

**Supplemental information**

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

**Damon A. Hofman, Jorge Ruiz-Orera, Ian Yannuzzi, Rakesh Murugesan, Adam Brown, Karl R. Clauser, Alexandra L. Condurat, Jip T. van Dinter, Sem A.G. Engels, Amy Goodale, Jasper van der Lugt, Tanaz Abid, Li Wang, Kevin N. Zhou, Jayne Vogelzang, Keith L. Ligon, Timothy N. Phoenix, Jennifer A. Roth, David E. Root, Norbert Hubner, Todd R. Golub, Pratiti Bandopadhyay, Sebastiaan van Heesch, and John R. Prensner**

## Supplemental Information

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

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

#### Author affiliations

<sup>1</sup>Princess Máxima Center for Pediatric Oncology, Heidelberglaan 25, 3584 CS, Utrecht, the Netherlands

<sup>2</sup>Cardiovascular and Metabolic Sciences, Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC), 13125 Berlin, Germany

<sup>3</sup>Broad Institute of MIT and Harvard, Cambridge, MA, 02142, USA

<sup>4</sup>Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, MA 02215, USA

<sup>5</sup>Department of Pathology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, 02215, USA

<sup>6</sup>Department of Pathology, Brigham and Women's Hospital, Boston, MA, 02215, USA

<sup>7</sup>Department of Pathology, Boston Children's Hospital, Boston MA 02115

<sup>8</sup>Division of Pharmaceutical Sciences, James L. Winkle College of Pharmacy, University of Cincinnati, Cincinnati, OH, 45229, USA

<sup>9</sup>Charité-Universitätsmedizin, 10117 Berlin, Germany.

<sup>10</sup>German Centre for Cardiovascular Research, Partner Site Berlin, 13347 Berlin, Germany.

<sup>11</sup>Division of Pediatric Hematology/Oncology, Boston Children's Hospital, Boston, MA, 02115, USA

<sup>12</sup>Departments of Pediatrics, Division of Pediatric Hematology/Oncology, and Biological Chemistry, University of Michigan Medical School, Ann Arbor, MI 48109, USA

<sup>13</sup>Current address: Arbor Biotechnologies, Cambridge, MA, 02140, USA

<sup>14</sup>Current address: Kaiser Permanente Bernard J. Tyson School of Medicine, Pasadena, CA, 91101, USA

<sup>15</sup>These authors contributed equally.

<sup>16</sup>Lead contact

\*Correspondence: [s.vanheesch@prinsesmaximacentrum.nl](mailto:s.vanheesch@prinsesmaximacentrum.nl) (S.v.H.), [prensner@umich.edu](mailto:prensner@umich.edu) (J.R.P.)

