrana, J.L., Taylor, M.D., Angers, S., and Kenney, A.M. (2016). YB-1 is elevated in medulloblastoma and drives proliferation in Sonic hedgehog-dependent cerebellar granule neuron progenitor cells and medulloblastoma cells. Oncogene *35*, 4256–4268.
10. Northcott, P.A., Buchhalter, I., Morrissy, A.S., Hovestadt, V., Weischenfeldt, J., Ehrenberger, T., Gröbner, S., Segura-Wang, M., Zichner, T., Rudneva, V.A., et al. (2017). The whole-genome landscape of medulloblastoma subtypes. Nature *547*, 311–317.
11. Robinson, G., Parker, M., Kranenburg, T.A., Lu, C., Chen, X., Ding, L., Phoenix, T.N., Hedlund, E., Wei, L., Zhu, X., et al. (2012). Novel mutations target distinct subgroups of medulloblastoma. Nature *488*, 43–48.
12. Valentin-Vega, Y.A., Wang, Y.D., Parker, M., Patmore, D.M., Kanagaraj, A., Moore, J., Rusch, M., Finkelstein, D., Ellison, D.W., Gilbertson, R.J.,

et al. (2016). Cancer-associated DDX3X mutations drive stress granule assembly and impair global translation. Sci. Rep. *6*, 25996.

13. Barna, M., Pusic, A., Zollo, O., Costa, M., Kondrashov, N., Rego, E., Rao, P.H., and Ruggero, D. (2008). Suppression of Myc oncogenic activity by ribosomal protein haploinsufficiency. Nature *456*, 971–975.
14. Zielke, N., Vähärautio, A., Liu, J., Kivioja, T., and Taipale, J. (2022). Upregulation of ribosome biogenesis via canonical E-boxes is required for Myc-driven proliferation. Dev. Cell *57*, 1024–1036.e5.
15. Ruggero, D. (2009). The role of myc-induced protein synthesis in cancer. Cancer Res. *69*, 8839–8843.
16. Forget, A., Martignetti, L., Puget, S., Calzone, L., Brabetz, S., Picard, D., Montagud, A., Liva, S., Sta, A., Dingli, F., et al. (2018). Aberrant ERBB4-SRC signaling as a hallmark of Group 4 medulloblastoma revealed by integrative phosphoproteomic profiling. Cancer Cell *34*, 379–395.e7.
17. Archer, T.C., Ehrenberger, T., Mundt, F., Gold, M.P., Krug, K., Mah, C.K., Mahoney, E.L., Daniel, C.J., LeNail, A., Ramamoorthy, D., et al. (2018). Proteomics, post-translational modifications, and integrative analyses reveal molecular heterogeneity within medulloblastoma subgroups. Cancer Cell *34*, 396–410.e8.
18. Leprivier, G., Remke, M., Rotblat, B., Dubuc, A., Mateo, A.R., Kool, M., Agnihotri, S., El-Naggar, A., Yu, B., Somasekharan, S.P., et al. (2013). The eEF2 kinase confers resistance to nutrient deprivation by blocking translation elongation. Cell *153*, 1064–1079.
19. Rivero-Hinojosa, S., Lau, L.S., Stampar, M., Staal, J., Zhang, H., Gordish-Dressman, H., Northcott, P.A., Pfister, S.M., Taylor, M.D., Brown, K.J., et al. (2018). Proteomic analysis of Medulloblastoma reveals functional biology with translational potential. Acta Neuropathol. Commun. *6*, 48.
20. Kuzuoglu-Ozturk, D., Aksoy, O., Schmidt, C., Lea, R., Larson, J.D., Phelps, R.R.L., Nasholm, N., Holt, M., Contreras, A., Huang, M., et al. (2023). N-myc-mediated translation control is a therapeutic vulnerability in medulloblastoma. Cancer Res. *83*, 130–140.
21. Mudge, J.M., Ruiz-Orera, J., Prensner, J.R., Brunet, M.A., Calvet, F., Jungreis, I., Gonzalez, J.M., Magrane, M., Martinez, T.F., Schulz, J.F., et al. (2022). Standardized annotation of translated open reading frames. Nat. Biotechnol. *40*, 994–999.
22. Johnstone, T.G., Bazzini, A.A., and Giraldez, A.J. (2016). Upstream ORFs are prevalent translational repressors in vertebrates. EMBO J. *35*, 706–723.
23. Chen, J., Brunner, A.D., Cogan, J.Z., Nuñez, J.K., Fields, A.P., Adamson, B., Itzhak, D.N., Li, J.Y., Mann, M., Leonetti, M.D., et al. (2020). Pervasive functional translation of noncanonical human open reading frames. Science *367*, 1140–1146.
24. van Heesch, S., Witte, F., Schneider-Lunitz, V., Schulz, J.F., Adami, E., Faber, A.B., Kirchner, M., Maatz, H., Blachut, S., Sandmann, C.L., et al. (2019). The translational landscape of the human heart. Cell *178*, 242–260.e29.
25. Sandmann, C.L., Schulz, J.F., Ruiz-Orera, J., Kirchner, M., Ziehm, M., Adami, E., Marczenke, M., Christ, A., Liebe, N., Greiner, J., et al. (2023). Evolutionary origins and interactomes of human, young microproteins and small peptides translated from short open reading frames. Mol. Cell *83*, 994–1011.e18.
26. Prensner, J.R., Enache, O.M., Luria, V., Krug, K., Clauser, K.R., Dempster, J.M., Karger, A., Wang, L., Stumbraite, K., Wang, V.M., et al. (2021). Noncanonical open reading frames encode functional proteins essential for cancer cell survival. Nat. Biotechnol. *39*, 697–704.
27. Jayaram, D.R., Frost, S., Argov, C., Liju, V.B., Anto, N.P., Muraleedharan, A., Ben-Ari, A., Sinay, R., Smoly, I., Novoplansky, O., et al. (2021). Unraveling the hidden role of a uORF-encoded peptide as a kinase inhibitor of PKCs. Proc. Natl. Acad. Sci. USA. *118*.
28. Sendoel, A., Dunn, J.G., Rodriguez, E.H., Naik, S., Gomez, N.C., Hurwitz, B., Levorse, J., Dill, B.D., Schramek, D., Molina, H., et al. (2017). Translation from unconventional 5′ start sites drives tumour initiation. Nature *541*, 494–499.
29. Cloutier, P., Poitras, C., Durand, M., Hekmat, O., Fiola-Masson, É., Bouchard, A., Faubert, D., Chabot, B., and Coulombe, B. (2017). R2TP/Prefoldin-like component RUVBL1/RUVBL2 directly interacts with ZNHIT2 to regulate assembly of U5 small nuclear ribonucleoprotein. Nat. Commun. *8*, 15615.
30. Vainberg, I.E., Lewis, S.A., Rommelaere, H., Ampe, C., Vandekerckhove, J., Klein, H.L., and Cowan, N.J. (1998). Prefoldin, a chaperone that delivers unfolded proteins to cytosolic chaperonin. Cell *93*, 863–873.
31. Siegert, R., Leroux, M.R., Scheufler, C., Hartl, F.U., and Moarefi, I. (2000). Structure of the molecular chaperone prefoldin: unique interaction of multiple coiled coil tentacles with unfolded proteins. Cell *103*, 621–632.
32. Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science *324*, 218–223.
33. D'Lima, N.G., Ma, J., Winkler, L., Chu, Q., Loh, K.H., Corpuz, E.O., Budnik, B.A., Lykke-Andersen, J., Saghatelian, A., and Slavoff, S.A. (2017). A human microprotein that interacts with the mRNA decapping complex. Nat. Chem. Biol. *13*, 174–180.
34. Anderson, D.M., Anderson, K.M., Chang, C.L., Makarewich, C.A., Nelson, B.R., McAnally, J.R., Kasaragod, P., Shelton, J.M., Liou, J., Bassel-Duby, R., et al. (2015). A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. Cell *160*, 595–606.
35. Ouspenskaia, T., Law, T., Clauser, K.R., Klaeger, S., Sarkizova, S., Aguet, F., Li, B., Christian, E., Knisbacher, B.A., Le, P.M., et al. (2022). Unannotated proteins expand the MHC-I-restricted immunopeptidome in cancer. Nat. Biotechnol. *40*, 209–217.
36. Stein, C.S., Jadiya, P., Zhang, X., McLendon, J.M., Abouassaly, G.M., Witmer, N.H., Anderson, E.J., Elrod, J.W., and Boudreau, R.L. (2018). Mitoregulin: A lncRNA-encoded microprotein that supports mitochondrial supercomplexes and respiratory efficiency. Cell Rep. *23*, 3710–3720.e8.
37. Calvo, S.E., Pagliarini, D.J., and Mootha, V.K. (2009). Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proc. Natl. Acad. Sci. USA. *106*, 7507–7512.
38. Whiffin, N., Karczewski, K.J., Zhang, X., Chothani, S., Smith, M.J., Evans, D.G., Roberts, A.M., Quaife, N.M., Schafer, S., Rackham, O., et al. (2020). Characterising the loss-of-function impact of 5′ untranslated region variants in 15,708 individuals. Nat. Commun. *11*, 2523.
39. Neville, M.D.C., Kohze, R., Erady, C., Meena, N., Hayden, M., Cooper, D.N., Mort, M., and Prabakaran, S. (2021). A platform for curated products from novel open reading frames prompts reinterpretation of disease variants. Genome Res. *31*, 327–336.
40. Khan, Y.A., Jungreis, I., Wright, J.C., Mudge, J.M., Choudhary, J.S., Firth, A.E., and Kellis, M. (2020). Evidence for a novel overlapping coding sequence in POLG initiated at a CUG start codon. BMC Genet. *21*, 25.
41. Loughran, G., Zhdanov, A.V., Mikhaylova, M.S., Rozov, F.N., Datskevich, P.N., Kovalchuk, S.I., Serebryakova, M.V., Kiniry, S.J., Michel, A.M., O'Connor, P.B.F., et al. (2020). Unusually efficient CUG initiation of an overlapping reading frame in POLG mRNA yields novel protein POLGARF. Proc. Natl. Acad. Sci. USA. *117*, 24936–24946.
42. Yosten, G.L.C., Liu, J., Ji, H., Sandberg, K., Speth, R., and Samson, W.K. (2016). A 5′-upstream short open reading frame encoded peptide regulates angiotensin type 1a receptor production and signalling via the β-arrestin pathway. J. Physiol. *594*, 1601–1605.
43. Huang, N., Li, F., Zhang, M., Zhou, H., Chen, Z., Ma, X., Yang, L., Wu, X., Zhong, J., Xiao, F., et al. (2021). An upstream open reading frame in phosphatase and tensin homolog encodes a circuit breaker of lactate metabolism. Cell Metab. *33*, 128–144.e9.
44. Rathore, A., Chu, Q., Tan, D., Martinez, T.F., Donaldson, C.J., Diedrich, J.K., Yates, J.R., 3rd, and Saghatelian, A. (2018). MIEF1 microprotein regulates mitochondrial translation. Biochemistry *57*, 5564–5575.
45. Slavoff, S.A., Mitchell, A.J., Schwaid, A.G., Cabili, M.N., Ma, J., Levin, J.Z., Karger, A.D., Budnik, B.A., Rinn, J.L., and Saghatelian, A. (2013).

Peptidomic discovery of short open reading frame-encoded peptides in human cells. Nat. Chem. Biol. *9*, 59–64.