## **Table of Contents**

### **Supplementary Figures**

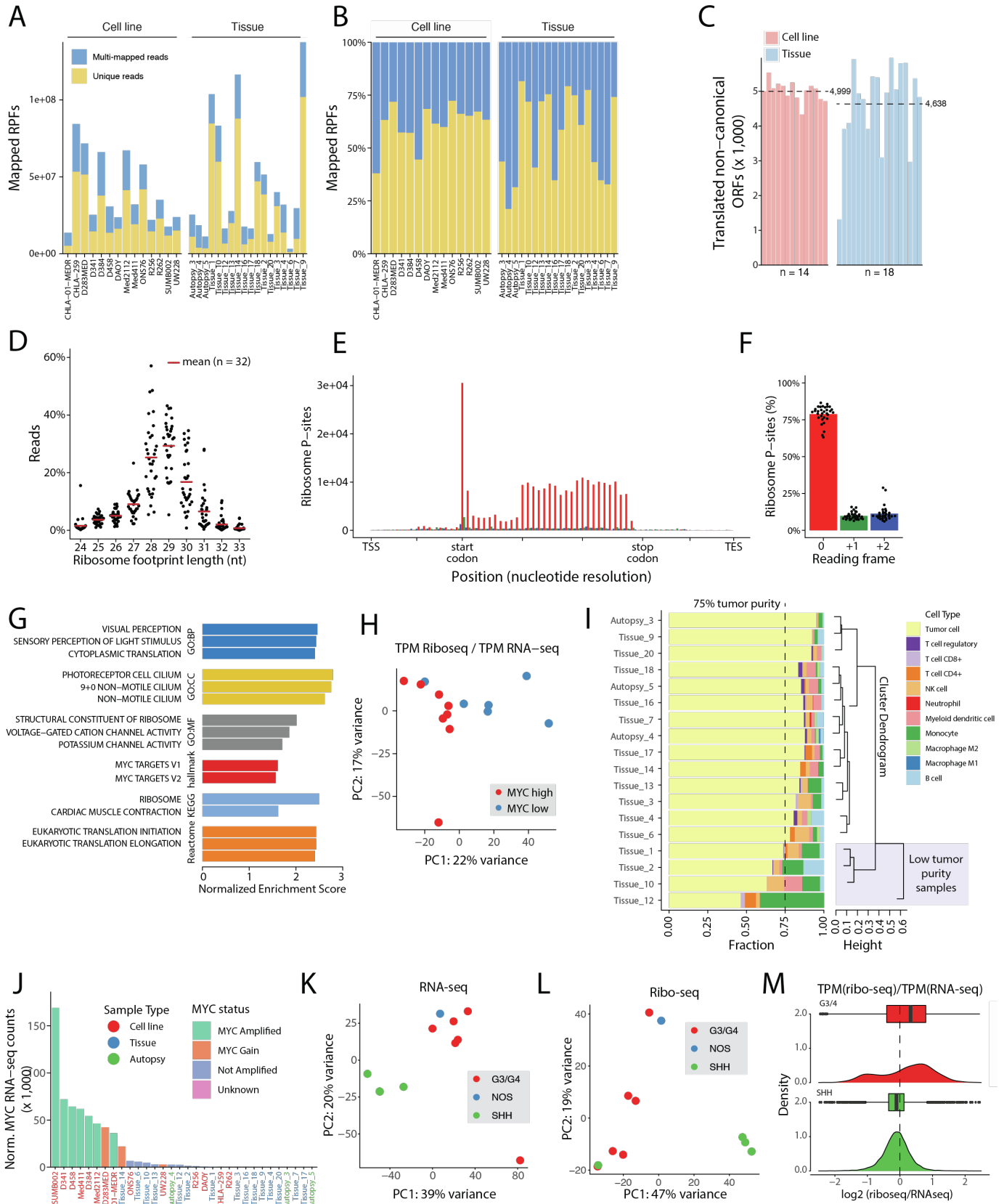
Figure S1: Profiling RNA translation in medulloblastoma, related to Figure 1.

Figure S2: Genomic perturbation of non-canonical ORFs in medulloblastoma to reveal ORF dependencies, related to Figure 2.

Figure S3: Characterization of ASNSD1-uORF as a genetic dependency in medulloblastoma, related to Figure 3.

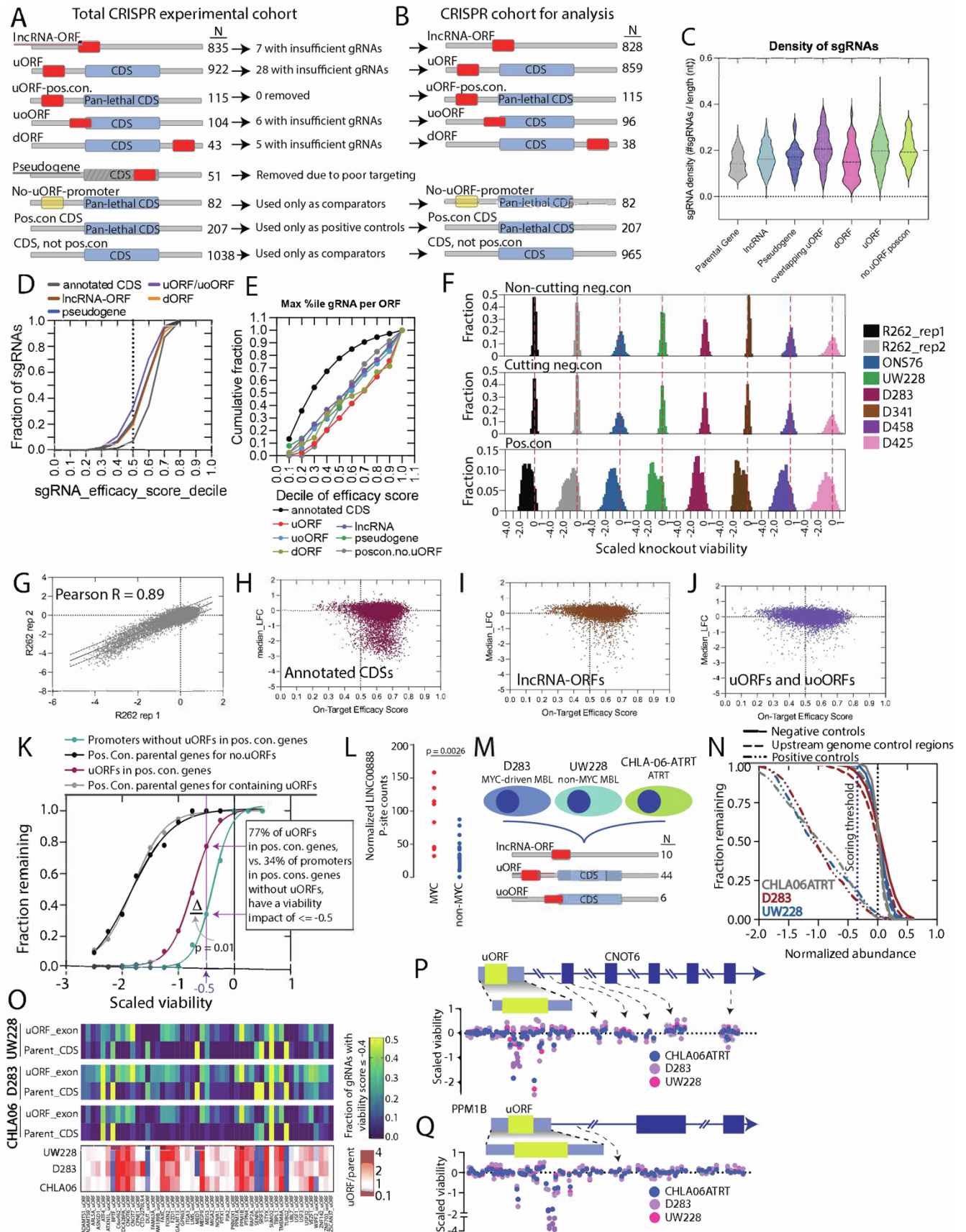
Figure S4: Association of ASNSD1-uORF to the prefoldin-like complex in medulloblastoma, related to Figure 4.

Figure S5: Signatures of RNA-protein discordance in ASNSD1-uORF high medulloblastoma, related to Figure 5.



**Figure S1: Profiling RNA translation in medulloblastoma, related to Figure 1.**

- A) The absolute number of mapped ribosome protected footprints across each sample. Unique reads are indicated in yellow, and multi-mapping reads are indicated in blue.
- B) The relative proportion of unique and multi-mapping reads for each sample.
- C) The number of translated non-canonical ORFs (defined as P-sites per million > 1) across all samples. Dashed line indicates the mean number of translated non-canonical ORFs across all samples.
- D) Ribosome footprint sizes isolated for ribosome profiling. Each dot reflects a sample. The X axis shows the footprint size in nucleotides and the Y axis indicates the percentage of reads for each sample.
- E) A summarized plot of all ribosome profiling data, showing the in-frame P site periodicity across annotated protein-coding sequences.
- F) The average percentage of in-frame P-site reads, indicating >75% periodicity cumulatively for the dataset.
- G) Biological signatures enriched in genes that demonstrate differential translational efficiency in MYC-driven vs. non-MYC-driven medulloblastoma cell lines.
- H) A principal component analysis of the translational efficiency of non-canonical ORFs based on MYC-driven or non-MYC-driven status of medulloblastoma cell lines.
- I) Estimated immune cell fractions per sample, inferred from RNA-seq data using immune cell type deconvolution. Colors indicate different types of immune cells. 'Uncharacterized cell' indicates non-immune, or tumor cell fraction.
- J) Normalized MYC RNA-seq counts across all cell line and tumor samples with matching ribo-seq data, sorted from high to low expression. Bar color indicates MYC status of the corresponding sample; sample name color indicates sample type (Cell line, Tissue, or Autopsy).
- K) A principal component analysis (PCA) based on RNA-seq data of tissue samples. Each dot indicates one sample. Dot color denotes the molecular subgroup of the corresponding sample.
- L) A principal component analysis (PCA) based on ribo-seq data of tissue samples. Each dot indicates one sample. Dot color denotes molecular subgroup of the corresponding sample.
- M) A density plot showing the distribution of translational efficiency values for each canonical CDS in tissue samples, split by molecular subgroup. Boxplots show lower quartile, median, and upper quartile values, with whiskers extending to highest and lowest observations.