46. Ma, J., Ward, C.C., Jungreis, I., Slavoff, S.A., Schwaid, A.G., Neveu, J., Budnik, B.A., Kellis, M., and Saghatelian, A. (2014). Discovery of human sORF-encoded polypeptides (SEPs) in cell lines and tissue. J. Proteome Res. *13*, 1757–1765.
47. Omenn, G.S., Lane, L., Overall, C.M., Corrales, F.J., Schwenk, J.M., Paik, Y.-K., Van Eyk, J.E., Liu, S., Snyder, M., Baker, M.S., et al. (2018). Progress on identifying and characterizing the human proteome: 2018 metrics from the HUPO human proteome project. J. Proteome Res. *17*, 4031–4041.
48. Malynn, B.A., de Alboran, I.M., O'Hagan, R.C., Bronson, R., Davidson, L., DePinho, R.A., and Alt, F.W. (2000). N-myc can functionally replace c-myc in murine development, cellular growth, and differentiation. Genes Dev. *14*, 1390–1399.
49. Murphy, D.M., Buckley, P.G., Bryan, K., Das, S., Alcock, L., Foley, N.H., Prenter, S., Bray, I., Watters, K.M., Higgins, D., et al. (2009). Global MYCN transcription factor binding analysis in neuroblastoma reveals association with distinct E-box motifs and regions of DNA hypermethylation. PLoS One *4*, e8154.
50. Mukherjee, B., Morgenbesser, S.D., and DePinho, R.A. (1992). Myc family oncoproteins function through a common pathway to transform normal cells in culture: cross-interference by Max and trans-acting dominant mutants. Genes Dev. *6*, 1480–1492.
51. Henriksson, M., and Lüscher, B. (1996). Proteins of the Myc network: essential regulators of cell growth and differentiation. Adv. Cancer Res. *68*, 109–182.
52. Nusinow, D.P., Szpyt, J., Ghandi, M., Rose, C.M., McDonald, E.R., 3rd, Kalocsay, M., Jané-Valbuena, J., Gelfand, E., Schweppe, D.K., Jedrychowski, M., et al. (2020). Quantitative proteomics of the cancer cell line encyclopedia. Cell *180*, 387–402.e16.
53. Cloutier, P., Poitras, C., Faubert, D., Bouchard, A., Blanchette, M., Gauthier, M.S., and Coulombe, B. (2020). Upstream ORF-encoded ASDURF is a novel prefoldin-like subunit of the PAQosome. J. Proteome Res. *19*, 18–27.
54. Wang, P., Zhao, J., Yang, X., Guan, S., Feng, H., Han, D., Lu, J., Ou, B., Jin, R., Sun, J., et al. (2015). PFDN1, an indicator for colorectal cancer prognosis, enhances tumor cell proliferation and motility through cytoskeletal reorganization. Med. Oncol. *32*, 264.
55. Dai, Y.-H., Li, F., Kong, W.-J., Zhang, X.-Q., Wang, M., Ma, H.-L., and Wang, Q. Identification of prognostic biomarkers and independent indicators among PFDN1/2/3/4/5/6 in liver hepatocellular carcinoma. Preprint at Research Square. https://www.researchsquare.com/article/rs-725619/v1
56. Zhou, C., Guo, Z., Xu, L., Jiang, H., Sun, P., Zhu, X., and Mu, X. (2020). PFND1 predicts poor prognosis of gastric cancer and promotes cell metastasis by activating the Wnt/β-catenin pathway. Onco Targets Ther. *13*, 3177–3186.
57. Simons, C.T., Staes, A., Rommelaere, H., Ampe, C., Lewis, S.A., and Cowan, N.J. (2004). Selective contribution of eukaryotic prefoldin subunits to actin and tubulin binding. J. Biol. Chem. *279*, 4196–4203.
58. Cao, X., Khitun, A., Luo, Y., Na, Z., Phoodokmai, T., Sappakhaw, K., Olatunji, E., Uttamapinant, C., and Slavoff, S.A. (2021). Alt-RPL36 downregulates the PI3K-AKT-mTOR signaling pathway by interacting with TMEM24. Nat. Commun. *12*, 508.
59. Na, Z., Dai, X., Zheng, S.J., Bryant, C.J., Loh, K.H., Su, H., Luo, Y., Buhagiar, A.F., Cao, X., Baserga, S.J., et al. (2022). Mapping subcellular localizations of unannotated microproteins and alternative proteins with MicroID. Mol. Cell *82*, 2900–2911.e7.
60. Pauli, A., Norris, M.L., Valen, E., Chew, G.L., Gagnon, J.A., Zimmerman, S., Mitchell, A., Ma, J., Dubrulle, J., Reyon, D., et al. (2014). Toddler: an embryonic signal that promotes cell movement via apelin receptors. Science *343*, 1248636.
61. Zheng, C., Wei, Y., Zhang, P., Xu, L., Zhang, Z., Lin, K., Hou, J., Lv, X., Ding, Y., Chiu, Y., et al. (2023). CRISPR/Cas9 screen uncovers functional translation of cryptic lncRNA-encoded open reading frames in human cancer. J. Clin. Invest. *133*, e159940.
62. Zhang, H., Wang, Y., Wu, X., Tang, X., Wu, C., and Lu, J. (2021). Determinants of genome-wide distribution and evolution of uORFs in eukaryotes. Nat. Commun. *12*, 1076.
63. Tsherniak, A., Vazquez, F., Montgomery, P.G., Weir, B.A., Kryukov, G., Cowley, G.S., Gill, S., Harrington, W.F., Pantel, S., Krill-Burger, J.M., et al. (2017). Defining a cancer dependency map. Cell *170*, 564–576.e16.
64. Delaidelli, A., Negri, G.L., Jan, A., Jansonius, B., El-Naggar, A., Lim, J.K.M., Khan, D., Oo, H.Z., Carnie, C.J., Remke, M., et al. (2017). MYCN amplified neuroblastoma requires the mRNA translation regulator eEF2 kinase to adapt to nutrient deprivation. Cell Death Differ. *24*, 1564–1576.
65. Dassi, E., Greco, V., Sidarovich, V., Zuccotti, P., Arseni, N., Scaruffi, P., Tonini, G.P., and Quattrone, A. (2015). Translational compensation of genomic instability in neuroblastoma. Sci. Rep. *5*, 14364.
66. Schepke, E., Löfgren, M., Pietsch, T., Olsson Bontell, T., Kling, T., Wenger, A., Ferreyra Vega, S., Danielsson, A., Dosa, S., Holm, S., et al. (2022). DNA methylation profiling improves routine diagnosis of paediatric central nervous system tumours: a prospective population-based study. Neuropathol. Appl. Neurobiol. *48*, e12838.
67. Patel, S.K., Hartley, R.M., Wei, X., Furnish, R., Escobar-Riquelme, F., Bear, H., Choi, K., Fuller, C., and Phoenix, T.N. (2020). Generation of diffuse intrinsic pontine glioma mouse models by brainstem-targeted in utero electroporation. Neuro Oncol. *22*, 381–392.
68. Krueger, F., James, F., Ewels, P., Afyounian, E., Weinstein, M., Schuster-Boeckler, B., and Hulselmans, G. (2023). FelixKrueger/TrimGalore, v0.6.10 - add default decompression path. Zenodo. https://doi.org/10.5281/zenodo.5127899.
69. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal *17*, 10–12.
70. Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at http://www.bioinformatics.babraham.ac.uk/projects/fastqc
71. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.
72. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930.
73. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.
74. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. PNAS *102*, 15545–15550.
75. Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M.N., and Sergushichev, A. (2021). Fast gene set enrichment analysis. bioRxiv. https://www.biorxiv.org/content/10.1101/060012v3.
76. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.
77. Calviello, L., Sydow, D., Harnett, D., and Ohler, U. (2019). Ribo-seQC: comprehensive analysis of cytoplasmic and organellar ribosome profiling data. bioRxiv. https://doi.org/10.1101/601468.
78. Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods *14*, 417–419.
79. Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., et al. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat. Biotechnol. *34*, 184–191.

80. Bhuva, D., Smyth, G., and Garnham, A. (2023). msigdb: an ExperimentHub package for the molecular signatures database (MSigDB). R package version 1.10.0. https://bioconductor.org/packages/msigdb.

81. Clement, K., Rees, H., Canver, M.C., Gehrke, J.M., Farouni, R., Hsu, J.Y., Cole, M.A., Liu, D.R., Joung, J.K., Bauer, D.E., et al. (2019). CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat. Biotechnol. *37*, 224–226.

82. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

83. Kawauchi, D., Ogg, R.J., Liu, L., Shih, D.J.H., Finkelstein, D., Murphy, B.L., Rehg, J.E., Korshunov, A., Calabrese, C., Zindy, F., et al. (2017). Novel MYC-driven medulloblastoma models from multiple embryonic cerebellar cells. Oncogene *36*, 5231–5242.

84. Palomar-Siles, M., Heldin, A., Zhang, M., Strandgren, C., Yurevych, V., van Dinter, J.T., Engels, S.A.G., Hofman, D.A., Öhlin, S., Meineke, B., et al. (2022). Translational readthrough of nonsense mutant TP53 by mRNA incorporation of 5-fluorouridine. Cell Death Dis. *13*, 997.

85. McGlincy, N.J., and Ingolia, N.T. (2017). Transcriptome-wide measurement of translation by ribosome profiling. Methods *126*, 112–129.

86. Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J.P. (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics *27*, 1739–1740.

87. Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The molecular signatures database (MSigDB) hallmark gene set collection. Cell Syst. *1*, 417–425.

88. Morales, J., Pujar, S., Loveland, J.E., Astashyn, A., Bennett, R., Berry, A., Cox, E., Davidson, C., Ermolaeva, O., Farrell, C.M., et al. (2022). A joint NCBI and EMBL-EBI transcript set for clinical genomics and research. Nature *604*, 310–315.

89. Blighe, K., Rana, S., and Lewis, M. (2023). EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. doi:10.18129/B9.bioc.EnhancedVolcano. R package version 1.20.0. http://bioconductor.org/packages/EnhancedVolcano/.

90. Sturm, G., Finotello, F., and List, M. (2020). Immunedeconv: an R package for unified access to computational methods for estimating immune cell fractions from bulk RNA-sequencing data. Methods Mol. Biol. *2120*, 223–232.

91. Piccioni, F., Younger, S.T., and Root, D.E. (2018). Pooled lentiviral-delivery genetic screens. Curr. Protoc. Mol. Biol. *121*, 32.1.1–32.1.21.

92. Richter, M.F., Zhao, K.T., Eton, E., Lapinaite, A., Newby, G.A., Thuronyi, B.W., Wilson, C., Koblan, L.W., Zeng, J., Bauer, D.E., et al. (2020). Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. Nat. Biotechnol. *38*, 883–891.

93. Madeira, F., Pearce, M., Tivey, A.R.N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., and Lopez, R. (2022). Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic Acids Res. *50*, W276–W279.

94. Meehan, T.F., Conte, N., West, D.B., Jacobsen, J.O., Mason, J., Warren, J., Chen, C.K., Tudose, I., Relac, M., Matthews, P., et al. (2017). Disease model discovery from 3,328 gene knockouts by the International Mouse Phenotyping Consortium. Nat. Genet. *49*, 1231–1238.

95. Koscielny, G., Yaikhom, G., Iyer, V., Meehan, T.F., Morgan, H., Atienza-Herrero, J., Blake, A., Chen, C.K., Easty, R., Di Fenza, A., et al. (2014). The international mouse phenotyping consortium web portal, a unified point of access for knockout mice and related phenotyping data. Nucleic Acids Res. *42*, D802–D809.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit Monoclonal anti V5 (D3H8Q) (1:2500)	Cell Signaling Technology	Cat#13202S; RRID: AB_2687461
Mouse Monoclonal anti V5 [SV5-Pk1] (1:2500)	Abcam	Cat#ab27671; RRID: AB_471093
Rabbit Polyclonal anti GFP (1:2500)	Cell Signaling Technology	Cat#2555S; RRID: AB_10692764
Rabbit Polyclonal anti PFDN1 (1:1000)	Millipore Sigma	Cat#HPA006499; RRID: AB_1079596
Rabbit Polyclonal anti PFDN2 (1:500)	Millipore Sigma	Cat#HPA028700; RRID: AB_10603983
Rabbit Polyclonal anti PFDN5 (1:500)	Millipore Sigma	Cat#HPA008587; RRID: AB_1079597
Rabbit Polyclonal anti PFDN6 (1:1000)	Millipore Sigma	Cat#HPA043032; RRID: AB_2678278
Mouse Monoclonal anti Alpha-tubulin (1:2000)	Abcam	Cat#ab7291; RRID: AB_2241126
Rabbit Polyclonal anti Beta-tubulin (1:2000)	Cell Signaling Technology	Cat#2146S; RRID: AB_2210545
Rabbit Polyclonal anti Gamma-tubulin (1:1000)	Bethyl Laboratories	Cat#A302-631A-M; RRID: AB_2780661
Rabbit Monoclonal anti HSP90 (1:1000)	Cell Signaling Technology	Cat#4877S; RRID: AB_2233307
Rabbit Monoclonal anti Vinculin (1:1000)	Abcam	Cat#ab219649; RRID: AB_2819348
Rabbit Monoclonal anti GAPDH (1:2000)	Cell Signaling Technology	Cat#2118L; RRID: AB_561053
Rabbit Polyclonal anti Beta Galactosidase (1:2000)	Abcam	Cat#Ab616; RRID: AB_305327
Mouse Monoclonal anti Beta-Actin (1:4000)	Sigma-Aldrich	Cat#A5316; RRID: AB_476743
Goat anti-mouse secondary (1:5000)	LI-COR	Cat#926-32210; RRID: AB_621842
Goat anti-rabbit secondary (1:5000)	LI-COR	Cat#926-68021; RRID: AB_10706309
Mouse Monoclonal anti Beta-Actin (1:4000)	Sigma-Aldrich	Cat#A5316; RRID: AB_476743
Goat anti-mouse secondary (1:5000)	LI-COR	Cat#926-32210; RRID: AB_621842
Goat anti-rabbit secondary (1:5000)	LI-COR	Cat#926-68021; RRID: AB_10706309
eGFP antibody	Aves	#GFP1020; RRID: AB_10000240
Normal rabbit IgG	Cell Signaling Technology	Cat#2729S; RRID: AB_1031062
THE™ Anti-GST Monoclonal Antibody (1:2000)	GenScript	Cat#A00865; RRID: AB_914654
Anti-6X His tag® antibody (1:2000)	Abcam	Cat#ab9108; RRID: AB_307016
Bacterial and virus strains		
T7 Express Competent E. coli	New England BioLabs	Cat#C2566H
Plx_311 Cas9 lentivirus	Addgene	Plasmid 96924; RRID: Addgene_96924
Chemicals, peptides, and recombinant proteins		
DMEM	Invitrogen, Carlsbad, CA	10569044
FBS (Fetal Bovine Serum)	Invitrogen, Carlsbad, CA	16140071
Penicillin-Streptomycin-Glutamine	Invitrogen, Carlsbad, CA	10378016
RPMI 1640	Invitrogen, Carlsbad, CA	61870036
IMDM	Invitrogen, Carlsbad, CA	31980030
DMEM/F12	Invitrogen, Carlsbad, CA	10565018
Neurobasal-A medium	Invitrogen, Carlsbad, CA	10888022
HEPES (1M)	Invitrogen, Carlsbad, CA	15630130
Sodium Pyruvate	Invitrogen, Carlsbad, CA	11360070
MEM non-essential amino acids	Invitrogen, Carlsbad, CA	11140050
GlutaMax	Invitrogen, Carlsbad, CA	35050079
B27 supplement	Invitrogen, Carlsbad, CA	17504044
Human EGF	StemCell Technologies	78006.1

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Human FGF-basic-154	StemCell Technologies	78003
Heparin solution (2ug/mL)	StemCell Technologies	07980
RIPA lysis buffer	Sigma-Aldrich, St. Louis, MO	R0278
Bicinchoninic acid (BCA)	Thermo Fisher Scientific, Waltham, MA	23225
Sample loading buffer	LI-COR	928–40004
Tris-Glycine 10-20% SDS-PAGE gels	Thermo Fisher Scientific, Waltham, MA	XP1020A
Bis-Tris 4-12% SDS-PAGE gels	Thermo Fisher Scientific, Waltham, MA	NW04120BOX
Qiazol	Qiagen	79306
DNase I	Qiagen	79254
Superscript III	Invitrogen, Carlsbad, CA	12574026
Random primers	Invitrogen, Carlsbad, CA	4819001
Turbo-DNase I	Invitrogen, Carlsbad, CA	AM2238
CircLigase Buffer	Lucigen	CL4115K
SYBR-Green mastermix	ThermoFisher Scientific	4367659
Phusion HiFi master mix	New England Biolabs	M0531S
T4 PNK	Lucigen	P0503K
R4 RNA ligase 2 Deletion mutant	Lucigen	LR2D11310K
T4 RNA ligase I	New England Biolabs	M0204S
5′ deadenylase	New England Biolabs	M0331S
Rec J Exonuclease	New England Biolabs	M0264S
EpiScript RT enzyme	Lucigen	ERT12925K
Exonuclease I	Lucigen	X40520K
Rnase I	Lucigen	N6901K
Hybridase	Lucigen	H39500
CircLigase I	Lucigen	CL4115K
Trizol	Thermofisher Scientific	15596026
Puromycin	Thermofisher Scientific	J68236-XF
SimplyBlue Coomassie stain	Thermo Fisher Scientific	LC6065
SybrGold 10000x	Thermo Fisher Scientific	S11494
Magnetic anti-V5 beads	MBL International	M167-11
Anti-FLAG(R) M2 Magnetic Beads	Sigma-Aldrich	M8823
Hygromycin	Thermo Fisher Scientific	10687010
Fugene HD	Promega	E2311
AMPur beads	Beckman Coulter	A63880
OptiMem	Thermo Fisher Scientific	31985062
V5 peptide	Sigma-Aldrich	V7754
FLAG peptide	ApexBio	A6001
LDS sample buffer	Thermo Fisher Scientific	NP0007
Sample-reducing agent	Thermo Fisher Scientific	NP0004
EZview Red Protein A bead affinity gel slurry	Sigma-Aldrich	P6486
EZview Red Protein G bead affinity gel slurry	Sigma-Aldrich	E3403
Carbenicillin	Thermo Fisher Scientific	10177012
IPTG	Thermo Fisher Scientific	15529019
Polybrene	Sigma Aldrich	TR-1003
Halt™ Protease Inhibitor Cocktail	Thermo Fisher Scientific	78429

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical commercial assays		
Lonza MycoAlert assay	Lonza	LT07-701
miRNeasy Kit	Qiagen, Hilden, Germany	217004
Human RiboPool kit	siTOOLS Biotech, Germany	dp-R096-000042
Direct-zol RNA miniprep kit	Zymo	R2050
RNA Clean and Concentrator kit	Zymo	R1013
Oligo Clean and Concentrator Kit	Zymo	D4060
HS DNA High Sensitivity Kit	Agilent	5067-4626
Qiagen RNAeasy kit	Qiagen	74104
Roche Kapa mRNA Hyper Prep kit	Roche, Basel, Switzerland	KK8581
Roche Kapa DNA Hyper Prep kit	Roche, Basel, Switzerland	KK8504
Plasmid Plus MIDI kit	Qiagen	12941
Cell-Titer Glo assay	Promega	G7570
P3 Primary Cell 4D X Kit	Lonza	V4XP-3032
Lenti-X Concentrator	Takara Bio	631232
Pierce™ GST Protein Interaction Pull-Down Kit	Thermo Fisher Scientific	Cat#21516
Qiagen DNeasy Blood & Tissue Kit	Qiagen	69504
BrdU Cell Proliferation ELISA kit	Abcam	cat#ab126556
Qiagen Gel Extraction kit	Qiagen	28704
Deposited data		
Human reference genome Ensembl release 102, GRCh38	Ensembl	https://ftp.ensembl.org/pub/release-102/
Custom code	This paper	Zenodo: https://zenodo.org/records/8319309
Immunoblot images	This paper	Mendeley: https://data.mendeley.com/datasets/5vx9fk85zt/1
Immunoprecipitation and immunoblot images	This paper	Mendeley: https://data.mendeley.com/datasets/d63f7yzk3j/1
Identifiable human sequencing data from 4 medulloblastoma patients	This paper	European Genome-phenome Archive (EGA) EGAS00001007426
Identifiable human sequencing data from North American medulloblastoma patients	This paper	dbGAP phs003446
Non-identifiable human sequencing data from medulloblastoma cell lines	This paper	NCBI Short Read Archive bioproject PRJNA957428
GENCODE Phase1 Ribo-seq ORFs	GENCODE	https://www.gencodegenes.org/pages/riboseq_orfs/
GTEx RNA-seq data	Broad Institute of MIT and Harvard	https://www.gtexportal.org/home/
Allen Institute Developing Brain Atlas	Allen Institute	https://www.brainspan.org/
Cancer Cell Line Encyclopedia proteomics data	Nusinow et al.[52]	Table S2 in the original manuscript
Cancer Cell Line Encyclopedia RNA-Seq Data	Broad Institute of MIT and Harvard	https://portals.broadinstitute.org/ccle
DepMap_public_21Q2 release	Broad Institute of MIT and Harvard	https://depmap.org/portal/download/all/
ASNSD1-uORF knockout experiments RNA-seq data	This paper	NCBI Short Read Archive bioproject PRJNA95742
Archer et al, Medulloblastoma patient tissue RNAseq data	Archer et al.[17]	European Genome-phenome Archive EGAS00001001953
Archer et al, Medulloblastoma patient tissue mass spectrometry data	Archer et al.[17]	ftp://massive.ucsd.edu/MSV000082644
Experimental models: Cell lines		
CHLA-259	Children's Oncology Group	Cat#CHLA-259; RRID: CVCL_M148
MB002	Bandopadhayay lab	RRID: CVCL_VU79

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
H9-NSCs	Invitrogen	Cat#N7800-100; RRID: CVCL_IU37
HEK293T	ATCC	Cat#CRL-3216; RRID: CVCL_0063
D283Med	ATCC	Cat#HTB-185; RRID: CVCL_1155
D341	ATCC	Cat#HTB-187; RRID: CVCL_0018
JIMT1	CCLE	RRID: CVCL_2077
D425	Bandopadhayay lab	RRID: CVCL_1275
D458	Bandopadhayay lab	RRID: CVCL_1161
D384	CCLE	RRID: CVCL_1157
DAOY	ATCC	Cat#HTB-186; RRID: CVCL_1167
R262	CCLE	RRID: CVCL_VU83
R256	CCLE	RRID: CVCL_DG09
UW228	CCLE	RRID: CVCL_8585
RPE10	CCLE	RRID: CVCL_4388
MCF7	ATCC	Cat#HTB-22; RRID: CVCL_0031
MDA-MB-231	ATCC	Cat#CRM-HTB-26; RRID: CVCL_0062
HCC1806	ATCC	Cat#CRL-2335; RRID: CVCL_1258
HCC1954	ATCC	Cat#CRL-2338; RRID: CVCL_1259
HCC95	CCLE	RRID: CVCL_5137
HCC15	CCLE	RRID: CVCL_2057
A549	ATCC	Cat#CRM-CCL-185; RRID: CVCL_0023
Jurkat	ATCC	Cat#TIB-152; RRID: CVCL_0065
ES2	ATCC	Cat#CRL-1978; RRID: CVCL_AX39
MIAPACA2	ATCC	Cat#CRM-CRL-1420; RRID: CVCL_0428
SNU503	CCLE	RRID: CVCL_5071
HT29	ATCC	Cat#HTB-38; RRID: CVCL_0320
KYSE410	CCLE	RRID: CVCL_1352
KYSE510	CCLE	RRID: CVCL_1354
ONS76	CCLE	RRID: CVCL_1624
A375	ATCC	Cat#CRL-1619; RRID: CVCL_0132
HS294T	ATCC	Cat#HTB-140; RRID: CVCL_0331
LOXIMVI	Millipore Sigma	Cat#SCC201; RRID: CVCL_1381
CHLA-02-ATRT	ATCC	Cat#CRL-3020; RRID: CVCL_B045
CHLA-05-ATRT	ATCC	Cat#CRL-3037; RRID: CVCL_AQ41
CHLA-06-ATRT	ATCC	Cat#CRL-3038; RRID: CVCL_AQ42
CHLA-01-MED	ATCC	Cat#CRL-3021; RRID: CVCL_B044
CHLA-01-MEDR	ATCC	Cat#CRL-3034; RRID: CVCL_N534
Med2112-mCherry-Luc	Brain Tumor Resource Lab	https://www.seattlechildrens.org/research/centers-programs/childhood-cancer/our-labs/jim-olson-lab/btrl/
Med411-GFP-Luc	Brain Tumor Resource Lab	https://www.seattlechildrens.org/research/centers-programs/childhood-cancer/our-labs/jim-olson-lab/btrl/
Experimental models: Organisms/strains		
Mouse: NSG	The Jackson Laboratory, Bar Harbor, ME	RRID:IMSR_JAX:005557
Oligonucleotides		
CAGAGTTCTACAGTCCGACGAT	This paper	JRP_qPCR-ribo-F2 primer
AGACGTGTGCTCTTCCGATCT	This paper	JRP_qPCR-ribo-R2 primer
AATGATACGGCGACCACCGAGATCTAC ACGTTCAGAGTTCTACAGTCCGACG	This paper	JRP_ribo-seq_lib_forward

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Barcoded reverse primers used for Ribo-seq, see Table S5	This paper	JRP_ribo-seq_lib_reverse
gRNAs used in primary and validation CRISPR screens, see Table S2	This paper	see Table S2
Primers used for qPCR of prefoldin-like complex members, see Table S5P	This paper	see Table S5P
gRNAs used for base editing, see Table S2M	This paper	see Table S2M
sgRNAs used for ASNSD1-uORF knockout experiments, see Table S5Q	This paper	see Table S5Q
AACGGCGGATTGACCGTAAT	This paper	LacZ control gRNA
GGTGTGCGTATGAAGCAGTGG	This paper	Chr2-2 cutting control gRNA
GCTTAGATCCTCCTTGTGTG	This paper	ASNSD1-uORF #1 gRNA
TAAAGAACAAAAAATTGTGG	This paper	ASNSD1-uORF #2 gRNA
Recombinant DNA		
ASNSD1-uORF and MYCN cDNAs	This paper	see Table S5O
PiggyBac transposase DNA plasmids with luciferase and IRES-GFP	Patel et al.[67]	https://academic.oup.com/neuro-oncology/article/22/3/381/5602255
pCAG-PBase transposase plasmid	Patel et al.[67]	https://academic.oup.com/neuro-oncology/article/22/3/381/5602255
pET-23a(+) vector (T7 promoter, 6xHis-tag)	GenScript	https://www.genscript.com/expression-vector-selection-guide.html?page_no=1&position_no=1&sensors=googlesearch
pET-GST vector (T7 promoter, GST tag)	GenScript	https://www.genscript.com/expression-vector-selection-guide.html?page_no=1&position_no=1&sensors=googlesearch
Software and algorithms		
Original code used for analyses	This paper	https://doi.org/10.5281/zenodo.8319308
TrimGalore v0.6.6	Krueger et al.[68]	https://github.com/FelixKrueger/TrimGalore
Cutadapt v3.4	Martin[69]	https://github.com/marcelm/cutadapt
FastQC v0.11.9	Andrews et al.[70]	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
STAR v2.7.8a	Dobin et al.[71]	https://github.com/alexdobin/STAR
featureCounts v2.0.2	Liao et al.[72]	https://subread.sourceforge.net/featureCounts.html
R v4.0.3	R Project	https://www.r-project.org/
DESeq2	Love et al.[73]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
msigdb R package	Subramanian et al.[74]	https://bioconductor.org/packages/release/data/experiment/html/msigdb.html
fgsea R package	Korotkevich et al.[75]	https://bioconductor.org/packages/release/bioc/html/fgsea.html
Bowtie2 v2.4.2	Langmead and Salzberg[76]	https://github.com/BenLangmead/bowtie2
RiboseQC R package	Calviello et al.[77]	https://github.com/ohlerlab/RiboseQC
Salmon v1.8.0	Patro et al.[78]	https://salmon.readthedocs.io/en/latest/salmon.html
CRISPick	Doench et al.[79]	https://portals.broadinstitute.org/gppx/crispick/public
ClustralOmega package	Madeira et al[80]	https://www.ebi.ac.uk/Tools/msa/clustalo/
GraphPad PRISM v10.0.02	GraphPad software	https://www.graphpad.com/guides/prism/latest/user-guide/index.htm
CRISPResso	Clement et al.[81]	http://crispresso.pinellolab.partners.org

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Spectrum Mill v.7.09	Broad Institute Proteomics Platform	https://proteomics.broadinstitute.org/millhtml/SM_slides/SpectrumMillOverview.pdf
NCBI DAVID Bioinformatics platform	NCBI	https://david.ncifcrf.gov/tools.jsp
STRING-db tool	STRING-db	www.string-db.org
Bedtools v2.25.0	Quinlan and Hall[82]	https://github.com/arq5x/bedtools2

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, John Prensner (prensner@umich.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

All raw sequencing data and custom code are publicly available.

Ribo-seq and RNA-seq data for medulloblastoma cell lines, including RNA-seq following ASNSD1-uORF and PFDN2 knockout in D425 and D283 cells, are available through the NCBI Short Read Archive through BioProject ID PRJNA957428. Proteomics data for ASNSD1-uORF and PFDN2 knockout in D425 and D283 cells are available in PRIDE/ProteomeXchange as PXD046091. Ribo-seq and RNA-seq data for patient tissue samples from the Dana-Farber Cancer Institute are submitted to the NCBI dbGaP as phs003446.v1.p1. Human Ribo-seq and RNA-seq data for patient tissue samples from the Princess Máxima Center are submitted to the European Genome-Phenome Archive (EGA) and are available under accession number EGAS00001007426. Original western blots are available at Mendeley Data at https://data.mendeley.com/datasets/d63f7yzk3j/1 and https://data.mendeley.com/datasets/5vx9fk85zt/1.

All original code for RNA-seq and Ribo-seq analyses is available through GitHub at https://github.com/damhof/hofman_et_al_2023_seq and has been deposited at Zenodo. DOIs are listed in the key resources table.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Mouse models

For mouse xenografting experiments, our sample size of mice was predetermined based on the optimum number of animals needed to attain statistical significance of $p<0.05$ with a power level of 80 percent.