**Figure S2: Genomic perturbation of non-canonical ORFs in medulloblastoma to reveal ORF dependencies, related to Figure 2.**

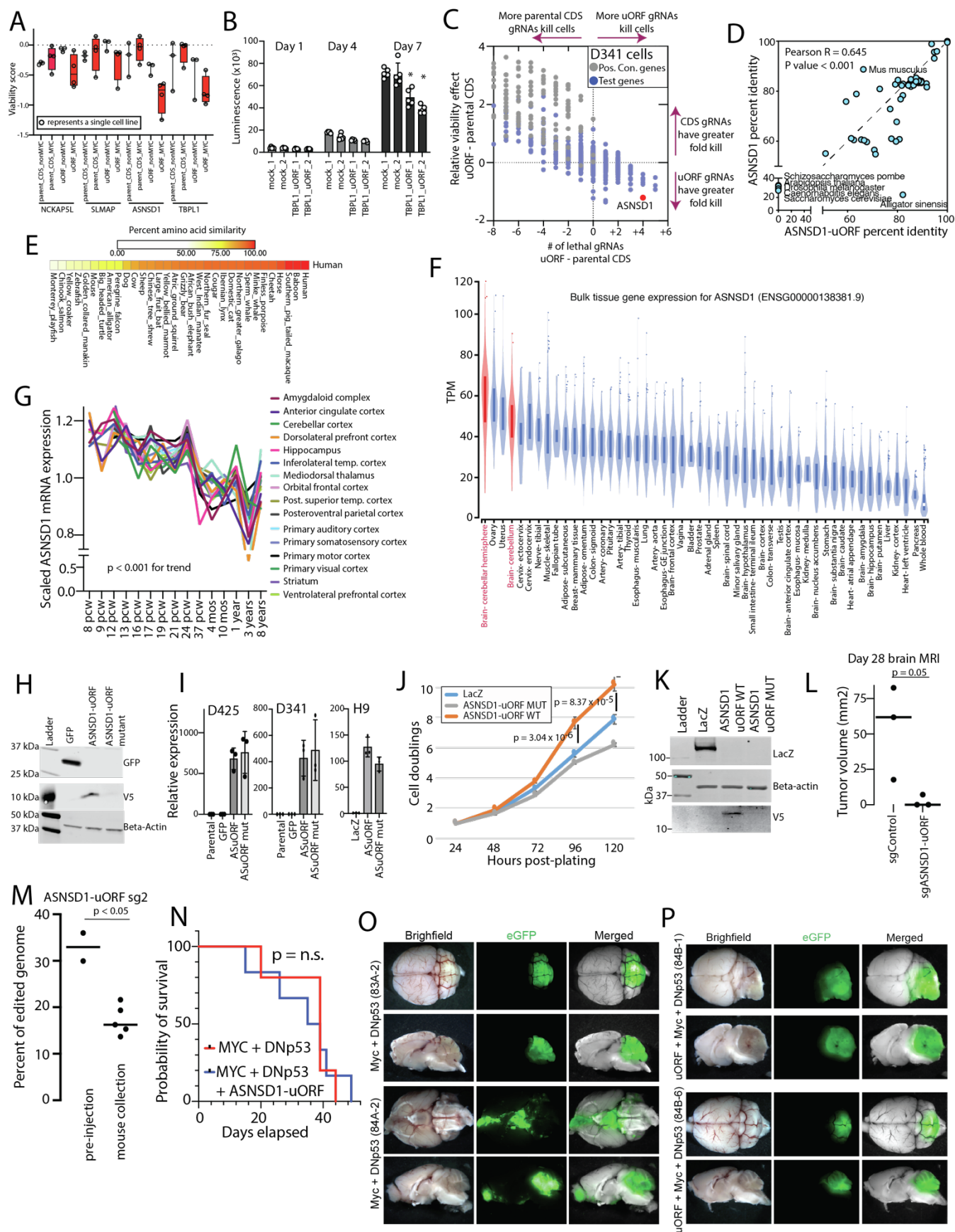
- A) A schematic detailing all ORFs targeted by the custom CRISPR library and the number of cases which were removed from analysis.
- B) A schematic showing the final cohort of ORFs analyzed by CRISPR screening.
- C) A violin plot showing the density of gRNA design per length of ORF.
- D) The fraction of gRNAs that achieve each decile of gRNA efficacy score. Results are displayed for each of the indicated ORF categories.
- E) The cumulative fraction of gRNAs targeting a given ORF biotype compared to the decile of the efficacy score for the gRNA with the least favorable characteristics.
- F) A histogram showing the scaled knockout viability effect for gRNAs targeting positive control CDSs (n = 1,654 gRNAs) compared to non-cutting gRNA controls (n = 503 gRNAs) or genome cutting gRNA controls (n = 497 gRNAs). Each cell line is shown in the indicated color.
- G) The correlation of gRNA knockout viability phenotypes for the two replicates of R262. A Pearson correlation is shown.
- H) A scatter plot for annotated CDSs, showing the gRNA on-target efficacy score compared to the median log<sub>2</sub> fold change in gRNA representation at the Day 14 timepoint for all cell lines.
- I) A scatter plot for lncRNA-ORFs, showing the gRNA on-target efficacy score compared to the median log<sub>2</sub> fold change in gRNA representation at the Day 14 timepoint for all cell lines.
- J) A scatter plot for uORFs and uoORFs, showing the gRNA on-target efficacy score compared to the median log<sub>2</sub> fold change in gRNA representation at the Day 14 timepoint for all cell lines.
- K) An analysis of gRNAs targeting promoters of positive control genes without uORFs compared to gRNAs targeting uORFs found in positive control genes. The X axis is the scaled viability for the gRNAs and the Y axis is the fraction of gRNAs achieving that viability threshold. P value by a Kolmogorov-Smirnov test.
- L) The abundance of P-site counts for an ORF in *LINC00888* across the medulloblastoma dataset. P value by a two-tailed Student's t-test.
- M) A schematic representation of the cell lines and ORF types targeted in the secondary CRISPR screen for gRNA saturation.
- N) Verification of controls in the secondary tiling CRISPR screen, showing the fraction of positive control or negative control gRNAs (on the Y axis) achieving the indicated scaled viability threshold on the X axis.
- O) *Top*, a heatmap for each of the 3 cell lines tested in the secondary tiling screen showing the fraction of gRNAs with a viability score of  $\leq -0.4$  for each pair of a parental CDS and the matched uORF or uoORF. *Bottom*, a heatmap showing the fold change in fraction of gRNAs with a viability score of  $\leq -0.4$  for each of the