Murine orthotopic xenograft experiments

Animal experiments were performed after approval by the Broad Institute and the Dana-Farber Institutional Care and Use Committee (IACUC) and were conducted as per NIH guidelines for animal welfare. Animals were housed and cared for according to standard guidelines with free access to water and food. All experiments were performed on n=10 seven weeks-old female NSG mice (NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ, The Jackson Laboratory, Bar Harbor, ME). Mice were allocated to two groups: sgControl (n = 9 mice, after one death during the procedure) and sgASNSD1uORF (n = 10 mice). Mice were euthanized as they developed neurological symptoms. To perform xenografting experiments, animals were injected intraperitoneally with the analgesic buprenorphine 0.05 mg/kg and then anesthetized with isoflurane 2–3% mixed with medical air and placed on a stereotactic frame. Next, a small incision and a small burr hole was made with a 25-gauge needle and D425 cells (60,000 cells in 1 μL PBS) were injected stereotactically into the cerebellum (stereotactic coordinates zeroed on bregma: -1.0 mm X (ML), -7.0 mm Y (AP) and -2.5 mm Z (DV)) of 7 weeks-old female NSG mice at rate of 1 μL/min with use of an infusion pump before the incision was closed. Mice were then checked daily for signs of distress, including seizures, weight loss, or tremors, and euthanized as they developed neurological symptoms, including head tilt, seizures, sudden weight loss, loss of balance, and/or ataxia. Mouse brains collected at the survival endpoint were either fixed in 4% paraformaldehyde for 24 hours and subsequently stored in 70% ethanol and stored at room temperature, or snap-frozen on dry ice and stored at -80 °C.

Murine magnetic resonance imaging

MRI was performed using a Bruker BioSpec 7T/30 cm USR horizontal bore Superconducting Magnet System (Bruker Corp.). This system provides a maximum gradient amplitude of 440 mT/m and slew rate of 3,440 T/m/s and uses a 23 mm ID birdcage volume radiofrequency (RF) coil for both RF excitation and receiving. Mice were anesthetized with 1.5% isoflurane mixed with 2 L/min air flow and positioned on the treatment table using the Bruker AutoPac with laser positioning. Body temperature of the mice was maintained at 37 °C using a warm air fan while on the treatment table, and respiration and body temperature were monitored and regulated using the SAII (Sa Instruments) monitoring and gating system, model 1025T. T2-weighted images of the brain were obtained using a fast spin echo (RARE) sequence with fat suppression. The following parameters were used for image acquisition: repetition time (TR) = 6,000 ms, echo time (TE) = 36 ms, field of view (FOV) = 19.2 x 19.2 mm2, matrix size = 192 x 192, spatial resolution = 100 x 100 μm2, slice thickness = 0.5 mm, number of slices = 29, rare factor = 16, number of averages = 8, and total acquisition time 7:30 min. Bruker Paravision 6.0.1 software was used for MRI data acquisition, and tumor volume was determined from MRI images processed using a semiautomatic segmentation analysis software (ClinicalVolumes).

Murine *in utero* electroporation experiments

For *in utero* electroporation, our sample size of 2-3 pregnant female mice to produce 10 electroporated murine pups per cohort reflects the known penetrance of tumor formation with cMYC and DNp53 with this technique,[67,83] and a sample size of 10 mice per cohort was designed to enable a statistical significance of $p < 0.05$ with a power level of 80 percent. Murine experiments were randomized by alternating treatments between successive mice and the investigators were blinded to allocation during experiments and outcome assessment. *In utero* electroporation (IUE) experiments were performed as previously described.[67,83] Briefly, mouse medulloblastomas are formed by the introduction of cDNAs expressing MYC and dominant negative p53 (DNp53). PiggyBac transposase DNA plasmids have luciferase and an IRES-GFP site for continuous GFP expression. We tested two conditions: DNp53 + MYC and DNp53 + MYC + ASNSD1-uORF. Both conditions included the pCAG-PBase transposase plasmid to stably integrate cDNA expression constructs. Specifically, 1 μg of concentrated DNA plasmid mixtures (1 μg/μL containing 0.05% Fast Green (Sigma)) was injected into the 4th ventricle of E13.5 mouse embryos using a pulled glass capillary pipette. Following DNA injection, embryos were electroporated by applying 5 pulses (45 V, 50 ms pulses with 950 ms intervals) with a 3 mm tweezer electrode positioned at the upper rhombic lip and cerebellar ventricular zone. Once born, pups were imaged via IVIS for luciferase at 1-2 weeks of age to identify successfully electroporated offspring. Mice were monitored every 3 days for new tumor-related neurologic symptoms (e.g. hydrocephalus, altered gait, lethargy, weight loss). Mice with symptoms were then euthanized according to IACUC guidelines. Tumor burden was be confirmed with GFP immunohistochemistry, using 50 uM tissue sections that are blocked in PBS + 0.5% Triton X-100 + 10% normal donkey serum prior to incubation with an antibody for eGFP (Aves, #GFP1020) and Hoechst (Thermo Fisher) for cell nuclei. 10 IUE tumor-bearing offspring were used per condition. The primary endpoint of time-to-death was analyzed using Kaplan-Meier curves with a log-rank test with a two-sided $p<0.05$ being significant. IUE experiments were performed under the University of Cincinnati IACUC approval protocol #16-07-06-01.

Cell lines and reagents

All parental cell lines were obtained directly from the American Type Culture Collection (ATCC, Manassas, VA), from the Bandopadhayay lab (MB002, D425, D458), Broad Institute Cancer Cell Line Encyclopedia (JIMT1, D384, R262, R256, UW228, HCC95, HCC15, SNU503, KYSE410, KYSE510, ONS76, RPE10-1), the Straehla lab (Med2112 and Med411) or from the Children's Oncology Group (CHLA-259). H9-derived neural stem cells were obtained from Invitrogen (Invitrogen, cat# N7800-100). Cas9-derived cell lines were obtained from the Broad Institute. Cell lines were maintained according to established tissue culture media and conditions. HEK293T, D283Med (D283), D341, D384, D425, D458, DAOY, R262, R256, UW228, RPE10-1, JIMT1, MCF7, MDA-MB-231, HCC1806, HCC1954, HCC95, HCC15, A549, JURKAT, ES2, and MIAPACA2 cells were maintained in DMEM supplemented with 10% FBS (Invitrogen, Carlsbad, CA) and 1% penicillin-streptomycin (Invitrogen, Carlsbad, CA) in a 5% CO2 cell culture incubator. SNU503, HT29, KYSE410, KYSE510, ONS76, A375, HS294T, and LOXIMVI cells were maintained in RPMI 1640 (Invitrogen, Carlsbad, CA) supplemented with 10% FBS and 1% penicillin-streptomycin in a 5% CO2 cell culture incubator. CHLA-259 cells were maintained in IMDM (Invitrogen, Carlsbad, CA) supplemented with 20% FBS and 1% penicillin-streptomycin in a 5% CO2 cell culture incubator. CHLA-02-ATRT, CHLA-05-ATRT, CHLA-06-ATRT, CHLA-01-MED, CHLA-01-MEDR, H9-derived NSCs, Med2112 (expressing mCherry and luciferase), Med411 (expressing GFP and luciferase) and MB002 cells were maintained in Tumor Stem Media comprised of DMEM/F12 (1:1) with Neurobasal-A medium (Invitrogen, Carlsbad, CA) and supplemented with HEPES (1M, 0.1% final concentration; Invitrogen, Carlsbad, CA), sodium pyruvate (1mM final concentration; Invitrogen, Carlsbad, CA), MEM non-essential amino acids (0.1mM final concentration; Invitrogen, Carlsbad, CA), GlutaMax (1x final concentration; Invitrogen, Carlsbad, CA), B27 supplement (1x final concentration; Invitrogen, Carlsbad, CA), human EGF (20ng/mL; StemCell Technologies), human FGF-basic-154 (20ng/mL; StemCell Technologies), and heparin solution 0.2% (2ug/mL final concentration, StemCell Technologies). H9-derived NSC cells were cultured on GelTrex-coated tissue culture plates (ThermoFisher). Cell lines were routinely verified via STR genotyping and tested for mycoplasma contamination using the Lonza MycoAlert assay (Lonza). Details of cell lines, including genotypic sex and culture media, are listed in Table S5N.

Primary tissue samples

21 human medulloblastoma tissue samples were obtained from the Boston Children's Hospital BioBank and the Dana-Farber Harvard Cancer Center Neuro-oncology Program and Tumor BioBank. Patient samples were acquired with the informed consent of DFCI protocol 10-417. Four human medulloblastoma tissue samples were obtained from the Princess Máxima Center biobank under approval from the Medical Ethics Committee of the Erasmus Medical Center (ID number, MEC-2016-739). All samples were de-identified prior to use for research. Molecular subtyping of the tissue samples was based on de-identified surgical pathology reports, including histopathology, immunohistochemistry, copy number and mutational profiling, and potentially DNA methylation arrays for more recent samples, where indicated in Table S1A. Except for four tissue samples from the Princess Máxima Center with available sex information, data on sex, gender, ancestry, race, ethnicity, socioeconomic status, and age was not provided. Combined, for 18 out of these 25 tissue samples matched Ribo-seq and RNA-seq data could be obtained and these 18 samples were therefore included in this manuscript. Samples for which Ribo-seq failed due to inadequate sample material (n = 7) were not included.

METHOD DETAILS

Immunoblot Analysis

Cells were grown to 70-80% confluence, collected by scraping the tissue culture dish and washed once in 1x PBS. They were then lysed in RIPA lysis buffer (Sigma-Aldrich, St. Louis, MO) with 1x HALT protease inhibitor (Thermo Fisher Scientific, Waltham, MA) and homogenized by chilling them on ice for 15 minutes. Cellular proteins were separated by centrifugation for 15 minutes at 13,200 RPM and supernatant was saved. Protein lysate yields were determined using bicinchoninic acid (BCA), and appropriate volumes of lysate were prepared for immunoblotting by boiling in a 1x sample loading buffer at 95C for 5 minutes. Tris-Glycine 10-20% or Bis-Tris 4-12% SDS-PAGE gels were run at 4°C and proteins were transferred onto nitrocellulose membranes using 15 Volts for 7 minutes via the iBlot-2 system (Thermo Fisher Scientific, Waltham, MA). The membrane was then blocked for 1 hour in LICOR Odyssey blocking buffer and incubated at 4°C with the appropriate antibody overnight. The blot was then washed 4 times with 1x TBS with 0.1% Tween20 and incubated with fluorophore-specific IRDye secondary antibodies (LI-COR, Lincoln, NE) and imaged on a LI-COR Odyssey machine. The full list of immunoblot antibodies used in this study can be found in the key resources table, with details on staining conditions listed in Table S5O.

RNA isolation and cDNA synthesis

Total RNA was isolated using Qiazol and an miRNeasy Kit (Qiagen, Hilden, Germany) with DNase I digestion according to the manufacturer's instructions. RNA integrity was verified on an Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA). cDNA was synthesized from total RNA using Superscript III (Invitrogen, Carlsbad, CA) and random primers (Invitrogen, Carlsbad, CA). Quantitative Real-time PCR (qPCR) was performed using Power SYBR Green Mastermix (Applied Biosystems, Foster City, CA) on a Thermo QStudio FLX Real-Time PCR System (Thermo Fisher Scientific, Waltham, MA). The relative quantity of the target gene was completed for each sample using the ΔΔCt method by the comparing mean Ct of the gene to the average Ct of the geometric mean of the indicated housekeeping genes (GAPDH, beta-actin, HMBS). The primer sequences are listed in Table S5P.

Ribosome profiling

Ribo-seq for human tissue samples was performed according to the protocol described in Palomar-Siles et al.[84] Ribo-seq for cancer cell lines was performed based upon the protocol by McGlincy et al.[85] with modifications as described below. Briefly, cells were grown to 60-70% confluence prior to collection. After collection, all cell pellets were washed once in 1x PBS, re-pelleted by centrifugation, and lysed in lysis buffer (20mM Tris HCl, 150mM NaCl, 5mM MgCl2, 1mM dithiotrietol, 0.05% NP-40, 25U/mL Turbo-DNase I (Invitrogen), 2ug/mL cycloheximide). After clearing the lysate and recovering the supernatant, RNA abundance was determined by measuring the A260. 2.5U/ug of RNase I was added to an appropriate volume of lysate and incubated at 22C for 45 minutes without shaking. The RNase I was then quenched with 1U/uL of Superase RNase Inhibitor (Ambion). RNA from ribosome protected fragments were recovered using a 1M sucrose cushion with ultracentrifugation (55,000 RPM, 4C, 2 hours), and rRNA was depleted using the siTOOLS human RiboPool kit according to manufacturer's instructions (siTOOLS Biotech, Germany). Ribosome protected fragments were then denatured using a 1:1 mixture with 2x sample loading buffer (98% v/v formamide, 10mM EDTA, 300ug/mL bromophenol blue) at 95C for 3 minutes, and further purified using size selection from a 15% TBE-Urea gel (200V for 65 minutes). The 26 – 32 nucleotide band was cut from the gel, RNA extracted by freezing gel slices in 400uL RNA gel extraction buffer (300mM NaOAc, 1mM EDTA, 0,25% v/v SDS), and rotating at room temperature for 5-6 hours. RNA was precipitated with 500uL isopropanol and 2.0uL GlycoBlue at -20C overnight; pellets were washed once in chilled 70% ethanol, and subjected to end-repair with T4 PNK (Lucigen, 37C for 1 hr). End-repaired RNA was cleaned up with the RNA Clean and Concentrator kit (Zymo), ligated to a 3′ linker (sequence below, 6.67% w/v PEG-8000, 6.67 mM dithiotrietol, 1x T4 RNL2 Truncation buffer, 6.67 U/uL R4 RNA ligase 2 Deletion mutant, 0.33 U/uL T4 RNA ligase I) for 3 hours at room temperature. Linker reactions were removed with 5′ deadenylase (New England Biolabs) and Rec J Exonuclease (NEB), and cDNA was generated with EpiScript RT enzyme (Lucigen, 50C for 30 minutes) followed by reaction clean up with exonuclease I (Lucigen, 37C for 30 minutes), RNase I/Hybridase (Lucigen, 55C for 5 minutes) and the Oligo Clean and Concentrator Kit (Zymo). cDNA was mixed 1:1 with 2x sample loading buffer, boiled, and purified with a 10% TBE-Urea gel (70 minutes, 175V). The product between 70 – 90 nucleotides was excised from the gel, and DNA was extracted with 450uL DNA extraction buffer

(300mM NaCl, 10mM Tris, 1mM EDTA, 0.02% SDS) with a flash-freeze on dry ice (30 minutes) and rotation at 22C for 6 hours. DNA was precipitated with 700uL isopropanol and 2uL GlycoBlue at -80C overnight followed by centrifugation at 14,500 RPM for 45 minutes at 4C. DNA pellets were washed once in 80% ethanol and pellets were air-dried and dissolved in 11uL of water, which was then circularized with the addition of 9uL of CircLigase I mix (1M betaine, 1x CircLigase Buffer (Lucigen), 2.5mM MnCl2, 50uM ATP, 5U/uL CircLigase I (Lucigen)) at 60C for 3 hours with heat inactivation at 80C for 10 minutes. Circularized cDNA was quantified using quantitative real-time PCR (10uL of 2x SYBR-Green mastermix (Thermo), 2uL of cDNA, 6uL water, 1uL forward and reverse primer each) for twenty cycles, using the following PCR primers (JRP_qPCR-ribo-F2 primer: CAGAGTTCTACAGTCCGACGAT; JRP_qPCR-ribo-R2 primer: AGACGTGTGCTCTTCCGATCT). Library PCR amplification was performed with 10uL of 2x Phusion HiFi master mix (New England Biolabs), 8uL of cDNA sample, 1uL of the forward library primer (AATGATACGGCGACCACCGAGATC TACACGTTCAGAGTTCTACAGTCCGACG) and 1uL of the appropriate barcorded reverse primer (Table S5Q). PCR reactions were run with the following cycle conditions: 98C for 1 minute, followed by 12-15 cycles of 94C for 16 seconds, 55C for 6 seconds, and 65C for 11 seconds, with a final extension of 65C for 1 minute. PCR products were mixed with 6x gel loading buffer and size-selected on a 8% TBE gel, 100V for 75 minutes. The product at ~150 bps was gel-excised, placed in 400uL of DNA extraction buffer, flash frozen on dry ice for 30 minutes, thawed at 22C for 6 hours on a rotating platform, and DNA was precipitated with 700uL of isopropanol with 2uL of GlycoBlue overnight at -80C. Samples were then centrifuged at 14,500 RPM for 45 minutes; DNA pellets were washed once in 80% ethanol, air-dried, and dissolved in 18uL of 5mM Tris. Samples were quantified by DNA Qubit (Thermofisher) and library size was confirmed using an Agilent Bioanalyzer HS DNA High Sensitivity Kit (Agilent). Libraries were sequenced at the Dana-Farber Molecular Biology Core Facility on an Illumina NovaSeq 6000.

RNA sequencing

Matched RNA sequencing for all samples was performed by removing 1/3rd of the sample lysate from the ribosome profiling sample and placing it in 400uL Trizol. RNA was then extracted using the Qiagen RNAeasy kit (Qiagen) according to the manufacturer's instructions. RNA abundance was quantified using spectrophotometry via Nanodrop as well as RNA Qubit (Thermofisher). RNA samples were submitted to the Dana-Farber Molecular Biology Core Facility for mRNA sequencing using the Roche Kapa mRNA Hyper Prep kit (Roche, Basel, Switzerland) with samples sequenced on an Illumina NextSeq or NovaSeq. RNA samples from the Princess Maxima Center were processed through the Princess Maxima Center Diagnostics core facility according to institutional protocols.

RNAseq sample clustering and pathway analysis

The raw RNA-seq reads from cell lines and tissue samples were subjected to quality control and read trimming using TrimGalore v0.6.6,[68] which internally employs Cutadapt v3.4[69] for adapter removal and FastQC v0.11.9 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) for quality assessment. using standard parameters for paired-end reads.

Trimmed and filtered reads were aligned to human reference genome hg38 using STAR v2.7.8a[71] in the two-pass mapping mode, with genome annotation provided in GTF format (Ensembl release 102). Default STAR settings were used, with the following modified parameters: *–outFilterType BySJout –outSAMunmapped Within –outSAMattributes NH HI AS nM NM MD jM jI MC ch –outSAMstrandField intronMotif –outSAMtype BAM Unsorted –outFilterMismatchNmax 6 –alignSJoverhangMin 10 –outFilterMultimapNmax 10 –outFilterScoreMinOverLread 0.75*.

Counts for annotated CDS regions were obtained using featureCounts v2.0.2[72] with genome annotation provided in GTF format (Ensembl release 102), and CDS regions used as the counting feature in paired-end mode. To improve read counting for junctions, the *–J* option was used with reference sequences for transcripts provided in FASTA format (GRCh38, Ensembl release 102).

CDS read counts from the cell line samples (annotated as either MYC high or MYC low) were used as input for DESeq2[73] to perform principal component analysis and differential expression analysis, using the default DESeq2 workflow and MYC status (MYC high vs MYC low) as contrasting variable.

Gene ontology (GO), hallmark, KEGG, and Reactome gene sets were obtained from the MSigDB[86,87] database using the msigdb R package,[80] and were used as query gene sets. A list of $\log_2$ fold change values, obtained from the DESeq2 output, was used as input for gene set enrichment analysis using the fgsea R package.[75] Gene set enrichment analysis was performed separately for each of the gene set categories (GO:CC, GO:BP, GO:MF, Hallmark, Reactome, KEGG). Gene sets with an adjusted P-value < 0.05 and a normalized enrichment score > 0 were considered significantly over-enriched in MYC-driven compared to non-MYC-driven samples.

Obtaining gene-level RNA-seq read counts

To facilitate comparison with ribo-seq data and calculate gene-level translational efficiency values, the RNA-seq reads were reprocessed using different alignment and filtering parameters as described below.

The raw RNA-seq reads were subjected to quality control and read trimming using TrimGalore. Only the first reads of the read pairs were used, to imitate single-end ribosome profiling reads. The RNA-seq reads were hard-trimmed to 29-mers using Cutadapt with the *–hardtrim5* option. Then, TrimGalore was run on the trimmed reads with options set to remove Ns (*–trim-n*) and retain reads with a minimum length of 25 bp (*–length 25*). FastQC was executed within TrimGalore to remove low quality reads.

To eliminate reads corresponding to contaminants such as tRNA, rRNA, snRNA, snoRNA, and mtDNA, Bowtie2 (v2.4.2)[72] was executed with standard parameters and option *–seedlen=25* to align the reads to a custom reference database containing sequences of these contaminants. The unaligned reads, i.e., those not mapping to any of the contaminants, were output to a gzipped FASTQ file for further processing.

The filtered reads were aligned to reference genome GRCh38 using STAR v2.7.8a with options *–outFilterMismatchNmax 2 –outFilterMultimapNmax 20 –outSAMattributes All –outSAMtype BAM SortedByCoordinate –quantMode GeneCounts –limitOutSJcollapsed 10000000 –outFilterType BySJout –alignSJoverhangMin 1000*, using the MANE Select v1.0[88] transcript annotation, supplied in a GTF file, as reference annotation.

To quantify reads aligning to annotated CDS features, featureCounts was used with the options *–J, –t "CDS", –g "gene_id",* resulting in CDS counts summarized on gene-level. Annotations and sequences for reference transcripts for GRCh38 / Ensembl release 102 were provided in FASTA and GTF files, respectively.

Ribo-seq read alignment and processing

Raw ribosome profiling reads were trimmed and filtered using TrimGalore with the following options: *–gzip –length 25 –trim-n*. Contaminant reads were filtered out with Bowtie2 with the option *–seedlen=25*, using a custom index containing tRNA, rRNA, snRNA, snoRNA, and mtDNA sequences. Filtered ribo-seq reads were aligned to reference genome GRCh38 using STAR v2.7.8a with options *–outFilterMismatchNmax 2 –outFilterMultimapNmax 20 –outSAMattributes All –outSAMtype BAM SortedByCoordinate --quantMode GeneCounts –limitOutSJcollapsed 10000000 –outFilterType BySJout –alignSJoverhangMin 1000*, using GRCh38 / Ensembl release 102 reference annotation provided in GTF file. Annotated CDS features were quantified using featureCounts with the options *–J –t "CDS" –g "gene_id",* with Ensembl release 102 annotation provided in GTF format and GRCh38 / Ensembl release 102 transcript sequences provided in FASTA format. We then used RiboseQC[77] provided with Ensembl release 102 transcript annotation in GTF format to assess data quality and quantify P-site positions in the aligned ribo-seq reads in all samples. For clustering, ribo-seq read counts for annotated CDS regions were used as input for DESeq2[69] to perform principal component analysis and differential expression analysis, using the default DESeq2 workflow and MYC status (MYC high vs MYC low) as contrasting variable.

Calculating translational efficiency values

Translational efficiency values for annotated genes were calculated using gene-summarized RNA-seq and Ribo-seq CDS read counts in cell line samples. To ensure that the genes used for TE calculation showed robust expression in both ribo-seq and RNA-seq data, genes with fewer than 128 read counts on average across all samples in either RNA-seq or ribo-seq were removed. To make the RNA-seq and ribo-seq read counts comparable, they were first converted to TPM values. The TE for each gene was then calculated as the ratio of TPM(ribo-seq) over TPM(RNA-seq). Non-real values resulting from divisions by zero were set to "NA". To plot the densities of the translational efficiency values for all genes in MYC-driven and non-MYC samples, the TE values were $\log_2$-transformed and centered by subtracting the TE value of each gene in each sample by the median TE of that gene across all samples.

Quantifying ORF-level P-sites

To quantify ribo-seq P-sites on an ORF level, we generated BED files that contain all possible P-site positions for annotated as well as non-canonical ORFs. We used a GTF file containing MANE Select transcript definitions[79] (matching the Ensembl annotations in version hg38) to obtain annotations for annotated CDS regions, and a custom GTF file containing merged definitions from GENCODE Phase 1 ORFs[21] and our prior custom cancer ORFeome[21,26] for non-canonical ORFs. A custom Python script was used to generate 'reference' BED files containing the coordinates of all potential P-site positions for each ORF, annotated by frame (p0, p1, or p2) in each codon. Incomplete proteins were excluded using provided annotation files (see Data Availability statement).

P-site coordinates and counts in each sample were extracted from RiboseQC output files and stored in BED files. Bedtools intersect v2.25.0[82] was used to overlap detected P-sites with the 'reference' P-sites using the options '-wa -wb -header -f 1.00 -s'. For each sample, the resulting BED files contained the P-site coordinates, counts, and ORF names (annotated and non-canonical) of overlapping 'reference' P-sites.

The resulting intersected BED were then used to generate a matrix of P-site counts per ORF in each sample. To construct this matrix, we first calculated the frame with the highest P-site fraction for each ORF in a given sample. We then added the total P-site count of the dominant frame of each ORF to the P-site count matrix.

To identify translated ORFs, P-site counts were converted to TPM-like count values (P-sites per million, or PPM). First, P-sites for each ORF were divided by the ORF length in kb to calculate P-sites per kb (PPK). Per-million scaling factors for each sample were calculated by dividing the sum of each sample's PPK values by 1,000,000. Each ORF's PPM value was then calculated by dividing the ORF's PPK by the sample's scaling factor. To define a PPM cutoff for determining translation, the density of $\log_2$-transformed PPM values was plotted and visually inspected. There was a clear bimodal distribution, so we selected a cutoff value between the low and high distributions, which corresponded to a PPM value of 1. Translated ORFs were then defined as ORFs with a PPM > 1 in at least 5 samples.

Identifying differentially translated ORFs

The matrix with raw ORF P-site counts for the cell line samples was loaded into R and used as input for DESeq2 to perform principal component analysis and differential expression analysis, using the default DESeq2 workflow, and using MYC status (MYC-driven vs non-MYC) as contrasting variable. The volcano plot showing differentially translated ORFs between MYC-driven and non-MYC samples was generated using the EnhancedVolcano R package.[89] ORFs were sorted by p-value, and top 5 upregulated (log_2 fold change > 2) and top 5 downregulated (log_2 fold change < -2) were highlighted.

ORF-level translational efficiency analysis

To obtain ORF-level RNA-seq read counts, we used Salmon v1.8.0,[78] with Bowtie2-filtered raw RNA-seq reads as input (see section processing of RNA-seq data for gene-level translational efficiency calculation). A custom Salmon index was generated based on a custom GTF file containing the merged set of annotated MANE transcripts as well as non-canonical GENCODE Phase1 ORFs[21] (https://www.gencodegenes.org/pages/riboseq_orfs/) and ORFeome definitions (Table S1P). Briefly, CDS regions were extracted from the custom GTF file and stored in a separate, cleaned up GTF file with transcript IDs set to match ORF IDs, since Salmon uses transcript IDs to differentiate between features. We ran Salmon with the following parameters: *'salmon quant –libtype "A" –validateMappings –gcBias –numGibbsSamples 30'.*

We loaded the matrices with ORF-level RNA-seq counts and P-site counts for the cell line samples into R, and removed ORFs with fewer than 4 counts on average across all samples in either RNA-seq or ribo-seq. We calculated TPM and PPM values for the remaining ORFs. ORF lengths for TPM and PPM calculations were based on the annotated CDS and non-canonical ORF definitions in the merged MANE+GENCODE Phase 1+ORFeome GTF file (Table S1P). Translational efficiency for each ORF was calculated as the ratio of TPM(Ribo-seq) over TPM(RNA-seq). Non-real values resulting from divisions by zero were set to "NA". TE values were log_2-transformed and scaled to perform principal component analysis. The full code can be found at: https://github.com/damhof/hofman_et_al_2023_seq

Tissue sample RNA-seq and Ribo-seq analyses

Sequencing data were processed as described above. Only tissue samples with matching RNA-seq and ribo-seq data were included. Autopsy samples were excluded from analysis due to low RNA quality. Tumor purity was estimated using the immunedeconv R package.[90] Samples with < 75% tumor purity were excluded from the analyses, resulting in a total of 11 remaining tissue samples. Lacking samples with true MYC amplifications, samples were grouped by molecular subtype (Group 3, Group 4, Group 3/4, NOS, and SHH). To increase statistical power, Group 3, Group 4, and Group 3/4 samples were grouped together as 'Group 3/4'. RNA-seq and ribo-seq counts for canonical CDS sequences were used for sample clustering and translational efficiency calculations, as described above.