three cell lines, calculated as (Fraction of uORF gRNAs with a viability score  $\leq -0.4$ ) / (Fraction of CDS gRNAs with a viability score  $\leq -0.4$ ).

- P) A graphical representation of the tiling CRISPR screen data for *CNOT6*. Each dot reflects a gRNA. gRNAs are colored according to each of the three cell lines. The Y axis reflected scaled viability for each gRNA knockout. The X axis reflects genomic position of the gRNA relative to the shown gene structure.
- Q) A graphical representation of the tiling CRISPR screen data for *PPM1B*. Each dot reflects a gRNA. gRNAs are colored according to each of the three cell lines. The Y axis reflected scaled viability for each gRNA knockout. The X axis reflects genomic position of the gRNA relative to the shown gene structure.





**Figure S3: Characterization of ASNSD1-uORF as a genetic dependency in medulloblastoma, related to Figure 3.**

- A) A boxplot showing the loss of viability associated with knockout of NCKAP5L, SLMAP, ASNSD1, and TBPL1 parental and uORF genes in MYC-driven and non-MYC-driven cell lines. Each dot reflects a different cell line. Boxplots show lower quartile, median, and upper quartile viability scores per condition.
- B) D425 cells were electroporated with gRNAs targeting the TBPL1 uORF start codon or two genomic controls along with the adenine base editor ABE8e-NRCH. Cell viability was measured at the indicated timepoints. Each dot represents a biological replicate (N=5 per condition). An asterisk indicates a P value of < 0.001. Statistical significance is by a two-tailed T test comparing against mock\_1. Error bars represent standard deviation.
- C) A scatter plot comparing the magnitude of viability phenotype of uORF knock-out relative to parental CDS knock-out in D341 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 4 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.
- D) A heat map showing percent amino acid similarity between human ASNSD1-uORF and the amino acid sequences of its homolog in the indicated species.
- E) A scatter plot showing the percent amino acid similarity of homologs to ASNSD1-uORF to the human sequence (X axis) compared to the percent amino acid similarity of homologs of ASNSD1 to the human sequence (Y axis). Several species are highlighted if strongly discordant between the two proteins.
- F) *ASNSD1* mRNA expression levels in the GTEx consortium. Cerebellar tissue is highlighted in red. Bulk tissue gene expression for ENSG00000138381.9 is shown.
- G) Normalized gene expression for *ASNSD1* mRNA across human brain development. Data were obtained for *ASNSD1* mRNA (ENSG00000138381.9) from the Allen Institute Developing Brain Atlas. P value by a two-sided ANOVA test. pcw, post-conception week; mos, months.
- H) Western blot analysis of overexpression of V5-tagged ASNSD1-uORF in D341 cells.
- I) qPCR expression values for ASNSD1-uORF plasmid constructs expressed in D425, D341 and H9 cells. Dots reflect technical replicates of qPCR. Error bars represent standard deviation.
- J) The impact of ectopic expression of ASNSD1-uORF on cell growth in H9 neural stem cells. N=6 biological replicates per data point. P values by a two-tailed Student's T test.
- K) Western blot analysis of overexpression of V5-tagged ASNSD1-uORF in H9 cells.

- L) Orthotopic xenograft tumor volume for D458 medulloblastoma cells on Day 22 after cerebellar injection. Tumor volume determined by MRI. P value by a two-tailed Student's T-test.
- M) Assessment of *ASNSD1-uORF* knockout efficiency and persistence of knockout in D458 murine xenograft experiments. Knockout efficiency was determined by CRISPR sequencing of the gRNA cut site. P value by a Student's T-test.
- N) Kaplan-Meier survival curves for in utero electroporation experiments testing mouse survival and medulloblastoma formation with cerebellar injection of cDNAs encoding MYC with a dominant-negative p53 (DNp53), either with or without addition injection of a cDNA encoding ASNSD1-uORF. P value by a log-rank test. n.s., non-significant.
- O) Whole brain images with GFP fluorescence for two mice with MYC and DNp53 induced medulloblastomas.
- P) Whole brain images with GFP fluorescence for two mice with MYC, DNp53 and ASNSD1-uORF medulloblastomas.



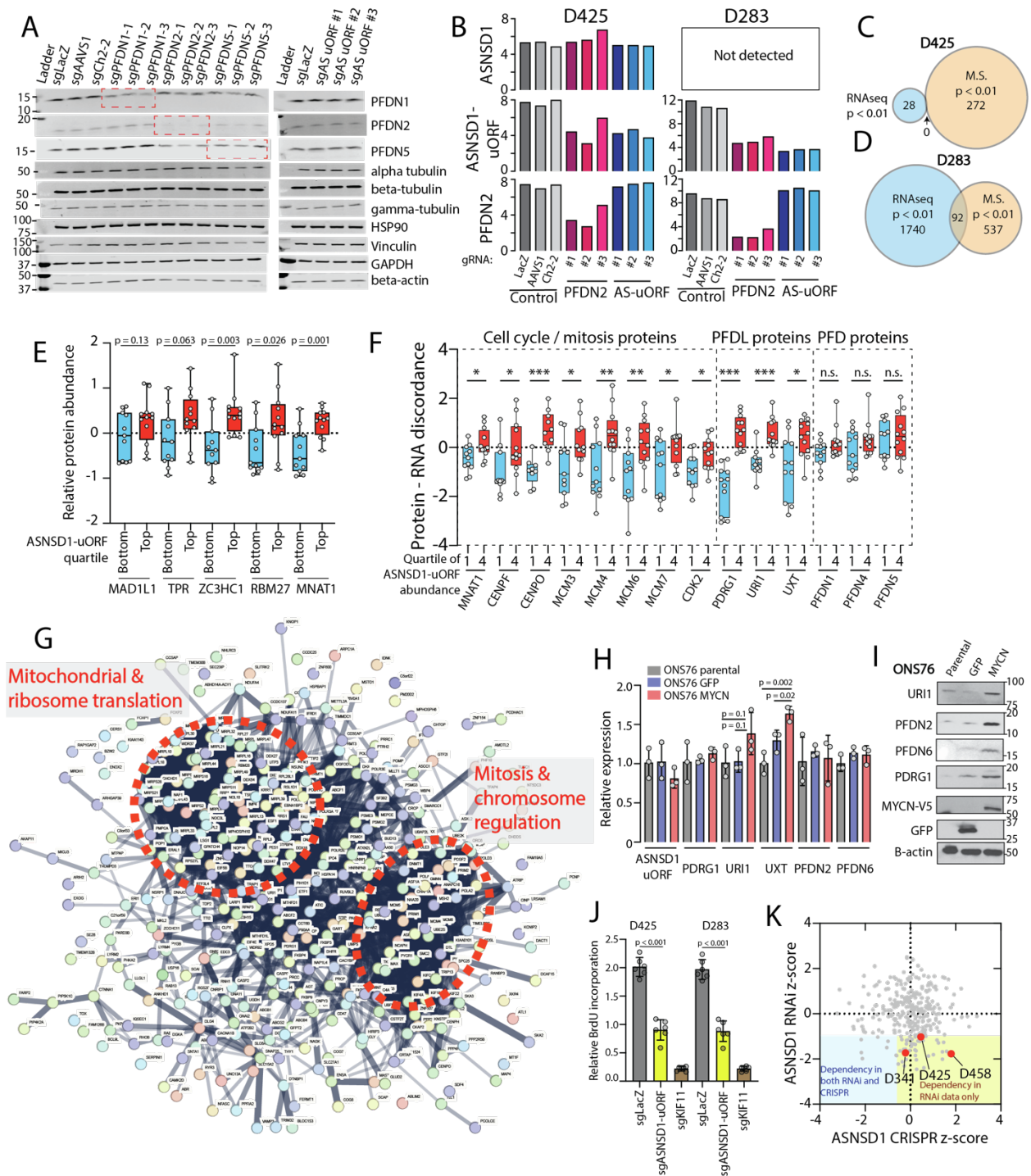
**Figure S4: Association of ASNSD1-uORF to the prefoldin-like complex in medulloblastoma, related to Figure 4.**