Lentiviral transduction for CRISPR screens

Optimal infection conditions were determined in each cell line in order to achieve 30-50% infection efficiency, corresponding to a multiplicity of infection (MOI) of ~0.5 - 1. Spin-infections were performed in 12-well plate format with 3 x 10e6 cells each well. Optimal conditions were determined by infecting cells with different virus volumes with a final concentration of 4 ug/mL polybrene. Cells were spun for 2 hours at 1000 g at 30 degrees. Approximately 24 hours after infection, cells were trypsinized and approximately 2x10e5 of R262, UW228, ONS76, D458, D425, D283, or D341 cells from each infection were seeded in 2 wells of a 6-well plate, each with complete medium, one supplemented with 1.5ug/mL of puromycin. Cells were counted 4-5 days post selection to determine the infection efficiency, comparing survival with and without puromycin selection. Volumes of virus that yielded ~30 - 50% infection efficiency were used for screening.

Primary and validation CRISPR screens

For the primary CRISPR screen, 528 previously described non-canonical ORFs from ref.[26] were included as an internal basis of comparisons across screens. In addition, new non-canonical ORF amino acid sequences were selected by focusing on the following considerations:

- One ORF per gene was selected, with bias for the longest ORF to enable sufficient gRNA representation, if multiple ORFs were present. Several exceptions included several genes with two ORFs that were present in (Prensner NBT insert citation): CTD-261913.14, LINC00665, LINC02081, LOC401320, PIK3R1, RP1-261D10.2, ZNF788.
- For ORFs with >=2 exons, all gRNAs could not come from the same exon.
- Minimum of 3 gRNAs successfully designed, with gRNA features as described below.
- Exclude intORFs, doORFs or uoORFs that have >=25% overlap with the main CDS.
- Minimum ORF size of 12 amino acids.

The lentiviral barcoded library used in the primary screen contains 26,819 sgRNAs and the validation library contains 6,557 gRNAs targeting selected regions of the ORFs, which were designed using the CRISPick program (https://portals.broadinstitute.org/gppx/

crispick/public) from Broad Institute Genomic Perturbation Platform, using settings for the reference genome Human GRCh38 (Ensembl v.108) for "CRISPRko" with enzyme "SpyoCas9 (NGG)" with the following modifications:

- Each ORF and parental CDS were targeted by up to 8 gRNAs where possible. A distribution of the number of gRNAs per target is displayed in Table S2A.
- For ORFs with >= 2 exons, the best gRNA design was selected for each exon to a maximum of 8 gRNAs. For ORFs with >2 but <8 exons, the remaining gRNAs were selected as the top picks from any exon.
- The spacing requirement for gRNA separation was reduced to 1% across the total target length for ORFs and maintained at 5% for parental CDSs.
- A 2:1 on-target to off-target ratio was employed.
- For the validation library, ORFs were targeted with a maximum of 24 gRNAs per exon, 5′UTR and 3′UTRs with a maximum of 12 gRNAs per UTR region, up to 3 introns with 6 gRNAs per intron, the upstream genome promoter region with 6 gRNAs (defined as within 1000 basepairs of the transcript start site), and up to 3 parental CDS exons with 8 gRNAs per exon.
- Both libraries employed a common set of 503 non-targeting gRNAs without genome cutting, and 497 non-targeting gRNAs with genome cutting for negative controls. The primary library had 1694 positive control pan-lethal gRNAs. The validation library had 527 positive control pan-lethal gRNAs.

Genome-scale infections were performed in three replicates with the predetermined volume of virus in the same 12-well format as the viral titration described above, and pooled 24 h post-centrifugation. Infections were performed with enough cells per replicate, in order to achieve a representation of at least 500 cells per gRNA (for primary screen) or 1000 cells per gRNA (for validation screen) following puromycin selection (~1.5x10e7 surviving cells). Approximately 24 hours after infection, all wells within a replicate were pooled and were split into T225 flasks. 24 hours after infection, cells were selected with puromycin for 7 days to remove uninfected cells. After selection was complete, 1.5-2x10e7 of cells were harvested for assessing the initial abundance of the library. Cells were passaged every 3-4 days and harvested ~14 days after infection. For all genome-wide screens, genomic DNA (gDNA) was isolated using Midi or Maxi kits for the validation screens gDNA was isolated using Midi kits according to the manufacturer's protocol (Qiagen). PCR and sequencing were performed as previously described.[79,91] Samples were sequenced on a HiSeq2000 or NextSeq (Illumina). For analysis, the read counts were normalized to reads per million and then $\log_2$ transformed. The $\log_2$ fold-change of each sgRNA was determined relative to the initial time point for each biological replicate.

Analysis of CRISPR screening data

CRISPR data was transformed into $\log_2$ fold change values computed between the day 14 timepoint and the input plasmid DNA. All values were then normalized to the positive control gRNAs in the following way: for each cell line, the gRNAs targeting parental_poscon genes were averaged. This geometric mean of the poscons was scaled to equal -1. This was accomplished by dividing individual gRNA values by the poscon mean, multiplied by -1 to retain a negative value to represent gRNA drop-out. The equation is as follows: (gRNA/average_poscon)*-1. A "hit" was defined as a non-canonical ORF that had at least 2 gRNAs with a normalized abundance of less than or equal to -1.0 at the day 14 timepoint in the primary screen. For uORFs, uoORFs, and dORFs, the comparison between the non-canonical ORF and the parental CDS should demonstrate a differential effect (delta_ORF-CDS effect) of less than or equal to -0.3 to yield a potential differential dependency. uORFs, uoORFs and dORFs were further assessed by comparing the absolute number of gRNAs with a normalized abundance of less than or equal to -1.0 to the absolute number of parental CDS gRNAs with a normalized abundance of less than or equal to -1.0.

Assessment of Cas9 toxicity at gene promoters

To assess Cas9 toxicity when targeting uORFs located near to the gene promoter, the primary screen further targeted 120 pan-lethal positive control genes known to have a uORF as well as 82 pan-lethal positive control genes with no known uORF. For the latter, a 150 bp segment of the gene 5′UTR was targeted with gRNAs. The data were analyzed as described above to estimate the potential impact of Cas9 genome toxicity at the promoters of genes. Figure S2K provides additional details.

Defining medulloblastoma-specific CRISPR hits

We compared the CRISPR screen data for 7 medulloblastoma cell lines in this study (ONS76, R262, UW228, D425, D458, D341, D283) with publicly-available data for CRISPR screen data for 553 non-canonical ORFs across 8 non-medulloblastoma cell lines (HELA, A549, HT29, HEPG2, HA1E, A375, PC3, MCF7). There were 528 non-canonical ORFs tested in both studies. We calculated the average loss-of-viability score (as calculated above) for both sets of cell lines. Statistical significance for differential vulnerability was first calculated with a two-tailed Student's T test, and candidate hits were determined by considering an FDR-corrected Q value using the Benjamini, Krieger and Yekutieli method. There were 22 non-canonical ORFs with a corrected Q value of <0.01 when comparing between groups, of which 14 showed substantially increased loss-of-viability in the medulloblastoma cells. An increased loss-of-viability was defined as the delta of <= -0.25 for the average loss of viability in medulloblastoma cells minus the average loss of viability in non-medulloblastoma cells: delta = (ave_MBL) - (ave_nonMBL).

Analysis of CRISPR validation screen

The validation screen targeted 44 uORFs, 6 uoORFs, 10 lncRNA-ORFs, and their associated parental CDS and genomic regions (Table S2I). The validation screen was performed on the CHLA06ATRT, D283, and UW228 cell lines, and data for each cell line were normalized to the 527 positive control pan-lethal gRNAs as described above. In the secondary screen, because the number of gRNAs for each gene varied, a scoring candidate was defined as a gene in which at least 30% of the gRNAs achieved a normalized abundance of less than or equal to -0.4. This threshold reflected the point that >95% of all negative control gRNAs failed to achieve in all 3 cell lines but >75% of all positive control gRNAs successfully achieved in all 3 cell lines. gRNAs were then grouped into their respective genomic region (e.g. UTR, ORF exon, adjacent gene exon, intron). Genes were then classified in the following manner according to the viability effect of the gRNAs: ''selective uORF dependency'' if only the ORF region gRNAs reached that threshold; ''uORF and adjacent nucleotides'' if the ORF gRNAs and gRNAs to only one other region of the RNA transcript scored; ''uORF and CDS'' if the ORF and an annotated adjacent protein coding gene both scored; ''weak phenotype'' if none of the cell lines showed a phenotype for that ORF.

Base editing

gRNAs for base editing were manually designed to target the start codon of the uORF or associated parental CDS. The targeted nucleotide was positioned between basepairs 3 and 9 on the gRNA. gRNAs were synthesized via a commercial vendor (Synthego) with standard modifications (2'-O-Methyl at 3 first and last bases, 3′ phosphorothioate bonds between first 3 and last 2 bases). For base editing experiments, 200,000 D425 cells per reaction were centrifuged (1200RPM for 5 minutes), washed once in PBS, centrifuged again (1200 RPM for 5 minutes), and resuspended in 15uL of Nucleofector solution from the P3 kit (Lonza) in a 1.5mL microcentrifuge tube. Concurrent, a plasmid mix was prepared consisting of 1uL of Electroporation Enhancer (100uM, Lonza), 1.5 uL of 2ug/uL ABE8e-NRCH ribonucleoprotein editor,[92] 1uL of base editor primer (50uM stock) and 3.6uL of Nucleofector supplement (Lonza). The ABE8e-NRCH base editor was a kind gift from Dr. David Liu's lab at the Broad Institute. This 7.1uL of plasmid mix was added to the 15uL of cells in Nucleofector solution and samples were transferred to the Nucleocuvette vessels, ensuring that no bubbles were introduced in transfer. Cells were then electroporated using the Lonza DN-100 program. Afterwards, cells were recovered with the addition of 80uL of cell culture media. A cell count was repeated using a Beckman Coulter ViCell to ensure equal cell numbers and viability, and cells were transferred to 96 well poly-lysine coated plates at 2500 cells per well. Unused cells were plated on a 6 well poly-lysine coated plate and harvested for genomic DNA on day 4. Cell viability was measured at day 4 and day 6 using the Cell-Titer Glo assay (Promega). Viability data was analyzed by comparing the relative viability change between base editing with the uORF gRNA and the associated parental CDS gRNA. Negative controls were biological triplicate mock nucleofections.

gRNA sequences used for base editing, including details such as PAM sites and target start codons, are listed in Table S5R.

Nomination of ASNSD1-uORF

To compare the overall impact for knockout of uORFs, uoORFs, and dORFs across molecular disease subtypes, the differential dependency for each ORF was assessed across each individual cell line. Individual values were averaged as the geometric mean across cell line subtypes as follows: MYC_medulloblastoma (D341, D283, D425, D458) and nonMYC (UW228, R256, ONS76) The distributions of differential dependency scores were compared across groups using a two-sided Student's T test. For individual outlier uORFs, the weighted average of the differential dependency scores for uORFs and uoORFs for D283 and D341 were compared to those of UW228 and ONS76. Additionally, for each cell line, individual uORF outliers were assessed by calculating the delta differential dependency score between the uORF and the parental CDS and comparing this to the difference in the number of gRNAs that scored for the uORF compared to the parental CDS.

Prefoldin in medulloblastoma DepMap data

The DepMap_public_21Q2 release of CRISPR DepMap CERES scores and cell line gene expression data was downloaded from https://depmap.org/portal/download, reflecting the current data release at the time of the performed analysis. Data for PFDN1, PFDN2, PFDN4, PFDN5, PFDN6, VBP1, PDRG1, UXT, and URI1 were extracted for all annotated medulloblastoma cell lines in the dataset. CERES scores were z-scored. Data were visualized in the Morpheus platform hosted by the Broad Institute (https://software.broadinstitute.org/morpheus/) using a hierarchical cluster via the ''one minus the pearson correlation'' metric. MYC expression was obtained via the corresponding CCLE gene expression data file.

ASNSD1-uORF evolutionary analysis

The amino acid sequence for ASNSD1-uORF (UniProt ID L0R819 isoform 1) and for the parental ASNSD1 CDS (UniProt ID Q9NWL6 isoform 1) were analyzed using the NCBI ProteinBlast feature (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) using default parameters against the ''non-redundant protein sequences (nr)'' database and the ''model organisms (landmark)'' database. All identified non-human amino acid sequences were downloaded and analyzed for similarity to either ASNSD1-uORF of ASNSD1 respectively using the ClustralOmega package[93] (https://www.ebi.ac.uk/Tools/msa/clustalo/).

ASNSD1 gene expression analysis

Processed RNA expression data for ASNSD1 mRNA expression (ENSG00000138381.9) were downloaded from GTeX for bulk RNA sequencing data (https://www.gtexportal.org/home/) and the Allen Institute Developing Brain Atlas (https://www.brainspan.org). In cell lines, ASNSD1 expression was evaluated through Cancer Cell Line Encyclopedia data for ASNSD1 (ENSG00000138381.9). CCLE data was downloaded from https://portals.broadinstitute.org/ccle. Data were analyzed in GraphPad Prism as shown.