- A) Protein abundance of small protein controls in a cohort of MYC-driven (n=10) or non-MYC (n=6) medulloblastoma cell lines. P values by a two-tailed Student's T test.
- B) Protein abundance of ASNSD1-uORF compared MYC in Group 3 medulloblastoma tissue samples only. Each dot reflects a sample. The Pearson correlation is shown.
- C) Protein abundance of ASNSD1-uORF compared MYC in Group 4 medulloblastoma tissue samples only. Each dot reflects a sample. The Pearson correlation is shown.
- D) RNA expression of *MYC* compared to *MYCN* in the medulloblastoma tissue cohort. Data obtained from<sup>13</sup>. Red dots reflect Group 3 medulloblastoma patients. A Pearson correlation is shown.
- E) Western blot analysis of ONS76 cells ectopically expressing a *GFP*, *MYCN* or *MYCN R393H* transgene.
- F) Quantification of ASNSD1-uORF endogenous protein levels in ONS76 cells with the indicated transgene expression, as well as endogenous ASNSD1-uORF peptide levels in MYC-high and MYC-low medulloblastoma cell lines. Data represent peptide abundance of ASNSD1-uORF detected by mass spectrometry, normalized by input protein amount. P value by a two-tailed Student's T test. Error bars represent standard deviation.
- G) The correlation of ASNSD1-uORF protein abundance with *ASNSD1* mRNA levels. The Pearson correlation is shown.
- H) The correlation of *ASNSD1* mRNA level with sample MYC/MYCN protein abundance. The Pearson correlation is shown.
- I) A scatter plot showing the magnitude of viability loss after *ASNSD1-uORF* knockout (X axis) with the normalized ASNSD1-uORF protein abundance determined by mass spectrometry for 32 cell lines. Red dots are MYC-driven medulloblastoma cell lines. Blue dots are nonMYC medulloblastoma cell lines. Gray dots are non-medulloblastoma cell lines.
- J) Protein abundance of ASNSD1-uORF across the proteomics dataset for the Cancer Cell Line Encyclopedia (CCLE). The X axis shows the normalized protein abundance and the Y axis shows the fraction of samples within a given cell line lineage with the corresponding protein abundance. The red line indicates neuroblastoma cell lines, and gray lines reflect other cancer lineages. The blue line is the average of the dataset.
- K) Raw numbers of total peptides for co-immunoprecipitation experiments for endogenous PFDN6. Two replicates are shown.
- L) A table showing mouse germline knockout phenotypes for the indicated prefoldin

or prefoldin-like complex proteins.

- M) Upregulation of prefoldin-like proteins in medulloblastoma tissue samples with high MYC/MYCN levels. The Y axis shows up regulation of the indicated protein compared to the mean. Size and color of the circles indicates degree of statistical significance of upregulation. P values by an ANOVA test.
- N) The translational efficiency of prefoldin-like complex members across the set of MYC-driven and non-MYC-driven medulloblastoma cell lines.
- O) The representation of cancer cell line lineages across the 484 cell lines in the pooled knockout experiment.
- P) A scatter plot showing correlation of individual ASNSD1-uORF gRNAs used in the pooled knockout experiment. Each dot reflects a cell line. A Pearson correlation is shown for scaled viability values obtained on Day 10 after knockout.
- Q) The relative CRISPR dependency score for genetic knockout of *PFDN1*, *PFDN2*, *PFDN4*, *PFDN5*, *PFDN6*, *VBP1*, *URI1*, *UXT*, and *PDRG1* across medulloblastoma cell lines in the 21\_Q2 release of the Cancer Dependency map. Dependency score data were normalized as Z-scores across the medulloblastoma cell lines and subjected to hierarchical clustering. The *MYC* mRNA level is additionally shown.
- R) An immunoblot showing GST-pull down assays for ASNSD1-uORF with members of the PFD and PFDL complex. ASNSD1-uORF-GST was used as bait to interact with a pooled protein mixture of PFD or PFDL proteins. The sample flow-through and eluate are shown.





**Figure S5: Signatures of RNA-protein discordance in ASNSD1-uORF high medulloblastoma, related to Figure 5.**

A) Western blot analysis of proteins reported to be regulated by the prefoldin complex following knockout of prefoldin proteins or ASNSD1-uORF in D425 medulloblastoma cells.

B) Quantification of ASNSD1-uORF and PFDN2 protein knockout in D425 and D283



cells in shotgun mass spectrometry experiments. *ASNSD1* parent CDS protein levels are also shown.

- C) The overlap and number of genes identified as regulated in mass spectrometry or RNAseq following *ASNSD1-uORF* knockout in D425 cells. The p values refer to the thresholds used to determine regulated genes in either mass spectrometry or RNAseq data.
- D) The overlap and number of genes identified as regulated in mass spectrometry or RNAseq following *ASNSD1-uORF* knockout in D283 cells. The p values refer to the thresholds used to determine regulated genes in either mass spectrometry or RNAseq data.
- E) Medulloblastoma patient tissue abundance for five top proteins observed in *ASNSD1-uORF* knockout proteomics experiments. Abundance of proteins is shown in medulloblastoma patient tissues stratified into the top and bottom quartile of *ASNSD1-uORF* protein level. P value by a two-tailed Student's t-test. Boxplots show lower quartile, median, and upper quartile protein abundance per group.
- F) Medulloblastoma patient tissue protein-RNA discordance score for representative top proteins observed as significantly differentially-regulated in the proteome of *ASNSD1-uORF* high medulloblastoma samples (top quartile) compared to *ASNSD1-uORF* low medulloblastoma samples (bottom quartile). Example proteins are grouped as proteins involved in the PFDL complex, PFD complex, or cell cycle / mitosis-implicated proteins. P value by a two-tailed Student's t-test. Boxplots show lower quartile, median, and upper quartile protein abundance per group. \*p<0.05; \*\*p<0.01; \*\*\*p<0.001
- G) Interactome analysis of proteins (N=383) showing significant protein-RNA discordance in *ASNSD1-uORF* high (top quartile) compared to *ASNSD1-uORF* low (bottom quartile) medulloblastoma tissues. Interactome visualization is shown using the STRING tool ([www.string-db.com](http://www.string-db.com)). Two main clusters of protein networks are observed, relating to chromosome regulation and translation-associated genes (ribosome and mitochondrial translation).
- H) qPCR data for prefoldin-like complex members in isogenic ONS76 medulloblastoma cells expressing either GFP control or MYCN cDNA. P values represent a two-tailed T test. Error bars represent standard deviation.
- I) Western blot analysis of prefoldin-like complex members in isogenic ONS76 medulloblastoma cells expressing either GFP control or MYCN cDNA.
- J) Bromouridine (BrdU) incorporation in D283 and D425 medulloblastoma cells following knockout with a control gRNA (*LacZ*), knockout of *ASNSD1-uORF*, or knockout of *KIF11*. P values by a two-tailed Student's T-test. Error bars represent standard deviation.
- K) A scatter plot showing loss-of-viability data for *ASNSD1* in 313 cell lines tested by both RNA interference screening from Project Achilles<sup>59</sup> and CRISPR in the cancer

Dependency Map ([www.depmap.org](http://www.depmap.org)). Three medulloblastoma cell lines are highlighted.