ASNSD1-uORF overexpression experiments

The indicated ASNSD1-uORF or MYCN cDNAs were synthesized using a commercial vendor (GenScript) and cloned into the pLX_307 or pLX_313 mammalian expression vector (Table S5S for sequences). pLX_307 and pLX_313 are Gateway-compatible expression vectors where E1a is the promoter of the ORF and SV40 is the puromycin resistance gene with either puromycin (pLX_307) or hygromycin (pLX_313) resistance (details at https://portals.broadinstitute.org/gpp/public/resources/protocols). Lentivirus was produced in HEK293T cells as previously described,[26] using the Lenti-X Concentrator (Takara Bio) to achieve a 50x virus concentration. For overexpression experiments, H9-derived NSC and D341 cells were transduced with lentivirus and stably-expressing cells were selected with either puromycin (0.5 ug/mL, plx_307 lentivirus) or hygromycin (300ug/mL, plx_313 lentivirus) for 72 hours prior to transitioning back to standard culture media. In 96 well plates (GelTrex pre-coated for H9-derived NSC or poly-lysine for D341), 4000-5000 cells per well were plated. For H9-derived NSC experiments, cell viability was monitored daily using the CellTiter Glo reagent. For D341 experiments, cells were infected with the indicated gRNA lentivirus 4-6 hours after plating. 16 hours after infection, cells were selected with 1ug/mL puromycin for 48 hours and grown for 7 days prior to cell viability analysis using CellTiter-Glo reagent.

ASNSD1-uORF knockout experiments

Cells were plated in 96-well plates and allowed to grow for 4-8 hours prior to infection with the indicated sRNA or treatment condition. 1,000 - 5,000 cells per well were plated depending on the cell line. gRNAs were obtained from the Broad Institute Genomic Perturbation Platform (Broad Institute, Cambridge, MA, USA) or from direct synthesis into the BRDN0003 or BRDN0023 backbone via commercial vendor (GenScript, Piscataway, NJ). sgRNA sequences are listed in Table S5T.

All sgRNAs were sequenced and verified. After sequence verification, constructs were transfected with packaging vectors into HEK-293T with Fugene HD (Sigma-Aldrich, St. Louis, MO). After plating, cells were then infected with sgRNA lentivirus to achieve maximal knockout but without viral toxicity. 16 hours after infection, cells were selected with 2ug/uL puromycin (Invitrogen, Carlsbad, CA) for 48 hours. Cell viability was measured CellTiter-Glo reagent (Promega, Madison, WI) was measured at 16 hours post-transfection for a baseline assessment, and additional timepoints as needed. For stable knockout cell lines, cells were plated at equal densities and cell viability was measured by CellTiter-Glo every 24 hours as indicated.

Analysis of cell line knockout data

Cell line knockout data was normalized as previously described.[26] Briefly, data for each cell line were standardized such that the average of the positive controls was equal to -1 and the average of the negative controls was equal to 0.

Pooled ASNSD1-uORF knockout

Pooled knockout screens in the PRISM cell line set were performed as previously described.[26] Briefly, we used a pool of 486 barcoded human cancer cell lines, which were collectively grown in RPMI1640 media supplemented with 10% FBS. gRNAs used were non-cutting LacZ control (AACGGCGGATTGACCGTAAT), cutting control Chr2-2 (GGTGTGCGTATGAAGCAGTGG), ASNSD1-uORF #1 (GCTTAGATCCTCCTTGTGTG), and ASNSD1-uORF #2 (TAAAGAACAAAAAATTGTGG). Briefly, on Day 0, the cell pool was plated at 400,000 cells per well in a 6 well plate with a cell pellet collected for a "no infection" control. On Day 1, cells were transduced with gRNA and Cas9 using an all-in-one plasmid with lentiviral titer at an MOI of 10 and 4ug/mL polybrene. On Day 4, cell culture media was changed to include 1ug/mL puromycin for 72 hours, after which antibiotic-free media was used. Cells were then passaged every 72 hours and a cell pellet (2e6 cells) was collected for DNA on day 6, 10 and 15. For genomic DNA extraction, cell pellets were washed in PBS and then processed using the DNA Blood and Tissue Kit according to manufacturer's instructions (Qiagen, Hilden, Germany).

For determination of individual cell line representation, DNA from each time point was amplified by PCR with universal barcode primers, and PCR products were confirmed on a 2% agarose gel for size. Then, PCR products were pooled and purified with AMPur beads (Beckman Coulter, Brea, CA), and DNA concentration was measured via Qubit fluorometric quantification (Thermo Fisher Scientific, Waltham, MA). DNA was sequenced on a NovaSeq (Illumina, San Diego, CA) at the Genomics Platform at the Broad Institute.

Analysis of pooled ASNSD1-uORF knockout data

484 of 486 cell lines were detectable at the day 15 time point and were used for data analysis. Cell line abundance was determined by RNA expression of each cell line's barcode using RNA-sequencing as previously described. Data analysis was performed as previously described[26] with the following modifications: cell lines with a detected number of reads but with fewer than 12 reads were included in the analysis. Following calculation of reads, the $\log_2$ fold change abundance in each cell line was determined by comparing the day 15 abundance with the input plasmid pool. For lineage analysis of ASNSD1-uORF knockout across cancer types,

we integrated the average $\log_2$ fold change of ASNSD1-uORF gRNA #1 and ASNSD1-uORF gRNA #2 with cancer cell line metadata from the DepMap database (www.depmap.org). For correlation of ASNSD1-uORF knockout phenotype with prefoldin complex knockout phenotypes, we used the Cancer Dependency Map release 21Q2 data to obtain gene-level knockout effects for 17643 human genes. A total of 389 cell lines were shared between the pooled ASNSD1-uORF knockout dataset and the Dependency Map dataset. For these 389 cell lines, the Pearson correlation was calculated for the knockout phenotypes relative to ASNSD1-uORF or members of the prefoldin and prefoldin-like complexes (PFDN1, PFDN2, PFDN4, PFDN5, PFDN6, URI1, UXT, PDRG1, VBP1) along with FDR-corrected Q values. The Pearson coefficients for each comparison were then permuted into a percentile rank and plotted as such. For evaluation of ASNSD1-uORF knockout with gene expression, the averaged ASNSD1-uORF knockout phenotype was compared to ASNSD1 mRNA expression (ENSG00000138381.9) using RNA-seq data values made available through the CCLE data at https://portals.broadinstitute.org/ccle.

CRISPR-seq

The indicated cell lines were transduced with lentivirus for Ch2-2 or LacZ gRNA negative controls, ASNSD1-uORF gRNA #1 or ASNSD1-uORF gRNA #2. After selection of puromycin-resistant cells with 1 ug/mL puromycin for 48 hours, cells were grown until 96 hours post-transduction. Genomic DNA was then isolated from cells using the Qiagen DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructors. 100ng of DNA was amplified by PCR with the following thermocycler conditions: 94C for 2 minutes, followed by 30 cycles of 94C for 30 seconds, 52C for 30 seconds, and 68C for 1 minute; final elongation was 68C for 7 minutes. PCR products were confirmed for specificity with a 1% agarose gel and then gel-purified with a Qiagen Gel Extraction kit according to manufacturer's instructions. DNA was diluted to a concentration of 25ng/uL and submitted to the Massachusetts General Hospital Center for Computational and Integrative Biology (CCIB) DNA Core for sequencing. FASTQ sequencing files were analyzed using CRISPResso[81] (http://crispresso.pinellolab.partners.org) according to default parameters. Primers used for CRISPR-seq are listed in Table S5P.

ASNSD1-uORF protein level in cell lines

Cancer cell lines were grown in standard tissue culture as previously described to a confluency of ~80%. For ONS76-GFP, ONS76-ASNSD1-uORF and ONS76-ASNSD1-uORF mutant cell lines, cells were first transduced with lentivirus for the indicated plasmid and selected with antibiotics as above. Cells were then washed three times in 1x ice-cold PBS, pelleted, and lysed using RIPA buffer. 35ug of cleared cell lysate was loaded per cell line on a 10-well, 10-20% Tris-Glycine gel and ran for 90 minutes at 125V. In each gel, samples were separated by an empty well. Then, the gels were washed 3x with deionized water at room temperature and stained with SimplyBlue Coomassie stain (Thermofisher) for 90 minutes at room temperature to ensure equal loading of protein. Gels were then washed 5 times with deionized water, 1 hour per wash at room temperature. Gel bands corresponding to the gel slice between 10 – 15 kDa were cut out using a sterile razor, started in 1mL of RNase/DNase free water, and then subjected to mass spectrometry analysis at the Taplin Mass Spectrometry facility at Harvard Medical School as previously described.[26] Mass spectrometry data were first normalized by standardizing the input protein amount for gel analysis and secondarily normalized for individual proteins by calculating the fraction of that protein's abundance relative to all proteins that were detected in that size range. The process was standardized using triplicate measurements for the D458 cell line. Additional cell lines were run in single replicates.

ASNSD1-uORF in proteomics and RNA-seq

ASNSD1-uORF abundance was determined in publicly-available medulloblastoma mass spectrometry data[17] as previously described.[26] Briefly, a fasta database containing the amino acid sequence of ASNSD1-uORF was appended to a reference protein database (UCSC, RefSeq) and used to search peptide mass spectra from the Clinical Proteomics Tumor Analysis Consortium (CPTAC) downloaded from the following repository: ftp://massive.ucsd.edu/MSV000082644. Raw mass spectrometry data were analyzed in Spectrum Mill MS Proteomics Workbench v.7.09 (https://proteomics.broadinstitute.org) employing a target-decoy-based false discovery rate (FDR) estimation for non-canonical ORFs. An FDR of <0.01 was used for statistical significance. We additionally minimized potential false-positive identifications by requiring a minimal Spectrum Mill PSM score of 8 for ASNSD1-uORF PSMs. Next, individual protein abundances were correlated to ASNSD1-uORF abundance using Pearson correlation coefficients and statistical significance of each correlation was corrected for multiple hypothesis testing by calculation of a q-value. Full values are available in Table S4C. For comparison of ASNSD1-uORF, PFDN1, PFDN2, PFDN4, PFDN5, PFDN6, VBP1, URI1, UXT, and PDRG1 abundance to MYC and MYCN levels, the maximum value of MYC or MYCN protein abundance was used, given their mutual exclusivity (Figure S4D). Then, samples were divided into quartiles based upon the maximum MYC or MYCN protein abundance for the 45 mass spectrometry samples, with N=11 samples in Quartiles 1, 2 and 3 and N=12 samples in Quartile 4. Data were normalized across the average of all samples to define the fold upregulation of Quartile 4 compared to all samples. For correlation of *ASNSD1* transcript to ASNSD1-uORF protein, the matched publicly-available RNA-seq expression values for the medulloblastoma tissue samples were extracted from Table S2 of the original publication.[17] RNA-seq data was correlated to ASNSD1-uORF protein level of MYC/MYCN level as described above. Skew in protein levels was statistically determined using a 1-way ANOVA p value on GraphPad PRISM.

ASNSD1-uORF immunoprecipitation

HEK293T cells were transiently transfected with ASNSD1-uORF-V5, ASNSD1-uORF-FLAG, ASNSD1-uORF deletion mutants (V5-tagged), GFP-V5 or GFP-FLAG fusion proteins using OptiMem and Fugene HD (Sigma-Aldrich). Forty-eight hours later, cells were washed once in ice-cold PBS and collected by centrifugation at 1,500 RPM for 5 minutes. Cells were lysed in lysis buffer (50 nM Tri-HCl pH 8.0, 150 nM NaCl, 2 mM EDTA pH 8.0, 0.2% NP-40 and 1 μg ml–1 PMSF protease inhibitor) for 20 minutes on ice and then cell debris was removed with centrifugation at 13,500 RPM for 15 minutes. Cell lysates were quantified using the BCA method and 2 mg of protein was used for input. Next, lysates were cleared with Pierce magnetic A/G beads (Thermo Fisher Scientific) for 1 h while rotating at 18–20 RPM. Beads were then discarded, and 10% of the medium was removed as an input sample and kept at 4 °C without freezing. The remaining culture medium was then treated with 50 μl of magnetic anti-V5 beads (MBL International) or 50uL of Anti-FLAG(R) M2 Magnetic Beads (Sigma-Aldrich) and rotated at 18–20 RPM overnight at 4°C. The following day, the supernatant was discarded and beads were washed four times in immunoprecipitation wash buffer (50 nM Tri-HCl pH 8.0, 150 nM NaCl, 2 mM EDTA pH 8.0, 0.02% NP-40 and 1 μg/ml PMSF protease inhibitor) with rotation for 10 min per wash. After the final wash, beads were gently centrifuged and residual wash buffer was removed. Then, proteins were eluted twice with 2 μg/μl V5 peptide in water (Sigma-Aldrich) or 1 μg/μl 3x FLAG peptide (ApexBio) at 37 °C for 15 min with shaking at 1,000 RPM The two elution fractions were pooled and samples were prepared with 4x LDS sample buffer and 10× sample-reducing agent (Thermo Fisher Scientific), followed by boiling at 95 °C for 5 min. One-third of the eluate was then run on a 10–20% Tris-glycine SDS–PAGE gel and stained with SimplyBlue Coomassie stain (Thermo Fisher Scientific) for 2 h. Gels were destained with a minimum of three washes in water for at least 2 h per wash. Bands were visualized using Coomassie autofluorescence on LI-COR Odyssey in the 800-nM channel. Gel lanes were then cut into six equal-sized pieces using a sterile razor under sterile conditions, and stored in 1 ml of RNase/DNAse-free water before LC-MS/MS analysis.

PFDN6 co-immunoprecipitation

D425 medulloblastoma cells were grown to 80% confluency to ∼90 million cells. Cells were collected and washed twice in ice-cold PBS. Cells were lysed in endogenous IP lysis buffer (50 nM Tri-HCl pH 8.0, 150 nM NaCl, 2 mM EDTA pH 8.0, 0.2% NP-40, 2.5% Glycerol v/v 2.5%, Rnase I (1U/10uL) and Turbo DNase (25U/10uL), 1 μg/ml PMSF protease inhibitor). Lysis occurred for 15 minutes at room temperature and then 10 minutes on ice. Samples were centrifuged at 14000 RPM at 4C for 12 minutes to clear the lysates. Protein concentration was determined using the BCA method, and 200ug of input protein was saved for the input samples. 2.5 mg of protein was aliquoted as the input for control IP and PFDN6 IP tubes, and samples were adjusted to 600uL with additional endogenous IP lysis buffer. Samples were mixed with 200uL of pre-washed slurry of a 1:1 mix of EZview Red Protein G and EZview Red Protein A bead affinity gel slurry (Sigma-Aldrich) and rotated at 4C for 1hr. Prior to usage, the protein A/G slurry was pre-washed 2x in endogenous IP lysis buffer. Samples were centrifuged at 250g x 4 minutes at 4C and supernatant was removed and kept in a new tube, with the beads discarded. This was performed twice to increase purity. Then, 20 uL of PFDN6 (Sigma-Aldrich #HPA043032) or normal rabbit IgG (Cell Signaling Technology #2729S) antibody was added to the appropriate tube, and samples were rotated at 18 RPM at 4C overnight. After overnight rotation, samples were incubated with 100uL EZview Red protein A/G bead slurry (1:1 mixture as above, pre-washed twice in IP wash buffer) for 2 hrs at 4C with 18 RPM Samples were centrifuged at 250g x 4 minutes at 4C and supernatant was removed, with the beads left behind. Beads were washed three times for 10 minutes each in ice-cold IP wash buffer with glycerol (50 nM Tri-HCl pH 8.0, 150 nM NaCl, 2 mM EDTA pH 8.0, 0.02% NP-40, 2.5% glycerol v/v and 1 μg/ml PMSF protease inhibitor). During each wash, samples were rotated at 18 RPM at 4C, and after each wash samples were centrifuged at 250g x 4 minutes at 4C and supernatant was removed. Samples were then eluted in 100uL of 1x sample loading buffer and boiled for 5 min at 95C. For mass spectrometry analysis, samples were run on a 16% Tris-Glycine gel at 125V for 100 minutes, then rinsed with deionized water and stained with SimplyBlue Coomassie stain (Thermo Fisher Scientific) for 2 hr. Gels were destained with a minimum of three washes in water for at least 2 h per wash. A gel slice corresponding to the band between 10 - 20 kDa was removed using a sterile razor under sterile conditions, and stored in 1 ml of RNase/DNAse-free water before LC-MS/MS analysis at the Taplin Mass Spectrometry Facility at Harvard Medical School. Experiments were performed in biological duplicate.

Expression of GST constructs

The protein domains of the prefoldin/prefolin-like complex members were cloned into the pET-23a(+) vector containing a T7 promoter and 6xHis-tag by a commercial vendor (GenScript). The ASNSD1-uORF coding sequence was cloned into the pET-GST vector containing a T7 promoter and Glutathionine S Transferase (GST) tag by a commercial vendor (GenScript). Expression plasmids were transformed into T7 Express Competent E. coli (C2566H, New England BioLabs) following the manufacturer's protocol. The next day, colonies were picked and a 10 mL culture of LB containing 100 ug/mL carbenicillin was inoculated and cultured overnight at 37°C and 225 rpm. The following morning, the 10 mL LB overnight cultures were added to 250 mL of LB containing 100 ug/mL carbenicillin in a 1L flask and cultured for ∼6-8 hours at 37°C and 225 rpm, until OD600 reached 0.8. Next, IPTG was added to bacterial culture for a final concentration of 1 mM for induction. The expression cultures were then incubated at 30°C and 225 rpm overnight. Bacteria was pelleted by centrifugation at 3000 rpm for 10 minutes. Pellets were washed once with TBS, repeat pelleted, and stored at -80°C until lysis.

GST-Pull Down Assay

We utilized the Pierce™ GST Protein Interaction Pull-Down Kit (#21516, Thermo Fisher Scientific). Expression culture pellets were thawed on ice. Pellets were resuspended in 5 mL of TBS and Halt™ Protease Inhibitor Cocktail was added (Thermo Fisher Scientific). Next, 5 mL of Pull-Down Lysis Buffer was added. Samples were incubated on ice for 30 mins with periodical inversions. Samples were then centrifuged and the clarified E. coli lysates were extracted and stored on ice. To equilibrate Glutathione Agarose, the Glutathione Agarose resin was thoroughly resuspended using a vortex mixer and ~50 uL of Glutathione Agarose resin was added to the spin column. 400 uL of a wash solution containing a 1:1 of TBS to Pull-Down Lysis Buffer was added to the spin columns, gently inverted five times, and the wash buffer was discarded by centrifugation. This wash step was repeated for a total of four washes. To immobilize the bait protein, 800 uL of the ASNSD1-uORF GST-tagged lysate was added to the spin columns and they were rotated gently on a shaker at 4°C for 30 minutes. The bait lysate was then removed by centrifugation and the spin column was washed for a total of four washes. For prey protein capture, 800 uL of purified prey lysate was added to the spin columns. The spin columns were then gently rotated on a shaker at 4°C for 1 hour. The prey lysate was removed by centrifugation and the spin columns were washed for a total of four washes. For bait-prey elution, 1mL of 10mM Glutathione Elution Buffer was prepared by adding 3.1 mg of Glutathione to 1 mL of TBS. 250 uL of the Elution Buffer was added to the spin columns. The spin columns were incubated for 5 minutes with gentle rocking on a rotating platform. Spin columns were centrifuged and elution collected. Samples were then prepared for SDS-PAGE analysis. For visualization, we used the THE™ Anti-GST Monoclonal Antibody (A00865, GenScript) and Anti-6X His tag® antibody (#2365S, Cell Signaling Technology).

Identification of downstream targets

1.5 million D425 cells or 2.0 million D283 cells were plated in each well of poly-lysine coated 6 well plates. Cells were allowed to attach for 3 hours and then subsequently transduced with 30uL of 10x concentrated lentivirus with 4ug/mL polybrene. Transductions were done in biological triplicate. Cells were grown for 24 hours and the 1.5ug/mL of puromycin was added. Cells were antibiotic-selected for 48 hours and then fresh media was added. Cells were grown for an additional 48 hours. At the 120 hour time point, cell media was aspirated and cells were washed in ice-cold PBS four times. Cells were scraped, counted, and aliquoted into 1 million cells for RNA-seq and 3 million cells for mass spectrometry. Cells were pelleted; PBS was removed and cells were flash-frozen in liquid nitrogen. RNA was isolated as above and mRNA sequencing was performed at the Dana-Farber Cancer Institute Molecular Biology Core Facility as above. RNA-seq read processing, alignment and quantification was performed as above. CDS read count normalization and differential expression analysis between knockout and control conditions was performed separately for each cell line using DESeq2. Cell pellets reserved for mass spectrometry were transferred to the Harvard Medical School ThermoFisher Center for Multiplexed Proteomics (TCMP) for total proteome analysis using TMT 10-plex or 15-plex. Protein lysates were subject to quantification, reduction and alkylation, precipitation and digestion followed by peptide quantification, TMT-labeling, LC-MS3 label check, basic reverse-phase HPLC fractionation (bRP-HPLC), LC-MS3 analysis of 12 bRP-LC peptide fractions, database searching, filtering to 1% FDR at protein level, TMT reporter quantification, and data analysis accord to standard TCMP core facility pipelines as previously described.[26] To identify downstream targets, significantly differentially-abundant proteins with a $p < 0.01$ were considered. Proteins that had statistically-significant changes in both PFDN2 and ASNSD1-uORF knockouts were tested for gene network modules using the NCBI DAVID Bioinformatics platform (https://david.ncifcrf.gov/tools.jsp) on default settings.

Stratifying ASNSD1-uORF proteomic targets

Processed proteomics data were used for N=45 medulloblastoma samples with mass spectrometry data from Archer et al. as above.[17] Then, samples were divided into quartiles based upon the ASNSD1-uORF protein abundance, with N=11 samples in Quartiles 1, 2 and 3 and N=12 samples in Quartile 4. Protein abundances for MNAT1, RBM27, MAD1L1, TPR and ZC3HC1 were compared between Quartile 4 and Quartile 1 using a two-tailed Student's t-test.

Calculation of protein-RNA discordance score

Processed data N=39 medulloblastoma samples with matched mass spectrometry and RNAseq data were acquired from Table S2 of the original publication.[17] Data from the original source reflected z-score normalized values for each data type. Samples were stratified into quartiles based on ASNSD1-uORF protein abundance using the metrics applied in the prior section (ASNSD1-uORF abundance and correlations in mass spectrometry and RNA-seq datasets). For each gene in each sample, the protein-RNA discordance index was calculated by subtracting the RNA-seq z-score value from the mass spectrometry z-score value. To ensure sufficient data for analysis, genes that exhibited >=10 samples with unquantified or missing values were discarded. This filtering step decreased the number of genes in the analysis from 11,711 to 9,571 genes. Then, the average Protein-RNA discordance was calculated for each gene in each quartile of ASNSD1-uORF abundance. The ASNSD1-uORF signature was determined by calculating two-tailed Student's T-test for these Protein-RNA discordance scores for each gene between tier 4 (highest quartile of ASNSD1-uORF abundance) and tier 1 (lowest quartile of ASNSD1-uORF abundance). These p-values were converted to FDR-corrected q value using the Benjamini, Krieger, and Yekutieli approach. Genes were then ranked as a percentile rank where 1 equals the most statistically significant gene. To perform Gene Ontology enrichments, we considered the 383 genes whose Protein-RNA discordance had an uncorrected p value of <= 0.01. Interacting protein modules for these proteins were visualized using the STRING-db tool (www.string-db.org). For biological pathways, these 383 genes were uploaded to www.geneontology.org and analyzed for Biological Process, Molecular

Function, and Cellular Component using default parameters. For visualization, representative non-overlapping concepts were selected based on the maximum FDR and fold enrichment for the set of overlapping concepts. To assess ASNSD1-uORF proteomic targets in the medulloblastoma tissue Protein-RNA discordance score, the 637 (D283) and 276 (D425) proteins displaying significant regulation of an p value of <= 0.01 in the ASNSD1-uORF knock-out experiments (as above) were considered, resulting in 790 unique proteins. As a comparator, 8781 additional proteins were detected in the Protein-RNA discordance analysis, which were used as a negative control. The distribution of the 790 ASNSD1-uORF target proteins of 8781 control proteins were calculated across the percentile rank for the Protein-RNA discordance score. This was done by calculating the number of proteins present in each unit of 5 percentage points, i.e. within percentile rank of 0.000 - 0.0499, 0.050 - 0.099, 0.100 - 0.1499, etc). These distributions were compared using a Kolmogorov–Smirnov test with a p value < 0.05 being considered statistically significant.

Bromouridine (BrdU) quantification

D283 and D425 cells were plated in a 96 well plate at a density of 5,000 (D283) or 2,500 (D425) cells per well. Cells were transduced with Cas9 lentivirus and the indicated gRNA (LacZ control, sgASNSD1-uORF gRNA #3, or KIF11 gRNA). 24 hours later, stably-transduced cells were selected using 1.5ug/mL of puromycin. 96 hours later, cells were processed for BrdU incorporation using the BrdU Cell Proliferation ELISA kit (Abcam, cat# ab126556) according to the manufacturer's instructions. N=6 biological replicates were used per condition. Statistical significance was measured by a two-tailed Student's P value.

Prefoldin complex lethality in murine embryos

Each subunit of the prefoldin and prefoldin-like complex was queried for mouse embryonic phenotypes using the information provided by the International Mouse Phenotyping Consortium.[94,95] Data were downloaded from https://www.mousephenotype.org and phenotypes observed in the homozygous knockout setting are reported.

Comparing CRISPR data with Project Achilles

The ASNSD1 gene was evaluated for cell line phenotypes using the DepMap_public_19Q4 release of CRISPR DepMap data and the Achilles RNA interference screens using the file "Achilles_logfold_change" (available at https://depmap.org/portal/download). Knockout phenotypes for 313 cell line assessed by both CRISPR and RNAi were z-scored and compared to each other.

QUANTIFICATION AND STATISTICAL ANALYSIS

Experimental schematics in the Graphical Abstract and Figures 1A, 4D, 4G, 4K, 5A, and 5I were generated using BioRender.com. Figures were generated and statistical analyses were performed using custom scripts in R v4.3.0 and GraphPad PRISM v10.0.2. A detailed list of software used for data processing, quantification and analysis is stated in the respective STAR Methods sections and the key resource table. Statistical parameters such as the value of n, mean/median, standard deviation and significance level (including the statistical tests used) are reported in the STAR Methods, figures, and/or figure legends. All data are expressed as means ± standard deviation. All experimental assays were performed in duplicate or triplicate. Unless stated otherwise, a p value <0.05 was considered statistically significant. Values of 'n' represent numbers of human or cell line samples (experimental model and study participant details; Figures 1B, 1C, 3A, 5A, S1C, S4A, S4O), numbers of ORFs (translation of non-canonical ORFs is common in medulloblastoma; Figures 1I, 2A, 2G, S2A, S2B), numbers of proteins (Figures 5B, 5C, 5E) and numbers of genes (Figure S5C). Statistical parameters used to indicate differential expression were derived from DESeq2 (STAR Methods). Type of statistical test used (Student's T-test, ANOVA, Kolmogorov-Smirnov test, log-rank P value, or other as indicated) is indicated in the figure legend and in the STAR Methods segment specific to each analysis. Data normality was tested using the Shapiro-Wilk test. For non-normally distributed data, the Wilcoxon/Mann-Whitney U test was used. For RNA-seq and ribo-seq analysis of tissue samples, autopsy samples were excluded from analysis due to low RNA quality. Tissue samples with low tumor purity were also excluded (tissue sample RNA-seq and Ribo-seq analyses). For the primary CRISPR screen, intORFs, doORFs or uoORFs with > 25% overlap with the main CDS, and ORFs shorter than 12 amino acids, were excluded.